

LaBRI



Laboratoire Bordelais
de Recherche en Informatique
UMR 5800
CNRS - Université Bordeaux 1 -
ENSEIRB
Partenaire INRIA Futurs



RAPPORT D'ACTIVITÉS SCIENTIFIQUES

2002-2004

(Équipes de Recherche)



Laboratoire Bordelais de Recherche en Informatique

351, Cours de la Libération

33405 Talence - France

Tél. : +33 (0) 5 4000 6900 - Fax : +33 (0) 5 4000 6669

URL : <http://www.labri.fr/>

SOMMAIRE

| | | |
|----------|--|-----------|
| 1 | Combinatoire et Algorithmique | 11 |
| 1.1 | Composition de l'équipe | 12 |
| 1.1.1 | Membres permanents | 12 |
| 1.1.2 | Doctorants | 12 |
| 1.1.3 | Membre non permanents | 13 |
| 1.2 | Présentation de l'équipe | 14 |
| 1.3 | Thèmes de recherche de l'équipe | 17 |
| 1.3.1 | Combinatoire énumérative et algébrique | 17 |
| | Introduction | 17 |
| | Description des activités et principaux résultats | 18 |
| | Projets et perspectives 2005-2008 | 25 |
| 1.3.2 | Graphes & Applications | 28 |
| | Introduction | 28 |
| | Mots clés | 28 |
| | Description des activités de recherche et des résultats | 29 |
| 1.3.3 | Algorithmique Distribuée | 40 |
| | Introduction | 40 |
| | Description des activités et principaux résultats | 40 |
| 1.3.4 | Modélisation | 40 |
| | Les questions étudiées et les principaux résultats obtenus | 41 |
| | Analyse d'algorithmes | 44 |
| | Projets et perspectives 2005-2008 | 45 |
| 1.4 | Publications de l'équipe | 47 |
| 1.4.1 | Publications d'audience internationale | 47 |
| 1.4.2 | Publications d'audience nationale | 61 |
| 1.4.3 | Autres publications | 62 |
| 1.4.4 | Formation par la recherche | 64 |
| 1.5 | Collaborations internationales | 66 |
| 1.5.1 | Invités | 66 |
| 1.5.2 | Invitations | 67 |
| 1.6 | Liste des contrats et subventions | 69 |
| 1.6.1 | Contrats internationaux et européens | 69 |
| 1.6.2 | Contrats avec des entreprises | 69 |
| 1.6.3 | Contrats publics avec des instances nationales | 69 |
| 1.6.4 | Contrats publics avec instances locales ou régionales | 69 |
| 1.7 | Animation de la recherche | 70 |
| 1.7.1 | Comités de rédaction | 70 |
| 1.7.2 | Comités de programmes de colloques | 70 |

| | | |
|----------|--|------------|
| 1.7.3 | Organisation de colloques, écoles jeunes chercheurs, ... | 70 |
| 1.7.4 | Administration de la Recherche | 70 |
| 2 | Logiques, Automates, Algorithmes et Applications (L3A) | 73 |
| 2.1 | Composition de l'équipe | 74 |
| 2.1.1 | Membres permanents | 74 |
| 2.1.2 | Membres non permanents | 74 |
| 2.2 | Présentation de l'équipe | 76 |
| 2.3 | Thèmes de recherche de l'équipe | 79 |
| 2.3.1 | Logiques et Graphes | 79 |
| | Méthodes algébriques | 79 |
| | Méthodes logiques | 80 |
| | Vérification | 80 |
| | Preuves mathématiques assistées | 81 |
| | Synthèse | 82 |
| | Projets et perspectives 2005-2008 | 82 |
| 2.3.2 | Langages et Automates | 84 |
| | Introduction | 84 |
| | Description des activités et principaux résultats | 84 |
| | Projets et perspectives 2005-2008 | 87 |
| 2.3.3 | Informatique et Linguistique | 89 |
| | Introduction | 89 |
| | Description des activités et principaux résultats | 90 |
| | Projets et perspectives 2005-2008 | 91 |
| 2.4 | Publications de l'équipe | 92 |
| 2.4.1 | Publications d'audience internationale | 92 |
| 2.4.2 | Publications d'audience nationale | 99 |
| 2.4.3 | Autres publications | 100 |
| 2.4.4 | Formation par la recherche | 101 |
| 2.5 | Collaborations internationales | 102 |
| 2.5.1 | Invités | 102 |
| 2.5.2 | Invitations | 102 |
| 2.6 | Liste des contrats et subventions | 104 |
| 2.6.1 | Contrats internationaux et européens | 104 |
| 2.6.2 | Contrats avec des entreprises | 104 |
| 2.6.3 | Contrats publics avec des instances nationales | 105 |
| 2.6.4 | Contrats publics avec instances locales ou régionales | 106 |
| 2.7 | Animation de la recherche | 107 |
| 2.7.1 | Comités de rédaction | 107 |
| 2.7.2 | Comités de programmes de colloques | 107 |
| 2.7.3 | Organisation de colloques, écoles jeunes chercheurs, ... | 108 |
| 2.7.4 | Administration de la Recherche | 108 |
| 3 | Modélisation, Vérification et Test des systèmes informatiques (MVTsi) | 111 |
| 3.1 | Composition de l'équipe | 112 |
| 3.1.1 | Membres permanents | 112 |
| 3.1.2 | Doctorants | 112 |
| 3.1.3 | Membres non permanents | 112 |
| 3.2 | Présentation de l'équipe | 113 |

| | | |
|-------|---|-----|
| 3.3 | Thèmes de recherche de l'équipe | 114 |
| 3.3.1 | Théorie et pratique de la vérification | 114 |
| | Introduction | 114 |
| | Description des activités et principaux résultats | 114 |
| | Projets et perspectives 2005-2008 | 122 |
| 3.3.2 | Test de systèmes informatisés | 127 |
| | Introduction | 127 |
| | Description des activités et principaux résultats | 127 |
| | Projets et perspectives 2005-2008 | 130 |
| 3.3.3 | Modélisation du vivant | 132 |
| | Introduction | 132 |
| | Description des activités et principaux résultats | 132 |
| | Projets et perspectives 2005-2008 | 133 |
| 3.4 | Publications de l'équipe | 135 |
| 3.4.1 | Publications d'audience internationale | 135 |
| 3.4.2 | Publications d'audience nationale | 139 |
| 3.4.3 | Autres publications | 140 |
| 3.4.4 | Formation par la recherche | 140 |
| 3.5 | Collaborations internationales | 142 |
| 3.5.1 | Invitations | 142 |
| 3.6 | Liste des contrats et subventions | 143 |
| 3.6.1 | Contrats internationaux et européens | 143 |
| | ESACS [Enhanced Safety Assessment for Complex Systems] | 143 |
| | ISAAC [Improvement of Safety Activities on Aeronautical Complex Systems] | 143 |
| | IDEAS [Interoperability Developments for Enterprise Application and Software] | 143 |
| | TAROT [Training And Research On Testing] | 143 |
| 3.6.2 | Contrats avec des entreprises | 144 |
| | AIRBUS - ATOS ORIGIN | 144 |
| | EUROCONTROL | 144 |
| | EUROCONTROL | 144 |
| 3.6.3 | Contrats publics avec des instances nationales | 144 |
| | Morse | 144 |
| | Persée | 144 |
| | Calife | 145 |
| | Platonis | 145 |
| | Averroes | 145 |
| | AS 23 | 145 |
| | AS 111 | 145 |
| | AS 155 | 146 |
| | ATIP | 146 |
| 3.6.4 | Contrats publics avec instances locales ou régionales | 146 |
| 3.7 | Animation de la recherche | 147 |
| 3.7.1 | Comités de rédaction | 147 |
| 3.7.2 | Comités de programmes de colloques | 147 |
| 3.7.3 | Organisation de colloques, écoles jeunes chercheurs, ... | 147 |
| 3.7.4 | Administration de la Recherche | 147 |

| | | |
|----------|--|------------|
| 4 | Image et Son | 149 |
| 4.1 | Composition de l'équipe | 150 |
| 4.1.1 | Membres permanents | 150 |
| 4.1.2 | Membres non permanents | 150 |
| 4.2 | Présentation de l'équipe | 152 |
| 4.3 | Thèmes de recherche de l'équipe | 155 |
| 4.3.1 | Analyse et Indexation Vidéo | 155 |
| | Introduction | 155 |
| | Description des activités et principaux résultats | 156 |
| | Projets et perspectives 2005-2008 | 162 |
| 4.3.2 | Modélisation du son et de la musique | 164 |
| | Introduction | 164 |
| | Description des activités et principaux résultats | 164 |
| | Projets et perspectives 2005-2008 | 168 |
| 4.3.3 | Modélisation et synthèse d'images | 172 |
| | Introduction | 172 |
| | Description des activités et principaux résultats | 172 |
| | Projets et perspectives 2005-2008 | 179 |
| 4.3.4 | Structuration et analyse d'image | 182 |
| | Introduction | 182 |
| | Description des activités et principaux résultats | 183 |
| | Projets et perspectives 2005-2008 | 188 |
| 4.4 | Publications de l'équipe | 191 |
| 4.4.1 | Publications d'audience internationale | 191 |
| 4.4.2 | Publications d'audience nationale | 203 |
| 4.4.3 | Autres publications | 204 |
| 4.4.4 | Formation par la recherche | 206 |
| 4.5 | Collaborations internationales | 207 |
| 4.5.1 | Invités | 207 |
| 4.5.2 | Invitations | 207 |
| 4.6 | Liste des contrats et subventions | 208 |
| 4.6.1 | Contrats internationaux et européens | 208 |
| 4.6.2 | Contrats avec des entreprises | 208 |
| 4.6.3 | Contrats publics avec des instances nationales | 208 |
| 4.6.4 | Contrats publics avec instances locales ou régionales | 210 |
| 4.7 | Animation de la recherche | 212 |
| 4.7.1 | Comités de rédaction | 212 |
| 4.7.2 | Comités de rédaction de colloques | 212 |
| 4.7.3 | Présidence de session de colloques | 213 |
| 4.7.4 | Organisation de colloques, écoles jeunes chercheurs, ... | 213 |
| 4.7.5 | Administration de la Recherche | 213 |
| 5 | Informatique Parallèle et Distribué (PARADIS) | 215 |
| 5.1 | Composition de l'équipe | 216 |
| 5.1.1 | Membres permanents | 216 |
| 5.1.2 | Membre non permanents | 216 |
| 5.2 | Présentation de l'équipe | 218 |
| 5.3 | Thèmes de recherche de l'équipe | 220 |
| 5.3.1 | Comet | 220 |

| | | |
|-------|--|-----|
| | Introduction | 220 |
| | Description des activités et principaux résultats | 220 |
| | Projets et perspectives 2005-2008 | 225 |
| 5.3.2 | Phoenix | 226 |
| | Introduction | 226 |
| | Description des activités et principaux résultats | 226 |
| | Projets et perspectives 2005-2008 | 230 |
| 5.3.3 | Runtime | 233 |
| | Introduction | 233 |
| | Description des activités et principaux résultats | 237 |
| | Projets et perspectives 2005-2008 | 239 |
| 5.3.4 | ScAlApplix | 242 |
| | Introduction | 242 |
| | Description des activités et principaux résultats | 242 |
| | Projets et perspectives 2005-2008 | 252 |
| 5.3.5 | Sod | 254 |
| | Introduction | 254 |
| | Description des activités et principaux résultats | 254 |
| | Projets et perspectives 2005-2008 | 255 |
| 5.4 | Publications de l'équipe | 257 |
| 5.4.1 | Publications d'audience internationale | 257 |
| 5.4.2 | Publications d'audience nationale | 265 |
| 5.4.3 | Autres publications | 267 |
| 5.4.4 | Production de logiciels | 268 |
| 5.4.5 | Formation par la recherche | 269 |
| 5.5 | Collaborations internationales | 271 |
| 5.5.1 | Invités | 271 |
| 5.5.2 | Invitations | 271 |
| 5.6 | Liste des contrats et subventions | 272 |
| 5.6.1 | Contrats internationaux et européens | 272 |
| | Amigo – Informatique ubiquitaire pour le réseau domestique | 272 |
| | Serveurs de stockage distribués sur réseaux hautes performances | 272 |
| | Préconditionneurs parallèles robustes | 272 |
| 5.6.2 | Contrats avec des entreprises | 273 |
| | Plate-forme programmable de service de communications sur Windows | 273 |
| | Service .NET Programmable | 273 |
| | Routeur parallèle tolérant aux pannes | 273 |
| | Bibliothèque de multithreading sur machines CC-NUMA | 273 |
| | Parallélisation du code CALDER d'interaction Laser/Plasma | 274 |
| | Parallélisation du code CODDEX de dynamique des dislocations | 274 |
| | Simulation de la propagation de la fissuration de la silice due au laser | 274 |
| | Parallélisation du code MIRO | 274 |
| | Chaîne logicielle EMILIO – Solveur direct hautes performances | 275 |
| | Solveur hybrides hautes performances | 275 |
| | Téléphonie et nomadique | 275 |
| 5.6.3 | Contrats publics avec des instances nationales | 275 |
| | SWAN – Self-aWare mAnagement | 275 |
| | Composition et raffinement de systèmes sûrs | 275 |
| | Noyau d'infrastructure répartie adaptable | 276 |

| | |
|---|------------|
| RMI | 276 |
| Protocoles hautes performances, Stockage à grande échelle, Indexation distribuée | 276 |
| Méthodologie de programmation des grilles, Action d'animation | 277 |
| ALTA : Asynchronous Loss Tolerant Algorithms | 277 |
| EPSN – Environnement pour le Pilotage de Simulations Numériques distribuées | 277 |
| RedGrid – Redistribution de données sur la grille | 277 |
| TLSE – Test pour grands systèmes d'équations | 277 |
| GRID2 – Action d'animation scientifique | 278 |
| Simulation informatique des systèmes enzymatiques | 278 |
| Sécurité Java Card | 278 |
| 5.6.4 Contrats publics avec instances locales ou régionales | 279 |
| Modélisation, conception et spécification d'algorithmes et de systèmes distribués | 279 |
| Pôle de Recherche en Informatique | 279 |
| Plate-forme pour le développement d'applications multimédia fiables sur terminaux communicants | 279 |
| Développement de services de téléphonie robustes | 280 |
| Composants systèmes à haute performance | 280 |
| GRINTA – Grappes et grilles de recherche pour l'interactivité et les applications scientifiques | 280 |
| Mise en œuvre et sécurisation d'un environnement distribué pour la modélisation et la simulation de structures en biologie cellulaire | 280 |
| 5.7 Animation de la recherche | 281 |
| 5.7.1 Comités de rédaction | 281 |
| 5.7.2 Comités de programmes de colloques | 281 |
| 5.7.3 Organisation de colloques, écoles jeunes chercheurs, ... | 282 |
| 5.7.4 Administration de la Recherche | 282 |
| 6 Projets transversaux | 283 |
| 6.1 Projet VISIDIA : Visualisation de structures complexes évolutives et d'algorithmes distribués sur ces structures | 283 |
| 6.1.1 Introduction | 283 |
| 6.1.2 Description des activités et principaux résultats | 284 |
| 6.1.3 Projets et perspectives 2005-2008 | 288 |
| 6.1.4 Publications d'audience internationale | 289 |
| 6.1.5 Publications d'audience nationale | 291 |
| 6.1.6 Autres publications | 291 |
| 6.1.7 Formation par la recherche | 292 |
| 6.2 Algorithmes et Méthodes pour la Bioinformatique | 292 |
| 6.2.1 Introduction | 292 |
| 6.2.2 Description des activités et principaux résultats | 293 |
| 6.2.3 Projets et perspectives 2005-2008 | 298 |
| 6.2.4 Publications d'audience internationale | 300 |
| 6.2.5 Publications d'audience nationale | 300 |
| 6.2.6 Autres publications | 300 |
| 6.2.7 Formation par la recherche | 302 |
| 6.3 Visualisation de Structures Combinatoires | 302 |

| | | |
|-------|---|-----|
| 6.3.1 | Introduction | 302 |
| 6.3.2 | Description des activités et principaux résultats | 302 |
| 6.3.3 | Projets et perspectives 2005-2008 | 308 |
| 6.3.4 | Publications d'audience internationale | 310 |
| 6.3.5 | Publications d'audience nationale | 311 |
| 6.3.6 | Autres publications | 311 |
| 6.3.7 | Formation par la recherche | 313 |

Combinatoire et Algorithmique

Responsable(s) Y. Métivier

Combinatoire énumérative et algébrique

Responsable(s) M. Bousquet-Mélou

Mots-clés : énumération, génération aléatoire, série génératrice, bijection, carte, groupe, fonction symétrique et quasi-symétrique

Graphes et applications

Responsable(s) A. Raspaud

Mots-clés : Routage dans les réseaux, structures de données compactes, hamiltonisme, coloration, indice de transmission

Algorithmique distribuée

Responsable(s) Y. Métivier

Mots-clés : calculs locaux, élection, détection de terminaison, reconnaissance, fibration, revêtement, algorithme probabiliste, système à agents mobiles

1.1 Composition de l'équipe

1.1.1 Membres permanents

AMAR, Denise, MdC, Bx1
AUBER, David, MdC, Bx1, depuis le 1er janvier 2004,
AVAL, Jean-Christophe, CR, CNRS, depuis le 1er octobre 2004,
BAUDON, Olivier, MdC, Bx1
BÉTRÉMA, Jean, MdC, Bx1
BIANCHI, Giuliana, MdC, Bx1
BONICHON, Nicolas, MdC, Bx1, depuis le 1er octobre 2004,
BOUSQUET-MÉLOU, Mireille, Dr, CNRS
CORI, Robert, Pr, Bx1
DELEST, Maylis, Pr, Bx1
DELMAS, Olivier, MdC, Bx1
DUCHON, Philippe, MdC, ENSEIRB
DULUCQ, Serge, Pr, Bx1
DUTOUR, Isabelle, MdC, Bx1
FERRARO, Pascal, MdC, Bx1
GAVOILLE, Cyril, Pr, Bx1
GUIBERT, Olivier, MdC, Bx1
HANUSSE, Nicolas, CR, CNRS
LAFON, Pierre, MdC, Bx1
LALANNE, Jean-Claude, MdC, Bx1
LOVEJOY, Jérémy, CR, CNRS
MARCUS, Michel, PRAG, Bx1
MÉTIVIER, Yves, Pr, ENSEIRB
NARBEL, Philippe, MdC, Bx1
PÉCHER, Arnaud, MdC, Bx1
PENAUD, Jean-Guy, Pr, Bx1
RASPAUD, André, Pr, Bx1
ROBSON, Mike, Pr, Bx1
SAHEB-DJAHROMI, Nasser, MdC, Bx1
SOPENA, Éric, Pr, Bx1
VIENNOT, Xavier, Dr, CNRS
ZEMMARI, Akka, MdC, Bx4
ZVONKINE, Alexandre, Pr, Bx1.

1.1.2 Doctorants

ANDREN, Daniel, doct(Europe-ACE)
AUILANS, Pascal, doct(CIFRE), depuis le 1er octobre 2000, jusqu'au 28 février 2005,
BARRIOT, Roland, doct(MRT), depuis le 1er octobre 2001
BAZZARO, Fabrice, doct(MRT), depuis le 1er octobre 2001, jusqu'au 4 février 2005,
BERNARDI, Olivier, doct(AC), depuis le 1er octobre 2003
CHALOPIN, Jérémie, doct(AC), depuis le 1er octobre 2003
CHEVALIER, Fany, doct(MRT), depuis le 1er octobre 2004
COULONGES, Sylvain, doct(MRT), depuis le 1er octobre 2004

DERBEL, Bilel, doct(MRT), depuis le 1er octobre 2002
 DON, Anthony, doct(MRT), depuis le 1er octobre 2003
 DOURISBOURE, Yon, doct(MRT), depuis le 1er octobre 2001, jusqu'au 15 décembre 2003,
 EL HIBAOU, Abdelaziz, doct(financement propre), depuis le 1er octobre 2002
 FERCHAUD, Frédéric, doct(contrat), depuis le 1er octobre 2003
 GONCALVES, Daniel, doct(AC), depuis le 1er octobre 2002
 HERRBACH, Claire, doct(Région), depuis le 1er octobre 2003
 HOSSEINI, Mohammad, doct(sfere), depuis le 1er octobre 2001, jusqu'au 4 février 2005,
 JAMES, William, doct(Australie), depuis le 1er octobre 2004
 LABOUREL, Arnaud, doct(ACI), depuis le 1er octobre 2004
 LE BORGNE, Yvan, doct(AC), depuis le 1er octobre 2000, jusqu'au 13 décembre 2004,
 LESUR, Isabelle, doct(Caltech,USA), depuis le 1er octobre 2000
 MONTASSIER, Mickael, doct(MRT), depuis le 1er octobre 2002
 OCHEM, Pascal, (doct), depuis le 1er octobre 2003
 OSSAMY, Rodrigue, doct(gouv-Gabon), depuis le 1er octobre 2002
 OUANGRAOUA, Aïda, doct(ACI), depuis le 1er octobre 2004
 OUCHTERLONY, Erik, doct(Europe-ACE), depuis le 1er février 2005, jusqu'au 30 juin 2005
 PELAT-ALLOIN, Sylvain, doct(MRT), depuis le 1er octobre 2002
 PINLOU, Alexandre, doct(MRT), depuis le 1er octobre 2003
 PONETI, Madalena, doct(co-tutelle), depuis le 1er octobre 2003
 SALVADORI, Nicolas, doct(co-tutelle), depuis le 1er octobre 2001
 TICHIT, Laurent, doct(MRT), depuis le 1er octobre 1999, jusqu'au 19 septembre 2003,
 TOLI, Ilia, doct(Europe-ACE), depuis le 1er mars 2004, jusqu'au 30 juin 2004
 VANDENBOGAERT, Mathias, doct(MRT), depuis le 1er octobre 2000, jusqu'au 4 mars 2004,
 WILL, James, doct(Australie), depuis le 1er octobre 2002

1.1.3 Membre non permanents

DUKES, Mark, post-doct(Europe-ACE)
 EISENKÖLBL, Theresia, post-doct(Europe-ACE)
 MISHNA, Marni, post-doc(Canada-CRSNG)

1.2 Présentation de l'équipe

L'informatique définit, étudie et manipule des structures finies ou discrètes à travers des algorithmes. Un cadre formel naturel et important pour abriter cette activité est la combinatoire. Ainsi la recherche conduite au sein de cette équipe mêle étroitement l'informatique et la combinatoire : beaucoup d'algorithmes et les analyses ou les preuves associées utilisent des objets, des constructions et des analyses combinatoires et en retour sont à l'origine de problèmes combinatoires nouveaux.

Parmi les objets ou les structures les plus utilisées, on peut citer les mots, les arbres, les graphes, les cartes, les chemins, les partitions, les permutations et les séries génératrices.

Les exemples classiques d'opérations ou de manipulations définies ou étudiées sont : les bijections, les codages, l'énumération, la génération aléatoire, les dessins (la visualisation), le coloriage et plus généralement le réétiquetage.

La recherche d'algorithmes, c'est à dire de procédures effectives permettant de répondre à des questions ou bien de calculer la valeur d'une fonction en un nombre fini d'étapes, est particulièrement active autour des thèmes suivants :

- la génération aléatoire qui a pour objectif d'engendrer des objets combinatoires non pas de façon exhaustive mais selon une loi de probabilité donnée,
- la manipulation et la visualisation efficace de gros graphes issus par exemple d'internet ou de la biologie,
- le calcul formel,
- le traitement algorithmique (modélisation et comparaison) des séquences et des structures macromoléculaires,
- les problèmes de communication dans les réseaux d'interconnexion,
- la modélisation et l'étude de problèmes issus de l'algorithmique distribuée,
- les algorithmes probabilistes.

Les questions combinatoires suscitées par l'algorithmique peuvent être de nature différente :

- énumération et comptage d'objets permettant en particulier d'effectuer des analyses exactes et en moyennes d'algorithmes,
- représentation et manipulation efficaces d'objets mathématiques variés (graphes, noeuds, groupes, séries formelles...),
- calculs explicites sur ces objets,
- mise à jour de correspondances entre ces objets à travers des bijections,
- études purement combinatoires des propriétés de certains objets ou familles d'objets comme les mots, les arbres, les tas de sables ou bien encore les graphes.

De fait, l'activité de notre équipe se structure autour des thèmes et des projets suivants :

- Combinatoire énumérative et algébrique,
- Graphes et applications,
- Algorithmique distribuée,
- Informatique et génome,
- Visualisation de l'information.

La **Combinatoire énumérative et algébrique** a pour activité première l'énumération d'objets discrets, qui peuvent avoir diverses origines : l'algorithmique, la théorie des langages, diverses branches des mathématiques (probabilités, algèbre...) ou même la physique statistique. Les méthodes combinent des approches récursives et bijectives, l'utilisation de séries formelles, et différents théorèmes "de structure" applicables à des gammes de problèmes. Ces mêmes approches permettent souvent de construire des algorithmes efficaces pour la génération aléatoire d'objets de grande taille. Le champ varié des applications de

l'énumération explique que ce thème entretienne des relations étroites avec les autres thèmes de l'équipe. Par exemple, il "partage" une ACI Masses de données avec le thème Graphes et applications. La combinatoire algébrique est une autre facette de nos activités : un travail est effectué d'une part sur la combinatoire de structures algébriques classiques (groupes, algèbres, polynômes...) et d'autre part sur les liens entre groupes et cartes.

La recherche du thème **Graphes et applications** porte d'une part sur la théorie des graphes en tant que telle (et ses conjectures classiques), et d'autre part sur des problèmes issus des communications dans les réseaux qui sont modélisés par des graphes. En théorie des graphes nous travaillons sur les problèmes de coloration. Nous développons de nouveaux concepts qui débouchent sur un vaste champ d'investigation. C'est ici que l'on peut souligner l'importance de la notion de coloration orientée. Notre recherche concerne aussi la cyclabilité des graphes et les graphes circulaires parfaits. Dans le cadre des problèmes à motivation plus appliquée, nous travaillons sur la conception de réseaux tolérants aux pannes et étudions les réseaux de grande taille (les réseaux petits-mondes, le graphe du Web,...); l'aspect algorithmique est ici une priorité. Un des membres de ce groupe de recherche a été responsable de l'Action Spécifique "Algorithmique des Grands Graphes". Dans ce type de problématique une question importante est le codage des graphes. Il faut par exemple proposer des schémas fournissant des adresses les plus courtes possibles, un décodage le plus efficace et un calcul des adresses le plus rapide possible. Ce codage doit permettre de calculer des paramètres comme par exemple la distance entre deux sommets. Dans le même ordre d'idée nous étudions le routage dans les réseaux de grande taille, il est alors important d'avoir ici des tables de routages de taille réduite pour le passage à l'échelle. Ce travail se fait en partie dans le cadre d'une ACI Masse de données consacrée à l'étude des réseaux Pair à Pair. Nous nous intéressons aussi à la visualisation d'informations au moyen de graphes. C'est une problématique majeure actuellement. Nous avons travaillé sur les réseaux sémantiques et plus particulièrement sur le fractionnement des graphes. Les algorithmes de décomposition et de dessins de graphe sont donc de première importance.

Le thème **Algorithmique distribuée** se propose de fournir des outils permettant de coder, d'étudier et d'enseigner différents aspects de l'algorithmique distribuée. Un système distribué peut être codé par un graphe étiqueté : les sommets correspondent aux processus (processeurs), les arêtes aux liens de communication ou d'interaction et les étiquettes associées codent l'état du processus ou du lien. Une règle de calcul s'exprime alors sous la forme d'une modification locale des étiquettes en fonction de valeurs locales : calcul interne à un processus, émission d'un message, réception d'un message, lecture, écriture, calcul local. Un axe important que l'on développe dans ce thème est l'étude de la frontière entre l'existence et la non existence d'algorithmes distribués permettant de résoudre certains problèmes. La non existence d'algorithmes distribués déterministes pour résoudre certains problèmes nous conduit à proposer et à étudier des algorithmes distribués probabilistes dont l'expression est souvent simple mais l'analyse difficile. Un autre aspect considéré est l'étude des liens entre les trois modèles : calculs locaux, communication par messages et communication par mémoires et la comparaison de leurs puissances.

Des membres de l'équipe sont largement présents dans les **projets transversaux** autour de la **visualisation**, de la **bio-informatique** et de l'**algorithmique distribuée**.

En guise de conclusion de cette présentation, on peut noter que les problèmes posés aux informaticiens et par les informaticiens s'enrichissent régulièrement : par exemple les problèmes s'enracinant dans la biologie ou bien les problèmes issus de l'algorithmique dans les réseaux. L'équipe combinatoire et algorithmique essaye à la fois de développer ses axes fondamentaux et, à partir de son expertise et de ses outils, de contribuer aux développements de certains thèmes émergents. C'est ainsi que la composante combinatoire algébrique a été

renforcée, que les activités bio-informatique, visualisation de gros graphes et algorithmique des réseaux sont apparues dans le paysage du LaBRI. Le berceau commun de l'ensemble des thèmes et projets de cette équipe explique une grande unité des outils et méthodes mis en oeuvre. Ces activités n'ont pas de frontières étanches et donc de nombreux chercheurs contribuent à plusieurs thèmes. Enfin, on peut noter qu'il existe de nombreuses interactions et coopérations avec les autres équipes du LaBRI qui se traduisent en particulier par des projets communs ou des publications communes.

1.3 Thèmes de recherche de l'équipe

1.3.1 Combinatoire énumérative et algébrique

Introduction

Rappelons pour commencer à quoi ressemble un problème générique de combinatoire énumérative : étant donnée une famille d'objets discrets, dotés d'une *taille* de valeur entière, on s'interroge sur le nombre a_n (supposé fini) d'objets de taille n dans cette famille. Par exemple : combien y a-t-il de permutations sur n éléments ? de groupes d'ordre n ? d'arbres planaires à n sommets ? Bien sûr, on peut raffiner cette question en s'intéressant simultanément à plusieurs paramètres : combien y a-t-il de graphes planaires à n sommets et k arêtes ? Les nombres auxquels on s'intéresse sont en général "rangés" dans une *série génératrice*, c'est-à-dire dans la série formelle $\sum_n a_n t^n$.

Mais au fait... Pourquoi énumérer ? Imaginons une famille d'objets discrets, surgie de l'informatique, des mathématiques, d'un modèle de physique ou de biologie, ou de tout autre domaine. Chercher à énumérer ses éléments oblige à élucider au mieux la structure de ces objets. Pour prendre un exemple précis, une relation de récurrence sur le nombre a_n d'objets de taille n traduit en général l'existence d'une certaine description récursive des objets : c'est cette description qui nous éclaire sur leur structure. Au terme du travail, on obtient, bien entendu, une information *quantitative* sur les objets. Dans la forme du résultat transparaissent encore souvent certaines propriétés structurelles. Des estimations asymptotiques des nombres a_n permettent également de *comparer* différentes familles d'objets dans une même échelle. Enfin, l'énumération selon plusieurs paramètres simultanément, combinée à l'étude asymptotique du résultat, fournit des informations sur la forme générale des objets (un arbre planaire à n sommets a, en moyenne, $n/4$ feuilles). Dans les cas les plus favorables, on parvient à dégager des lois limites (la distribution du nombre de feuilles de ces arbres, convenablement normalisée, converge vers une loi gaussienne), et à établir un pont avec des processus aléatoires continus.

Au delà de ces motivations générales, certains problèmes d'énumération ont des justifications naturelles plus précises, provenant notamment de deux domaines qui nous intéressent particulièrement : l'analyse d'algorithmes et la physique statistique.

Analyse d'algorithmes Tout, dans un algorithme, est discret : l'entrée, la sortie, les structures de données... Énumérer selon leur taille et le coût de leur traitement (en temps, en espace) les différentes entrées possibles permet de calculer le coût moyen de l'algorithme sur les entrées de taille n . L'analyse asymptotique des résultats se révèle alors particulièrement significative. La fructueuse interaction entre la combinatoire énumérative et l'analyse d'algorithmes peut être illustrée par exemple par les livres de D. Knuth, déjà historiques, et, en France et maintenant, par les activités de Philippe Flajolet et de son équipe de l'INRIA-Rocquencourt.

Par ailleurs, il faut souligner la *grande parenté* entre certaines techniques combinatoires et des structures d'informatique théorique : ainsi, énumérer une famille d'objets par la méthode des "matrices de transfert", c'est coder ces objets par les mots d'un langage rationnel et construire un automate pour ce langage. De même, la méthodologie dite "de Schützenberger" consiste à coder des objets par les mots d'un langage algébrique (non ambigu) : toute grammaire engendrant ce langage reflète alors les propriétés de structure des objets codés, et permet de surcroît de les énumérer.

Physique statistique La physique statistique cherche à rendre compte de phénomènes macroscopiques par la description des forces et des interactions à l'échelle microscopique : par exemple, comment expliquer qu'un aimant, chauffé, perde ses propriétés magnétiques, à partir de la description des forces qui règnent entre les atomes le constituant ? Bien sûr, la structure microscopique réelle des corps étudiés et les interactions entre leurs composants élémentaires sont en général bien trop compliquées pour se prêter à une étude mathématique. Ceci amène à définir des modèles simplifiés (et probabilistes) : les composants élémentaires (appelons-les atomes) sont situés sur les sommets d'une grille régulière (cubique par exemple), et un atome n'interagit qu'avec ses voisins. Chaque atome peut-être dans un nombre fini d'états, et la liste des états des atomes est une *configuration* possible de l'objet macroscopique. Cette configuration apparaît avec une certaine probabilité, qui est fonction des états des paires d'atomes voisins.

Résoudre un tel modèle, c'est calculer sa *fonction de partition*, c'est-à-dire énumérer les configurations selon leur taille (le nombre d'atomes) et leur probabilité d'apparition. Comme on s'intéresse aux objets de grande taille, ce sont surtout, là encore, les propriétés asymptotiques des nombres cherchés qui sont significatives.

Combinatoire algébrique La combinatoire algébrique vise, comme la combinatoire énumérative, à mieux comprendre la structure d'objets discrets. Mais les objets qu'elle étudie sont d'essence algébrique, et requièrent donc un certain savoir-faire algébrique.

Son champ d'action est très vaste. Citons entre autres l'étude des groupes et des algèbres, la théorie des représentations, la théorie des invariants, et des liens importants avec la topologie, la géométrie algébrique et... la combinatoire énumérative.

Pour souligner la parenté entre ces deux disciplines, prenons l'exemple des fonctions symétriques et des tableaux de Young, intimement liés aux représentations du groupe symétrique. La combinatoire des tableaux s'incarne de façon exemplaire dans la correspondance de Robinson-Schensted qui établit une bijection entre paires de tableaux standard de taille n et de même forme et permutations de taille n . Ce résultat semble avant tout énumératif. Mais dès lors que les fonctions symétriques de Schur, définies algébriquement, se réalisent par le théorème de Richardson comme somme de monômes associés à certains tableaux de Young, cette correspondance devient l'outil fondamental de la combinatoire des fonctions symétriques.

Mots-Clés Énumération, Génération aléatoire, Fonctions symétriques et quasi-symétriques, Séries génératrices, Bijections, Cartes, Groupes.

Description des activités et principaux résultats

Lorsque sonne l'heure du rapport d'activité, un vieux dilemme se représente : faut-il présenter les résultats obtenus en les classant selon le type d'objets, ou selon la méthode employée ?

La première solution est conforme aux motivations données dans l'introduction, et bien adaptée si l'on cherche à dompter une famille particulière d'objets. Cependant, au développement de techniques nouvelles adaptées à un problème spécifique répond un autre type de travail, transversal : il consiste à bien comprendre comment fonctionne une technique, éventuellement à en unifier plusieurs, et surtout à bien cerner la nature des problèmes auxquels elles peuvent s'appliquer. Bref, à faire un peu de rangement dans le désordre (sans doute réel) des problèmes d'énumération. Ainsi, une technique initialement développée pour résoudre un problème purement mathématique pourra être recyclée pour une analyse d'algorithme, ou inversement. Tout nouvel outil d'énumération est intéressant, et on ne peut prédire à l'avance

quelle sera sa prochaine utilisation.

Par souci de masquer un peu les aspects les plus techniques, nous avons privilégié dans ce rapport une classification selon le type des objets étudiés. Avant de rentrer dans le détail des résultats, accordons donc quelques mots à la description des méthodes employées.

Pour résoudre des problèmes d'énumération, notre équipe a, au fil des ans, élaboré un certain nombre de techniques, construit une sorte de "boîte à outils" largement applicable. L'un de ces outils est l'usage fréquent de codages, de *bijections* entre classes d'objets, permettant de relier plusieurs questions d'énumération — le plus souvent, de convertir un problème difficile en une question plus classique, qu'on saura résoudre par une méthode standard. La méthodologie de Schützenberger, qui consiste à coder des objets par les mots d'un langage algébrique, apparaît ainsi comme une démarche typique de l'approche bijective. Cependant, certains problèmes résistent à ces techniques et persistent à ne pas livrer leur structure combinatoire fine : il est alors (parfois) possible d'utiliser une information naïve sur la structure pour écrire des équations fonctionnelles régissant leur énumération. Naturellement, la naïveté combinatoire se paye par la complexité des équations : nous sommes ainsi amenés à nous intéresser à certaines classes non triviales d'équations fonctionnelles d'origine combinatoire.

Nos travaux sont soutenus par l'idée selon laquelle il est intéressant et informatif de savoir à quelle *classe de séries formelles* appartient la série génératrice $\sum_n a_n t^n$. Les trois classes auxquelles on s'intéresse principalement sont, par généralité croissante, les séries rationnelles, les séries algébriques (solutions d'une équation polynomiale), et les séries différentiellement finies (solutions d'une équation différentielle linéaire à coefficients polynomiaux). Intuitivement, la première classe correspond aux langages réguliers, la deuxième aux langages algébriques, tandis que l'intuition fait un peu défaut pour la troisième classe.

Les objets auxquels s'appliquent notre expertise sont très variés : graphes (arbres, cartes), figures diverses tracées sur des réseaux (chemins, polygones), permutations, partitions... Potentiellement, tous les objets combinatoires classiques sont concernés.

Parmi eux, insistons sur les cartes combinatoires : ce sont, historiquement, les premiers objets familiers de l'équipe d'informatique théorique bordelaise. Une carte combinatoire donne une description précise d'un graphe dessiné sur une surface. Ce dessin permet souvent, en utilisant la représentation d'une carte par un couple de permutations, d'obtenir des algorithmes performants sur les graphes. Cette représentation, d'abord conçue comme outil de codage et de manipulation des cartes, s'est avérée plus tard très liée à d'autres structures mathématiques (surfaces de Riemann, corps de nombres, groupes de permutations, caractères, etc). Tous ces aspects du sujet interagissent souvent de manière remarquable. Dans une monographie très récente de S. Lando (Moscou) et A. Zvonkine (LaBRI) sont exposés de nouveaux aspects de la théorie des cartes combinatoires et ses divers liens, notamment : les dessins d'enfants de Grothendieck et les liens entre les cartes et la théorie de Galois ; les modèles matriciels et les liens entre les cartes et la gravitation quantique ; les espaces des modules des surfaces de Riemann et la preuve de la conjecture de Witten (théorème de Kontsevich) ; les espaces de Hurwitz et la classification topologique des polynômes et des fonctions méromorphes ; enfin, les invariants de Vassiliev des nœuds. Tous ces domaines sont traités pour la première fois sous forme d'un livre.

Nous détaillons ci-dessous quelques familles de résultats obtenus récemment par notre groupe.

Chemins sur réseau L'énumération des chemins à sommets de coordonnées entières, partant d'un point donné, formés de pas choisis dans un ensemble fini prescrit à l'avance, et évitant une zone donnée de l'espace, est un sujet très classique en combinatoire, mais

encore imparfaitement compris. Notons qu'il existe une approche générique, basée sur une construction des chemins pas à pas : elle donne en général une équation fonctionnelle, mais sa résolution peut comporter bien des mystères.

Un exemple désormais bien clarifié est celui des chemins vivant sur une demi-droite : ce problème mène à des équations linéaires dites "à une variable catalytique". La méthode du noyau, familière en probabilités, et développée pour la combinatoire en collaboration avec M. Petvokšek (Slovénie), permet d'obtenir systématiquement leur série génératrice, qui est toujours algébrique. Cette méthode et ce résultat s'étendent aux chemins du plan contraints de rester dans un demi-plan : là encore, les séries obtenues sont toujours algébriques.

La question naturelle qui se pose ensuite est de compter les chemins du plan contraints d'appartenir à l'intersection de *deux* demi-plans. Quitte à déformer l'espace, on peut toujours se ramener à des chemins vivant dans le quart de plan. Nous avons montré que la situation est bien plus compliquée : certes, on sait dans tous les cas écrire une équation linéaire "à deux variables catalytiques", mais la série sous-jacente n'est plus algébrique en général, ni même différentiellement finie. Nous avons même donné un exemple où elle comporte un nombre infini de singularités (alors que toute série D-finie n'en a qu'un nombre fini). Il est possible d'associer à chaque ensemble de pas donné un certain groupe : intuitivement, si ce groupe est fini, on a de meilleures chances de résoudre l'équation. Nous avons même l'espoir d'établir une équivalence entre la finitude du groupe et la D-finitude de la série génératrice. Si on ne considère que des "petits" pas (c'est-à-dire variant de 0, +1 ou -1 dans chacune des directions du plan), il n'y a alors qu'un nombre fini de cas à étudier, ce que nous avons fait en collaboration avec M. Mishna (Canada). À l'heure actuelle, parmi les modèles menant à un groupe fini, un seul résiste encore. En particulier, nous avons donné la première preuve constructive d'un vieux résultat de Kreweras (1965), resté jusque-là mystérieux, correspondant à un cas où la série est algébrique, et montré qu'un deuxième modèle, apparenté, était aussi de nature algébrique.

Par ailleurs, nous avons obtenu des résultats d'algébricité imprévu sur les chemins du plan évitant une demi-droite fixée. Le dénombrement asymptotique de ces chemins avait déjà été fait (dans le cas du réseau carré), mais nos résultats ont créé une petite surprise car personne ne pensait que ce modèle pouvait être résolu exactement.

Énumération de cartes. L'énumération des cartes planaires, c'est-à-dire des graphes plans plongés dans la sphère, est un sujet "ancien" bordelais, qui a connu depuis une petite dizaine d'années des rebondissements très intéressants. Par ailleurs, nous avons entamé des échanges avec certains physiciens théoriciens, qui s'intéressent aux cartes comme modèles de *gravitation quantique*.

Il y a une dizaine d'années, on disposait de deux approches pour résoudre des problèmes de comptage de cartes : l'approche d'origine de Tutte, basée sur la destruction récursive d'une arête distinguée, menait à certaines équations fonctionnelles, résolues (ou pas) au cas par cas ; l'approche "physique", en général non rigoureuse, mais dotée d'un très vaste champ d'applications, reposait sur l'évaluation d'intégrales de matrices.

La thèse de G. Schaeffer (LaBRI, puis Paris) a été le point de départ d'une troisième approche, très combinatoire (donc élégante!), presque purement bijective, et dont on explore encore les applications possibles. Par exemple, en collaboration avec G. Schaeffer, nous avons résolu de manière combinatoire deux types de modèles de physique sur les cartes planaires de degré borné : le modèle d'Ising et le modèle des particules dures. Nos travaux expliquent combinatoirement, en établissant un lien avec certaines familles d'arbres, pourquoi on trouve toujours des séries algébriques dans ces problèmes, et, en prime, fournissent la première solution rigoureuse de ces modèles.

Par ailleurs, en revenant à l'approche originale de Tutte, nous avons examiné de plus près le type d'équations auxquelles elle mène. On ne savait jusqu'alors les résoudre systématiquement que lorsqu'elles étaient quadratiques. Dans un travail commun avec un collègue de l'A2X (Bordeaux), A. Jehanne, nous avons prouvé qu'en fait la solution de toutes ces équations était toujours algébrique, et donné une façon de les résoudre. Cette théorie implique que de vastes classes de problèmes d'énumération de cartes vont toujours avoir une solution algébrique : cartes à degrés contraints, cartes biparties à degrés contraints, etc. Encore plus récemment, nous avons prouvé que l'approche de Tutte pouvait être appliquée à des cartes contraintes à la fois par le degré des faces et celui des sommets (c'est une nouveauté, car en général ce type de problème est très dur) et prouvé que les triangulations non-séparables dont les sommets ont degré au moins 3 (resp. 4, 5) ont une série génératrice algébrique.

Mentionnons pour finir ce qui est peut-être la naissance d'une deuxième approche bijective pour les cartes, en lien avec les orientations contraintes de cartes. Nous avons trouvé une bijection naturelle permettant de compter les cartes planaires munies d'un arbre couvrant. Ce résultat d'énumération était connu, mais la simplicité de sa forme n'avait jusque-là pas été expliquée.

Les empilements de pièces, en combinatoire et en physique Les empilements de pièces, version géométrique des monoïdes partiellement commutatifs (Cartier-Foata), ou des monoïdes de traces proposés comme modèles de parallélisme (Mazurkiewicz), ont été introduits au LaBRI dans les années 80. Cette notion s'est avérée très féconde depuis.

Une de ses premières applications a été de rendre facile l'énumération de certains objets issus de la physique statistique, appelés animaux dirigés. La possibilité de *factoriser* dans ce monoïde a rendu transparente l'algébricité de leur série génératrice. Plus récemment, en collaboration avec A. Rechnitzer (Melbourne), nous avons étendu la correspondance de base entre empilements de dominos et animaux dirigés, en franchissant quelques uns des (nombreux) pas qui mènent des animaux dirigés aux animaux généraux.

Les empilements de dominos ont fait récemment une deuxième apparition en physique, sous une forme déguisée, en "gravitation quantique Lorentzienne". C'est un sujet très actif de la physique théorique, consistant à réunir deux théories réputées incompatibles : la relativité générale et la mécanique quantique. Les physiciens Ambjorn et Loll ont introduit un nouveau modèle discret (les triangulations quantiques Lorentziennes) et ont montré que le modèle était résoluble en dimension deux. Cette étude fut prolongée par les physiciens Di Francesco, Guitter et Kristjansen en introduisant une notion de courbure des triangulations de l'espace-temps.

Avec W. James (Melbourne), nous avons expliqué tous ces résultats exacts en montrant le lien profond existant entre ces triangulations Lorentziennes et les empilements de pièces. Nous avons étendu certains résultats des physiciens au cas général des triangulations Lorentziennes arbitraires (sans conditions aux bords). La théorie est liée à celle des animaux multidirigés mentionnés ci-dessus.

Combinatoire des nombres de Strahler Le nombre de Strahler d'un arbre binaire a été introduit en hydrogéologie pour l'étude morphologique des bassins fluviaux. Il se retrouve dans des problèmes d'optimisation en informatique, en infographie, ainsi qu'en biologie moléculaire dans l'étude des structures secondaires de molécules du genre ARN. Nous avons donné une preuve bijective (constructive) du fait que la distribution des arbres planaires selon le paramètre ordre (venant de la biologie moléculaire) était la même que celle de certains chemins du plan (chemins de Dyck) selon leur hauteur logarithmique, laquelle distribution est bien connue pour être celle des arbres binaires selon le nombre de Strahler.

Tout récemment, nous avons introduit un nouvel objet appelé *tours de Kepler*, énuméré par les nombres de Catalan (nombre d'arbres binaires) et faisant apparaître d'une manière originale la distribution nombre de Strahler des arbres binaires. Une construction bijective entre les chemins de Dyck et les tours de Kepler, laissant invariante la distribution nombre de Strahler a été donnée, en collaboration avec D. Knuth.

Théorie des polynômes orthogonaux et fractions continues Notre travail sur les nombres de Strahler des arbres binaires et la hauteur logarithmique des chemins de Dyck est intimement lié à la théorie combinatoire des polynômes orthogonaux et des fractions continues développée depuis 25 ans par Flajolet et Viennot. De nouvelles considérations sur les contractions dans les chemins de Dyck donnent des interprétations des formules classiques permettant de contracter des fractions continues. Nous avons introduit de nouvelles constructions bijectives, permettant de contracter des empilements de dominos, en liaison avec la combinatoire des animaux dirigés.

Permutations à motifs exclus. L'un des centres d'intérêt importants en combinatoire des mots est la recherche et l'analyse de régularités dans les mots, et, de façon duale, la recherche de mots ne comportant pas certaines régularités. Souvent, les propriétés de régularité que l'on recherche ou que l'on essaie d'éviter s'expriment en termes de facteurs ou de sous-mots.

Nous nous sommes intéressés à un domaine particulier de cette problématique : la combinatoire des permutations à motifs exclus, c'est-à-dire des permutations ne comportant pas de sous-suites (sous-mots) d'un type donné. De telles familles de permutations apparaissent très souvent dans différents domaines de la combinatoire. Par exemple, les permutations n'admettant pas de sous-suite croissante de longueur supérieure ou égale à k sont les permutations ne comportant pas le motif $123\dots k$; elles sont en correspondance avec les paires de tableaux de Young de même forme et de hauteur inférieure à k .

Plus généralement, l'énumération des permutations à motifs exclus a tout le charme des problèmes durs de formulation élémentaire. La question de base est la suivante : étant donnée une permutation τ , appelée motif, combien de permutations de longueur n ne contiennent aucun sous-mot de forme τ ? En d'autres termes, à quel point est-il restrictif d'interdire le sous-mot τ ? Cette question a suscité ces dernières années de très nombreuses publications dans la communauté combinatoire, notamment nourries par les deux (ex-)conjectures suivantes : la première, récemment démontrée, dit que le nombre de permutations de taille n évitant un motif donné croît au plus exponentiellement. La deuxième n'a été ni prouvée ni infirmée, mais son auteur n'y croit plus lui-même : elle portait sur la nature de la série génératrice des permutations évitant un motif donné, et prédisait que cette série serait toujours différentiellement finie, c'est-à-dire solution d'une équation différentielle linéaire à coefficients polynomiaux. Certes, les rares cas où cette série est connue sont tous différentiellement finis, mais ils sont encore si peu nombreux qu'on ne peut pas vraiment en déduire une conjecture solide.

La difficulté de la question "de base" posée ci-dessus a orienté la recherche vers trois directions : la première consiste à trouver des méthodes d'énumération valides pour plusieurs familles de motifs ; la seconde (et la plus répandue) consiste à étudier des variations de la question, par exemple en interdisant plusieurs motifs simultanément, ou en restreignant par des conditions naturelles la famille de permutations de départ ; la troisième, purement combinatoire, s'attache à prouver que plusieurs classes de permutations à motifs interdits sont dénombrées par la même suite. Notre groupe a contribué sur ces trois points.

Certaines classes de permutations à motifs interdits sont suffisamment structurées pour qu'il soit possible de les engendrer récursivement, en ne "gardant en mémoire" qu'un nombre borné de paramètres. Les séries génératrices associées satisfont des équations à variables *catalytiques*. Il y a dans ces équations autant de variables catalytiques qu'il faut garder de paramètres en mémoire. Nous avons désormais bien compris la résolution des équations à une variable catalytique. Elle mène toujours à une série algébrique (et donc différentiellement finie). Nous avons fait de récents progrès sur les équations à deux variables catalytiques : même si notre approche n'est pas entièrement systématique, elle a permis d'énumérer de façon systématique toutes les classes de permutations à deux variables catalytiques qui existaient dans la littérature (et faisaient jusque-là l'objet de solutions ad hoc).

Nous avons par ailleurs obtenu plusieurs résultats d'énumération sur les *involutions* à motifs exclus. Par exemple, avec T. Mansour (Haïfa), nous avons pu énumérer les involutions évitant le motif 312 et un deuxième motif τ , et prouver que le résultat est indépendant de τ dès que τ appartient à une certaine classe de motifs. Le résultat fait intervenir les polynômes de Chebycheff, qui apparaissent déjà dans d'autres problèmes de permutations à motifs exclus. De même, avec E. Steingrímsson (Göteborg, Reykjavik), nous avons montré, par une approche purement bijective, que pour tout motif τ sur l'alphabet $\{k+1, k+2, \dots, k+m\}$, le nombre d'involutions de longueur n évitant $12 \cdots k\tau$ est égal au nombre d'involutions de longueur n évitant $k \cdots 21\tau$. En d'autres termes, retourner le préfixe $12 \cdots k$ ne change rien à l'énumération. Un résultat semblable avait été prouvé il y a quelques années pour les permutations. Le fait que cet énoncé "passe" aux involutions repose sur le fait qu'une certaine transformation commute avec l'inversion des permutations, ce qui reste une propriété assez mystérieuse.

Tas de sable. Le modèle du tas de sable, introduit par les physiciens Bak, Tang et Wiesenfeld, est un système dynamique simulant la diffusion de grains de sable sur les sommets d'un graphe et induisant un "phénomène critique auto-organisé". En régime stationnaire, ce système ne peut atteindre qu'un sous-ensemble des distributions de grains, qui sont dites récurrentes et sont en bijection avec les arbres couvrants du graphe de base, comme l'a montré Dhar. Nous avons donné une preuve bijective de ce résultat, qui prouve de surcroît que le nombre de grains de la configuration récurrente est distribué comme l'activité externe de l'arbre. Ce dernier paramètre est essentiel dans la définition du polynôme de Tutte du graphe. Ce polynôme à deux variables contient beaucoup d'informations sur le graphe ; par exemple, une de ses spécialisations donne le polynôme chromatique.

Ce nouveau résultat s'ajoute aux nombreuses propriétés du modèle du tas de sable : existence d'une structure de groupe abélien sur les configurations récurrentes (et donc sur les arbres couvrants), équiprobabilité des configurations récurrentes... Le calcul explicite du groupe et son lien avec la structure du graphe est une question émergente du domaine. Nous avons proposé, en collaboration avec D. Rossin (Liafa), une preuve élémentaire d'une propriété de l'identité de ce groupe dans le cas des grilles fortement rectangulées.

Partitions d'entiers Une partition d'un entier n est une suite décroissante dont la somme des entrées est égale à n . Ces objets sont classiques et ont été étudiés par de nombreux mathématiciens (Euler, Sylvester, MacMahon, Andrews...), en lien avec la théorie des nombres et la physique. De nombreux points de vue différents (mais inter-reliés) existent sur les partitions.

Identités de partitions Ces "identités de partitions" sont parmi les résultats les plus appréciés dans la théorie des partitions. La plus connue est due à Rogers et Ramanujan : le nombre de partitions de n telles que les parts sont 1 ou 4 modulo 5 est égal au nombre

de partitions dans lesquelles la différence entre parts consécutives est au moins deux. Bien que les preuves des telles identités soient rarement bijectives, nous avons, en collaboration avec S. Corteel (Versailles), trouvé une grande famille d'identités de partitions en n'utilisant que la combinatoire bijective. Cette famille comprend des identités célèbres dues à Schur et Andrews.

Identités entre q -séries Beaucoup d'identités entre q -séries restent à comprendre combinatoirement. Nous avons récemment trouvé une bijection entre certains ensembles de partitions ordinaires et de partitions “à la Frobenius”, qui implique la fameuse somme ${}_1\Psi_1$ de Ramanujan ainsi que le q -analogue de la formule de Gauss. Ce sont les premières preuves combinatoires simples et naturelles pour ces sommes.

La combinatoire des surpartitions Une surpartition est une partition dans laquelle la première occurrence d'une part peut être surlignée. Beaucoup de techniques et résultats pour les partitions ont des analogues pour les surpartitions. Notamment, il y a des liens entre les surpartitions et les corps quadratiques réels et il y a des “partition theorems” comme ceux de Rogers-Ramanujan pour les partitions ordinaires. Il reste certainement beaucoup à découvrir sur les surpartitions.

Les partitions et les formes modulaires Dès que Ramanujan a découvert certaines congruences de la fonction donnant le nombre de partitions, il y a eu beaucoup de recherches sur les propriétés arithmétiques des fonctions combinatoires. Nous avons étendu certaines familles classiques de congruences dues à Ramanujan en utilisant la théorie des formes modulaires de poids demi-entier et des résultats de Shimura.

Les partitions, les corps quadratiques, et le lemme de Bailey En 1988, Andrews, Cohen, Dyson, et Hickerson ont étudié une q -série qui a été importante dans les travaux de Ramanujan. Même si cette q -série ressemble à une forme modulaire de poids 1, elle vient en fait d'un caractère de Hecke sur un corps quadratique réel, et elle est liée à la théorie des “Maass waveforms”. Récemment nous avons découvert qu'un ingrédient de leurs preuves, le lemme de Bailey, peut être utilisé pour prouver qu'il y a beaucoup de q -séries qui sont des fonctions génératrices pour certaines partitions et qui ont des coefficients qui viennent des corps quadratiques réels, sans être des formes modulaires. Il existe ainsi plusieurs applications des techniques de Bailey dans les q -séries et nous avons décrit une structure de treillis pour les paires de Bailey, qui devrait permettre de lier beaucoup de q -séries à une grande variété d'objets provenant de la théorie des nombres, tels que les formes modulaires, les fonctions thêta, les corps quadratiques, et les formes quadratiques ternaires.

Génération aléatoire Nous avons, en collaboration avec F. Flajolet (Rocquencourt), G. Louchard (Bruxelles) et G. Schaeffer (Paris), proposé une nouvelle méthode générique pour la génération de structures “décomposables”. Cette approche, dite boltzmannienne, permet grosso modo de traiter les mêmes familles de structures que des méthodes plus classiques à base de décomposition et d'énumération ; mais, en échange d'une incertitude sur la taille des structures obtenues, les algorithmes de génération aléatoires opèrent en temps linéaire ou quasi-linéaire, et ne nécessitent le précalcul que de petits nombres de valeurs de séries génératrices – contrairement aux méthodes classiques qui nécessitent des précalculs importants.

Combinatoire énumérative, combinatoire algébrique et physique statistique : le modèle des promeneurs méchants Ce travail fait parfaitement la jonction entre les activités énumératives et les activités algébriques de notre groupe.

Il s'agit d'une étude en collaboration avec A. J. Guttmann (Melbourne) et C. Krattenthaler (Lyon). En 1984, M. Fisher a introduit le modèle des promeneurs dits "méchants" (*vicious walkers*) comme modèle pour les transitions de phase "mouillées". Ces promeneurs ont des trajectoires qui ne se rencontrent pas (ce qui les dispense d'exercer les uns sur les autres leur méchanceté). Nous résolvons le problème des promeneurs méchants confinés entre deux barrières absorbantes. En appliquant la méthodologie dite de Lindström-Gessel-Viennot, le problème revient à calculer des sommes de déterminants, sommation que nous arrivons à résoudre en ramenant les calculs à des considérations de combinatoire algébrique. Nous introduisons un lien avec la combinatoire des caractères symplectiques de forme rectangulaire et des caractères orthogonaux impairs du groupe symétrique et utilisons des formules connues sur ces caractères.

Fonctions symétriques et quasi-symétriques Les polynômes symétriques de Macdonald unifient, dans un cadre à deux paramètres q, t , les deux q -analogues classiques des fonctions de Schur, à savoir les polynômes de Hall-Littlewood et les polynômes de Jack. De nombreuses propriétés combinatoires de ces polynômes ont pour cœur la conjecture $n!$ énoncée par Garsia et Haiman. Nous avons travaillé, en collaboration avec F. Bergeron (Montréal) et N. Bergeron (Toronto), sur l'obtention de résultats explicites. Ce travail s'appuie sur l'étude d'opérateurs "de sauts" agissant sur des diagrammes plans associés aux polynômes, et permet de traiter plusieurs cas particuliers (cas des partitions en équerre, cas d'un alphabet), et de présenter certaines généralisations (partitions avec un ou plusieurs "trous").

Les fonctions quasi-symétriques ont été définies par Gessel pour l'énumération de certains objets (les P -partitions, introduites par Stanley). L'étude de leurs propriétés algébriques se révèle aussi intéressante. Nous avons étudié, en collaboration avec F. Bergeron (Montréal) et N. Bergeron (Toronto), la structure de l'idéal engendré par ces polynômes. Nous avons obtenu une base de Gröbner explicite de cet idéal, et montré que le quotient de l'espace des polynômes par cet idéal a pour dimension les nombres de Catalan $C_n = \frac{1}{n+1} \binom{2n}{n}$ (une base explicite est décrite en bijection avec les chemins de Dyck). Nous avons ensuite étendu cette étude, en particulier en plusieurs alphabets.

Classification topologique des polynômes Deux polynômes P et Q , considérés comme des applications de la sphère complexe vers elle-même, sont topologiquement équivalents s'il existe deux homéomorphismes f, g de la sphère tels que $g \circ P = Q \circ f$. Le *groupe de monodromie* est un invariant de cette équivalence. Grâce aux résultats récents provenant de la classification des groupes finis simples, on est parvenu à résoudre complètement le problème de classification pour les groupes primitifs différents de S_n (le groupe symétrique) et A_n (le groupe alterné).

Nous avons aussi obtenu quelques premiers résultats vers une classification de revêtements de genre supérieur.

Projets et perspectives 2005-2008

Voici quelques projets, d'envergure variable, auxquels nous proposons de nous attaquer dans les années qui viennent. Certains sont déjà entamés, et/ou sont dans le prolongement naturel des résultats présentés dans la section précédente.

Chemins sur un réseau Un programme minimal consiste à achever la preuve du résultat suivant : les chemins contraints à rester dans le quart de plan, formés de pas au plus unitaires, et associés à un groupe fini, ont toujours une série génératrice différentiellement finie (voir plus haut pour le sens de ces notions). Il reste en fait un seul cas à traiter (mais il résiste...) pour arriver à cet énoncé. L'étape suivante dans la classification des chemins du quart de plan est de prouver que les chemins, même à pas unitaires, associés à un groupe *infini*, ont une série génératrice non-D-finie.

La méthode du noyau en grande dimension Voici une petite infraction à la règle selon laquelle on parlerait toujours dans ce rapport d'objets combinatoires plutôt que de méthodes... Cette méthode du noyau, au départ, permet de résoudre des équations dites "linéaires à une variable catalytique". Ses applications comprennent l'énumération systématique des chemins confinés dans un demi-plan, par exemple. Son extension à des équations à deux variables, sur laquelle nous avons travaillé ces dernières années, a permis de résoudre des problèmes variés (chemins dans le quart de plan, permutations à motifs exclus...). Cette extension ne fonctionne (et encore) que dans le cas où un certain groupe, associé au "noyau" de l'équation, est fini. On frémit d'abord en pensant à ce que pourrait être une généralisation à un nombre arbitraire de variables catalytiques, mais quelques signes prometteurs sont venus récemment ranimer la confiance : certains problèmes d'énumération mènent à un groupe très bien structuré (le groupe symétrique par exemple, ou des groupes de réflexions en général). Par ailleurs, le principe de réflexion d'André, généralisé par Gessel et Zeilberger, peut-être vu comme une instance de la méthode du noyau en dimension arbitraire. Nous projetons d'étendre cette méthode au delà du principe de réflexion. Un tout premier problème sur lequel nous travaillons est une formule pour les chemins "osculateurs" conjecturée en 1997, et restée sans preuve à ce jour.

Configurations de boucles compactes Une configuration de boucles compactes (en anglais : fully-packed loops) est un sous-graphe d'une grille carrée de taille n , dans lequel tous les sommets sont de degré 2. Certaines conditions aux bords assurent que ces configurations sont en bijection avec les célèbres matrices à signes alternants. Ces conditions définissent, pour chaque configuration, un couplage sans croisements des $2n$ arêtes sortantes. Une conjecture due à Razumov et Stroganov affirme qu'une action de l'algèbre de Temperley-Lieb sur les couplages fait apparaître des coefficients qui sont ceux de l'énumération des configurations de boucles compactes correspondant à chaque couplage. Nous tentons de démontrer, de manière directement combinatoire, cette conjecture, qui porte sur différentes familles d'objets traditionnellement bien connus dans l'équipe (chemins sur réseau carré, couplages comptés par les nombres de Catalan...).

Matrices à signes alternants C'est un domaine qui mobilise la communauté combinatoire depuis une vingtaine d'années. Il y avait la fameuse conjecture des matrices à signes alternants, qui donnait une formule simple pour le nombre de telles matrices. Il y a eu la preuve (apocalyptique) de D. Zeilberger, ramenant le problème à celui du dénombrement des partitions totalement symétriques auto-complémentaires, autre problème célèbre et résolu par le tour de force de G. Andrews, puis la preuve de G. Kuperberg montrant l'équivalence avec le "modèle de la glace" et ramenant le problème au calcul d'un certain déterminant "quantique". Bien que ces problèmes aient été résolus, ainsi que la plupart des conjectures et autres formules tournant autour de ces matrices et de ces partitions, comme exposé dans le livre récent de D. Bressoud consacré entièrement au sujet, la question d'une explication bijective ou d'une meilleure compréhension combinatoire reste entièrement ouverte. Une surprise

récente est arrivée en physique quantique où ce nombre de matrices à signes alternants est apparu mystérieusement dans le calcul de fonctions d'onde du modèle de spins de Heisenberg XXZ.

Modèle des hexagones durs Le modèle des hexagones durs est un modèle de gaz avec exclusion mutuelle sur réseau triangulaire. Il a été résolu par R. Baxter il y a 25 ans. Sa solution, faisant intervenir des q -séries, et en particulier les célèbres identités de Rogers-Ramanujan, est un grand classique de la physique statistique des modèles résolubles. La théorie des empilements de pièces, développée par le groupe combinatoire bordelais, permet de donner une interprétation combinatoire de la densité du gaz. La solution de Baxter permet de montrer que cette série est algébrique. Nous nous proposons de donner une explication combinatoire de ce fait. Cette approche bijective, qui contient comme cas particulier le problème classique des animaux dirigés, est complètement ouverte et extrêmement difficile. Une conséquence serait de pouvoir aborder avec de nouveaux outils le problème voisin dit des "carrés durs", resté complètement ouvert.

Permutations à motifs exclus Le domaine de recherche des permutations à motifs exclus est très actif (pour preuve, un colloque international annuel spécifique existe maintenant depuis trois ans) et nombre de problèmes sont encore à considérer. Par exemple, énumérer l'ensemble des permutations excluant le motif 1324 reste à ce jour un problème ouvert. De même, nous avons commencé à travailler sur un ensemble de permutations excluant quatre motifs auquel se sont intéressés V. Gasharov et V. Reiner.

Outre les permutations, on peut également s'intéresser aux involutions, dérangements, permutations alternantes, permutations doublement alternantes, etc. Nous avons déjà ébauché l'étude des dérangements. Actuellement, nous nous intéressons aux huit ensembles de conjugaison de permutations excluant un motif de longueur 5 ainsi qu'à certaines des permutations doublement alternantes excluant un motif de longueur 4. De même, une preuve combinatoire reste à trouver pour unifier les nombreux résultats obtenus sur les involutions excluant un motif de longueur 3 et excluant ou contenant exactement une occurrence d'un autre motif.

Des travaux ont montré le lien entre certains ensembles de permutations à motifs exclus et d'autres familles d'objets combinatoires (tableaux de Young, cartes, mots, etc.). Par exemple, nous travaillons sur une identité (qui n'a pas eu à ce jour d'explication combinatoire) due à Wilf sur les tableaux standards et des permutations excluant un motif particulier. D'autre part, nous nous intéressons au mélange de tableaux de Young.

Idéaux de fonctions quasisymétriques Nous voulons poursuivre l'étude des idéaux de polynômes quasisymétriques, vus comme invariants de différents groupes et algèbres. Nous venons d'achever ce travail dans le cas des polynômes B -quasisymétriques, introduits par Baumann et Hohlweg, et souhaitons étendre ceci à tous les groupes de Coxeter. En parallèle, et en collaboration avec N. Bergeron et C. Hohlweg (Toronto), nous comptons essayer d'obtenir, dans l'esprit du théorème de Chevalley sur les groupes de réflexions, une classification des groupes tels que la dimension du quotient de l'idéal des quasi-invariants est égale à celle de l'algèbre agissant fidèlement sur ces polynômes.

Classification topologique des polynômes Les polynômes peuvent être vus comme des revêtements de la sphère par elle-même ayant un seul pôle. Cette classe de revêtements possède deux généralisations naturelles : d'une part les fonctions rationnelles (c'est-à-dire

les revêtements à plusieurs pôles), d'autre part les fonctions méromorphes définies sur des surfaces de Riemann de genre supérieur mais ayant toujours un seul pôle. En fait, pour la classification topologique, la deuxième classe semble être la plus prometteuse car : les revêtements de ce type sont plus faciles à énumérer ; les groupes de permutations primitifs ayant un grand cycle sont complètement classifiés ; et les logiciels développés pour classifier les polynômes sont directement applicables à ce cas-là. Par contre, une étude de compositions peut présenter beaucoup de nouveautés. On envisage donc d'entamer une classification des revêtements de ce type.

En 1995, P. Müller, en se basant sur le grand théorème de classification des groupes finis simples et sur les résultats, à l'époque incomplets, sur une classification des groupes primitifs ayant un grand cycle, a trouvé une liste exhaustive des groupes primitifs de monodromie de polynômes, différents de S_n et A_n . Nous envisageons de faire le même travail pour les revêtements de petit genre ayant un seul pôle.

1.3.2 Graphes & Applications

Introduction

La théorie des graphes est un outil puissant d'investigation des structures combinatoires complexes. Notre recherche porte sur le développement de cette théorie et sur ses applications dans le domaine des communications dans les réseaux.

En théorie des graphes nous nous intéressons à des problèmes classiques dans leurs développements les plus récents et nous introduisons de nouveaux axes de recherche. Notre expertise en théorie des graphes se situe en particulier dans les problèmes liés à l'hamiltonisme, aux colorations, et au calcul d'invariants. Il faut rappeler que la notion de *coloration orientée* a été introduite et développée au LaBRI et qu'elle est actuellement largement étudiée dans la communauté internationale. Notre travail porte aussi sur l'analyse de paramètres de graphes généraux et de graphes plus fortement structurés (planaires, de largeur arborescente bornée) ; calcul de tables de routage compactes, d'étiquetage de distance, algorithmes pour les communications ; codage des graphes. Les recherches du groupe Graphes et Applications sont très productives, elles sont finalisées par de nombreuses publications et collaborations internationales. Le groupe de recherche Graphes et Applications est impliqué dans un projet européen COMBSTRU du 5ème PCRD (RTN 2002-2006), dans deux ACI Masses de données (Pair à Pair et Navgraphe (2003-2006)), dans un programme d'actions intégrées Stéfanik avec l'université Comenius de Bratislava (2004-2005). Nous avons aussi été impliqués dans deux autres PAI : Barrande (2002) avec l'Université Charles de Prague et Picasso (2002-2003) avec l'Université Polytechnique de Catalogne (Barcelone). Nous avons aussi dirigé l'Action Spécifique "Algorithmique des grands Graphes" du CNRS, dept. STIC (RTP 23 No 94, 2003). De plus nous animons le groupe de travail Graphes et Algorithmes du GDR ALP.

Mots clés

Graphes, cyclabilité, hamiltonisme, coloration, coloration orientée, flots, coloration d'incidence, coloration par liste, coloration totale, coloration circulaire, polytope des stables, dessin de graphes, décomposition arborescente, étiquetage, communication, assignation de fréquences, routage compact, codage compact, petit-monde, réseau, tolérance aux pannes, pair-à-pair, réseau sémantique, graphe de grande taille, navigation, visualisation.

Description des activités de recherche et des résultats

Nous donnons maintenant une description détaillée de nos recherches. Comme nous avons un large champ d'investigation nous avons de nombreux résultats à considérer. On peut dire qu'il y a un grand pôle coloration, et un autre grand pôle routage, codage, étiquetage. Nous avons aussi des recherches à motivations appliquées comme la tolérance aux pannes, les grands graphes (petit-monde), la visualisation et les dessins de graphes. La présentation contient cinq paragraphes :

- Propriétés structurelles
- Coloration
- Structures de données compactes
- Modélisation et communications
- Dessin et visualisation

Nous retrouverons ce découpage dans la section Projets et perspectives 2005-2008.

Propriétés structurelle

Cyclabilité dans les graphes

Si $G = (V(G), E(G))$ est un graphe, on dit qu'un sous-ensemble S de sommets ou un sous-ensemble M d'arêtes indépendantes (donc un couplage) est cyclable s'il existe un cycle du graphe contenant tous les sommets de S , resp. toutes les arêtes du couplage M .

Les conditions d'existence de ces cycles qui ont été étudiées portent sur la connexité et sur le degré minimum ou sur la somme des degrés de tout stable à 2 ou 3 sommets dans les graphes généraux, les graphes bipartis, des graphes sans $K_{1,3}$; ces conditions peuvent souvent être renforcées pour assurer qu'il existe un cycle hamiltonien du graphe contenant tous les sommets de S , resp. toutes les arêtes de M .

Les résultats ont été obtenus en collaboration avec des chercheurs d'Orsay ou de Cracovie, dans le cadre du projet "Polonium".

Décomposition arborescente

La décomposition des graphes en arbre comme celle définie par Robertson et Seymour est un outil algorithmique très puissant. Cette décomposition permet notamment de résoudre en temps polynomial toute une catégorie de problèmes NP-difficiles si le graphe est décomposable en parties de petite taille. Une observation essentielle est que pour certains problèmes (calcul de distance, routage compact, graphes couvrants peu denses, *etc.*) une décomposition en partie de petit diamètre est suffisante. Nous avons développé ce type d'optimisation, c'est-à-dire la minimisation du diamètre des parties lors d'une décomposition arborescente de Robertson et Seymour. Nous avons obtenu les résultats suivants :

- Le calcul du diamètre optimal (appelé *longueur-arborescente* du graphe) pour plusieurs graphes ou classes de graphes (graphes planaires extérieurs, grilles, *etc.*), et plusieurs algorithmes d'approximations.
- Construction de nouveaux schémas de routage pour les graphes possédant une décomposition de longueur-arborescente bornée, de même que l'établissement de nouvelles bornes sur le nombre d'arêtes d'un graphe couvrant approximant additivement les distances (on parle d'*additive k-spanner*) pour les graphes triangulés et leurs généralisations, ceci a été fait en collaboration avec F. Dragan (USA).

Colorations

Colorations orientées.

Notre travail sur les colorations des graphes orientés initié au LaBRI est maintenant un domaine de recherche de nombreux collègues dans le monde. Cette problématique reste une de nos préoccupations. Il s'agit de colorier les sommets d'un graphe orienté à l'aide d'un nombre minimum de couleurs de telle façon que si un arc a pour origine un sommet colorié par i et pour extrémité un sommet colorié par j , d'une part $i \neq j$ et d'autre part il n'existe pas d'arc ayant pour origine un sommet colorié j et d'extrémité un sommet colorié i . Les résultats obtenus dans le passé ont mis en évidence des problèmes et conjectures non résolues. Il existe cependant peu de travaux liés à l'étude de la complexité des problèmes apparaissant dans ce cadre. Nous nous sommes ainsi intéressés à la détermination de la complexité de plusieurs de ces problèmes, tant dans le cadre des colorations orientées que d'autres types de colorations corrélées, telles que les colorations impropres ou les colorations acycliques. Nous avons donc caractérisé les classes de complexité des problèmes de coloration orientée ou de coloration (propre ou impropre) acyclique pour certaines classes de graphes.

Coloration d'incidence.

La notion de coloration d'incidence a été introduite en 1993 par Brualdi et Massey. Une incidence d'un graphe correspond à une demi-arête, deux incidences étant adjacentes lorsqu'elles sont incidentes à deux sommets adjacents. Une coloration d'incidence est naturellement une fonction associant une couleur aux incidences d'un graphe de façon telle que deux incidences adjacentes ont des couleurs distinctes.

Nous nous intéressons d'une part à la détermination du nombre chromatique d'adjacence de certaines familles de graphes et au lien entre ce paramètre et d'autres paramètres de graphes. Nous avons obtenu des encadrements du nombre chromatique d'incidence de certaines familles de graphes telles que les graphes planaires (ou planaires extérieurs) sans triangles, les graphes de Halin, les graphes d'excès k , les graphes k -dégénérés où certains types de graphes de distances.

Coloration par liste

Avec des chercheurs de Novosibirsk, nous avons travaillé sur la coloration acyclique des graphes planaires (un cycle ne peut pas être bicolorié) en utilisant des couleurs issues de listes de couleurs affectées aux sommets. Pour les graphes planaires nous avons prouvé que des listes de 7 couleurs suffisent (pour toute affectation de liste), il est conjecturé que 5 suffisent. Nous avons amélioré ce résultat dans le cas où le graphe planaire possède une maille donnée (taille du plus petit cycle).

Coloration $(d, 1)$ -totale

Nous développons de nouvelles recherches sur le problème de l'attribution de fréquences dans les réseaux de communication. En termes de coloration, il s'agit d'un problème de coloration totale (sommets et arêtes) d'un graphe avec des entiers. Cette coloration est appelée coloration $(d, 1)$ -totale. Les couleurs de deux sommets voisins doivent être différentes, de même pour des arêtes adjacentes. De plus la couleur d'un sommet doit être à distance d de la couleur d'une arête incidente. Nous avons des bornes intéressantes pour le nombre de

couleurs dans le cas des graphes planaires de grandes mailles et des graphes de degrés moyen maximum faibles. Ces bornes sont optimales pour des petites valeurs de d .

Coloration circulaire

La notion de coloration circulaire a été introduite par Vince en 1992. Etant donné deux entiers p et q ($p \geq 2q$), il s'agit de colorier les sommets d'un graphe avec des entiers $\{0, \dots, p-1\}$ de telle façon que deux sommets voisins aient des couleurs dont la différence en valeur absolue est supérieure ou égale à q et inférieure ou égale à $p - q$. Pour un graphe G donné, le minimum de p/q pour les (p, q) -colorations circulaires de G est le nombre chromatique circulaire de G , noté $\chi_c(G)$. Ce nombre est compris entre le nombre chromatique usuel et le nombre chromatique moins un. Cette notion est plus fine que la notion de nombre chromatique usuel et donne donc plus d'information sur la structure du graphe. Depuis une dizaine d'années de nombreux travaux ont été réalisés sur cette coloration.

Le nombre circulaire chromatique $\chi_c(G)$ d'un graphe G permet de mesurer la performance d'un système distribué : chaque sommet représente un processus et deux sommets sont adjacents si les processus correspondants sont en conflit, i.e. désirent accéder à une ressource non-partageable (registre...). En moyenne, le nombre maximal de processus pouvant opérer est égal à $\frac{\text{nombre de processus}}{\chi_c(G)}$.

Dans le cadre de cette coloration, nous avons exhibé la première famille infinie de graphes critiques dont le nombre chromatique circulaire diminue d'une unité, lorsqu'on enlève n'importe quel sommet.

Polytope des stables

Une coloration propre des sommets d'un graphe est une décomposition de l'ensemble des sommets en sous-ensembles indépendants (stables). Il est donc naturel de s'intéresser au polytope des stables d'un graphe.

Les webs sont des graphes circulants, dont la description du polytope des stables possède un intérêt particulier vis-à-vis de deux sur-classes : les graphes "quasi-ligne" et les graphes sans-griffes.

Établir la liste des facettes des polytopes des stables des graphes sans griffes est un problème ouvert depuis une quinzaine d'années, et même la restriction aux webs est largement ouverte. Avant nos travaux, il était seulement connu que les polytopes des stables des webs avec un nombre de clique au plus 3 n'ont que des facettes de rang (rank facet), et il y avait quelques exemples de facettes qui ne sont pas de rang, pour le polytope des stables de quelques webs avec un nombre de clique au moins 5.

Dans un premier temps, nous avons démontré qu'il n'y a qu'un nombre fini de webs avec un nombre de clique 4, dont les polytopes des stables n'ont que des facettes de rang.

Nous avons donné une construction pour obtenir des facettes du polytope des stables de webs, qui ne sont pas de rang. Nous avons utilisé cette construction pour montrer qu'il n'y a qu'un nombre fini de webs avec un nombre de clique $\omega \geq 4$, dont les polytopes des stables n'ont que des facettes de rang.

Flot antisymétrique

Dans l'optique d'avoir une notion duale pour la notion de coloration orientée, nous avons défini la notion de coloration forte correspondant à une coloration des sommets par des éléments d'un groupe, ce qui par dualité dans le cas des graphes planaires donne la notion de

flot antisymétrique. Nous avons prouvé que les graphes 3-arête-connexes possèdent un flot antisymétrique dont la valeur maximale est bornée. Ce domaine se révèle être un important champ d'investigation, et nous donne un nouveau point de vue pour l'étude des colorations orientées des graphes planaires.

Les applications flot- et tension-continues

Un *cycle* d'un graphe G est un ensemble $C \subseteq E(G)$ tel que tout sommet du graphe $(V(G), C)$ soit de degré pair. Soient G, H deux graphes, on dit qu'une application $\phi : E(G) \rightarrow E(H)$ est *cycliquement continue* si la préimage de tout cycle de H est un cycle de G . Une fascinante conjecture de Jaeger dit qu'il existe une application cycliquement continue de tout graphe sans isthme dans le graphe de Petersen. Jaeger a prouvé que si cette conjecture est vraie alors la conjecture de la 5-double couverture par des cycles et la conjecture de Fulkerson sont vraies.

La notion d'application cycliquement continue conduit naturellement à la définition d'un quasi-ordre \succ dans la classe des graphes finis. Plus précisément on dira que $G \succ H$ s'il existe une application cycliquement-continue de G à H . Nous avons étudié ce quasi-ordre ainsi que d'autres quasi-ordres reliés. Nous exhibons des liens entre les propriétés structurelles de ces quasi-ordres et les problèmes classiques de flot et de coloration. Ce contexte conduit aussi à un certain nombre de nouvelles questions qui nous permettent d'aborder des conjectures classiques sous un nouvel angle. Ce travail a été réalisé en collaboration avec J. Nešetřil (Tchèque) et M. DeVos (USA).

Structures de données compactes

Codage compact de graphes

Ce sujet porte sur un aspect algorithmique de la théorie des graphes : comment coder la structure d'un graphe avec un nombre optimal de bits ? Les applications sont nombreuses : en synthèse d'images ou en visualisation d'information, la structure et la topologie des objets peuvent comporter des millions de points ou sommets. Pour économiser l'espace mémoire, le codage optimal ou compressé s'avère primordial.

D'un point de vue théorique, l'obtention d'un codage optimal est satisfaisant mais le praticien objectera que la compression et la décompression mettent tellement de temps qu'il peut devenir totalement inutilisable. Pour cette raison, nous rencontrons dans la littérature considérer deux familles de codage compact : des codages optimaux (ou quasi) et depuis peu, des codages efficaces. L'efficacité concerne deux paramètres : l'espace mémoire et le temps requis pour effectuer des opérations élémentaires sans avoir à décompresser. Dans le cas des graphes, il est intéressant de tester l'adjacence en temps constant, calculer le degré en temps linéairement proportionnel au nombre de voisins, ... Ce souci d'efficacité dans le domaine du codage compact est relativement récent (voir Munro et Raman 2001).

Nos résultats portent sur trois familles de graphes : les graphes planaires, les graphes planaires extérieurs (une sous-famille des graphes planaires), les plongements de graphes en k pages, pour lesquelles nous proposons de nouvelles bornes ou formules exactes d'énumération, des algorithmes de codage optimal ou efficace, des algorithmes de génération aléatoire.

Nous avons prouvé que le codage des graphes k pages à n sommets et m arêtes nécessite $kn - o(kn)$ bits. Ce résultat implique que le codage de Munro et Raman de $4kn + 2n - k^2 + o(n)$

bits est optimal à un facteur 4 près. Nous avons proposé un nouveau codage en $2m\log(k) + 4m$ bits qui s'avère plus efficace pour de nombreux graphes k pages (dès que le graphe est peu dense). De plus, notre codage est efficace car il supporte des requêtes d'adjacence et de degré en temps $O(k)$.

Pour les graphes planaires, nous avons montré qu'un graphe planaire à n sommets pouvait être codé avec $5.03n$ bits et que le logarithme en base 2 du nombre de graphes planaires est plus petit que $5.007n$. Il est connu que le nombre de graphes planaires biconnexes a un logarithme binaire égal à $4.71n$. Il s'en suit que notre codage est quasi-optimal et améliore le codage en $8n + o(n)$ bits de Chiang et al. Notre technique est basée sur la décomposition entre 3 arbres des triangulations du plan (appelée en anglais *realizer*) et des techniques sophistiquées de codages à l'aide de mots de parenthèses.

Nous avons défini une décomposition de Cartes Planaires Extérieures Enracinées CPEE en un arbre et une liste d'arêtes. Cette décomposition, constructible en temps linéaire, implique l'existence d'une bijection entre les CPEE à n sommets et les arbres bicoloriés enracinés à n sommets ayant comme dernière branche des sommets unicoloriés. Comme conséquence, nous avons montré :

- une formule d'énumération nouvelle qui compte les CPEE à n sommets ;
- un codage optimal et efficace en $3n$ bits qui admet sans décompression des requêtes d'adjacence et de degré en temps constant et donne la liste des d voisins en temps $O(d)$;
- un algorithme de génération aléatoire uniforme qui fonctionne en temps linéaire.

Etiquetage

D. Peleg en 1999 a introduit la notion de *schéma d'étiquetage informatif* qui consiste à étiqueter les sommets d'un graphe de telle sorte qu'une fonction P fixée à l'avance et définie sur des sommets puisse être calculée à partir des étiquettes, sans aucune autre connaissance du graphe. Un exemple typique de fonction P est : $P(x, y) =$ la distance dans un graphe entre les sommets x et y . Ici il s'agira donc d'étiqueter les sommets d'un graphe tel que si L_1 et L_2 sont les étiquettes de deux sommets pris dans un même graphe G , alors il existe une fonction $f(L_1, L_2)$ (appelée *décodeur*), fonction indépendante de G , donnant la distance dans G entre les sommets correspondants. Dans le cas précis où P représente la distance, on parle d'*étiquetage de distance*.

L'objectif principal est de proposer des schémas fournissant les étiquettes les plus courtes possible, un décodage (le calcul de la fonction f) le plus efficace possible et enfin un calcul des étiquettes le plus rapide possible.

Nous avons les résultats suivants :

- Un survol sur les structures de données compactes et distribuées, c'est-à-dire des étiquetages, publié à l'occasion du 20e anniversaire du symposium *Principles of Distributed Computing (PODC)*, a été rédigé.
- Une série de résultats généraux sur l'étiquetage de distance. Notamment on prouve que tout graphe peut être étiqueté avec $O(n)$ bits, $O(n \log n)$ bits étant la borne triviale. On montre également que certains graphes possèdent un étiquetage théorique en $O(\log n)$ bits, cependant si les étiquettes sont compressées en $o(n)$ bits alors un temps exponentiel (en fait une tour d'exponentielles !) sera requis pour décoder les distances.
- Un algorithme en $O(n)$ calculant un étiquetage de distance pour les graphes d'intervalles à n sommets, ainsi que pour les graphes d'intervalles circulaires. Les étiquettes sont de taille $O(\log n)$ bits.
- Un algorithme également en $O(n)$ calculant un étiquetage de distance pour les graphes de

permutations avec des étiquettes d'au plus $O(\log n)$ bits. Une borne inférieure absolue de $\Omega(n^{1/3})$ sur la taille des étiquettes montre qu'il n'y a aucun espoir d'étendre le résultat aux graphes de comparabilité d'ordre partiel de dimension ≥ 3 (les graphes de permutation étant les graphes de comparabilité des ordres partiels de dimension < 3).

- Un algorithme en $O(n)$ qui calcule le plus petit ancêtre commun entre deux sommets d'un arbre avec des étiquettes de taille au plus $O(\log n)$.
- De nombreux schémas d'étiquetage approximant la distance pour de nombreuses familles de graphes.

Routage compact

Dans les réseaux de grande taille il est important que le routage point-à-point (entre deux routeurs) puisse être effectué à l'aide de tables les plus *simples* possibles. Il n'est en effet pas concevable d'archiver la totalité de la structure du réseau dans chacun des routeurs. Ce point de vue est particulièrement critique pour les applications récentes comme le pair-à-pair (ici le réseau est un réseau logique). Il a été suggéré à la fin des années 1970 qu'il existe un compromis entre la *taille* des tables de routage et le *facteur d'étirement*, le ratio maximum entre la longueur de la route et la longueur de la route la plus courte. Ce compromis est très général puisqu'il s'applique à tous les graphes à n sommets. *Grosso modo*, il énonce qu'il est toujours possible de construire des tables de routage de $\tilde{O}(n^{1/k})$ bits/sommets¹ pour un étirement en $O(k)$, où k est un paramètre entier dépendant du niveau de compression choisi.

En fait ces bornes ne sont pas fines du tout, une partie de la difficulté du problème étant caché dans la notation O . Par exemple, pour $k = 2$, c'est-à-dire avec une borne mémoire de $\tilde{O}(\sqrt{n})$ bits/sommets, les premiers schémas des années 1990 produisaient un étirement de $1024k = 2048$. De plus la borne $\tilde{O}(\sqrt{n})$ sur la mémoire n'était atteinte qu'en moyenne, certain sommet pouvant se voir assigner $\Omega(n)$ bits. Parfois, les résultats ne s'appliquaient que pour des graphes valués polynomialement (la valeur du poids sur les arêtes étant bornée par un polynôme en n). Ainsi un graphe valué arbitrairement avec un nombre borné de sommets pouvait se voir attribuer des tables de routage de taille non bornée! Enfin, la plupart des schémas, pendant l'initialisation et la construction des tables, ne conservent pas le nom original des sommets. On parle alors de schémas étiquetés, dans le cas contraire on parle de schémas avec *indépendance des noms*. Cette variante est bien évidemment plus délicate à gérer et produit des tables moins compactes, puisque le "nom" de la destination n'apporte aucune information implicite sur sa localisation (coordonnées, niveau hiérarchique dans une décomposition, etc.) ce que permettent les schémas étiquetés. Nos résultats sont les suivants :

- Un schéma garantissant $\tilde{O}(\sqrt{n})$ bits/sommets avec indépendance des noms et un facteur d'étirement de 3, ce qui est optimal à la fois pour le facteur d'étirement et le nombre de bits.
- Une extension du résultat précédent pour des tables plus compactes avec $\tilde{O}(n^{1/k} \log D)$ et un facteur $O(k)$ avec encore indépendances des noms, où D est le diamètre valué du graphe. Malheureusement le schéma n'est efficace que pour les graphes polynomialement valués, $D = 2^{\Omega(n)}$ étant possible en théorie.
- Un résultat sur la taille des tables de routage pour les grilles aléatoires (une grille avec une probabilité d'existence des arêtes = $1/2$), basé sur la méthode du routage par intervalles.

¹La notation \tilde{O} est similaire à la notation O à un facteur poly-logarithmique près.

Modélisation et communication dans les réseaux

Les graphes petits-mondes

Récemment, il a été montré que de nombreux graphes modélisant des données réelles issues de domaines aussi variés que la topologie de réseaux, les réseaux d'interaction protéines-protéines ou les réseaux sociaux partagent de nombreuses propriétés communes : densité globale faible, densité locale forte, petit diamètre, ... Les graphes possédant ces propriétés sont nommés "petits-mondes". Nous nous intéressons donc à des méthodes de construction ou de génération de telles familles de graphes ainsi qu'à ses implications algorithmiques pour des applications sur le routage ou le partitionnement de graphes. En effet, il n'existe quasiment pas jusqu'à ce jour d'algorithmes dédiés aux graphes petits-mondes.

Nos résultats sont les suivants :

- Construction déterministe d'une famille de graphes petits-mondes ;
- Etude du routage dédié aux graphes petits-mondes ;
- Proposition de méthodes pour transformer un graphe quelconque en graphe petit-monde.

Les deux derniers point traitent d'algorithmique de routage et de construction de graphes petits-mondes. En effet Kleinberg a montré en 2000 qu'il est facile et rapide de router dans une certaine famille de graphes petits-mondes, appelés petits-mondes "navigables". Dans ces graphes en supposant que chaque sommet ne connaisse que ses voisins dans le graphe ainsi que la distance aux autres sommets dans le graphe de base, un routage glouton trouve sa cible en un nombre polylogarithmique de sauts.

On peut construire une famille de graphes petits-mondes navigables en partant d'un graphe de base (ici une grille) et en ajoutant un ensemble d'arêtes construites de manière aléatoire (mais pas uniforme). Nous avons montré que cette construction est généralisable à de nombreux graphes de base. En effet, au lieu de prendre la grille, nous pouvons prendre n'importe quel graphe pour lequel la croissance des boules (sous-ensemble de sommets à distance au plus r d'un sommet donné) est modérée (croissance polynomiale). Nous avons décrit la loi de distribution des arêtes aléatoires à rajouter en considérant que nous augmentons le graphe de base à l'aide d'une arête aléatoire par sommet.

Réseaux et télécommunication

Le développement exponentiel des réseaux de communications (internet ou de télécommunications) et l'explosion du volume de données rendues ainsi disponibles drainent des problèmes pouvant être modélisés et explorés via des outils mathématiques tels que la théorie des graphes, l'optimisation combinatoire ou encore l'algorithmique. Notre recherche porte sur les points suivants :

- les réseaux dits « pair-à-pair » – Le groupe de recherche « Graphes et Applications » est associé au projet PairAPair² de l'ACI Masses de données. Ce projet a pour origine le constat de l'essor des réseaux pair-à-pair. Ces réseaux logiques ("overlay networks") peuvent en effet permettre le partage de ressources de tout type : fichiers, calcul, etc, et ce à très grande échelle. Cependant, malgré les énormes potentialités des réseaux pair-à-pair, les travaux de recherche sur le sujet n'en sont encore qu'à leur début.
- la conception de réseaux tolérants aux pannes – Certains satellites de télécommunications reçoivent des signaux qui sont routés à travers un réseau d'interconnexion embarqué, vers des amplificateurs qui réémettent ces signaux vers la terre. Pendant la durée de

²<http://gyroweb.inria.fr/pairapair/>

vie du satellite, des amplificateurs peuvent tomber en panne de façon définitives. Ces réseaux doivent alors posséder une propriété de tolérance aux pannes qui s'exprime de la façon suivante : quelque soit l'ensemble de k sorties en panne, les signaux entrant dans le satellite doivent toujours pouvoir être routés vers les amplificateurs valides restants et ceci en minimisant le nombre de commutateurs composant le réseau d'interconnexion embarqué dans le satellite. Ce problème, issu de l'industrie spatiale, nous fut proposé par ALCATEL SPACE INDUSTRIE. Sur ces problèmes de tolérance aux pannes, nous avons d'ores et déjà obtenu des résultats significatifs, notamment :

- des théorèmes constructifs de réseaux optimum ou quasi optimum (réseaux pouvant être réalisés). Certains satellites construits par Alcatel Space (basés sur nos premiers résultats, i.e. pour des réseaux composés de commutateurs de degré 4) ont été lancés en orbite géostationnaire ;
- la réalisation d'un logiciel de vérification de la validité des réseaux de petites tailles ;
- une meilleure compréhension des propriétés des réseaux, notamment via l'établissement de bornes inférieures.

Désormais nous cherchons à étendre nos résultats, afin d'obtenir des théorèmes, si possibles constructifs, encore plus généraux, i.e., quelque soit le nombre de pannes et quelque soit le degré des commutateurs. Des variantes à ce problème (elles aussi proposées par Alcatel Space) comme par exemple la présence de signaux prioritaires par rapport à d'autres, sont aussi considérées.

Dessin et visualisation de graphes

Dessin de graphes

Les algorithmes de dessins de graphes ont principalement deux applications : la visualisation d'information (par exemple interactions de protéines, graphes sémantiques, ...) et la conception de circuits électroniques. Suivant le domaine d'applications, plusieurs critères de qualité sont considérés pour les dessins des graphes : le nombre de croisements d'arêtes (lorsque le graphe n'est pas planaire), la taille de la grille (les sommets ayant coordonnées entières), la longueur des arêtes... Parallèlement certaines propriétés peuvent être également souhaitées : les arêtes doivent être des lignes droites ou alors être une suite de segments horizontaux ou verticaux, les faces doivent être des polygones convexes, ... Dans ce contexte, nous avons mis au point un algorithme permettant de dessiner n'importe quel graphe planaire sur des grilles de taille $(n - 2 - D) \times (n - 2 - D)$ (où D est un paramètre dépendant du graphe). De plus si le graphe est triconnexe, les faces sont des polygones convexes. Ce résultat améliore les algorithmes existants en termes de taille de grille. Cet algorithme est le fruit d'une étude profonde des propriétés des arbres de Schnyder des graphes plans triconnexe.

Réseaux sémantiques

Nous avons au cours de ces dernières années étudié les réseaux sémantiques, cartographiant les éléments d'un système d'information et leurs relations. Plus particulièrement, nous nous sommes intéressés à des réseaux sémantiques codés à l'aide de la norme Topic Map, en coopération avec la société Mondeca (Paris). D'un point de vue mathématique, une topic map peut être vue comme un hypergraphe, les noeuds représentant les sujets d'information et les hyperarêtes leurs relations.

Nous avons dans ce cadre tout d'abord étudié le problème du fractionnement des graphes, avec comme objectif la possibilité de traiter de grands graphes dans un temps raisonnable,

c'est à dire en utilisant des algorithmes linéaires ou quasi-linéaires. Trois algorithmes, adaptés d'algorithmes existant, ont ainsi été proposés et comparés. Cette première étude nous a amené à constater que la présence d'une couche "ontologie" dans les topic maps rendait moindre l'intérêt de méthodes classiques de fractionnement de graphes et nous a conduit à rechercher un modèle formel pour les hypergraphes issus des topics maps et les propriétés exprimées mathématiquement que doivent posséder ces hypergraphes pour être valides d'un point de vue sémantique.

Un modèle obtenu en coopération avec des chercheurs du CAMS (EHES) a été ainsi présenté à la communauté "Topic Maps" et des propriétés garantissant l'existence d'algorithmes polynomiaux, voire linéaires, tout en ayant un "sens" d'un point de vue sémantique proposé. La coopération avec la société Mondeca ayant pris fin, nous étudions actuellement le problème de la navigation dans un graphe, qui mêle des questions de dessin et de visualisation de graphes, mais aussi d'animation.

Projets et perspectives 2005-2008

Nous donnons les points que nous voulons développer dans les prochaines années dans nos différentes activités de recherche.

Propriétés structurelles

– *Décomposition arborescente*

Nous pensons poursuivre notre recherche sur les points suivants : établir la complexité exacte du calcul de la longueur-arborescente, établir des liens entre la longueur-arborescente et d'autres invariants de graphes ce basant sur la géométrie (graphes plongés sur des courbures négatives, graphes hyperboliques, *etc.*)

Colorations

– *Coloration d'orienté*

La notion de graphe représentatif des arcs donne de façon naturelle (c'est-à-dire comme dans le cas non orienté) une nouvelle notion d'arc-coloration des graphes orientés. Ce nouveau paramètre est lié à des paramètres déjà considérés dans la littérature, comme par exemple l'étoile-arboricité (ou étoile-coloration) des graphes orientés. Il serait naturellement intéressant d'exploiter notre expérience dans l'étude des colorations de graphes orientés pour étudier ce nouveau type de coloration de graphes d'une part, mais également pour aborder d'une façon originale plusieurs questions encore ouvertes à ce jour sur des paramètres qui lui sont reliés.

– *Coloration d'incidence*

De nombreuses questions restent posées concernant les colorations d'incidences. On peut notamment citer le problème de la caractérisation des familles de graphes dont le nombre chromatique d'incidence est majoré par le degré maximal du graphe à une constante près, ou l'étude du comportement du paramètre nombre chromatique d'incidence vis à vis de certaines opérations classiques de graphes (notamment les différents produits de graphes). L'étude de la complexité des problèmes de coloration d'incidence est également un domaine qui n'a pas encore été abordé.

– *Coloration par liste*

Les résultats que nous avons sur les colorations acycliques par listes nous induisent à proposer des nombreuses conjectures que nous essayons de résoudre. Par exemple une conjecture sur la quelle nous voulons travailler est la suivante : tout graphe planaire sans triangle peut être colorié acycliquement à partir de toute affectation aux sommets du graphe de listes de couleurs de taille 6. Nous avons montré que pour tout graphe planaire on peut le faire avec des listes de taille 7 et on conjecture qu'on peut le faire avec des listes de taille 5. Ce serait donc un premier pas vers la preuve de cette conjecture.

– *Coloration $(d, 1)$ -totale*

Nous avons attaqué la conjecture suivante de F. Havet : Tout graphe de degré maximum au plus 3 admet une coloration $(2, 1)$ -totale utilisant les couleurs $\{0, 1, 2, 3, 4, 5\}$. Nous avons déjà fait des calculs par ordinateur pour des graphes cubiques ayant peu de sommets (< 18). Les résultats obtenus nous laissent espérer la véracité de cette conjecture. La coloration $(d, 1)$ -totale est liée à la coloration $L(p, q)$ (p et q sont des entiers). Dans la coloration $L(p, q)$, deux sommets voisins ont des couleurs à distance p et deux sommets à distance 2 ont des couleurs à distance q . Nous voulons porter notre attention sur les colorations $L(p, q)$ orientées. Elles correspondent à des assignations de fréquences dans les réseaux ayant des liens de communications unidirectionnels.

– *Coloration circulaire et flot circulaire*

- La notion de flot circulaire est la notion duale de la notion de coloration circulaire. Nous nous intéressons à une conjecture de Jaeger qui peut se formuler en termes de flots circulaires. Il est conjecturé que tout graphe $4k$ - arête-connexe admet un $(2k + 1, k)$ -flot. Cette conjecture est considérée dans le cas des graphes planaires dans sa version duale de coloration circulaire. Nous nous intéressons à cette conjecture et voulons mettre en avant le lien qui existe entre flot circulaire et longueur de la couverture des graphes par des cycles. Avec X. Zhu nous avons déjà des résultats que nous comptons étendre aux matroïdes. Le lien entre ces travaux et le conjecture sur la double couverture par des cycles est tout a fait pertinent, nous allons porter nos efforts sur cette problématique.
- En 2001, X. Zhu a introduit la classe des graphes circulaires parfaits. Cette classe de graphes est une extension de la classe des graphes parfaits, une classe de graphes centrale dans le domaine de la théorie des graphes. Contrairement aux graphes parfaits, peu de choses sont connues sur les graphes circulaires parfaits, du fait de leur jeunesse. Nous avons commencé l'étude de ces graphes et cette problématique est un vaste champ d'investigation.

Flots antisymétriques

Les bornes que nous avons trouvées pour les flots antisymétriques des graphes planaires ont été améliorées. Mais l'optimalité n'est pas prouvée. Donc il serait intéressant de parfaire notre connaissance des bornes. De plus pour les graphes qui ne sont pas planaires nous sommes intéressés par les bornes pour certaines classes de graphes. Nous voulons aussi faire le lien entre flots antisymétriques et couverture des graphes par des cycles, en passant aux couvertures orientées.

Les applications flot- et tension-continues

Ce nouveau cadre mis en place nous permet de re-considérer les conjectures classiques sous un autre angle. Il y a de nombreuses questions ouvertes auxquelles nous voulons répondre. Une perspective intéressante est le prolongement de ce travail aux matroïdes.

Struture de données compactes– *Codage compact de graphes*

Nos constructions permettent d'envisager de nouveaux résultats sur d'autres familles de graphes : des sous-familles de graphes planaires comme les graphes planaires 4-connexes et les graphes k -planaires extérieurs. Ces travaux n'ont pas encore commencé mais nous sommes convaincus de leur faisabilité. D'autre part, nous désirons étendre le codage compact de graphes statiques aux graphes dynamiques. De très nombreuses applications sont alors envisageables car les structures déformables se retrouvent dans des thématiques liées au dessin de graphes et à la visualisation temps réel de graphes.

– *Étiquetage*

Nous poursuivons nos efforts pour donner de nouvelles bornes sur l'étiquetage d'adjacence (restriction de l'étiquetage de distance) pour certaines classes de graphes, notamment les graphes planaires et les graphes toriques où la taille optimale des étiquettes est encore inconnue. De nouvelles techniques, que nous avons développées, laissent envisager une borne supérieure en $2 \log n$ bits par étiquette, alors que les meilleures bornes actuellement connues sont d'au moins $3 \log n$.

– *Routage compact*

Nous envisageons un schéma de routage sans la dépendance (même logarithmique) de la valuation. Si ce résultat était prouvé, cela répondrait définitivement aux questions soulevées depuis 25 ans par le routage compact.

Modélisation et communication dans les réseaux– *Les graphes petits-mondes*

Nous espérons généraliser nos constructions de graphes petits-mondes et répondre en partie à la question : peut-on transformer tout graphe en petit-monde avec une opération d'augmentation qui ne modifie que de très peu le graphe de base ? D'autre part, nous désirons utiliser les graphes petit-monde comme structure d'indexation. L'usage est particulièrement pertinent si les objets manipulés ne peuvent pas être indexés de manière précise (par exemple une image) ou pour modéliser la dynamique d'un réseau pour lequel l'opération totale de ré-indexation (reconstruction de tables de routage par exemple) serait trop coûteuse.

Dessin et visualisation de graphes– *Dessin de graphes*

Un des problèmes ouverts qui anime la communauté du dessin de graphe depuis plusieurs années est le suivant. Peut-on dessiner n'importe quel graphe plan ayant n sommets sur

une grille $2n/3 \times 2n/3$. Il serait intéressant de voir si de telles dimensions peuvent déjà être atteintes en moyenne (sur l'ensemble des triangulations planes). L'algorithme de dessin que nous avons proposé peut-être un point de départ pour apporter une réponse à cette dernière question. Une autre perspective serait de pouvoir donner une classe, la plus large possible, de graphes pouvant être dessinés sur de telles grilles.

1.3.3 Algorithmique Distribuée

Introduction

Le thème **Algorithmique distribuée** se propose de fournir des outils permettant de coder, d'étudier et d'enseigner différents aspects de l'algorithmique distribuée. On s'intéresse plus particulièrement à certains problèmes comme : l'élection, le nommage, le calcul d'un arbre recouvrant, la détection de la terminaison, la synchronisation, le consensus ou bien la reconnaissance de certaines propriétés du réseau. Ces études sont faites d'une part en fonction des règles de calcul autorisées sur le réseau et d'autre part de la connaissance initiale ou de la configuration initiale du réseau. La non existence d'algorithmes distribués déterministes pour résoudre certains problèmes nous conduit à proposer et à étudier des algorithmes distribués probabilistes dont l'expression est souvent simple mais analyse difficile. Un autre aspect est l'étude des liens entre les trois modèles : calculs locaux, communication par messages et communication par mémoires et la comparaison de leurs puissances (idem pour les liens entre ces modèles et les systèmes à agents mobiles).

Mots-Clés Calculs locaux, élection, détection de la terminaison, reconnaissance, fibration, revêtement, algorithme probabiliste, système à agents mobiles, synchronisation.

Description des activités et principaux résultats

Les questions relatives aux systèmes distribués constituent un axe majeur de la recherche en informatique; citons par exemple la conception et le développement d'architectures distribuées, la définition et le développement d'environnements de programmation distribués, la description et la validation de programmes distribués ou bien encore l'étude des communications dans les systèmes distribués. Un moyen décisif est la maîtrise des mécanismes fondamentaux qui y interviennent; cela passe par :

- la définition et l'étude de différents modèles rendant compte de phénomènes qui s'y produisent,
- le développement d'outils permettant de visualiser, de tester et, plus généralement, d'aider à la mise au point et aux preuves de systèmes distribués,
- la mise en œuvre de plate-formes distribuées afin de vérifier et d'expérimenter.

Ces deux derniers points sont traités dans le projet VISIDIA.

1.3.4 Modélisation

Un système distribué est un ensemble de processeurs (processus) qui peuvent interagir. Il existe principalement 3 modèles d'interactions :

- le modèle avec communication par messages,
- le modèle avec communication par mémoires partagées,
- le modèle des calculs locaux.

Dans ces trois modèles, les processus sont représentés par les sommets d'un graphe et les interactions par des arêtes. Pour les modèles à base de communications, les processus

interagissent à travers des primitives de communication : des messages peuvent être envoyés le long des liens ou bien des opérations de lecture/écriture peuvent être exécutées dans des registres associés aux arêtes. Dans le modèle des calculs locaux les interactions sont définies par des réécritures de graphes. Ces différents modèles (et sous-modèles) codent différentes architectures de systèmes, différents niveaux de synchronisation et différents niveaux d'abstraction.

L'existence et l'expression des algorithmes distribués dépendent d'hypothèses techniques sur les réseaux concernant le mode de communication, la synchronisation partielle de processeurs voisins, l'anonymat des processeurs, l'état initial des processeurs ou bien encore de la connaissance initiale que l'on a sur le réseau. Les cadres classiques (comme par exemple CSP) rendent difficile la lisibilité de certains algorithmes, ils ne permettent pas toujours de comprendre leur puissance et leurs limites et de faire des preuves formelles de leurs propriétés.

Notre équipe étudie un formalisme original dans ce domaine : les réécritures de graphes. De fait, il constitue un prolongement des travaux menés par P. Rosenstiehl et al. dans les années 70, par D. Angluin dans les années 80 et plus récemment par Yamashita et Kameda concernant les propriétés locales et les propriétés globales dans les réseaux et les graphes. Un système distribué est codé par un graphe étiqueté : les sommets correspondent aux processeurs, les arêtes aux liens d'interactions et les étiquettes associées codant l'état du processeur ou du lien. Une règle de calcul est définie par la donnée d'un graphe connexe et de deux étiquetages de ce graphe, donc elle est appliquée localement.

Les principaux types de règle que nous étudions sont décrits sur la figure 1.

Un système de réécriture est défini par la donnée d'un ensemble dénombrable de telles règles. On considère des exécutions asynchrones : il n'y a pas d'horloge globale et deux pas de réécriture peuvent avoir lieu en parallèle si ils modifient des parties disjointes de leurs supports. Le comportement du réseau est défini par un étiquetage initial et par une suite de pas de calcul.

Remarque Sur la figure 1, les règles du type 7 correspondent au modèle étudié par Mazurkiewicz, celles du type 4 sont très proches du modèle étudié par Angluin et celles du type 5 sont un cas particulier du modèle étudié par Boldi et Vigna. Le modèle étudié par Rosenstiehl est synchrone et correspond aux réécritures du type 5.

Dans les cas où il n'existe pas d'algorithmes déterministes ou bien par souci d'efficacité, nous nous intéressons aux algorithmes probabilistes qui permettent d'apporter des réponses satisfaisantes d'un point de vue pratique. D'autre part, certains algorithmes distribués permettent d'obtenir un objet appartenant à un ensemble donné (par exemple, étant donné un graphe G on peut sous certaines hypothèses, calculer un arbre recouvrant T) il s'agit alors de déterminer la probabilité pour obtenir un objet particulier.

Ainsi nous sommes amenés à introduire et à analyser des algorithmes probabilistes ou bien à analyser la probabilité d'obtention d'une certaine solution.

Les questions étudiées et les principaux résultats obtenus

Étude et comparaison de différents modèles

Dans notre travail, on étudie le codage et les preuves des algorithmes distribués classiques ainsi que certains algorithmes de la théorie des graphes. L'une des principales questions considérées en algorithmique distribuée reste l'étude de la frontière entre les domaines du calculable et du non-calculable, selon les hypothèses faites sur le réseau ou la connaissance que l'on en a. Notre formalisme nous a permis d'obtenir des résultats nouveaux dans cette direction en considérant plus particulièrement 5 problèmes fondamentaux :

- le calcul d'un arbre recouvrant,

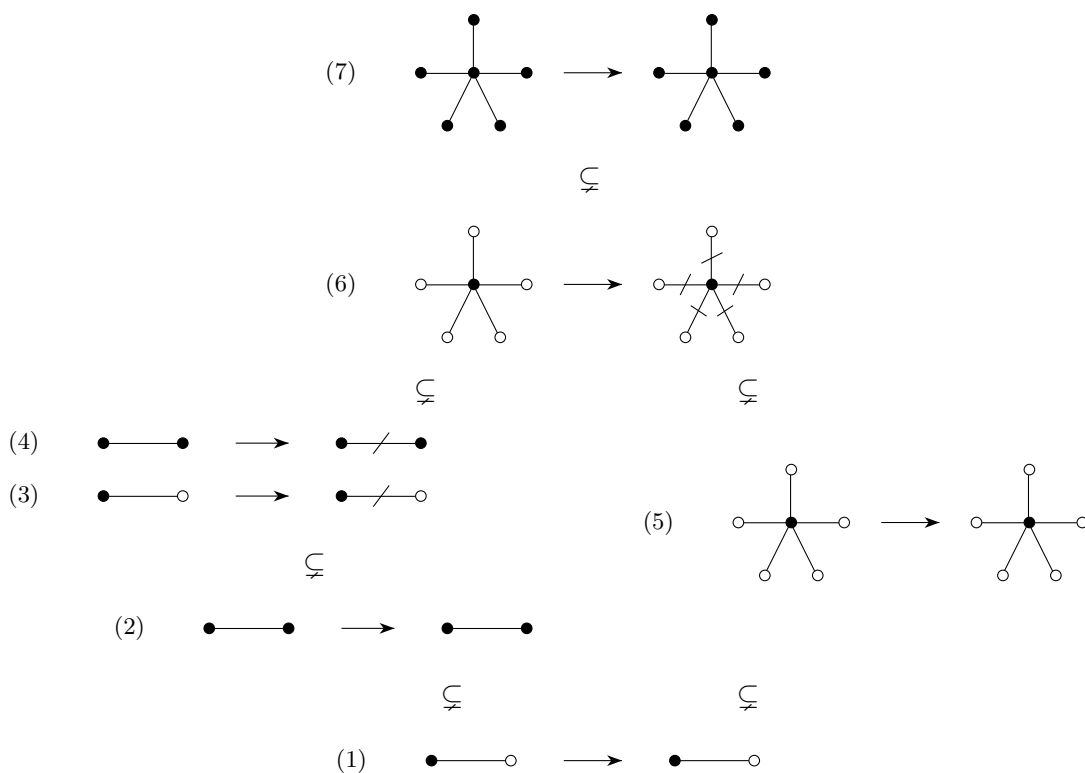


FIG. 1.1 Étant donnée une règle, les sommets (blancs ou noirs) et les arêtes formant le membre gauche de cette règle constituent le support de la règle. Les sommets noirs sont actifs : leurs étiquettes peuvent changer quand la règle est appliquée. Les sommets blancs sont passifs : leurs étiquettes conditionnent l'application de la règle mais ne peuvent être modifiées par la règle. Dans les modèles où les arêtes sont marquées, les arêtes sont étiquetées et l'application de la règle peut modifier leurs étiquettes. Cette figure montre la hiérarchie entre ces différents types de règles pour le problème de l'élection dans un réseau.

- le problème de l'élection, qui consiste à particulariser un processeur dans un réseau,
- le problème du nommage qui consiste à attribuer à chaque processeur une identité propre,
- le problème de la reconnaissance, qui consiste à déterminer certaines propriétés d'un réseau,
- le problème de la détection locale de la terminaison globale d'un algorithme.

Mazurkiewicz a donné une caractérisation des graphes pour lesquels il existe un algorithme d'élection avec des règles du type 7. Nous avons obtenu une caractérisation pour tous les autres types de règle aussi bien pour le problème de l'élection que pour le problème du nommage. Ces caractérisations reposent d'une part sur des adaptations non triviales de l'algorithme de Mazurkiewicz et d'autre part sur la définition de "bons morphismes" du type revêtement ou fibration. Nous avons obtenu également une caractérisation des familles de graphes pour lesquelles il existe un algorithme d'élection pour le modèle de Mazurkiewicz. Cette caractérisation utilise une notion nouvelle : les quasi-revêtements. Ce résultat fournit un algorithme d'élection pour des graphes en ne connaissant qu'une borne de leur taille (donc un algorithme plus robuste et dynamique). Pour ce modèle nous avons obtenu une caractérisation des familles de réseaux où une terminaison implicite (l'algorithme est terminé mais aucun sommet ne le sait) peut être transformée en terminaison explicite (l'algorithme est fini et au moins un sommet le sait). Nous avons proposé une notion de reconnaissance de familles de graphes par calculs locaux. Nous avons obtenu des caractérisations suivant que l'on a ou que l'on n'a pas d'informations sur les graphes ; des applications aux familles de graphes définis par mineurs exclus ont été données.

Nous avons établi un pont entre les résultats et les techniques utilisées par Yamashita et Kameda dans le modèle de communication asynchrone et les résultats et les techniques utilisées dans le modèle des calculs locaux.

Nous avons obtenu des caractérisations de connaissances équivalentes au sens où l'une peut être calculée à partir de l'autre. Nous avons donné sous forme de calculs locaux des structures de synchronisation dans les systèmes distribués.

Nous avons obtenu un résultat d'équivalence entre un algorithme distribué se déroulant sur un système distribué et un système à agents mobiles travaillant sur le même système distribué. Ce résultat se prouve simplement mais il permet de transporter tous les résultats obtenus dans le cadre des algorithmes distribués aux systèmes à agents mobiles : élection, rendez-vous, cartographie, reconnaissance...

Les systèmes à agents mobiles permettent d'aborder de façon naturelle des problèmes liés à l'existence d'ordre sur les étiquettes comme le montre l'exemple suivant. On considère une équipe internationale d'automobilistes dispersés dans un réseau autoroutier reliant différentes villes. Chaque automobiliste est initialement dans sa ville natale et ces villes sont reconnaissables par tous les automobilistes. On cherche à déterminer s'il existe un algorithme, que tous les automobilistes exécutent, et permettant à l'ensemble des automobilistes de se retrouver dans une seule ville.

Le seul moyen qu'un automobiliste a pour communiquer avec les autres est de laisser des messages dans chaque ville traversée. La difficulté étant que ces différents automobilistes sont tous de nationalités différentes et ne comprennent que leur langue maternelle. Néanmoins, lorsqu'ils voient un message écrit dans une langue qui n'est pas la leur, ils arrivent à reconnaître quelle langue a été utilisée, et ainsi comprendre quelle est la nationalité de l'automobiliste qui a laissé ce message.

De plus, dans chaque ville, les directions vers les autres villes sont écrites dans la langue du pays, et les automobilistes ne peuvent pas se mettre d'accord sur un ordre sur les différentes directions.

La configuration initiale peut être modélisée par un graphe étiqueté (G, λ) , où le graphe sous-jacent correspond au réseau autoroutier et où seules les villes de départ sont marquées (i.e., si un sommet v correspond à une ville de départ, alors $\lambda(v) = 1$ et $\lambda(v) = 0$ sinon). On montre qu'il existe un algorithme qui permet de résoudre le problème dans un tel environnement si et seulement si il n'existe pas d'automorphisme σ de (G, λ) et une constante $q > 1$ tels que pour tout sommet v , $|\{\sigma^k(v) \mid k \in \mathbb{N}\}| = q$.

De plus, on donne un algorithme qui pour toute instance de ce problème permet à chaque automobiliste de détecter qu'il est impossible de résoudre le problème, ou bien qui permet aux automobilistes de se rencontrer dans une unique ville.

Analyse d'algorithmes

Le résultat d'un algorithme distribué est un élément de l'ensemble S des solutions possibles (par exemple l'élection d'un sommet ou bien le calcul d'un arbre recouvrant) ; ceci donne naturellement naissance à des questions du type : étant donné un élément e de S , calculer la probabilité pour que le résultat de l'algorithme distribué soit e . Plus généralement on étudie comment contrôler localement le déroulement d'un algorithme distribué pour en guider le résultat. Ainsi, l'implémentation effective d'un algorithme d'élection dans les arbres passe par l'implémentation d'actions suivant une loi de probabilité : par exemple l'effeuillage d'un arbre conduisant à l'élection d'un sommet repose sur la disparition des feuilles, cette disparition peut être décidée localement par un jeu de pile ou face sur chaque feuille. Parmi les paramètres à calculer on peut citer : le temps moyen pour une élection, propriétés du sommet ayant la plus grande probabilité d'être élu. Nous avons défini un paramètre qui peut être calculé localement et qui permet de donner la même chance à tous les sommets d'un arbre d'être élus (alors qu'intuitivement on pense que les feuilles ont moins de chance d'être élues que le centre). Ces résultats ont été étendus aux familles des k -arbres et des polyomonoïdes.

Algorithmes probabilistes

L'implémentation automatique d'un algorithme décrit à l'aide de règles de réécriture nous amène, en particulier, à étudier les algorithmes probabilistes permettant de synchroniser deux sommets voisins ou bien un sommet avec l'ensemble de ses voisins et plus généralement un ensemble de sommets correspondant à une règle de réécriture. On considère un réseau asynchrone avec échange de messages en mode asynchrone :

- un processeur envoie un message à un autre processeur en le déposant dans le canal correspondant,
- il n'y a pas de borne supérieure sur le temps mis par le message pour atteindre son but.

Dans un système avec échange de messages en mode synchrone l'émetteur et le récepteur doivent être prêts pour communiquer. Une communication a lieu si les participants sont en attente pour cette communication : ils ont un rendez-vous.

Les travaux d'Angluin montrent qu'il n'existe pas d'algorithme déterministe pour implémenter un échange de messages en mode synchrone dans un système fonctionnant par échange de messages en mode asynchrone. Il faut donc utiliser un algorithme probabiliste.

De fait, de nombreux résultats reposent sur le modèle suivant. Pendant un pas de calcul, deux sommets voisins échangent leurs états et en calculent de nouveaux. Par exemple, un algorithme d'élection est donné par Angluin pour les graphes complets ou bien des algorithmes d'élection sont décrits dans le cas des anneaux ayant un nombre premier de sommets. L'implémentation de ces algorithmes, équivalente à l'implémentation de règles de réécriture faisant intervenir deux sommets liés par une arête, passe par l'implémentation du rendez-vous.

Nous avons proposé et analysé la procédure suivante :

Chaque sommet exécute en boucle les actions suivantes :

le sommet v choisit au hasard un de ses voisins $c(v)$;
 le sommet v envoie 1 à $c(v)$;
 le sommet v envoie 0 à ses voisins différents de $c(v)$;
 le sommet v reçoit les messages de tous ses voisins.
 (* Il y a un rendez-vous entre v et $c(v)$ si v reçoit 1 de $c(v)$ *)

Un modèle probabiliste de diffusion dans un réseau a été défini à partir de cette procédure ; l'analyse de son efficacité a été faite.

Des résultats ont été obtenus sur les procédures probabilistes suivantes permettant d'implémenter des synchronisations sur des étoiles.

Chaque sommet exécute en boucle les actions suivantes :

le sommet v choisit au hasard un entier $\text{rand}(v)$;
 le sommet v envoie $\text{rand}(v)$ à ses voisins ;
 le sommet v reçoit un entier de chacun de ses voisins ;
 (* Il y a une synchronisation centrée sur v si $\text{rand}(v)$ est strictement plus grand que les entiers reçus par v .*)

Chaque sommet exécute en boucle les actions suivantes :

le sommet v choisit au hasard un entier $\text{rand}(v)$;
 le sommet v envoie $\text{rand}(v)$ à ses voisins ;
 le sommet v reçoit un entier de chacun de ses voisins ;
 soit $m_{v/w}$ le maximum des entiers que v a reçu de ses voisins différents de w ;
 le sommet v envoie à chaque voisin w l'entier $m_{v/w}$;
 le sommet v reçoit un entier de chacun de ses voisins ;
 (* Il y a une synchronisation centrée sur v si $\text{rand}(v)$ est strictement plus grand que les entiers reçus par v .*)

Nous avons étudié un modèle de diffusion fondé sur les rendez-vous des robots. Dans ce modèle, chaque robot peut choisir au hasard un de ses voisins afin d'établir un contact. Un rendez-vous a lieu entre deux robots lorsqu'ils se choisissent mutuellement. L'efficacité de l'algorithme est mesurée en nombre moyens de "tours" nécessaires pour que tout le réseau soit informé, lorsqu'une information est à transmettre. Nous avons obtenu des bornes supérieures et des bornes inférieures uniformes pour différentes topologies de réseaux.

Projets et perspectives 2005-2008

Le travail relatif aux calculs locaux dans les graphes doit se développer suivant plusieurs directions.

Il faut poursuivre l'étude des algorithmes distribués classiques et leur codage à l'aide des calculs locaux. Ceci permettra de fournir un cadre général pour leur présentation (et donc pour leur enseignement) et leur preuve. Il faut également étudier les structures de contrôle et déterminer si les deux structures (priorités, contextes interdits) que nous avons introduites sont génériques. En ce qui concerne les réécritures de graphes, la caractérisation des familles de graphes que l'on peut reconnaître, en imposant certains critères sur les conditions finales à l'aide de formules logiques, est un problème ouvert ; il en est de même pour la reconnaissance avec détection locale de la terminaison et du résultat.

Pour le problème de l'élection il faut explorer la caractérisation des familles de réseaux pour lesquelles il existe un algorithme d'élection pour les règles du type 1, 2, 3, 4, 5 et 6.

Pour les algorithmes distribués, et donc les calculs locaux dans les graphes, deux types de terminaison existent :

- terminaison locale : l'état d'un sommet ou d'une arête est invariant,
- terminaison globale : les états associés à l'ensemble des sommets et des arêtes sont invariants.

Ainsi, étant donné un problème dans une famille de réseaux et un modèle, plusieurs questions peuvent se poser. Existe-t'il un algorithme distribué permettant de le résoudre? S'il existe un tel algorithme, est-ce qu'un sommet peut détecter la terminaison locale? Est-ce qu'un sommet peut détecter la terminaison globale? Quelles sont les hypothèses minimales sur la connaissance du réseau pour que de tels algorithmes existent? L'étude du problème de l'élection dans un réseau conduit à s'interroger sur le lien entre ce problème et la détection locale de la terminaison globale.

Un autre aspect de notre travail, lié au problème de l'élection, consiste à déterminer les paramètres d'un graphe que l'on peut calculer en bornant par une constante le nombre de pas de calcul associés à un sommet. Par exemple, dans les arbres on peut calculer en une seule étape : le nombre de sommets, la topologie ou bien un médian. La même question se pose pour d'autres familles de graphes comme, par exemple, la famille des graphes triangulés.

Les preuves de terminaison des systèmes de réécriture, codant des algorithmes distribués, passent par la mise en évidence de uplets d'entiers pour lesquels il existe un bon ordre compatible avec les réécritures. Ceci permet finalement de prouver la terminaison du système. Une étude systématique de l'existence de tels uplets doit être faite, elle peut permettre une aide automatique pour trouver les « bons uplets » et ainsi pour prouver la terminaison d'algorithmes distribués.

De la même façon que l'on a établi un pont entre le modèle avec communication par message et le modèle des calculs locaux, il faut étudier les liens avec le modèle avec communication par mémoires.

L'étude des calculs locaux à travers le paradigme des agents mobiles peut permettre d'obtenir des solutions résistantes à certaines pannes et fonctionnant sur des réseaux dynamiques.

Dans le cas des arbres, nous avons montré qu'il existait un paramètre localement calculable et assurant l'égalité des chances d'être élu pour les sommets. Nous avons des paramètres pour lesquels nous conjecturons que le sommet ayant la plus grande chance d'être élu est un centre ou bien un médian. On peut également se demander le type de distribution que l'on peut implémenter ainsi.

On est amené à effectuer deux types d'analyse. Certains algorithmes distribués sont valides si ils se déroulent sur un réseau vérifiant une propriété P (par exemple le réseau est un anneau premier). Ceci étant, dans certains cas, l'algorithme peut converger même si le réseau ne vérifie pas P ; il s'agit alors d'étudier la probabilité de convergence de l'algorithme. Par exemple, Mazurkiewicz a donné un algorithme d'élection dans la famille des graphes minimaux pour la relation « être revêtement » (on suppose que le nombre de sommets est connu). Si l'on applique cet algorithme à des graphes non minimaux pour la relation « être revêtement » on peut « parfois » obtenir un résultat valide. Le problème général qui se pose est d'étudier la probabilité d'obtenir un résultat valide.

La traduction automatique d'un système de réécriture en un algorithme distribué utilisant des messages ou bien des mémoires nécessite l'étude de la complexité des algorithmes ainsi générés.

1.4 Publications de l'équipe

1.4.1 Publications d'audience internationale

Reuves avec Comité de lecture

- [1] ALSTRUP (S), GAVOILLE (C), KAPLAN (H) et RAUHE (T). – Nearest common ancestors : A survey and a new algorithm for a distributed environment. *Theory of Computing Systems*, vol. 37, 2004, pp. 441–456.
- [2] AMAR (D), BRITO (D) et ORDAZ (O). – Hamiltonian cycles and hamiltonian bi-connectedness in bipartite digraphs. *Divulgaciones Matemáticas*, À paraître.
- [3] AMAR (D), FLANDRIN (E), GANCARZEWICZ (G) et WOJDA (A). – Bipartite graphs with every matching in a cycle. *Discrete Math.*, À paraître.
- [4] ANDREWS (G), LEWIS (R) et LOVEJOY (J). – Partitions with designated summands. *Acta Arith.*, vol. 105, 2002, pp. 51–66.
- [5] AUBER (D). – Using strahler numbers for real time visual exploration of huge graphs. *International Journal of Applied Mathematics and Computer Science*, 2002, p. à paraître.
- [6] AUBER (D) et DELEST (M). – A clustering algorithm for huge trees. *Advances in Applied Mathematics*, vol. 31, n1, 2003, pp. 46–60.
- [7] AVAL (J.-C). – On certain spaces of lattice diagram determinants. *Discrete Mathematics*, vol. 256, n3, 2002, pp. 557–575.
- [8] AVAL (J.-C). – Polynômes quasi-invariants et super-coinvariants pour le groupe symétrique généralisé. *Annales des Sciences Mathématiques du Québec*, vol. 27, No.2, 2003, pp. 111–121.
- [9] AVAL (J.-C), BERGERON (F) et BERGERON (N). – Spaces of lattice diagram polynomials in one set of variables. *Advances in Applied Mathematics*, vol. 28, n3-4, 2002, pp. 343–359. – Special issue in memory of Rodica SIMION.
- [10] AVAL (J.-C), BERGERON (F) et BERGERON (N). – Ideals of quasi-symmetric functions and super-covariant polynomials for s_n . *Advances in Mathematics*, vol. 181, No. 2, 2004, pp. 353–367.
- [11] AVAL (J.-C) et BERGERON (N). – Vanishing ideals of lattice diagram determinants. *J. Combin. Theory Ser. A*, vol. 99, n2, 2002, pp. 244–260.
- [12] AVAL (J.-C) et BERGERON (N). – Catalan paths and quasi-symmetric functions. *Proc. Amer. Math. Soc.*, vol. 131, n4, 2003, pp. 1053–1062 (electronic).
- [13] AVAL (J.-C) et BERGERON (N). – Schur partial derivative operators. *European Journal of Combinatorics*, À paraître.
- [14] BALOGH (J), OCHEM (P) et PLUHÁR (A). – On the interval number of special graphs. *J. Graph Theory.*, vol. 46 (4), 2004, pp. 241–253.
- [15] BANDERIER (C), BOUSQUET-MÉLOU (M), DENISE (A), FLAJOLET (P), GARDY (D) et GOUYOU-BEAUCHAMPS (D). – Generating functions for generating trees. *Discrete Math.*, vol. 246, n1-3, 2002, pp. 29–55.
- [16] BARRIOT (R), POIX (J), GROPPA (A), BARRE (A), GOFFARD (N), SHERMAN (D), DUTOUR (I) et DE DARUVAR (A). – New strategy for the representation and the integration of biomolecular knowledge at a cellular scale. *Nucl. Acids Res.*, vol. 32, n12, 2004, pp. 3581–3589.

- [17] BARTH (D), BAUDON (O) et PUECH (J). – Decomposable trees : a polynomial algorithm for tripodes. *Discrete Applied Mathematics*, vol. 119, n3, 2002, pp. 205–216.
- [18] BAZZARO (F), MONTASSIER (M) et RASPAUD (A). – $(d, 1)$ -total labelling of planar graphs with large girth and high maximum degree. *Discrete Mathematics*, 2005. – à paraître.
- [19] BERMOND (J.-C), DARROT (É) et DELMAS (O). – Design of fault-tolerant networks for satellites (TWTA redundancy). *Networks*, vol. 40, n4, 2002, pp. 202–207. – Wiley InterScience.
- [20] BONICHON (N). – A bijection between realizers of maximal plane graphs and pairs of non-crossing dyck paths. *Discrete Mathematics*, accepted. – FPSAC'02 Special Issue.
- [21] BONICHON (N) et MOSBAH (M). – Watermelon uniform random generation with applications. *Theoretical Computer Science*, vol. 307, n2, 2003, pp. 241–256.
- [22] BORODIN (O), FON-DER-FLAASS (D), KOSTOCHKA (A), RASPAUD (A) et SOPENA (E). – Acyclic list 7-coloring of planar graphs. *Journal of Graph Theory*, vol. 40, No. 2, 2002, pp. 83–90.
- [23] BORODIN (O), FON-DER-FLAASS (D), KOSTOCHKA (A), RASPAUD (A) et SOPENA (E). – Acyclic list 7-colouring of planar graphs. *J. Graph Theory*, vol. 40(2), 2002, pp. 83–90.
- [24] BORODIN (O), GLEBOV (A), RASPAUD (A) et SALAVATIPOUR (M). – Planar graphs without cycles of length from 4 to 7 are 3-colorable. *J. Combin. Theory Ser. B.*, vol. 93, 2005, pp. 303–311.
- [25] BORODIN (O) et RASPAUD (A). – A sufficient condition for planar graphs to be 3-colorable. *J. Combin. Theory Ser. B.*, vol. 88, 2003, pp. 17–27.
- [26] BOUSQUET-MÉLOU (M). – Four classes of pattern-avoiding permutations under one roof : generating trees with two labels. *Electron. J. Combin.*, vol. 9, n2, 2002/03, pp. Research paper 19, 31 pp. (electronic).
- [27] BOUSQUET-MÉLOU (M). – Algebraic generating functions in enumerative combinatorics, and context-free languages. *Lect. Notes Comput. Sci.*, vol. 3404, 2005, pp. 18–35.
- [28] BOUSQUET-MÉLOU (M). – Walks in the quarter plane : Kreweras' algebraic model. *Annals of Applied Probability*, à paraître.
- [29] BOUSQUET-MÉLOU (M) et PETKOVŠEK (M). – Walks confined in a quadrant are not always D-finite. *Theoret. Comput. Sci.*, vol. 307, n2, 2003, pp. 257–276.
- [30] BOUSQUET-MÉLOU (M) et RECHNITZER (A). – Lattice animals and heaps of dimers. *Discrete Math.*, vol. 258, n1-3, 2002, pp. 235–274.
- [31] BOUSQUET-MÉLOU (M) et RECHNITZER (A). – The site-perimeter of bargraphs. *Adv. in Appl. Math.*, vol. 31, n1, 2003, pp. 86–112.
- [32] BOUSQUET-MÉLOU (M) et SCHAEFFER (G). – Walks on the slit plane. *Probab. Theory Related Fields*, vol. 124, n3, 2002, pp. 305–344.
- [33] BOUSQUET-MÉLOU (M) et STEINGRIMSSON (E). – Decreasing subsequences in permutations and Wilf equivalence for involutions. *J. Algebraic Combinatorics*, à paraître.
- [34] BRLEK (S), MENDÈS FRANCE (M), ROBSON (J) et RUBEY (M). – Cantorian tableaux and permanents. *L'Enseignement Mathématique*, vol. 50, 2004, pp. 287–304.

- [35] CHALOPIN (J) et LEUNG (H). – On factorization forests of finite height. *Theoretical Computer Science*, vol. 310, 2004, pp. 489–499.
- [36] CHALOPIN (J), MÉTIVIER (Y) et ZIELONKA (W). – Local computations in graphs : the case of cellular edge local computations. *Fundamenta informaticae*, à paraître.
- [37] CHARRON-BOST (B) et CORI (R). – A note on linearizability and the global time axiom. *Parallel Processing Letters*, vol. 13, n1, 2003, pp. 19–24.
- [38] CHAUVE (C) et DULUCQ (S). – A geometric version of the Robinson-Schensted correspondence for skew oscillating tableaux. *Discrete Math.*, vol. 246, n1-3, 2002, pp. 67–81.
- [39] CHAUVE (C), DULUCQ (S) et RECHNITZER (A). – Enumerating alternating trees. *J. Combin. Theory Ser. A*, vol. 94, n1, 2002, pp. 142–151.
- [40] COHEN (J), FRAIGNIAUD (P) et GAVOILLE (C). – Recognizing Knödel graphs. *Discrete Mathematics*, vol. 250, 2002, pp. 41–62.
- [41] COMELLAS (F), FERTIN (G) et RASPAUD (A). – Recursive graphs with small-word scale-free properties. *Physical Review E*, vol. 69, 2004, pp. 037104–1,037104–4.
- [42] CORI (R) et LE BORGNE (Y). – The sand-pile model and Tutte polynomials. *Adv. in Appl. Math.*, vol. 30, n1-2, 2003, pp. 44–52. – Formal power series and algebraic combinatorics (Scottsdale, AZ, 2001).
- [43] CORI (R), ROSSIN (D) et SALVY (B). – Polynomial ideals for sandpiles and their Grobner bases. *Theoretical Computer Science*, vol. 276, n1-2, 2002, pp. 1–15.
- [44] CORI (R) et SCHAEFFER (G). – Description trees and Tutte formulas. *Theoretical Computer Science*, vol. 292, n1, 2003, pp. 165–183. – Selected papers in honor of Jean Berstel.
- [45] CORI (R) et POULALHON (D). – Enumeration of (p,q)-parking functions. *Discrete Mathematics*, vol. 256, n3, 2002, pp. 609–623.
- [46] CORTEEL (S) et LOVEJOY (J). – Frobenius partitions and the combinatorics of Ramanujan's ${}_1\psi_1$ summation. *J. Combin. Theory Ser. A*, vol. 97, 2002, pp. 177–183.
- [47] CORTEEL (S) et LOVEJOY (J). – An iterative-bijective approach to generalizations of schur's theorem. *Europ. J. Comb.*, à paraître.
- [48] CORTEEL (S), LOVEJOY (J) et YEE (A). – Overpartitions and generating functions for generalized frobenius partitions. *Mathematics and Computer Science III : Algorithms, Trees, and Probabilities*, 2004, pp. 15–24.
- [49] DE HERTOIGH (B), TALLA (E), TEKAIA (F), BEYNE (E), SHERMAN (D), BARET (P), DUJON (B) et GOFFEAU (A). – Novel transporters from hemiascomycetous yeasts. *J. Mol. Microbiol. Biotechnol.*, vol. 6, n1, 2003, pp. 19–28.
- [50] DELEST (M), DON (A) et BENOIS-PINEAU (J). – Dag based visual interfaces for navigation in index vidéo content. *Multimedia Tools and Applications*, 2005. – à paraître.
- [51] DEVOS (M), NEŠETŘIL (J) et RASPAUD (A). – Antisymmetric flows and edge-connectivity. *Discrete Mathematics*, vol. 276, 2004, pp. 161–167.
- [52] DOLAMA (M. H) et SOPENA (E). – On the oriented chromatic number of halin graphs. *Inform. Proc. Letters*, 2005. – à paraître.
- [53] DOLAMA (M. H), SOPENA (E) et ZHU (X). – Incidence coloring of k-degenerated graphs. *Discrete Math.*, vol. 283, 2004, pp. 121–128.

- [54] DOURISBOURE (Y) et GAVOILLE (C). – Tree-decompositions with bags of small diameter. *Discrete Mathematics*, 2005. – à paraître.
- [55] DUCHON (P), FLAJOLET (P), LOUCHARD (G) et SCHAEFFER (G). – Boltzmann samplers for the random generation of combinatorial structures. *Combinatorics, Probability and Computing*, vol. 13, n4-5, 2004, pp. 577–625.
- [56] DUCHON (P), HANUSSE (N), LEBHAR (E) et SCHABANEL (N). – Could any graph be turned into a small world? *Theoretical Computer Science*, à paraître.
- [57] DUCHON (P), HANUSSE (N), SAHEB (N) et ZEMMARI (A). – Broadcast in the rendezvous model. *Information and Computation*, à paraître.
- [58] DUJON (B), SHERMAN (D) et AL. – Genome evolution in yeasts. *Nature*, vol. 430, n6995, July 2004, pp. 35–44.
- [59] DULUCQ (S) et PENAUD (J.-G). – Interprétation bijective d’une récurrence des nombres de Motzkin. *Discrete Math.*, vol. 256, n3, 2002, pp. 671–676.
- [60] DULUCQ (S) et TICHIT (L). – RNA secondary structure comparison : exact analysis of the Zhang-Shasha tree edit algorithm. *Theoret. Comput. Sci.*, vol. 306, n1-3, 2003, pp. 471–484.
- [61] DULUCQ (S) et TOUZET (H). – Decomposition algorithms for the tree edit distance problem. *Journal of Discrete Algorithms*, 2005. – à paraître, available online September 2004.
- [62] DUTOUR (I) et FÉDOU (J.-M). – Object grammars and bijections. *Theoretical Computer Science*, vol. 290, n3, 2003, pp. 1915–1929.
- [63] EILAM (T), GAVOILLE (C) et PELEG (D). – Compact routing schemes with low stretch factor. *Journal of Algorithms*, vol. 46, 2003, pp. 97–114.
- [64] FERRARO (P) et GODIN (C). – An edit distance between quotiented graphs. *Algorithmica*, vol. 36, 2003, pp. 1–39.
- [65] FERRARO (P) et GODIN (C). – Optimal mappings with minimum number of connected components in tree-to-tree comparison problems. *Journal of Algorithms*, vol. 48, 2003, pp. 385–406.
- [66] FERRARO (P), GODIN (C) et PRUSINKIEWICZ (P). – Toward a quantification of self-similarity in plants. *Fractals*, vol. 13, n2, 2005, pp. 1–25.
- [67] FERTIN (G), GODARD (E) et RASPAUD (A). – Minimum feedback vertex set and acyclic coloring. *Inform. Process. Letters*, vol. 84, 2002, pp. 131–139.
- [68] FERTIN (G), GODARD (E) et RASPAUD (A). – Acyclic and k -distance coloring of the grid. *Inform. Process. Letters*, vol. 87, 2003, pp. 51–58.
- [69] FERTIN (G) et RASPAUD (A). – $L(p, q)$ labeling of d -dimensional grids. *Discrete Mathematics*, 2004. – à paraître.
- [70] FERTIN (G) et RASPAUD (A). – A survey of Knödel graphs. *Discrete Applied Mathematics*, vol. 137, n2, 2004, pp. 173–195.
- [71] FERTIN (G), RASPAUD (A) et REED (B). – Star coloring of graphs. *Journal of Graph Theory*, vol. 47, No.3, 2004, pp. 163–182.
- [72] FERTIN (G), RASPAUD (A) et ROYCHOWDHURY (A). – On the oriented chromatic number of grids. *Inform. Process. Letters*, vol. 85, 2003, pp. 261–266.
- [73] FON-DER-FLAASS (D), KOSTOCHKA (A), RASPAUD (A) et SOPENA (E). – Nilpotent families of endomorphisms of $(p(v)^+, \cup)$. *J. Combin. Theory Ser. B*, vol. 86(1), 2002, pp. 100–108.

- [74] FRAIGNIAUD (P) et GAVOILLE (C). – Header-size lower bounds for end-to-end communication in memoryless networks. *Computer Networks*, 2004. – Selected paper for the DISC '03 Special Issue.
- [75] GAVOILLE (C) et HANUSSE (N). – On compact encoding of pagenumbers k graphs. *Discrete Mathematics & Theoretical Computer Science*, 2004. – à paraître.
- [76] GAVOILLE (C) et NEHÉZ (M). – Interval routing in reliability networks. *Theoretical Computer Science*, vol. 333, n3, 2005, pp. 415–432.
- [77] GAVOILLE (C) et PAUL (C). – Distance labeling scheme and split decomposition. *Discrete Mathematics*, vol. 273, n1-3, 2003, pp. 115–130.
- [78] GAVOILLE (C) et PELEG (D). – Compact and localized distributed data structures. *Journal of Distributed Computing*, vol. 16, mai 2003, pp. 111–120. – PODC 20-Year Special Issue.
- [79] GAVOILLE (C), PELEG (D), PÉRENNÈS (S) et RAZ (R). – Distance labeling in graphs. *Journal of Algorithms*, vol. 53, n1, 2004, pp. 85–112.
- [80] GAVOILLE (C) et ZEMMARI (A). – The compactness of adaptive routing tables. *Journal of Discrete Algorithms*, vol. 1, n2, 2003, pp. 237–254.
- [81] GODARD (E) et MÉTIVIER (Y). – Deducible and equivalent structural knowledges in distributed algorithms. *Theory of Computing Systems*, vol. 36, 2003, pp. 631–654.
- [82] GODARD (E), MÉTIVIER (Y) et MUSCHOLL (A). – Characterizations of classes of graphs recognizable by local computations. *Theory of Computing Systems*, vol. 37, 2004, pp. 249–293.
- [83] GONÇALVES (D). – Caterpillar arboricity of planar graphs. *Discrete Math.*, 2003. – à paraître.
- [84] GUIBERT (O) et MANSOUR (T). – Some statistics on restricted 132 involutions. *Ann. Comb.*, vol. 6, n3-4, 2002, pp. 349–374.
- [85] HANUSSE (N), KRANAKIS (E) et KRIZANC (D). – Searching with mobile agents in networks with liars. *Discrete Appl. Math.*, vol. 137, n1, 2004, pp. 69–85. – 1st International Workshop on Combinatorics of Searching, Sorting, and Coding (COSSAC '01) (Ischia).
- [86] IRAGNE (F), NIKOLSKI (M), MATHIEU (B), AUBER (D) et SHERMAN (D). – Proviz : protein interaction visualization and exploration. *Bioinformatics*, vol. 21, n2, 2005, pp. 272–274.
- [87] JONES (G) et ZVONKIN (A). – Orbits of braid groups on cacti. *The Moscow Mathematical Journal*, vol. 2, n1, 2002, pp. 129–162.
- [88] KRATTENTHALER (C), GUTTMANN (A) et VIENNOT (X). – Vicious walkers, friendly walkers and Young tableaux III : betweentwo walls. *J. Stat. Physics*, vol. 110, 2003, pp. 1069–1086.
- [89] KUCHEROV (G), OCHEM (P) et RAO (M). – How many square occurrences must a binary sequence contain ? *Electr. J. Comb.*, vol. 10, 2003.
- [90] LE BORGNE (Y) et ROSSIN (D). – On the identity of the sandpile group. *Discrete Math.*, vol. 256, n3, 2002, pp. 775–790. – LaCIM 2000 Conference on Combinatorics, Computer Science and Applications (Montreal, QC).
- [91] LESUR (I) et CAMPBELL (J). – The transcriptome of prematurely aging yeast cells is similar to that of telomerase-deficient cells. *Mol. Biol. Cell.*, vol. 15, 2004, pp. 1297–1312.

- [92] LOVEJOY (J). – Lacunary partition functions. *Math. Res. Lett.*, vol. 9, 2002, pp. 191–198.
- [93] LOVEJOY (J). – Gordon’s theorem for overpartitions. *J. Combin. Theory Ser. A*, vol. 103, 2003, pp. 393–401.
- [94] LOVEJOY (J). – More lacunary partition functions. *Illinois J. Math.*, vol. 47, 2003, pp. 769–773.
- [95] LOVEJOY (J). – The number of partitions into distinct parts modulo powers of 5. *Bull. London Math. Soc.*, vol. 35, 2003, pp. 41–46.
- [96] LOVEJOY (J). – A bailey lattice. *Proc. Amer. Math. Soc.*, vol. 132, 2004, pp. 1507–1516.
- [97] LOVEJOY (J). – Overpartition theorems of the rogers-ramaujan type. *J. London Math. Soc.*, vol. 69, 2004, pp. 562–574.
- [98] LOVEJOY (J). – Overpartitions and real quadratic fields. *J. Number Theory*, vol. 106, 2004, pp. 178–186.
- [99] LOVEJOY (J). – Rank and conjugation for the frobenius representation of an overpartition. *Ann. Comb.*, à paraître.
- [100] LOVEJOY (J). – A theorem on seven-colored overpartitions and its applications. *Int. J. Number Theory*, à paraître.
- [101] LOVEJOY (J) et ONO (K). – Extension of Ramanujan’s congruences for the partition function modulo powers of 5. *J. Reine Angew. Math.*, vol. 542, 2002, pp. 123–132.
- [102] LOVEJOY (J) et ONO (K). – Hypergeometric generating functions for values of Dirichlet and other L -functions. *Proc. Nat. Acad. Sci. USA*, vol. 100, 2003, pp. 6904–6909.
- [103] MÉTIVIER (Y), SAHEB (N) et ZEMMARI (A). – Randomized local elections. *Information Processing Letters*, vol. 82, n6, 2002, pp. 313–320.
- [104] MÉTIVIER (Y), SAHEB (N) et ZEMMARI (A). – Analysis of a randomized rendez-vous algorithm. *Information and Computation*, vol. 184, n1, 2003, pp. 109–128.
- [105] MÉTIVIER (Y), SAHEB (N) et ZEMMARI (A). – Locally guided randomized elections in trees : the totally fair case. *Information and Computation*, à paraître.
- [106] MONTASSIER (M) et RASPAUD (A). – A note on 2-facial coloring of plane graphs. *Information Processing Letters*, 2005. – à paraître.
- [107] NARBEL (P) et LOPEZ (L.-M). – Languages, D0L-systems, sets of curves and surface automorphisms. *Information and Computation*, vol. 180, n1, 2003, pp. 30–52.
- [108] OCHEM (P). – Oriented colorings of triangle-free planar graphs. *Inf. Process. Lett.*, vol. 92, n2, 2004, pp. 71–76.
- [109] PECHER (A). – Partitionable graphs arising from near-factorizations of finite groups. *Discrete Math.*, vol. 269, n1-3, 2003, pp. 191–218.
- [110] PECHER (A). – Cayley partitionable graphs and near-factorizations of finite groups. *Discrete Math.*, vol. 276, n1-3, 2004, pp. 295–311. – 6th International Conference on Graph Theory.
- [111] PECHER (A) et WAGLER (A). – Almost all webs are not rank-perfect. *Math. prog.*, 2005. – à paraître.
- [112] PECHER (A) et WAGLER (A). – A construction for non-rank facets of stable set polytopes of webs. *European J. of Comb.*, 2005. – à paraître.

- [113] PECHER (A) et WAGLER (A). – On non-rank facets in stable set polytopes of webs with clique number four. *Discrete Applied Math.*, 2005. – à paraître.
- [114] PENAUD (J.-G) et ROQUES (O). – Génération de chemins de Dyck à pics croissants. *Discrete Mathematics*, vol. 256, n3, 2002, pp. 255–267.
- [115] PENAUD (J.-G) et ROQUES (O). – Tirage à pile ou face de mots de Fibonacci. *Discrete Mathematics*, vol. 256, n3, 2002, pp. 791–800.
- [116] ROBSON (J). – Constant bounds on the moments of the height of binary search trees. *Theoretical Computer Science*, vol. 276, 2002, pp. 435–444.
- [117] S.CORTEEL et LOVEJOY (J). – Overpartitions. *Trans. Amer. Math. Soc.*, vol. 356, 2004, pp. 1623–1635.
- [118] SHERMAN (D), DURRENS (P), BEYNE (E), NIKOLSKI (M) et SOUCIET (J. G. C). – Génolevures : comparative genomics and molecular evolution of hemiascomycetous yeasts. *Nucleic Acids Research*, vol. 32, January 2004, pp. D315–D3128. – Database issue.
- [119] SOPENA (E). – There exist oriented planar graphs with oriented chromatic number at least sixteen. *Inform. Proc. Letters*, vol. 81(6), 2002, pp. 309–312.
- [120] VIENNOT (X). – A Strahler bijection between Dyck paths and planar trees. *Discrete Mathematics*, vol. 246, 2002, pp. 317–329.

Colloques avec Comité de programme et Actes

- [121] ABRAHAM (I), GAVOILLE (C) et MALKHI (D). – Routing with improved communication-space trade-off. In : *18th International Symposium on Distributed Computing (DISC)*. pp. 305–319. – Springer, octobre 2004.
- [122] ABRAHAM (I), GAVOILLE (C), MALKHI (D), NISAN (N) et THORUP (M). – Compact name-independent routing with minimum stretch. In : *16th Annual ACM Symposium on Parallel Algorithms and Architecture (SPAA)*. pp. 20–24. – ACM Press, juillet 2004.
- [123] ALSTRUP (S), GAVOILLE (C), KAPLAN (H) et RAUHE (T). – Nearest common ancestors : A survey and a new distributed algorithm. In : *14th Annual ACM Symposium on Parallel Algorithms and Architecture (SPAA)*. pp. 258–264. – ACM Press, août 2002.
- [124] AUBER (D). – Using strahler numbers for real time visual exploration of huge graphs. In : *International Conference on Computer Vision and Graphics*, pp. 56–69. – 2002.
- [125] AUBER (D.), CHIRICOTA (Y), JOURDAN (F) et MELANCON (G). – Multiscale visualization of small world networks. In : *IEEE Symposium on Information Visualisation*. pp. 75–81. – IEEE Computer Society, 2003.
- [126] AUBER (D), DELEST (M) et CHIRICOTA (Y). – Strahler based graph clustering using convolution. In : *8th International Conference on Information Visualisation, IV*. pp. 44–51. – IEEE Computer Society, 2004.
- [127] AUBER (D), DELEST (M), DOMENGER (J), FERRARO (P) et STRANDH (R). – EVAT : Environment for visualization and analysis of trees. In : *IEEE Symposium on Information Visualisation Contest*. pp. 124–126. – IEEE Computer Society, 2003.
- [128] AUBER (D), DELEST (M), FEDOU (J), DOMENGER (J) et DUCHON (P). – New Strahler numbers for rooted plane trees. In : *Proceedings of the Third Colloquium on Mathematics and Computer Science*, éd. par Drmota (M), Flajolet (P), Gardy (D) et Gittenberger (B), pp. 203–215. – Wien, 2004.

- [129] AUBER (D) et JOURDAN (F). – Interactive refinement of multi-scale network clusterings. *In : 9th International Conference on Information Visualisation, IV.* p. à paraître. – IEEE Computer Society, 2005.
- [130] AUILLANS (P). – Toward topic maps processing and visualization. *In : Sixth International Conference on Information Visualisation (IV'02).* – London, England, July 2002.
- [131] AUILLANS (P), OSSONA DE MENDEZ (P), ROSENSTIEHL (P) et VATANT (B). – A formal model for topic maps. *In : Proceedings of the First International Semantic Web Conference on The Semantic Web.* pp. 69–83. – Springer-Verlag, 2002.
- [132] AVAL (J.-C). – Quasi-invariant and super-coinvariant polynomials for the generalized symmetric group. *In : Formal power series and algebraic combinatorics*, éd. par ERIKSSON (K) et LINUSSON (S). – Vadstena, Suède, 2003.
- [133] BARRIOT (R) et LAMIABLE (A). – Web services and client packages as a framework for data mining on the neighborhood of organized sets of biological sequences. *In : Proceedings of the 12th International Conference on Intelligent Systems for Molecular Biology and the 3rd European Conference on Computational Biology in Glasgow (UK).* – 2004.
- [134] BONICHON (N). – A bijection between realizers of maximal plane graphs and pairs of non-crossing dyck paths. *In : Proceedings of FPSAC'02*, pp. 123–132. – 2002.
- [135] BONICHON (N), FELSNER (S) et MOSBAH (M). – Convex drawings of 3-connected planar graphs - (extended abstract). *In : Graph Drawing : 12th International Symposium, GD 2004*, pp. 60–70. – 2005.
- [136] BONICHON (N), GAVOILLE (C) et HANUSSE (N). – Canonical decomposition of outerplanar maps and application to enumeration, coding and generation. *In : 29th International Workshop, Graph - Theoretic Concepts in Computer Science (WG).* pp. 81–92. – Springer-Verlag, juin 2003.
- [137] BONICHON (N), GAVOILLE (C) et HANUSSE (N). – An information-theoretic upper bound of planar graphs using triangulation. *In : 20th Annual Symposium on Theoretical Aspects of Computer Science (STACS).* pp. 499–510. – Springer, février 2003.
- [138] BONICHON (N), GAVOILLE (C), HANUSSE (N), POULALHON (D) et SCHAEFFER (G). – Planar graphs, via well-orderly maps and trees. *In : 30th International Workshop, Graph - Theoretic Concepts in Computer Science (WG).* – Springer, juin 2004. 270-284.
- [139] BOUSQUET-MÉLOU (M). – Counting walks in the quarter plane. *In : Mathematics and computer science, II*, éd. par Birkhäuser, pp. 49–67. – Versailles, 2002.
- [140] BOUSQUET-MÉLOU (M). – Walks in the quarter plane : a functional equation approach. *In : Formal power series and algebraic combinatorics*, éd. par Brak (R), Foda (O), Greenhill (C), Guttmann (A) et Owczarek (A), pp. 148–160. – Melbourne, Australie, 2002.
- [141] BOUSQUET-MÉLOU (M), GUTTMANN (T) et JENSEN (I). – Self-avoiding walks crossing a square. *In : Formal power series and algebraic combinatorics.* – Taormina, Sicile, 2005.
- [142] BOUSQUET-MÉLOU (M) et SCHAEFFER (G). – The degree distribution in bipartite planar maps : applications to the Ising model. *In : Formal power series and algebraic combinatorics*, éd. par Eriksson (K) et Linusson (S), pp. 312–323. – Vadstena, Suède, 2003.

- [143] BOUSQUET-MÉLOU (M) et STEINGRIMSSON (E). – Decreasing subsequences in permutations and Wilf equivalence for involutions. *In : Formal power series and algebraic combinatorics.* – Taormina, Sicile, 2005.
- [144] CHALOPIN (J). – Election and local computations on closed unlabelled edges (*extended abstract*). *In : Proc. of SOFSEM 2005*, pp. 81–90. – 2005.
- [145] CHALOPIN (J) et MÉTIVIER (Y). – Election and local computations on edges. *In : Foundations of System Specification and Computation Structures (FoSSaCS)*. pp. 90–104. – Lecture Notes in Computer Science 2987, Springer-Verlag, 2004.
- [146] CHALOPIN (J), MÉTIVIER (Y) et ZIELONKA (W). – Election, naming and cellular edge local computations. *In : International conference on graph transformation (ICGT) (EATCS best paper award)*. pp. 242–256. – Lecture Notes in Computer Science 3256, Springer-Verlag, 2004.
- [147] COMELLAS (F), FERTIN (G) et RASPAUD (A). – Vertex labeling and routing in recursive clique-trees, a new family of small-world scale-free graphs. *In : 10th International Colloquium on Structural Information & Communication Complexity (SIROCCO)*, éd. par Sibeyn (J). pp. 73–87. – Carleton University Press, Juin 2003.
- [148] CORI (R) et NOY (M). – Formal power series and algebraic combinatorics. pp. i–x and 1–361. – Amsterdam, 2002. Selected papers from the 11th International Conference (FPSAC'99) held in Barcelona, June 7–11, 1999, *Discrete Mathematics* **246** (2002), no. 1-3.
- [149] CORI (R) et POULALHON (D). – Enumeration of (p, q) -parking functions. *Discrete Mathematics*, vol. 256, n3, 2002, pp. 609–623. – LaCIM 2000 Conference on Combinatorics, Computer Science and Applications (Montreal, QC).
- [150] COULONGES (S), PECHER (A) et WAGLER (A). – On the imperfection ratio of a -perfect graphs and some consequences. *In : CTW'05.* – mai 2005.
- [151] DELEST (M), DON (A) et BENOIS-PINEAU (J). – Graph-based visual interfaces for navigation in indexed video content. *In : CBMI*, éd. par Gabbouj (M). pp. 49–55. – Rennes, France, Septembre 2003.
- [152] DELEST (M), DON (A) et BENOIS-PINEAU (J). – Intuitive color-based visualization of multimedia content as large graphs. *In : EI'04.* SPIE, pp. 65–74. – 2004.
- [153] DELEST (M), MUNZNER (T), AUBER (D) et DOMENGER (J). – Exploring infovis publication history with tulip. *In : InfoVis 2004.* IEEE. – 2004. Concours InfoVis - Deuxième place.
- [154] DERBEL (B) et MOSBAH (M). – Distributing the execution of a distributed algorithm over a network. *In : The 7th IEEE International Conference on Information Visualization, IV03-AGT.*, pp. 485–490. – 16-18 July, 2003.
- [155] DERBEL (B) et MOSBAH (M). – A fully distributed linear time algorithm for cluster network decomposition. *In : The 16th IASTED International Conference on Parallel and Distributed Computing and Systems, PDCS, MIT.* pp. 548–553. – MIT, 9-11 November, 2004.
- [156] DOURISBOURE (Y). – An additive stretched routing scheme for chordal graphs. *In : 28-th International Workshop on Graph-Theoretic Concepts in Computer Science (WG '02)*. pp. 150–163. – Springer-Verlag, juin 2002.
- [157] DOURISBOURE (Y). – Compact routing schemes for tree-length δ graphs and for k -chordal graphs. *In : 18-th International Symposium on Distributed Computing (DISC 2004)*. pp. 365–378. – Springer, octobre 2004.

- [158] DOURISBOURE (Y) et GAVOILLE (C). – Improved compact routing scheme for chordal graphs. *In : 16th International Symposium on Distributed Computing (DISC)*. pp. 252–264. – Springer, octobre 2002.
- [159] DOURISBOURE (Y) et GAVOILLE (C). – Tree-decomposition of graphs with small diameter bags. *In : 2nd European Conference on Combinatorics, Graph Theory and Applications (EUROCOMB)*, éd. par Fila (J), pp. 100–104. – septembre 2003.
- [160] DOURISBOURE (Y) et GAVOILLE (C). – Sparse additive spanners for bounded tree-length graphs. *In : 11th International Colloquium on Structural Information & Communication Complexity (SIROCCO)*. pp. 123–137. – Springer, juin 2004.
- [161] DUCHON (P), FLAJOLET (P), LOUCHARD (G) et SCHAEFFER (G). – Random sampling from Boltzmann principles. *In : Proceedings of the 29th ICALP*, éd. par Widmayer (P), Triguero (F), Morales (R), Hennessy (M), Eidenbenz (S) et Conejo (R). pp. 501–513. – Springer, 2002.
- [162] DUCHON (P), HANUSSE (N), SAHEB (N) et ZEMMARI (A). – Broadcast in the rendezvous model. *In : Proceedings of STACS 2004*, éd. par Diekert (V) et Habib (M). pp. 559–570. – Springer, 2004.
- [163] DUCHON (P), HANUSSE (N) et TIXEUIL (S). – Optimal randomized self-stabilizing mutual exclusion on synchronous rings. *In : Proceedings of DISC'2004*, éd. par Guerraoui (R). pp. 216–229. – Amsterdam, October 2004.
- [164] DULUCQ (S) et TOUZET (H). – Analysis of tree edit distance algorithms. *In : Combinatorial pattern matching, 14th Annual Symposium, CPM 2003*. pp. 83–95. – Berlin, 2003.
- [165] DUONG (V), FERCHAUD (F), GAVOILLE (C) et MOSBAH (M). – A new slot allocation for ATFM. *In : 7th IEEE International Conference on Intelligent Transportation Systems (ITSC)*, pp. 1075–1079. – octobre 2004.
- [166] DUONG (V), FERCHAUD (F), GAVOILLE (C) et MOSBAH (M). – Reducing disturbances by using absorption areas. *In : 1st International Conference on Research in Air Transportation (ICRAT)*, pp. 111–115. – novembre 2004.
- [167] DUONG (V), FERCHAUD (F), GAVOILLE (C) et MOSBAH (M). – Using absorption areas to improve ATFM. *In : 23rd IEEE International Digital Avionics Systems Conference (DASC)*. – octobre 2004.
- [168] EL HIBAOU (A), SAHEB (N) et ZEMMARI (A). – Polyominoids and uniform election. *In : FPSAC : 17th International Conference on Formal Power Series and Algebraic Combinatorics*. – 2005.
- [169] EL HIBAOU (A), SAHEB (N) et ZEMMARI (A). – A uniform probabilistic election algorithm in k-trees. *In : IMACS : 17th IMACS World Congress : Scientific Computation, Applied Mathematics and Simulation*. – 2005.
- [170] FERCHAUD (F), GAVOILLE (C) et MOSBAH (M). – Using absorption areas to improve ATFM. *In : ERC Innovative Research Workshop*. – décembre 2003.
- [171] FERRARO (P), GODIN (C) et PRUSINKIEWICZ (P). – A structural method for assessing self-similarity in plants. *In : Fourth International Workshop on Functional-Structural Trees Models*, pp. 56–61. – Montpellier, France, June 07-11 2004.
- [172] FERRARO (P), PRUSINKIEWICZ (P), MÜNDERMAN (L) et GODIN (C). – L-systems and MTGs : Integrating simulation and formal analysis of architectural plant models. *In : 16th European Simulation Multiconference, Modelling and Simulation*, pp. 418–423. – Darmstadt, Juin 2002.

- [173] FERTIN (G), RASPAUD (A) et SÝKORA (O). – No-hole $l(p,0)$ -labelling of cycles, grids and hypercubes. *In : 11th International Colloquium on Structural Information & Communication Complexity (SIROCCO)*, éd. par Kralovič (R) et Sýkora (O). pp. 138–148. – Springer-Verlag, Juin 2004.
- [174] FRAIGNIAUD (P) et GAVOILLE (C). – A space lower bound for routing in trees. *In : 19th Annual Symposium on Theoretical Aspects of Computer Science (STACS)*. pp. 65–75. – Springer, mars 2002.
- [175] FRAIGNIAUD (P) et GAVOILLE (C). – Lower bounds for oblivious single-packet end-to-end communication. *In : 17th International Symposium on Distributed Computing (DISC)*. pp. 211–223. – Springer, octobre 2003.
- [176] FRAIGNIAUD (P), GAVOILLE (C) et PAUL (C). – Eclecticism shrinks even small worlds. *In : 23rd Annual ACM Symposium on Principles of Distributed Computing (PODC)*. pp. 169–178. – ACM Press, juillet 2004.
- [177] GAINANT (G) et AUBER (D). – Arna : Interactive comparison and alignment of rna secondary structure. *In : IEEE Symposium on Information Visualisation*. p. à paraître. – IEEE Computer Society, 2004.
- [178] GAVOILLE (C) et NEHÉZ (M). – Interval routing in reliability networks. *In : 9th International Colloquium on Structural Information & Communication Complexity (SIROCCO)*. pp. 149–164. – Carleton University Press, juin 2003.
- [179] GAVOILLE (C) et NEHÉZ (M). – On interval routing in random meshes. *In : 2nd European Conference on Combinatorics, Graph Theory and Applications (EUROCOMB)*, éd. par Fila (J), pp. 147–150. – septembre 2003.
- [180] GAVOILLE (C) et PAUL (C). – Optimal distance labeling schemes for interval and circular-arc graphs. *In : 11th Annual European Symposium on Algorithms (ESA)*, éd. par Battista (G. D) et Zwick (U). pp. 254–265. – Springer, septembre 2003.
- [181] GODARD (E) et MÉTIVIER (Y). – A characterization of families of graphs in which election is possible. *In : Foundations of System Specification and Computation Structures (FoSSaCS) (EATCS best paper award)*. pp. 159–171. – Lecture Notes in Computer Science 2303, Springer-Verlag, 2002.
- [182] GODARD (E) et MÉTIVIER (Y). – Equivalence of structural knowledges in distributed algorithms. *In : International Colloquium on Structural Information and Communication Complexity (SIROCCO)*. pp. 149–164. – Carleton scientific press, 2002.
- [183] GODARD (E), MÉTIVIER (Y), MOSBAH (M) et SELLAMI (A). – Termination detection of distributed algorithms. *In : International conference on graph transformation (ICGT)*. pp. 106–119. – Lecture Notes in Computer Science 2505, Springer-Verlag, 2002.
- [184] GONÇALVES (D). – Edge partition of planar graphs into two outerplanar graphs. *In : Proceedings of STOC '05*. – 2005. à paraître.
- [185] GRIVET (S), AUBER (D), DOMENGER (J) et MELANCON (G). – Bubble tree drawing algorithm. *In : International Conference on Computer Vision and Graphics*, p. à paraître. – 2004.
- [186] GUIBERT (O) et MANSOUR (T). – Restricted 132-involutions. *In : 48^{ème} Séminaire Lotharingien de Combinatoire*, p. Article B48a. – 2002.

- [187] HANUSSE (N), KAVVADIAS (D), KRANAKIS (E) et KRIZANC (D). – Memoryless search algorithm in networks with faulty advice. *In : IFIP-WCC'2002 (World Computer Congress); Track - International Conference on Theoretical Computer Science.* pp. 206–216. – Kluwer Academic, 2002. Montreal.
- [188] ILIE (L), OCHEM (P) et SHALLIT (J). – A generalization of repetition threshold. *In : MFCS*, pp. 818–826. – 2004.
- [189] LE BORGNE (Y). – Counting upper interactions in dyck paths. *In : FPSAC 05.* – June 2005. Formal Power Series and Algebraic Combinatorics (FPSAC 05) Taormina, Sicile, Italie.
- [190] MÉTIVIER (Y), MOSBAH (M), OSSAMY (R) et SELLAMI (A). – Synchronizers for local computations. *In : International conference on graph transformation (ICGT).* pp. 271–286. – Lecture Notes in Computer Science 3256, Springer-Verlag, 2004.
- [191] MÉTIVIER (Y), SAHEB (N) et ZEMMARI (A). – A uniform randomized election in trees. *In : International Colloquium on Structural Information and Communication Complexity (SIROCCO).* pp. 261–274. – Carleton scientific press, 2003.
- [192] MONTASSIER (M). – Observability of recursive clique-trees. *In : Eurocomb03. European conference on Combinatorics, Graph Theory and Applications*, pp. 275–279. – Prague, September 8 - 12, 2003.
- [193] NARBEL (P). – Abstract abstract data types in modular and object-oriented programming. *In : FOOL 10, ACM SIGPLAN Intl. Workshop on Foundations of OO Languages.* – January 2003. New-Orleans, USA.
- [194] NIKOLSKI (M), FERRARO (P), DURRENS (P) et AIGLE (M). – *Saccharomyces siliceus*. *In : First International Workshop on Systems Biology of Yeast.* – St Louis, Novembre 2003.
- [195] OCHEM (P). – A generator of morphisms for infinite words. *In : Proceedings of the Workshop on Words avoidability, complexity and morphisms - Edited by G. Richomme.* – 2004.
- [196] PECHER (A) et WAGLER (A). – A construction for non-rank facets of stable set polytopes of webs. *In : Eurocomb'03*, pp. 301–305. – septembre 2003. article complet à paraître dans European J. of Comb.
- [197] PECHER (A) et WAGLER (A). – On non-rank facets in stable set polytopes of webs with clique number four. *In : Electronic Notes in Discrete Mathematics.* pp. 96–99. – 2nd Cologne Twente Workshop on Graphs and Combinatorial Optimization, 2003. article complet à paraître dans Discrete Applied Mathematics.
- [198] PECHER (A), WAGLER (A) et ZHU (X). – Three classes of minimal circular imperfect graphs. *In : GRACO'05.* – avril 2005.
- [199] VIENNOT (X. G). – Lorentzian quantum gravity and heaps of pieces. – Mars 2002. *46ème édition du Séminaire Lotharingien*, Ottrott, Alsace.
- [200] VIENNOT (X). – Heaps of segments, 2d lorentzian triangulations and curvature of the space-time. – Janvier 2005. 27th Nordic and 1st Franco-Nordic Congress of Mathematicians, Reykjavik, Island.

Conférences invitées

- [201] AVAL (J.-C), BERGERON (F) et BERGERON (N). – Diagonal temperley-lieb invariants and harmonics. – A paraître dans le journal du Séminaire Lotharingien de Combinatoire. FPSAC'05.

- [202] BOUSQUET-MÉLOU (M). – Walks in the quarter plane : a functional equation approach. *In : Congrès de la CMS (Canadian Mathematical Society)*. – Québec, Canada, 2002.
- [203] BOUSQUET-MÉLOU (M). – Enumerative combinatorics via generating functions. *In : École d'été du réseau européen de combinatoire algébrique (ACE)*. – Linköping, Suède, 2003.
- [204] BOUSQUET-MÉLOU (M). – The Ising model on planar maps : a combinatorial solution. *In : Statistical Mechanics of Polymer Models*. – Banff, Canada, 2003.
- [205] BOUSQUET-MÉLOU (M). – The Ising model on planar maps : a combinatorial solution. *In : 50è édition du Séminaire Lotharingien de Combinatoire*. – Ottrott, France, 2003.
- [206] BOUSQUET-MÉLOU (M). – The Ising model on planar maps : a combinatorial solution. *In : Journée Matrix integrals, combinatorics and probability*. – Paris, France, 2003.
- [207] BOUSQUET-MÉLOU (M). – Minimal transitive factorizations of permutations. *In : Recent Advances in Algebraic and Enumerative Combinatorics*. – Banff, Canada, 2003.
- [208] BOUSQUET-MÉLOU (M). – Are algebraic generating functions ordinary? *In : Third Colloquium on Mathematics and Computer Science Algorithms, Trees, Combinatorics and Probabilities*. – Vienne, Autriche, 2004.
- [209] BOUSQUET-MÉLOU (M). – Planar maps and algebraic series : polynomial equations with one catalytic variable. *In : Séminaire Lotharingien de Combinatoire*. – Ottrott, France, 2004.
- [210] BOUSQUET-MÉLOU (M). – Algebraic generating functions in enumerative combinatorics, and context-free languages. *In : STACS*. – Stuttgart, Allemagne, 2005.
- [211] BOUSQUET-MÉLOU (M). – Are algebraic generating functions ordinary? *In : 24th Nordic and 1st Franco-Nordic Congress of Mathematicians*. – Reykjavík, Islande, 2005.
- [212] BOUSQUET-MÉLOU (M). – Embedded trees. *In : Quantum Field Theory, Then and Now : A tribute to Claude Itzykson*. – Saclay, France, 2005.
- [213] BOUSQUET-MÉLOU (M). – The ising and hard-particle models on planar maps. *In : 2nd Workshop on Tutte polynomials and applications*. – Bellaterra, Espagne, 2005.
- [214] GAVOILLE (C). – Distributed data structures : a survey (invited lecture). *In : SI-ROCCO '05*. – mai 2005. à paraître.
- [215] SHABAT (G) et ZVONKIN (A). – Plane trees and algebraic numbers. *In : Mini-Workshop on Hurwitz Theory and Ramifications*, éd. par Kwak (J. H) et Mednykh (A). Combinatorial and Computational Mathematics Center, Pohang University, Korea, pp. 361–403. – 2003.
- [216] VIENNOT (X. G). – Lorentzian quantum gravity and heaps of pieces. – Juin 2002. *Congrès de la CMS (Canadian Mathematical Society)*, Québec, Canada.
- [217] VIENNOT (X. G). – From priority queues and binary search trees to sylvester monoid and loday-ronco algebra. – Mars 2004. *52ème édition du Séminaire Lotharingien, colloque en l'honneur de Alain Lascoux*, Ottrott, Alsace.
- [218] VIENNOT (X. G). – From kepler to catalan, strahler and computer graphics. – Juin 2005. *colloque en l'honneur de Adriano Garsia*, Taormina, Italie.
- [219] VIENNOT (X. G). – The nordic decomposition for multidirected animals and 2d lorentzian quantum gravity. – Juillet 2005. *Counting complexity : an international*

- workshop on statistical mechanics and combinatorics, colloque en l'honneur de Tony Guttmann*, Dunk Island, Australie.
- [220] VIENNOT (X. G). – A walk in the garden of bijections. – Avril 2005. *54ème édition du Séminaire Lotharingien, colloque en l'honneur de Xavier Viennot*, Lucelle, Alsace.
- [221] VIENNOT (X). – Combinatorics and physics : the use of heaps of pieces. – Novembre 2002. *Colloquium de Combinatoire*, Magdeburg, Allemagne.
- [222] VIENNOT (X). – Paths, braids, dimers, young tableaux and reduced decompositions of permutations. – Novembre 2002. Ecole doctorale des universités de Berlin et Zurich.
- [223] VIENNOT (X). – Bijective proofs for polymers statistical mechanics. – Mai 2003. *Statistical Mechanics of Polymer Models*, Banff, Canada, 2003.
- [224] VIENNOT (X). – Combinatorics and quantum gravity. – Avril 2003. 47ème Matematicke Kolokvium, Université Charles, Prague, République Tchèque.
- [225] VIENNOT (X). – Heaps of pieces in physics. – Mars 2003. *Number Theory and Combinatorics in Physics*, Gainesville, Florida, USA.
- [226] VIENNOT (X). – Enumerative combinatorics, algorithms and physics. – Janvier 2004. *Complex systems*, Valparaiso, Chili.
- [227] ZVONKIN (A). – Megamaps : construction and examples. *In : Mini-Workshop on Hurwitz Theory and Ramifications*, éd. par Kwak (J. H) et Mednykh (A). Combinatorial and Computational Mathematics Center, Pohang University, Korea, pp. 405–415. – 2003.
- [228] ZVONKIN (A). – Towards topological classification of univariate complex polynomials. *In : Mini-Workshop on Hurwitz Theory and Ramifications*, éd. par Kwak (J. H) et Mednykh (A). Combinatorial and Computational Mathematics Center, Pohang University, Korea, pp. 419–430. – 2003.

Livres d'enseignement et de recherche

- [229] AVAL (J.-C). – *Conjecture $n!$ et généralisations*. – Publications du LaCIM, 2004. 144 pages.
- [230] LANDO (S) et ZVONKIN (A). – *Graphs on Surfaces and Their Applications*. – Springer Verlag, 2004, *Encyclopaedia of Mathematical Sciences, Low-Dimensional Topology subseries*. XVI+455 pages.

Chapitres d'ouvrages

- [231] AUBER (D). – Tulip : A huge graph visualisation framework. *In : Graph Drawing Softwares*, éd. par Mutzel (P) et Jünger (M), pp. 105–126. – Springer-Verlag, 2003.
- [232] VIENNOT (X). – Énumérons ! (de la combinatoire énumérative classique aux nouvelles combinatoires : bijectives, algébriques, expérimentales, quantiques et ... magiques). *In : Leçons de mathématiques d'aujourd'hui (vol. 3)*, éd. par et Nicolas Nikolski (E. C). – Cassini, à paraître.

Posters et autres communications

- [233] ZVONKIN (A). – Braids and polynomials. – Septembre 2002. Journées de Tresses, Lacanau.

- [234] ZVONKIN (A). – Composition of polynomials and their topological classification. – Juillet 2002. Symmetries in Graphs, Maps, and Complexes, Aveiro, Portugal.
- [235] ZVONKIN (A). – Polynomials with exceptional monodromy groups. – Mars 2002. Séminaire Lotharingien de Combinatoire 48, Domaine St.-Jacques, France.

1.4.2 Publications d'audience nationale

Colloques avec Comité de programme et Actes

- [236] BARRIOT (R), DUTOUR (I) et SHERMAN (D). – Efficient generation of pertinent target sets for the blastsets system. *In : Actes des 6^{ème} Journées Ouvertes Biologie Informatique Mathématiques à Lyon (France)*. – 2005.
- [237] BARRIOT (R), POIX (J), GROUPI (A), BARRE (A), GOFFARD (N), SHERMAN (D), DUTOUR (I) et DE DARUVAR (A). – New strategy for the representation and the integration of biomolecular knowledge at a cellular scale. *In : Actes des 5^{ème} Journées Ouvertes Biologie Informatique Mathématiques à Montréal (Canada)*. – 2004.
- [238] BAZZARO (F) et MONTASSIER (M). – Assignment de fréquences et étiquetage $(d, 1)$ -total dans les topologies planaires. *In : 6^{ème} Rencontres Francophones sur les aspects Algorithmiques des Télécommunications (ALGOTEL 2004)*. pp. 27–31. – Batz-sur-mer, France, 26–28 mai 2004.
- [239] BAZZARO (F) et GAVOILLE (C). – Etiquetage de distance des graphes de permutation. *In : 6^{èmes} Journées Graphes et Algorithmes*, éd. par et S. Gravier (E. D). pp. 22–24. – Les Cahiers du laboratoire Leibniz, No 114, octobre 2004.
- [240] BERMOND (J.-C), DELMAS (O), HAVET (F), MONTASSIER (M) et PÉRENNES (S). – Réseaux de télécommunication minimaux embarqués tolérants aux pannes. *In : 5^{ème} Rencontres Francophones sur les aspects Algorithmiques des Télécommunications (ALGOTEL 2003)*. pp. 27–32. – Banyuls-sur-mer, France, 12-14 mai 2003. ISBN 2-7261-1246-3.
- [241] DUCHON (P), HANUSSE (N), LEBAHR (E) et SCHABANEL (N). – Could any graph be turned into a small-world? *In : Actes d'AlgoTel'2005 (conférence francophone sur les algorithmes de communications)*, pp. –. – Giens, Mai 2005.
- [242] DUCHON (P), HANUSSE (N) et TIXEUIL (S). – Protocoles auto-stabilisants synchrones d'exclusion mutuelle pour les anneaux anonymes et uniformes. *In : Actes d'AlgoTel 2004*. pp. 135–140. – Batz sur Mer, Mai 2004.
- [243] FERRARO (P), GODIN (C) et PRUSINKIEWICZ (P). – Une methode structurelle pour evaluer les proprietes d'autosimilarite des plantes. *In : Colloque Autosimilarite et Applications (Self Similarity And Applications Workshop)*. – Clermont-Ferrand (France), Mai 2002.
- [244] FERRARO (P), TICHIT (L) et DULUCQ (S). – Local similarity between trees. *In : 5^{èmes} Journées Ouvertes Biologie Informatique Mathématiques*, p. 19. – Montréal, Canada, June 2004.
- [245] FRAIGNIAUD (P), GAVOILLE (C) et PAUL (C). – Eclectisme dans les petits mondes. *In : 6^{èmes} Rencontres Francophones sur les Aspects Algorithmiques des Télécommunications (AlgoTel)*. – INRIA, mai 2004.
- [246] NARBEL (P). – Programmation paramétrée en ML et automates d'architecture. *In : JFLA 2004, INRIA J. Francoph. des Lang. Applicatifs*, pp. 169–184. – 2004.

- [247] OUANGRAOUA (A) et FERRARO (P). – Distance multi-échelles entre structures secondaires d'ARNs. *In : 5èmes Journées Ouvertes Biologie Informatique Mathématiques*, p. 81. – Montréal, Canada, June 28-30 2004.
- [248] PECHER (A) et ANNEGRET (W). – La plupart des webs ne sont pas rang-parfaits. *In : Journées Polyèdres et Optimisation Combinatoire JPOC*. pp. 32–35. – Clermont-Ferrand, 2003.
- [249] PECHER (A), ANNEGRET (W) et XUDING (Z). – Quelques familles de graphes minimaux imparfaits circulaires. *In : 6èmes Journées Graphes et Algorithmes*. pp. 82–84. – Les Cahiers du laboratoire Leibniz, Grenoble, 2004.

Conférences invitées

- [250] BOUSQUET-MÉLOU (M). – Sur la forme des arbres binaires. *In : Journées d'Informatique mathématique du GDR ALP et du RTP MathInfo*. – Paris, 2004.
- [251] VIENNOT (X). – Combinatoire de l'algèbre de Temperley-Lieb nilpotente et fonctions symétriques. – Septembre 2002. *Journées du GDR Tresses*, Lacanau.
- [252] VIENNOT (X). – Partitions d'entiers, combinatoire des q -séries et physique. – Octobre 2002. *Journées en l'honneur de Jean-Louis Nicolas*, Lyon.
- [253] VIENNOT (X). – Empilements de pièces en physique. – Juin 2003. *Journées du LIPN, Combinatoire, Informatique et Physique*, Université Paris Nord, Villetaneuse.
- [254] VIENNOT (X). – Symétries et pavages avec la magie des tableaux de Young. – Mars 2003. *14ème colloque Maths en Jean, l'esprit et le gout de la recherche*, Bordeaux.
- [255] VIENNOT (X). – Une histoire de la conjecture des matrices à signes alternants. – Juin 2003. *Journées du LIPN, Combinatoire, Informatique et Physique*, Université Paris Nord, Villetaneuse.
- [256] VIENNOT (X). – Des rivières, des molécules d'arn, des synthèses d'images d'arbres, etc : quelles mathématiques ? – Aout 2004. *Ecole d'été de l'IREM, Sciences mathématiques et modélisation*, Université Bordeaux I.
- [257] VIENNOT (X). – Les dessous mathématiques des empilements : des dominos à la gravitation quantique en passant par la physique statistique. – Aout 2004. *Ecole d'été de l'IREM, Sciences mathématiques et modélisation*, Université Bordeaux I.

Chapitres d'ouvrages

- [258] DURAND (B) et ZVONKIN (A). – Complexité de Kolmogorov. *In : L'héritage de Kolmogorov en mathématiques*, éd. par Charpentier (E), Lesne (A) et Nikolski (N), pp. 269–287. – Belin, 2004.
- [259] ZVONKIN (A). – Cartes et dessins d'enfants. *In : Graphes*, éd. par Berilne (N) et Sabbah (C), pp. 53–74. – Les éditions de l'École Polytechnique, 2004.

1.4.3 Autres publications

Notices d'emplois de logiciels

- [260] GODIN (C), GUÉDON (Y), BELLOUTI (S), NOUGUIER (C), FERRARO (P), DONES (N) et ADAM (B). – *Introduction and reference manual, AMAPmod version 1.5*. – Rapport technique, Cirad, 2002.

Rapports internes et autres publications

- [261] ABRAHAM (I), GAVOILLE (C) et MALKHI (D). – *Routing with Improved Communication-Space Trade-Off*. – Research Report nRR-1330-04, University of Bordeaux 1, 351, cours de la Libération, 33405 Talence Cedex, France, LaBRI, juillet 2004.
- [262] BAZZARO (F) et GAVOILLE (C). – *Localized and Compact Data-Structure for Comparability Graphs*. – Research Report nRR-1343-05, University of Bordeaux 1, 351, cours de la Libération, 33405 Talence Cedex, France, LaBRI, février 2005.
- [263] BONICHON (N), GAVOILLE (C) et HANUSSE (N). – *An Information Upper Bound of Planar Graphs Using Triangulation*. – Research Report nRR-1279-02, University of Bordeaux 1, 351, cours de la Libération, 33405 Talence Cedex, France, LaBRI, septembre 2002.
- [264] BONICHON (N), GAVOILLE (C) et HANUSSE (N). – *Canonical Decomposition of Outerplanar Maps and Application to Enumeration, Coding and Generation*. – Research Report nRR-1295-03, University of Bordeaux 1, 351, cours de la Libération, 33405 Talence Cedex, France, LaBRI, mars 2003.
- [265] BONICHON (N), GAVOILLE (C), HANUSSE (N), POULALHON (D) et SCHAEFFER (G). – *Planar Graphs, via Well-Orderly Maps and Trees*. – Research Report, University of Bordeaux 1, 351, cours de la Libération, 33405 Talence Cedex, France, LaBRI, septembre 2004.
- [266] BOUSQUET-MÉLOU (M). – *Limit laws for embedded trees. Applications to the integrated superBrownian excursion*. – Rapport technique, ArXiv math.CO/0501266, 2005.
- [267] BOUSQUET-MÉLOU (M). – *Three osculating walkers*. – Rapport technique, ArXiv math.CO/0504153, 2005.
- [268] BOUSQUET-MÉLOU (M) et JEHANNE (A). – *Polynomial equations with one catalytic variable, algebraic series, and map enumeration*. – Rapport technique, ArXiv math.CO/0504018, 2005.
- [269] DELMAS (O), HAVET (F), MONTASSIER (M) et PÉRENNES (S). – *Design of fault tolerant on-board networks*. – Research report nRR 1345-05, Laboratoire bordelais de recherche en informatique (LaBRI), Unité Mixte de Recherche CNRS (UMR 5800), 2005.
- [270] DERBEL (B) et MOSBAH (M). – *A linear time algorithm for graph decomposition*. – Rapport technique nRR-1315-04, LaBRI, 2004. <http://www.labri.fr/Labri/Publications/Rapports-internes/Publications.htm>.
- [271] DOURISBOURE (Y) et GAVOILLE (C). – *Improved Compact Routing Scheme for Chordal Graphs*. – Research Report nRR-1274-02, University of Bordeaux 1, 351, cours de la Libération, 33405 Talence Cedex, France, LaBRI, mai 2002.
- [272] DOURISBOURE (Y) et GAVOILLE (C). – *Small Diameter Bag Tree-Decompositions*. – Research Report nRR-1326-04, University of Bordeaux 1, 351, cours de la Libération, 33405 Talence Cedex, France, LaBRI, mai 2004.
- [273] DOURISBOURE (Y) et GAVOILLE (C). – *Sparse Additive Spanners for Bounded Tree-Length Graphs*. – Research Report nRR-1316-04, University of Bordeaux 1, 351, cours de la Libération, 33405 Talence Cedex, France, LaBRI, février 2004.

- [274] DUCHON (P), HANUSSE (N), SAHEB (N) et ZEMMARI (A). – *Broadcast in the Rendezvous Model*. – Rapport technique, Toronto University, 2003. Présentation courte à DISC'2003.
- [275] DULUCQ (S) et TOUZET (H). – *Algorithmes pour la comparaison de structures d'ARN*. – Rapport technique, LaBRI, mars 2002. Journées Algorithmes pour la Bio-informatique (AlBio).
- [276] FERRARO (P). – *Rapport de Recherche Post-doctoral 2001-2002*. – Rapport technique n 2, Cirad, 2002.
- [277] FRAIGNIAUD (P) et GAVOILLE (C). – *Lower Bounds for Oblivious Single-Message End-to-End Communication*. – Research Report n RR-1289-03, University of Bordeaux 1, 351, cours de la Libération, 33405 Talence Cedex, France, LaBRI, février 2003.
- [278] FRAIGNIAUD (P), GAVOILLE (C) et PAUL (C). – *Eclecticism Shrinks the World*. – Research Report n LRI-1376, Université Paris-Sud, 91405 Orsay cedex, France, LRI, octobre 2003.
- [279] GAVOILLE (C) et NEHÉZ (M). – *Interval Routing in Reliability Networks*. – Research Report n RR-1294-03, University of Bordeaux 1, 351, cours de la Libération, 33405 Talence Cedex, France, LaBRI, mars 2003.
- [280] GAVOILLE (C) et PAUL (C). – *Small Universal Distance Matrices*. – Research Report n RR-1290-03, University of Bordeaux 1, 351, cours de la Libération, 33405 Talence Cedex, France, LaBRI, février 2003.
- [281] GONÇALVES (D) et MONTASSIER (M). – *Acyclic choosability of graphs with small maximum degree*. – Rapport technique n RR-1348-05, LaBRI, 2005.
- [282] MONTASSIER (M). – *On acyclic choosability of planar graphs with large girth*. – Rapport technique n RR-1336-04, LaBRI, 2004.
- [283] MONTASSIER (M), OCHEM (P) et RASPAUD (A). – *On the acyclic choosability of graphs*. – Rapport technique n RR-1335-04, LaBRI, 2004.
- [284] MONTASSIER (M) et RASPAUD (A). – *(d,1)-total labelling of graphs with a given maximum average degree*. – Rapport technique n RR-1308-03, LaBRI, 2003.
- [285] PINLOU (A) et SOPENA (E). – *The acircuitic directed star arboricity of subcubic graphs is at most four*. – Research Report n RR-1338-04, LaBRI, octobre 2004.

1.4.4 Formation par la recherche

Thèses

- [286] AUBER (D). – *Outils de visualisation de larges structures de données*. – Thèse de PhD, Université Bordeaux 1, December 2002.
- [287] AUILLANS (A). – *Modélisation de réseaux sémantiques par des hypergraphes et applications*. – Thèse de PhD, Université Bordeaux I, 2005.
- [288] BAZZARO (F). – *Etiquetage de graphes : codage et coloration*. – Thèse de PhD, Université Bordeaux 1, 2005.
- [289] DOLAMA (M. H). – *Contribution à l'étude de quelques problèmes de coloration de graphes*. – Thèse de PhD, Université Bordeaux 1, 2005.
- [290] DOURISBOURE (Y). – *Routage compact et longueur arborescente*. – Thèse de PhD, Université Bordeaux 1, 2003.

- [291] GODARD (E). – *Réécritures de graphes et algorithmique distribuée*. – Thèse de PhD, Université Bordeaux 1, 2002.
- [292] JAMES (W). – *Heaps of pieces and Polygons in Combinatorics, Statistical mechanics and Quantum gravity*. – Thèse de PhD, Université Bordeaux 1 and Melbourne University, 2005. thèse soumise, soutenance prévue en 2005, sous la direction de Xavier VIENNOT et Tony GUTTMANN.
- [293] LE BORGNE (Y). – *Variations combinatoires sur des classes d'objets comptées par la suite de Catalan*. – Thèse de PhD, Université Bordeaux 1, Décembre 2004.
- [294] LESUR (I). – *Study of the transcriptome of the prematurely aging dna2-1 yeast mutant using a new system allowing comparative DNA microarray analysis*. – Thèse de PhD, Université Bordeaux 1, 2005.
- [295] TICHIT (L). – *Algorithmique des structures biologiques : l'édition d'arborescences pour la comparaison de structures secondaires d'ARN*. – Thèse de PhD, Université Bordeaux 1, septembre 2003.
- [296] VANDENBOGAERT (M). – *Algorithmes et mesures statistiques pour la recherche de signaux fonctionnels dans les zones de régulation*. – Thèse de PhD, Université Bordeaux 1, mars 2004.

1.5 Collaborations internationales

La naissance du réseau européen ACE (Algebraic Combinatorics in Europe) a renforcé nos liens avec plusieurs partenaires européens, à savoir les équipes de combinatoire de Marburg (Allemagne), Linköping (Suède), Vienne (Autriche), Rome (Italie), Jérusalem (Israël) (et les équipes françaises de Marne-la-Vallée et Lyon). Par ailleurs, nous avons depuis de nombreuses années une collaboration suivie avec le LACIM (Laboratoire de Combinatoire et d'Informatique Mathématique) à l'Université du Québec à Montréal, avec le groupe de combinatoire de l'Université de Florence, et avec le groupe de physique statistique de l'Université de Melbourne. Ces trois collaborations ont donné lieu à des publications communes bien sûr, mais aussi à des thèses encadrées en co-tutelle. Citons enfin une collaboration plus récente avec l'équipe de combinatoire algébrique de l'Université d'York (et du Fields Institute) à Toronto.

La théorie des graphes au LaBRI a pris son véritable essor dans les années 90. Nos travaux sur les colorations orientées ont retenu alors l'attention de nombreux chercheurs étrangers, ce qui a débouché sur des collaborations internationales très fortes en particulier avec des chercheurs de Prague (Charles University) et de l'Institut Sobolev de Novosibirsk. Ces collaborations nous ont permis d'élargir nos champs d'investigations avec d'autres collaborations (Simon Fraser University, Vancouver, UPC Barcelone, Université Comenius). Nos travaux sur les réseaux d'interconnexions ont été aussi porteurs de collaborations avec des chercheurs de l'académie des Sciences de Bratislava et l'UPC de Barcelone. L'arrivée de jeunes collègues, travaillant sur des problèmes de routages et d'optimisation, a contribué à ce développement international par des collaborations avec l'Institut Weizman, l'Université de Kent, l'Université de Carleton et le ZIB de Berlin.

Les monoïdes de commutation et l'algorithmique distribuée sont à l'origine de nombreuses collaborations ; les plus importantes et les plus régulières ont lieu avec :

- l'institut d'informatique de Varsovie (académie des sciences polonaises),
- le groupe d'informatique théorique de l'université de Stuttgart,
- le département d'informatique de l'université d'Utrecht.

1.5.1 Invités

Abraham, I., *Jérusalem - Israël*
 Alstrup, S., *Copenhague - Danemark*
 Andren, D., *Umea - Suède*
 Bergeron, N., *Toronto - Canada*
 Bona, M., *Gainsville - USA*
 Borodin, O., *Novosibirsk - Russie*
 Caselli, F., *Rome - Italie*
 Comellas, F., *Barcelone - Espagne*
 Diekert, V., *Stuttgart - Allemagne*
 Dragan, F., *Kent - USA*
 Dukes, M.
 Eisenkölbl, T., *Vienne - Autriche*
 Fon der Flaass, D., *Novosibirsk - Russie*
 Hell, P., *Vancouver - Canada*
 Kavvadias, D., *Patras - Grèce*
 Kaplan, H., *Tel-Aviv - Israël*
 Kostochka, A. V., *Champaign - USA*

Kranakis, E., *Ottawa - Canada*
 Krizanc, D., *Weysleyan - USA*
 Kostochka, A. V., *Champaign - USA*
 Macajova, E., *Bratislava - Slovaquie*
 Malkhi, D., *Jérusalem - Israel*
 Mans, B., *Sydney - Australie*
 Mansour, T., *Haïfa - Israel*
 Mazurkiewicz, A., *Varsovie - Pologne*
 Nehéz, M., *Bratislava - Slovaquie*
 Neseřil, J., *Prague - Tchèque*
 Nisan, N., *Jérusalem - Israel*
 Ouchterlony, E., *Linköping - Suède*
 Peleg, D., *Rehovot - Israel*
 Ponetti, M., *Florence - Italie*
 Rauhe, T., *Copenhague - Danemark*
 Raz, R., *Rehovot - Israel*
 Rubey, M., *Vienne - Autriche*
 Rubey, M., *Vienne - Autriche*
 Serra, O., *Barcelone - Espagne*
 Shabat, G., *Moscou - Russie*
 Skoviera, M., *Bratislava - Slovaquie*
 Steingrimsson, E., *Göteborg - Suède*
 Sykora, O., *Loughborough - Angleterre*
 Tel, G., *Utrecht - Pays Bas*
 Thapper, J., *Linköping - Suède*
 Thorup, M., *ATT Labs - USA*
 Toli, I., *Moscou - Russie*
 Vrto, I., *Bratislava - Slovaquie*
 Wagler, A., *Berlin - Allemagne*
 West, J., *Vancouver - Canada*
 Zhu, X., *Kaohsiung - Taiwan*

1.5.2 Invitations

Aval, J.-C., *York University, Toronto, 2003*
 Bousquet-Mélou, M., *Université de Göteborg, 2004*
 Bousquet-Mélou, M., *Institut Mittag-Leffler (Suède), 2005*
 Cori, R., *Université de Rome, 2003*
 Lovejoy, J., *Université de Wisconsin, 2003*
 Lovejoy, J., *Université de Melbourne, 2004*
 Lovejoy, J., *Université de Nankai, 2004*
 Pêcher, A., *IASI, Rome, 2002*
 Pêcher, A., *ZIB, Berlin, 2002*
 Raspaud, A., *U. Southern, Denmark, Odense, 2003*
 Raspaud, A., *Charles University, Prague, 2004*
 Raspaud, A., *Université de Montréal, 2005*
 Raspaud, A., *Sun Yat-sen University, Kaoshiung Taiwan, 2005*
 Sopena, E., *U.Southern, Denmark, Odense, 2003*
 Viennot, X., *Université libre de Berlin, 2002*

Viennot, X., *Université de Florence*, 2002

Viennot, X., *Tata Institute Mumbai, IIT Mumbai, IIT Chennai*, 2003

Viennot, X., *Chennai Mathematical, Université de Pondicherry*, 2003

Viennot, X., *Université Charles*, 2003

Viennot, X., *Centro de Modelamiento Matematico, Santiago (Chili)*, 2004

Viennot, X., *Institut Mittag-Leffler (Suède)*, 2005

1.6 Liste des contrats et subventions

1.6.1 Contrats internationaux et européens

Réseau européen ACE (Algebraic Combinatorics in Europe), du 5e PCRD (2002-2005).
PICS “Mathématiques et Informatique” entre le LaBRI et le LACIM (Montréal) (2005-2007).
PICS “Visualisation” entre Amsterdam et Bordeaux (2001-2003).
PAI (Programme d’actions intégrées) Barande (Prague - Bordeaux).
PAI Picasso (Barcelone - Bordeaux).
PAI Stefanik (Bratislava - Bordeaux).
Réseau COMBSTRU (5e PCRD).

1.6.2 Contrats avec des entreprises

1.6.3 Contrats publics avec des instances nationales

ACI Masse de données : Navgraphe (2003-2006).
ACI Masse de données : Pair à Pair (2003-2006).
ACI Nouvelles interfaces des mathématiques “Arbres et chemins : probabilités et algorithmes” (2003-2006).
ACI Masse de données “GeoComp” (2004-2006).
ACI Jeunes chercheurs “Partitions d’entiers, à l’interface de la combinatoire, des q-séries et de la théorie des nombres” (2004-2007).
Projet MathSTIC “Combinatoire probabiliste des structures algorithmiques fondamentales” (2002-2003).
AS CNRS : Visualisation d’informations (2001-2002).
AS CNRS : Algorithmique des grands graphes (2003).
ACI jeunes chercheur “Cube de données : construction et navigation interactive” (2004-2006).

1.6.4 Contrats publics avec instances locales ou régionales

Globulyss : Pôle EITICA région aquitaine et l’entreprise In-exVivo ; Globulyss : RIAM avec l’entreprise In-exVivo ; Navigation dans les bases de données vidéo (conseil régional d’aquitaine) (2002-2003).

1.7 Animation de la recherche

1.7.1 Comités de rédaction

Advances in Applied Mathematics,
 Journal of combinatorial theory (Series A),
 RAIRO Theoretical informatics and applications,
 Mathematical Reviews,
 The electronic journal of combinatorics,
 Séminaire lotharingien de combinatoire.

1.7.2 Comités de programmes de colloques

Rencontres francophones sur les aspects algorithmiques des télécommunications (Algotel) (2002, 2003, 2004, 2005).
 Colloquium on Structural Information and Communication Complexity (SIROCCO) (2002, 2003, 2004).
 IEEE International Parallel and Distributed Processing Symposium (IPDPS) (2002).
 Symposium on Theoretical Aspects of Computer Science (STACS) (2002, 2004).
 Conference on Formal Power Series and Algebraic Combinatorics (FPSAC) (2002, 2003, 2005).
 Discrete Random Walks Conference (DRW) (2003).
 European Conference on Combinatorics, Graph Theory, and Applications (2003).
 Graph Theory 2004 : a conference in memory of Claude Berge (2004).
 International Workshop on Distributed Computing (2005).
 International Conference on High Performance Computing (HiPC) (2005).
 International Conference on PARALLEL AND DISTRIBUTED COMPUTING AND NETWORKS (PDCN) (2005).
 ACM Symposium on Principles of Distributed Computing (PODC) (2005).

1.7.3 Organisation de colloques, écoles jeunes chercheurs, ...

GDR ALP (groupe de travail "Graphes et algorithmes", École jeunes chercheurs "Interactions et visualisation de l'information" (2004),
 Journées "Arbres et chemins" (2004),
 École d'été du réseau européen ACE "Combinatoire des groupes et des algèbres" (2004),
 membre du comité de pilotage des "Leçons de mathématiques et d'informatique d'aujourd'hui" (ces leçons sont publiées par les éditions Cassini).
 Organisation du 54e séminaire lotharingien de combinatoire (Lucelle, Alsace 2005).

1.7.4 Administration de la Recherche

L'équipe a fourni 16 membres de jury de thèses dont 11 rapporteurs. Elle a également fourni 5 membres de jury d'HDR dont 3 rapporteurs.
 Participation au comité de pilotage du réseau thématique du département CNRS-STIC "Mathématiques de l'informatique" 2002-2004).
 Expert auprès de la Mission Scientifique Universitaire (MSU) (Y. Métivier).
 Membres du CNU (M. Delest, S. Dulucq, E. Sopena).
 Arbitrage depuis 1999 de Research grants council of Hong Kong (A. Raspaud, 2/an).
 Expertise pour l'université franco-italienne (Y. Métivier).
 Expert auprès de la MSTP (M. Delest).

Membre du comité scientifique de l'ACI Masse de données (M. Delest depuis 2003 Co-présidente depuis 2004).

Expertise pour le Canada et les Pays Bas (M. Delest).

Expertise de projets ACI masse de données (M. Delest).

Logiques, Automates, Algorithmes et Applications (L3A)

Responsable(s) Pascal Weil

Logiques et graphes

Responsable(s) Bruno Courcelle et Igor Walukiewicz

Mots-clés : Décomposition de graphes, Reconnaissabilité, Logique du second-ordre monadique, Mu-calcul, Vérification, Synthèse, Assistants de preuves, Descriptions finies d'objets dénombrables

Langages et Automates

Responsable(s) Géraud Sénizergues et Pascal Weil

Mots-clés : Automates, Mots, Arbres, Traces, Reconnaissabilité, Automates à pile, Semigroupes, Equations, Groupes libres, Réécriture, Mu-calcul, Logiques locales, Topologie

Informatique et Linguistique

Responsable(s) Christian Retoré

Mots-clés : Linguistique Informatique, Traitement Automatique des Langues, Syntaxe et Sémantique du Langage Naturel, Grammaires Formelles, Logique des Ressources, Logique Linéaire, Calcul de Lambek, Lambda-calcul

2.1 Composition de l'équipe

2.1.1 Membres permanents

BAUDERON, Michel, Pr, Bx1, en détachement MAE depuis septembre 2004
 BILLAUD, Michel, MdC, Bx1
 CARRÈRE, Frédérique, MdC, Bx1
 CASTÉРАН, Pierre, MdC, Bx1, depuis janvier 2003, en détachement puis délégation INRIA en 2003-05
 COURCELLE, Bruno, Pr, Bx1
 DURAND, Irène, MdC, Bx1, en CRCT à l'Université de Stuttgart en 2004-05
 HENRY, Patrick, Ingénieur, CNRS, depuis mars 2004
 HUET, Gérard, DR, INRIA, de septembre 2002 à juin 2003
 JANIN, David, MdC, ENSEIRB
 LAPOIRE, Denis, MdC, ENSEIRB
 LE SAËC, Bertrand, Pr, ENSEIRB
 LY, Olivier, MdC, Bx1, depuis octobre 2003
 MAABOUT, Sofian, MdC, Bx4
 MARLET, Renaud, CR, INRIA, depuis décembre 2004
 MOOT, Richard, CR, CNRS, depuis octobre 2004
 MOSBAH, Mohamed, Pr, ENSEIRB
 NOVELLI, Noël, MdC, Bx1, jusqu'à septembre 2004
 RETORÉ, Christian, Pr, Bx1, depuis septembre 2002
 SÉNIZERGUES, Géraud, Pr, Bx1, en CRCT à l'Université de Stuttgart en 2004-05
 WALUKIEWICZ, Igor, CR, CNRS
 WEIL, Pascal, DR, CNRS, mis à disposition de l'Université du Nebraska en 2003-04

2.1.2 Membres non permanents

AMBLARD, Maxime, doct (MENRT), Bx1, depuis octobre 2003
 ANOUN, Houda, doct (MENRT), Bx1, depuis octobre 2003
 BERNET, Julien, doct (MENRT) puis ATER, Bx1, depuis octobre 2001
 BONATO, Roberto, doct (bourse italienne), Bx1, depuis octobre 2002
 BOUQUET, Alexis-Julien, doct (MENRT), Bx1, depuis octobre 2002
 CHEN, Rui, doct (gouvernement chinois), Bx1, depuis octobre 2002
 FRATANI, Séverine, doct (MENRT), Bx1, depuis octobre 2002
 HAMID, Brahim, doct (financement propre), Bx1, depuis octobre 2003
 LE NIR, Yannick, doct (ATER), Bx1, d'octobre 2002 à septembre 2004
 SELLAMI, Afif, doct (financement propre, puis ATER), Bx1, depuis octobre 2000, thèse soutenue en décembre 2004, ATER depuis 2003

BLUMENSATH, Achim, post-doc (GAMES), Bx1, de janvier 2004 à décembre 2004
 BOJAŃCZYK, Mikołaj, doctorant (GAMES), Bx1, d'octobre 2002 à décembre 2002 et de janvier 2004 à mars 2004
 LENZI, Giacomo, post-doc (Marie Curie), Bx1, jusqu'à avril 2002
 MOHALIK, Swarup, post-doc (MENRT), Bx1, d'octobre 2002 à mars 2004

MOOT, Richard, post-doc (ERCIM) puis ATER, Bx1, de septembre 2002 à octobre 2004

2.2 Présentation de l'équipe

Le cadre général des travaux de l'équipe peut être décrit sommairement de la façon suivante : il s'agit de pouvoir manipuler, efficacement le cas échéant, des objets en général infinis et discrets, au moyen de représentations finies. Cette ambition est l'un des défis auxquels l'informatique est confrontée depuis son apparition, et forme le fondement d'une branche très importante de l'informatique théorique.

Les objets en question sont a priori ceux qui répondent aux besoins de modélisation de l'informatique : à l'origine, il s'est agi d'ensembles de mots finis (langages), pour représenter d'abord les programmes informatiques, puis des processus séquentiels terminants ; très rapidement les informaticiens ont considéré des ensembles de mots infinis pour les processus qui ne se terminent pas ; des ensembles d'arbres finis ou infinis pour décrire des scénarios de branchement liés soit à des choix (non-déterminisme), soit à une forme de parallélisme ou de concurrence – ou pour décrire des représentations syntaxiques et sémantiques en linguistique. Aujourd'hui, les besoins en modélisation amènent aussi à considérer des langages d'ensembles partiellement ordonnés pour la modélisation du parallélisme, et dans la plus grande généralité des ensembles de graphes.

La représentation de systèmes complexes au moyen de ces objets vise à rendre possible l'analyse de ces systèmes. Typiquement, on souhaite représenter l'ensemble des comportements possibles d'un système et déterminer, sur la base de cette représentation, si le système respecte telle ou telle condition. La nécessité, face à la complexité des systèmes, de procéder à cette analyse de façon automatique ou semi-automatique impose que représentations et spécifications soient données de façon finie.

Les outils à notre disposition pour représenter tant les systèmes que leurs spécifications sont très variés : il s'agit des automates, dans leurs nombreuses variantes (automates finis, à pile, à pile de piles, temporisés, d'arbres, branchants, etc), des expressions rationnelles, du λ -calcul, des divers formalismes logiques (logique du premier ordre, logique monadique du second ordre, logique temporelle, μ -calcul, logique des ressources, etc). Dans certains cas, on utilise aussi des représentations de nature plus algébrique (systèmes d'équations dans un cadre algébrique approprié, monoïdes finis, etc).

Le détail des travaux menés dans l'équipe sur ces thèmes est rapporté selon 3 thèmes et un projet :

- Thème *Logiques et Graphes* (resp. B. Courcelle, I. Walukiewicz),
- Thème *Langages et Automates* (resp. G. Sénizergues, P. Weil),
- Thème *Informatique et linguistique* (resp. Ch. Retoré),
- Projet VISIDIA (resp. M. Mosbah).

Le thème *Logiques et Graphes* se concentre principalement sur la logique monadique du second ordre et le μ -calcul, sur les arbres et les graphes, en s'intéressant d'une part à leur pouvoir expressif, et d'autre part à leur utilisation pour l'obtention d'algorithmes efficaces. Ces travaux sont motivés par des applications en vérification, en synthèse et en évaluation de requêtes. Certaines applications à la vérification passent aussi par l'utilisation d'assistants de preuves. Les graphes sont traités ici dans une perspective et avec des outils issus de la théorie des langages et des grammaires et non avec ceux de la combinatoire, même s'il y a naturellement dialogue avec les thèmes correspondants de l'équipe *Combinatoire et Algorithmique*. Plus précisément, au delà des applications des graphes à la modélisation, nous utilisons des méthodes algébriques et logiques pour décrire la structure et les propriétés de classes de graphes. L'objectif est alors de rapprocher une problématique issue de la théorie des langages formels des notions pertinentes en théorie des graphes.

Le thème *Langages et Automates* rassemble les travaux sur la classification et la com-

plexité des propriétés algorithmiques des langages de mots, d'arbres et de traces. Les outils principaux sont les automates et certains invariants algébriques. C'est cet aspect, en particulier pour les langages d'arbres, qui distingue ces travaux de ceux du thème *Logiques et Graphes*. On développe aussi dans l'équipe des travaux de recherche sur les structures algébriques, soit parce qu'elles s'avèrent être des outils efficaces en théorie des langages (comme les semigroupes finis), soit parce que les méthodes de la théorie des automates y sont particulièrement utiles (équations dans les semigroupes, combinatoire des groupes libres, 3-variétés arborescentes).

Le thème *Informatique et Linguistique* s'intéresse à la modélisation formelle du langage naturel, et en particulier de sa syntaxe et de sa sémantique. Selon qu'il s'agit de modéliser la syntaxe, la sémantique logique ou l'intégration de ces deux aspects, il est fait appel au formalisme des grammaires (de mots ou d'arbres, catégorielles ou minimalistes, etc), ou à celui du λ -calcul typé, voire à celui de la λ -DRT. La langue des signes française constitue un champ d'application privilégié. Les travaux de ce thème donnent lieu à la production de logiciels sous licence LGPL.

Les travaux du projet VISIDIA (Visualisation de structures complexes évolutives et d'algorithmes distribués sur ces structures) sont rapportés dans une autre partie de ce rapport d'activité. On retiendra simplement ici que ce thème concerne aussi les travaux de l'équipe *Combinatoire et Algorithmique* et qu'il mêle techniques de transformation et réécriture de graphes et problématique d'algorithmique distribuée. Le projet s'articule autour du développement d'un environnement logiciel pour l'implémentation, la visualisation et l'expérimentation concernant les algorithmes distribués sur des structures complexes.

Soulignons le choix fait par l'équipe de ne pas donner de listes de membres des thèmes. De fait, certains membres de l'équipe contribuent de façon substantielle à plusieurs de ces thèmes, et ces contributions multiples sont encouragées. D'ailleurs, la distinction faite entre les thèmes *Logiques et Graphes* et *Langages et Automates*, si elle peut facilement être justifiée scientifiquement, cherche aussi à rendre plus lisible une activité qui porte sur un champ continu – de sorte que certains travaux charnières sont arbitrairement rapportés dans un thème plutôt que dans l'autre. L'identité scientifique du thème *Informatique et Linguistique* et celle du projet VISIDIA sont plus affirmées, mais ils se rattachent de façon naturelle à l'équipe L3A par leur intérêt commun pour des méthodes formelles (grammaires, réécritures, logiques) développées dans les autres thèmes, qu'ils appliquent ensuite à des champs particuliers, la linguistique d'une part et les algorithmes distribués d'autre part.

L'évolution de ces thèmes, par rapport au précédent rapport d'activité (2001), reflète l'évolution de l'équipe. En 2001, nous annonçons un projet en *Jeux, Logiques et Automates* (resp. D. Janin). Ce projet a pleinement réussi à faire connaître et à développer l'activité scientifique autour des jeux, qui émergeait à ce moment-là dans plusieurs laboratoires européens et il a été l'un des créateurs du projet européen GAMES (voir plus loin). L'étude et l'utilisation des jeux se sont avérées très importantes en vérification, théorique et pratique, et on les retrouve dans un bon nombre de publications et de sujets de thèses en cours, qui se rattachent aux thèmes *Langages et Automates* et *Logiques et graphes* de L3A, et au thème *Vérification* de l'équipe MVTSI.

Nous annonçons aussi en 2001 un projet en *Bases de données* (resp. S. Maabout). L'équipe venait de voir partir N. Bidoit (PU, mutation Paris-11) et de recruter N. Novelli (MdC). S. Maabout et N. Novelli ont en effet mené des travaux fructueux en algorithmique de la fouille de données (passage en temps et en espace linéaire du produit de partitions). Cependant N. Novelli a quitté le LaBRI en 2004 (mutation vers l'Université de la Méditerranée et le LIF). De son côté, S. Maabout a réorienté son activité, en mettant à profit ses compétences

dans le traitement et l'interrogation de données semi-structurées par des collaborations avec R. Castanet (MVTISI) concernant l'utilisation d'XML en milieu industriel, et avec le thème *Visualisation d'information* (M. Delest, Combinatoire et Algorithmique). Il a donc été décidé d'un commun accord de ne plus faire état pour le moment d'une activité de recherche en bases de données.

En revanche, le thème *Informatique et Linguistique* est une activité nouvelle au LaBRI. Impulsée par l'arrivée au laboratoire de Ch. Retoré, CR1 INRIA, et de Y. Le Nir, ATER (2002), puis confortée par le recrutement de Ch. Retoré comme professeur à Bordeaux-1 (2003), par le recrutement de R. Moot (CR2 CNRS, 2004) et l'arrivée de R. Marlet (CR1 INRIA, 2004), la recherche en linguistique au LaBRI a connu un développement rapide. Cette recherche s'effectue en collaboration avec des linguistes de Bordeaux-3 (UMR ERSS) et pour une partie importante dans le cadre d'un projet commun INRIA Futurs - LaBRI, le projet SIGNES. Il n'y a pas cependant coïncidence exacte entre ce thème et ce projet, d'une part parce que le projet SIGNES comporte des membres de l'UMR ERSS, et d'autre part parce que contribuent également au thème *Informatique et Linguistique* des travaux de logique concernant l'application à la linguistique de méthodes venues de la théorie de la démonstration (COQ).

La liste des membres de l'équipe a connu d'autres évolutions depuis le rapport d'activité 2001 que celles déjà évoquées (concernant les bases de données et la linguistique). Le recrutement d'I. Walukiewicz (CR1 CNRS, 2001) et celui d'O. Ly (MdC, 2003) ont commencé à rajeunir l'équipe. Les activités en logique et le lancement d'une activité en informatique et linguistique ont amené P. Castéran à changer d'équipe au sein du LaBRI et à nous rejoindre. On notera par contre l'éloignement de M. Bauderon, en détachement auprès du MAE (conseiller scientifique au consulat de France à Shanghai depuis 2004).

On trouvera plus loin le détail des activités d'animation de la recherche aux échelons national et international, ainsi que des contrats industriels, régionaux, nationaux et internationaux que nous avons obtenus. Au niveau national, on mentionnera en particulier une Action Spécifique du département STIC (resp. P. Weil) sur les modèles d'automates distribués et temporisés, à laquelle a succédé une ACI *Sécurité de l'Information* (resp. I. Walukiewicz) sur la vérification de systèmes distribués.

Le projet européen GAMES (RTN, resp. D. Janin) a joué un rôle essentiel dans la vie et les ressources de l'équipe. Faisant suite à un projet JEMSTIC sur les jeux (resp. D. Janin), GAMES a d'une part financé l'accueil de stagiaires (le LaBRI a bénéficié de 6 mois d'accueil pré-doctoral et 12 mois d'accueil post-doctoral) et l'organisation de conférences (dont l'une a réuni plus de quatre-vingt participants pendant 4 jours à Bordeaux). D'autre part, en dynamisant un certain nombre de collaborations internationales existantes et en suscitant de nouvelles, GAMES a fortement contribué à structurer l'activité du LaBRI sur les fondements théoriques de la vérification, notamment en associant des membres des équipes L3A et MVTISI.

Nous mentionnons enfin avec plaisir les distinctions reçues par certains membres de l'équipe. En particulier, B. Courcelle a reçu le *Prix français des Citations 2004*, décerné par l'INIST (CNRS) et Thomson-Scientific (il figurait parmi les 15 chercheurs de France les plus cités, toutes disciplines confondues). G. Sénizergues a reçu le *Prix Gödel* en 2002 (parmi les 32 lauréats de ce prix à ce jour, il est le seul Français), puis le *Prix Humboldt* en 2004. Enfin, I. Walukiewicz a reçu la médaille de bronze du CNRS en 2003.

2.3 Thèmes de recherche de l'équipe

2.3.1 Logiques et Graphes

Les arbres, et plus généralement les graphes sont des objets combinatoires auxquels les méthodes de la logique et de la théorie des langages formels peuvent être appliquées. En particulier, les graphes modélisent les processus de façon conceptuellement simple et pertinente. Notre travail se concentre sur la description des arbres, des graphes et de leurs propriétés, par des moyens algébriques et logiques. Les deux langages principaux sont la logique du second-ordre monadique (LSOM) et le μ -calcul. Au-delà de son intérêt fondamental, cette recherche est motivée par des applications à l'évaluation des requêtes et à la vérification et à la synthèse des processus.

Mots-Clés Décomposition de graphes, Reconnaissabilité, Logique du second-ordre monadique, Mu-calcul, Vérification, Synthèse, Assistants de preuves, Descriptions finies d'objets dénombrables

Méthodes algébriques

Le point de départ est la notion d'une signature d'opérations de graphes et la vision des graphes comme des objets construits avec ces opérations. Cela permet de généraliser les concepts fondamentaux de la théorie des langages tels que la reconnaissabilité, les automates *d'arbres* (qui opèrent sur des termes), les grammaires algébriques, et leurs propriétés telles que l'équivalence entre la reconnaissabilité d'un ensemble et sa définissabilité en LSOM. Les graphes étant beaucoup plus complexes que les mots, il est nécessaire d'utiliser plusieurs opérations au lieu de la seule concaténation pour les mots. Différents ensembles d'opérations donnent différentes classes d'ensembles reconnaissables et d'ensembles algébriques qui généralisent, respectivement, des langages reconnaissables et des langages algébriques (*context-free*). Nous avons étudié les rapports entre les trois principales signatures d'opérations de graphes désignées par les sigles HR, VR et MD. La première, HR, est liée aux grammaires de graphes basées sur des remplacements d'hyperarcs et donne une formulation algébrique des décompositions arborescentes et de la notion correspondante de largeur (*tree-width*). La seconde, VR, est liée aux grammaires de graphes basées sur des remplacements de sommets et fonde la *largeur de clique*, autre mesure de complexité des graphes. La signature MD fournit une formulation algébrique de la décomposition modulaire des graphes. Courcelle et Weil [309] ont prouvé que ces notions sont robustes en ce sens que beaucoup de variantes de chaque signature donnent les mêmes notions de reconnaissabilité. Continuant dans cette direction, Courcelle et Blumensath [432] ont montré la stabilité de la famille des ensembles reconnaissables par *MS-transductions* inverses (les MS-transductions sont des transformations de graphes définies en LSOM). Weil [338] a donné une autre preuve de la robustesse de la famille des ensembles reconnaissables en montrant, pour une sous-classe significative de MD, que la reconnaissabilité équivaut à la définissabilité en LSOM avec comptage de cardinalité modulo les entiers. Des décompositions des graphes finis de plusieurs types sont formalisées comme des termes algébriques écrits avec les signatures HR, VR et MD. Ces notions de décomposition se prolongent aux graphes dénombrables et peuvent être formalisées par des termes algébriques infinis. Ceci est fait par Courcelle [306] pour la largeur de clique et les termes infinis sur VR et par Courcelle et Delhommé [355] pour la décomposition modulaire. La décomposition modulaire d'un graphe dénombrable est définissable en LSOM et représentable par une structure de *faible degré*. D'autres types de décomposition des graphes ont été définis au moyen de la théorie des catégories par Bauderon et Carrère, voir la section

consacrée au projet VISIDIA. Christian Retoré a également utilisé certains types de graphes pour caractériser les preuves de la logique linéaire et du calcul de Lambek. ([400, 330], voir aussi la section relative au thème *Informatique et Linguistique*).

Méthodes logiques

Une autre partie importante du programme de recherche est l'étude de la puissance expressive et de la décidabilité de certains langages logiques notamment, de la LSOM et du μ -calcul. Tandis que leurs pouvoirs expressifs sont identiques pour les arbres binaires, ils se comportent différemment pour les graphes.

Les notions de largeurs, arborescente et de clique, sont importantes dans l'approche algébrique des grammaires de graphes et le sont également pour la décidabilité et la vérification. L'ensemble des graphes de largeur de clique au plus k a une théorie décidable pour la LSOM avec prédicats de cardinalité modulo des entiers. Courcelle et Oum [433] ont montré une réciproque : si un ensemble de graphes a cette propriété de décidabilité, il a une largeur de clique bornée. Ce résultat est proche d'une preuve de la Conjecture de Seese (qui porte sur la LSOM sans prédicat de cardinalité). Dans [307] Courcelle prouve les relativisations de cette conjecture à diverses classes de graphes ainsi que l'équivalence de différentes relativisations. Des résultats similaires s'appliquent aux graphes de largeur arborescente bornée et au langage LSOM₂, qui est l'extension de la LSOM utilisant des quantifications sur les ensembles d'arcs. Certaines propriétés de graphes (circuit Hamiltonien) sont exprimables dans la LSOM₂ et non dans la LSOM. Dans [305] Courcelle a prouvé que la LSOM₂ n'est pas plus puissante que la LSOM pour les ensembles de graphes de degré moyen borné (de graphes où le nombre d'arcs est linéairement borné par le nombre de sommets). Dans [303] Courcelle et Dussaux étudient la puissance expressive de la LSOM pour les cartes combinatoires qui formalisent les plongements des graphes dans des surfaces orientables.

Pour décrire les propriétés des calculs, on est souvent intéressé par le fragment de la LSOM invariant par bisimulation : ce sont les propriétés qui ne peuvent pas distinguer deux graphes *bisimilaires*, c'est-à-dire, qui ont le même *comportement* si on les considère comme représentant des processus. Le fragment invariant par bisimulation de la LSOM est équivalent au μ -calcul. Janin et Lenzi ont prouvé que l'équivalence vaut aussi entre le fragment Σ_2 de la LSOM et le $\nu\mu$ -fragment du μ -calcul [316] ainsi qu'entre Σ_1 et le ν -fragment [317]. Dawar et Janin [356] ont posé la question de la validité de cette équivalence si l'on se limite aux graphes finis, et ils ont donné un résultat positif pour les graphes unaires. Le μ -calcul a également une propriété d'interpolation. Dans [434] Agostino et Lenzi ont étudié la complexité de la formule d'interpolation en termes de hiérarchie de points fixes. Lenzi avec Berwanger et Graedel ont proposé d'étudier les fragments du μ -calcul limités à des variables en nombre fini, par analogie avec les fragments correspondants de la logique de premier ordre. Ils ont prouvé que la hiérarchie est stricte pour le μ -calcul sans modalités universelles [346]. Cette restriction est supprimée dans un article de Berwanger et Lenzi (2005).

Vérification

Le problème est de vérifier si un graphe a une propriété donnée. Les graphes peuvent être finis ou infinis, mais présentés d'une manière finitaire. Les propriétés sont le plus souvent exprimées dans une des logiques mentionnées ci-dessus. La vérification des propriétés de la LSOM est un problème de grande complexité sur la classe de tous les graphes finis. Cela justifie d'étudier ce problème pour des sous-classes, comme celle des graphes de largeur de clique bornée, pour laquelle la vérification peut être faite en temps polynomial déterministe. Courcelle et Vanicat [304] ont proposé une implantation des graphes de largeur de clique

bornée telle que des requêtes prédéfinies puissent être évaluées rapidement, sans parcourir le graphe mais en utilisant des données attachées aux sommets fournis comme arguments des requêtes.

Les graphes des configurations de *machines*, et tout particulièrement des automates à pile et de leurs extensions sont des graphes infinis intéressants. La LSOM et le μ -calcul sont décidables sur ces graphes. Bouquet, Serre et Walukiewicz [352] ont étendu ce résultat pour permettre également de vérifier des propriétés de pile telles que *au cours de l'exécution, la taille de la pile est bornée*. Une étape supplémentaire consiste à considérer des graphes définis par des schémas de programmes récursifs d'ordre supérieur (supérieur au sens des types des fonctions). Les schémas du premier ordre correspondent exactement aux graphes des automates à pile déterministes et les schémas *sûrs* d'ordre n (la *sûreté* est une condition technique liée à la localité des paramètres) correspondent aux automates à pile d'ordre n . Ceci implique la décidabilité du problème de la vérification pour de tels automates et pour les propriétés de la LSOM. Courcelle et Knapik [308] ont fourni de nouvelles preuves en prouvant que l'évaluation de la substitution est *compatible* avec la LSOM. Dans [376] Walukiewicz et al. ont montré comment traiter tous les schémas d'ordre 2, même non sûrs. Ils ont caractérisé ces schémas par des automates avec des piles du second-ordre et une nouvelle opération de dépilage. Ils ont également montré la décidabilité du problème de vérification pour la LSOM et les schémas récursifs de l'ordre 2.

Les graphes automatiques constituent une autre représentation finitaire de graphes infinis. Olivier Ly [319] a montré l'indécidabilité de propriétés topologiques de ces graphes telles que la connexité.

Une autre manière d'approcher le problème de la vérification est de considérer des programmes écrits dans un langage procédural d'ordre supérieur et de chercher à vérifier l'équivalence de deux programmes. La pertinence de cette approche dépend largement de l'équivalence utilisée. Dans [393] et [392] Walukiewicz, Murawski et Ong considèrent l'équivalence contextuelle du langage *Idealized Algol*, une abstraction de langage fonctionnel avec des effets de bord proposée par Reynolds. L'équivalence contextuelle indique que les programmes sont équivalents si et seulement si ils ne peuvent pas être distingués par un contexte. Le problème de l'équivalence contextuelle est indécidable si les variables s'étendent à des domaines infinis, mais pour des domaines finis les réponses varient. Les deux articles ci-dessus complètent la classification des fragments de l'Algol idéalisé pour lesquels l'équivalence contextuelle est décidable.

Preuves mathématiques assistées

Les assistants de preuves sont des logiciels qui permettent la construction interactive de preuves mathématiques. Dans les cas où la démonstration automatique est impossible, ils fournissent une aide au moyen de tactiques programmables, allant de la simple application de règles d'inférence à la construction et à la validation de procédures de décision pour les fragments décidables de la théorie considérée.

Les domaines d'application privilégiés sont les problèmes indécidables, ceux pour lesquels une preuve formelle est nécessaire (validation de protocoles, sécurité), et les problèmes à haut degré de paramétricité (pour lesquels des preuves devraient être faites pour de nombreuses instances).

Le système Coq (Calcul des Constructions Inductives) permet ce type de formalisation. Il fait l'objet d'un livre de Bertot et Castéran [409]. Avec des membres de l'équipe *Méthodes Formelles* de la société Axalto (anciennement Schlumberger Systèmes), Ly a exploré les possibilités qu'offre Coq dans le cadre de la certification sécuritaire de logiciels implantés

dans une carte à microprocesseur, et plus précisément, d'une machine virtuelle JavaCard. En particulier, ils ont élaboré une méthodologie formelle (requis pour atteindre le niveau maximal EAL7 de certification des critères communs) pour certifier le mécanisme d'isolation des applets sur une carte Java [419, 354, 341, 438, 339].

Synthèse

Le problème de la synthèse est de construire un *programme* à partir de spécifications données. De cette formulation très générale découlent de nombreuses manières de formaliser le problème. On peut varier la puissance des langages de spécification : propriétés de sûreté, propriétés exprimées en LSOM, propriétés de synchronisation. On peut également varier les types des *programmes* considérés : automates finis, systèmes synchronisés, automates à pile, systèmes distribués. L'un des cadres classiques pour la synthèse est celui de la synthèse discrète de contrôleur de Ramadge et Wonham. Walukiewicz, Vincent et Arnold [297] ont proposé une extension de ce cadre aux comportements infinis et à toutes les propriétés régulières de ces comportements. Ils ont également prouvé que le problème de synthèse distribué dans ce cadre est en général indécidable. Certains problèmes centralisés de synthèse peuvent être présentés comme consistant à trouver des stratégies pour des jeux opposant un système et un environnement. Bernet, Janin, Mohalik et Walukiewicz [345, 384] ont étudié les jeux distribués. Ils ont prouvé que dans ce modèle il est possible d'exprimer et de résoudre la plupart des problèmes de synthèse distribuée connus d'une manière unifiée. Walukiewicz, Neuman et Szepietowski [325] ont fourni des résultats de complexité pour trouver des stratégies de gain dans les jeux dits à *conditions d'acceptation faibles*. Lorsqu'un système synthétisé est obtenu à partir d'une stratégie dans un jeu, la qualité de la stratégie devient importante. Bernet, Janin et Walukiewicz [299] ont observé que seules les caractéristiques de sûreté admettent des stratégies maximales. Dans ce cadre plus général ils ont proposé une notion de la stratégie *laxiste* qui, en un sens, se rapproche des stratégies maximales. Ly [381] a étudié la construction d'un programme à partir d'un ensemble donné de procédures. Des spécifications des *graphes de commande* du programme et des procédures sont données en LSOM. Alors qu'en général le problème est indécidable à cause de l'indécidabilité de la LSOM, il montre que le problème devient décidable pour des graphes de commande de largeur arborescente bornée. La synthèse des systèmes avec contraintes temporelles souffre du fait qu'actuellement nous n'avons pas de formalisme décidable pour les contraintes temporelles ayant une puissance expressive raisonnable et de bonnes propriétés de fermeture. Walukiewicz et Lasota [379] ont fait un pas dans cette direction en présentant un modèle d'automates temporels alternants. Alors qu'en général le problème du vide pour ces automates est indécidable, ils prouvent qu'il devient décidable si les automates utilisent une seule horloge.

Projets et perspectives 2005-2008

Les décompositions hiérarchiques de graphes sont importantes pour la construction d'algorithmes polynomiaux, en particulier basés sur la LSOM, pour la compréhension de la structure des graphes de certaines classes et pour prouver des résultats par induction sur les arbres de décomposition. Beaucoup de notions de décomposition des graphes et d'autres objets combinatoires, comme des cartes et des matroïdes, ont été définies. La compatibilité de ces notions avec les opérations de graphes et la LSOM sont importantes en vue de la construction d'algorithmes, et aussi pour permettre de caractériser les ensembles algébriques reconnaissables de tels objets. Dans certains cas, les décompositions hiérarchiques fournissent des bijections entre objets de différentes natures : arbres et graphes, graphes et mots, mots et cartes pour prendre quelques exemples. Il serait intéressant de savoir si ces bijections sont

exprimables dans la LSOM. Il est également intéressant d'étendre les résultats obtenus du fini au dénombrable. Les applications algorithmiques de la LSOM sont handicapées par les tailles des automates finis obtenus à partir des formules. Il serait intéressant de trouver de bonnes méthodes de construction d'automates pour quelques fragments de la LSOM ayant un intérêt pratique.

La preuve de la conjecture de Seese n'est pas complète (malgré les progrès cités) et son extension aux hypergraphes est largement ouverte et mérite d'être étudiée. Plus généralement, nous nous proposons d'étudier les rapports entre la structure d'un ensemble de graphes ou d'hypergraphes et la décidabilité de certains fragments ou extensions de la LSOM. Un objectif différent mais néanmoins relié consiste à développer des concepts de structure de graphe en vue de leur application aux systèmes d'information géographique. À cette fin, la notion classique d'un graphe défini comme une relation entre des sommets est insuffisante. Une notion de graphe à *niveaux multiples* est envisagée, où un sommet à un certain niveau peut représenter un sous-graphe à un autre. Une notion de structuration de graphe, basée sur des valeurs attachées aux composants (arcs et sommets à chaque niveau) et non basées (seulement) sur des paramètres comme la largeur arborescente est aussi envisagée.

Dans le domaine des assistants de preuves mathématiques, nous étudierons l'utilisation de grands ordinaux dans des preuves de terminaison pour des réécritures de termes. Les ordinaux donnent des principes forts d'induction qui peuvent parfois considérablement simplifier des preuves.

Dans le domaine de la synthèse, nous rechercherons les classes décidables intéressantes pour le problème de synthèse dans les cadres distribués et temporels. La recherche portera sur les jeux distribués ainsi bien que sur les automates temporels alternants. Nous nous concentrons sur ces cadres parce que la complexité de la synthèse est très grande et nous pensons qu'on peut trouver des applications dans les secteurs où la vérification de modèle est difficile. Avec la même motivation nous considérerons le problème de résoudre des jeux pour différentes conditions de gain non-régulières. Cela permettra d'aborder les problèmes de la vérification et de la synthèse avec des conditions non-régulières. Nous considérerons le problème de la vérification pour des programmes récursifs avec des fonctions d'ordre encore supérieur. Nous espérons utiliser les résultats sur l'Algol idéalisé mentionnés ci-dessus. Il serait intéressant d'étudier quels genres de problèmes de vérification peuvent être codés dans ce cadre. Nous espérons, par exemple, traiter ainsi quelques problèmes de sécurité.

2.3.2 Langages et Automates

Introduction

Ce thème regroupe les travaux portant sur les langages de mots, d'arbres et de traces. Nous nous préoccupons de leur *classification* suivant des critères algébriques, logiques et grammaticaux et tentons de cerner la *complexité* algorithmique des propriétés de ces langages.

Mots-Clés Automates, Mots, Arbres, Traces, Reconnaissabilité, Automates à pile, Semi-groupes, Groupes libres, Réécriture, Logiques locales, Topologie

Description des activités et principaux résultats

Langages de mots On décrit ici les travaux concernant l'algorithmique des automates à pile, la classification des langages rationnels par des invariants algébriques, et l'application de la théorie des automates sur les mots à des problèmes algorithmiques d'origine algébrique, notamment en combinatoire du groupe libre et en réécriture.

Equivalence des automates à pile déterministes Nous avons approfondi et développé nos travaux antérieurs sur les problèmes d'équivalence d'automates à pile, dans plusieurs directions. Sénizergues a écrit une preuve *plus épurée* de son résultat de base [332] (la décidabilité de l'égalité de deux langages reconnus par des automates à pile déterministes).

Sénizergues a étudié la complexité du problème dans la sous-classe des automates finite-turn [396]. Le cœur de ce progrès algorithmique est en fait une nouvelle propriété algébrique du semigroupe des matrices à coefficients préfixes. Ce semigroupe intervient naturellement, par ailleurs, dans la comparaison des langages préfixes, indépendamment de toute classe particulière d'automates.

Il a étendu ce résultat "de base" au problème de la *bisimulation* des graphes équationnels de degré sortant fini [333] : étant donnés deux automates à pile A, B *non-déterministes* et dont les transitions spontanées sont déterministes, on peut décider si A et B sont *bisimilaires*. La décidabilité de ce problème rejoint la décidabilité du problème de l'isomorphisme des graphes équationnels (démontrée par Courcelle il y a une quinzaine d'années) et généralise divers résultats de décidabilité du problème de la bisimulation pour des sous-classes d'automates à pile non-déterministes (automates sans états, automates sans transitions spontanées, automates à un compteur sans transitions spontanées). Mentionnons un bref article de vulgarisation sur ces problèmes d'équivalence [437].

Avec Bruyère et Carton, Sénizergues a résolu un problème analogue, le problème de l'inclusion pour les automates finis reconnaissant des ensembles de mots sur des ensembles ordonnés linéaires éparpillés (qui généralisent les mots ordinaux) [353].

Automates à piles itérées Nous nous sommes engagés dans une étude de longue durée des automates à piles itérées, ou automates à piles d'ordre supérieur.

Blumensath a établi un lemme d'itération pour les langages algébriques d'ordre supérieur qui généralise celui établi par Ayashi dans les années 70, pour les langages d'index (qui correspondent au niveau 2 de la hiérarchie des langages algébriques d'ordre supérieur). Remarquons que ce lemme de Blumensath est, à l'heure actuelle, le seul point d'appui pour décider si un langage d'ordre supérieur est fini ou infini [347].

Fratani et Sénizergues ont établi un lien entre les automates à piles itérées et les suites récurrentes d'entiers ou de nombres rationnels [366]. En particulier, les suites \mathbb{Q} -rationnelles sont décrites par des automates de niveau 2 et les suites \mathbb{Q} -algébriques, ou plus généralement P -récurrentes, sont décrites par des automates de niveau 3. Ce travail débouche sur des

méthodes nouvelles pour établir des résultats de décidabilité pour des arithmétiques dites *faibles* et aussi pour étudier les suites récurrentes.

Automates et semigroupes Ces dernières années ont vu l'émergence d'un raffinement de la notion de semigroupe syntaxique, par l'adjonction sur ce semigroupe d'un ordre canonique (Pin). La notion de semigroupe syntaxique ordonné qui en résulte est fort utile pour raffiner les classifications classiques. Une série de trois articles de Pin, Pinguet et Weil ([328], [327], [329]) pose les fondements de la décomposition des semigroupes ordonnés en produits en couronne (y compris l'analogue ordonné de la catégorie dérivée d'un morphisme relationnel), et applique cet outil à la décomposition de langages.

La théorie des semigroupes et des monoïdes apparaît comme un domaine naturellement connexe à l'étude des langages reconnaissables. Avec Fountain et Pin, Weil [315] donne une théorie générale des revêtements de monoïdes, en tout cas des monoïdes qui satisfont une condition très légère concernant la densité des idempotents. Cette théorie généralise un ensemble de résultats acquis depuis les années 70, pour des classes de monoïdes plus restreintes.

Equations dans les semigroupes Dans certains semigroupes on peut décider si une équation donnée a au moins une solution : ainsi en est-il dans le semi-groupe libre ou le groupe libre (Makanin). Dans d'autres semigroupes, ce même problème est indécidable : c'est notamment le cas du monoïde inversif libre (Rozenblat). Deis, Meakin et Sénizergues ont établi la décidabilité du problème de l'*extension* d'une solution du groupe libre en une solution dans le monoïde inversif libre, ce qui conduit à la résolubilité de certaines équations dans le monoïde inversif libre [310]. Remarquons qu'ici, ce n'est pas l'algèbre qui vient en aide à la théorie des automates, mais au contraire la théorie des automates (d'arbres) qui fournit un argument décisif de décidabilité.

Sous-groupes du groupe libre La combinatoire du groupe libre est une cousine classique de la combinatoire des mots. Lorsqu'il s'agit des sous-groupes finiment engendrés du groupe libre, on peut utiliser des représentations par automates finis, qui donnent à la fois une meilleure compréhension des problèmes, et de bonnes propriétés algorithmiques pour leur solution. Meakin et Weil [322] donnent une preuve partielle de la conjecture forte de Hanna Neumann. Cette conjecture énonce une borne supérieure pour le rang de l'intersection de sous-groupes de rang fini, et ils montrent que cette conjecture est vraie lorsque l'un des sous-groupes est positivement engendré. Les autres cas connus concernent les situations où l'un des sous-groupes a rang au plus 3, ou bien où l'un des sous-groupes est d'indice fini (notre résultat généralise ce dernier cas). Un autre travail, de Coulbois, Sapir et Weil [302], exploite des travaux récents de Margolis, Sapir et Weil (2001) sur le calcul de la clôture profinie d'un sous-groupe de rang fini pour décrire les situations où un morphisme injectif entre groupes libres admet un prolongement par continuité entre les clôtures pro- V qui est aussi injectif. Ici, V est une pseudovariété de groupes finis.

Réécriture de mots Matiyasevich et Sénizergues ont étudié les problèmes de décision classiques concernant les systèmes semi-Thuéiens, en s'efforçant de cerner du plus près possible la frontière entre le décidable et l'indécidable [321]. Il s'avère que les problèmes de l'accessibilité et de la terminaison (notamment) sont déjà indécidables pour des systèmes semi-Thuéiens à trois règles. Ils obtiennent en corollaire l'indécidabilité du *problème de correspondance de Post* pour des ensembles de 7 paires de mots. Ce corollaire est le point de départ de plusieurs preuves d'indécidabilité dans les domaines de la combinatoire des mots ou des semigroupes de matrices (travaux de Blondel, Cassaigne, Karhumäki notamment). Nos preuves s'appuient sur la notion de *code* (dans le monoïde libre) ainsi que sur un treillis de mots pour la re-

lation d'ordre définie par la présentation classique du monoïde bi-cyclique par un système semi-Thuéien.

Langages d'arbres Il s'agit ici des travaux sur la reconnaissabilité des langages d'arbres et leur classification par des invariants algébriques, de l'étude de hiérarchies de langages d'arbres déterminés par des alternances de points fixes, et des applications de la théorie des langages d'arbres à la réécriture de termes et à des problèmes algorithmiques sur les surfaces topologiques.

Langages d'arbres et structures algébriques Si l'équivalence entre reconnaissabilité, définissabilité et acceptation par un automate est bien connue depuis les années 1960, il n'y a pratiquement aucun résultat de classification des langages d'arbres réguliers, et en particulier on ne sait pas décider si un langage d'arbres réguliers est définissable au premier ordre. Ésik et Weil ([365, 314]) proposent une nouvelle structure algébrique sur l'ensemble des arbres, qui mène pourtant à la même notion de langage d'arbres reconnaissable. Néanmoins, cette nouvelle structure algébrique, appelée *préclone*, permet d'associer à chaque langage d'arbres régulier un invariant algébrique nouveau, son préclone syntaxique. Le préclone syntaxique d'un langage est entièrement spécifié par l'automate minimal du langage. Le résultat principal est une caractérisation de la définissabilité au premier ordre par une propriété du préclone syntaxique. Malheureusement, nous n'avons pas encore pu transformer cette caractérisation en un algorithme de décision.

Par des méthodes comparables, Bojańczyk et Walukiewicz ont obtenu des résultats de décidabilité pour des sous-classes importantes des langages définissables au premier ordre, et en particulier pour un fragment de la logique CTL sans *until* [348]. Il est intéressant et encourageant de noter que ce résultat ainsi qu'un autre résultat récent de Benedikt et Ségoufin (STACS 2005), admettent une expression élégante en termes de préclones syntaxiques.

Deux articles de survol peuvent être mentionnés : l'un porte sur les méthodes profinies en théorie des semigroupes [337], et l'autre [408] met en perspective les différents résultats concernant la reconnaissabilité algébrique des langages, qu'il s'agisse de langages de mots, d'arbres, de posets ou de graphes.

Langages d'arbres et μ -calcul Un de nos objectifs est de montrer la décidabilité des hiérarchies des classes de langages d'arbres définis par points fixes. Ces hiérarchies sont définies par le nombre d'alternances de points fixes mais elles peuvent aussi être définies par la taille des conditions d'acceptation dans le cadre des automates avec condition de parité. On dispose de trois hiérarchies différentes qui correspondent aux automates déterministes, non-déterministes et alternants. Dans [404], Walukiewicz a montré la décidabilité pour les deux niveaux les plus bas des trois hiérarchies (ces niveaux sont communs aux trois hiérarchies). Dans [326], Niwiński et Walukiewicz ont étudié les propriétés de la hiérarchie correspondant aux automates déterministes. Ils ont mis en évidence des relations entre les hiérarchies déterministe et non-déterministe, et ils ont donné un algorithme polynomial pour décider le niveau Σ_2 de la hiérarchie non-déterministe. Dans des travaux plus récents (soumis), ils donnent une solution complète pour la hiérarchie des automates déterministes.

Langages d'arbres et réécriture La théorie des langages d'arbres joue bien sûr un rôle central dans l'étude des systèmes de réécriture de termes. Durand et Middeldorp ont obtenu de nombreux résultats théoriques concernant les stratégies de calcul *par nécessité* (*call by need*) des formes normales dans les systèmes de réécriture de termes [313]. Leur approche et leurs résultats s'étendent au cas du calcul de formes normales de tête. Une étude de la complexité des diverses classes obtenues a été réalisée et rédigée dans le cadre de l'habilitation de Durand, soutenue en juillet 2005.

Le logiciel AUTOWRITE, qui implémente de nombreux algorithmes définis dans le cadre des stratégies par nécessité [363], a été considérablement enrichi avec des outils permettant de manipuler directement des automates [364]. Ce logiciel s'avère plus performant que les divers autres logiciels existants.

Arbres et espaces topologiques En étudiant les variétés de dimension 3 et leur propriétés algorithmiques Ly a introduit une classe de 3-variétés appelées *3-variétés arborescentes*, qui contient les 3-variétés compactes, les 3-variétés de volume fini et les 3-variétés non compactes possédant une triangularisation HR-équationnelle. L'introduction de cette nouvelle classe est motivée par le résultat suivant de Ly [320] : les 3-variétés hyperboliques arborescentes sont caractérisées par leurs groupes fondamentaux, c'est-à-dire qu'elles sont rigides ; cela étend les travaux de Mostow et de Prasad, qui ont démontré cette propriété pour les cas compact et de volume fini respectivement. Ainsi, le problème de l'homéomorphisme des 3-variétés hyperboliques arborescentes peut être *discrétisé* en ce sens qu'il se réduit à un problème d'isomorphisme de groupes, groupes eux-même décrits de façon constructive en terme de langages rationnels : un tel groupe possède en effet une présentation infinie dont les générateurs et les relations peuvent être codés par les mots d'un langage rationnel. On peut noter également qu'un tel groupe est construit comme produit amalgamé d'une infinité de groupes hyperboliques (au sens de Gromov) de présentations finies agencés le long d'un arbre régulier.

Langages de traces Thiagarajan et Walukiewicz ont proposé dans [334] la première logique à la LTL sur les traces dont le pouvoir d'expression est le même que celui de la logique du premier ordre. Cette logique est dite globale parce qu'elle parle des configurations dans des traces. Plus tard, Walukiewicz a montré [336] que la complexité du problème de la satisfiabilité pour une grande classe de logiques globales est non-élémentaire. Ce résultat négatif a orienté son travail vers les logiques dites locales, qui parlent des événements dans les traces. Walukiewicz a proposé plusieurs variantes du μ -calcul sur les traces qui sont décidables dans PSPACE et ont un pouvoir d'expression égal à celui de la logique monadique du second ordre [335]. Muscholl et Walukiewicz [394] ont fait un premier pas pour circonvenir la trop grande complexité des logiques globales. Ils ont trouvé un nouveau fragment de LTL dont le problème de satisfiabilité est dans NP. On espère qu'une extension de ce fragment sur les traces peut donner une logique globale à complexité élémentaire.

Projets et perspectives 2005-2008

Les travaux exposés ci-dessus nous conduisent naturellement à développer nos activités dans leur prolongement. Tentons de donner, pour chacune de nos directions principales, une idée de ces développements futurs :

Equivalence d'automates

- approfondir, affiner et généraliser les méthodes permettant de comparer des automates déterministes à pile ; aborder ce problème pour les automates à piles itérées ;
- rapprocher les méthodes de comparaison de suites récurrentes (Zeilberger) des méthodes de comparaison d'automates, en utilisant l'approche de Fratani et Sénizergues.

Automates et semigroupes

- poursuivre l'étude algorithmique de la théorie combinatoire des groupes, à l'aide des méthodes de la théorie des automates et des langages ;
- développer des techniques de résolution d'équations dans les monoïdes et les groupes ;

- progresser dans l'étude des groupes fondamentaux des variétés arborescentes : problèmes du mot, de la conjugaison ; résoudre le problème de l'isomorphisme est un objectif à (très) long terme ;
- caractériser les langages algébriques à l'aide d'une famille particulière de *groupoïdes*.

Langages de structures discrètes, algèbre et logique

- continuer d'étendre les méthodes algébriques qui ont obtenu des succès dans le cas des mots à des objets discrets plus complexes : arbres, posets, traces ; développer les structures algébriques adaptées (préclones, par exemple) et les appliquer aux problèmes de décision classiques ; un problème-test dans ce domaine est de savoir si l'on peut décider si un langage d'arbres donné est définissable par une formule de la logique du premier ordre ;
- progresser vers la décidabilité de tous les niveaux des hiérarchies d'ensembles d'arbres ;
- étudier des stratégies de réécriture de termes conservant la rationalité (resp. l'algébricité) ; appliquer ces résultats au calcul de formes normales infinies (arbres de Boehm).

Langages de traces

- persévérer dans la quête d'une logique globale mais algorithmiquement raisonnable pour les traces ;
- appliquer les logiques sur les traces aux problèmes de vérification.

2.3.3 Informatique et Linguistique

Introduction

Ce thème traite de la modélisation formelle du langage naturel, et plus particulièrement de sa syntaxe et de sa sémantique. Les méthodes utilisées, langages formels, grammaires, logique placent cette thématique au sein de Logique Automates, Algorithmes et Applications du LaBRI. Ce thème se développe au LaBRI depuis 2003, en partenariat avec des linguistes de l'Université Bordeaux-3 (ERSS UMR5610 et TELANCO JE2385) ainsi que des collaborateurs hors-site (INRIA Rocquencourt, Université Grenoble-2, INJS), pour l'essentiel sous la forme du projet INRIA *Signes linguistiques, grammaire et sens : algorithmique logique de la langue* (SIGNES, Projet commun INRIA-Futurs, LaBRI et département des sciences du langage de l'Université Michel de Montaigne Bordeaux-3).

Dès les débuts de l'informatique, la linguistique a été à la fois un champ de motivations (les langues et le raisonnement naturels), un domaine d'application privilégié (traduction automatique), et la source de notions clefs (grammaires et théorie des langages formels). Plus récemment la logique est apparue comme un élément fédérateur capable de traiter de la sémantique de la langue, mais aussi de sa syntaxe. Nous souhaitons réaliser et articuler les modèles et algorithmes correspondant aux différents niveaux d'étude et de traitement informatique de la langue, l'accent étant placé sur l'interface entre syntaxe et sémantique.

Plus précisément, nous nous proposons de modéliser certains aspects de la langue par des méthodes formelles : théorie des langages, logique des ressources, lambda-calcul et logique d'ordre supérieur. Les aspects de la langue retenus et qu'il faut articuler sont la structure des mots (morphologie) et celle de la phrase aussi bien au niveau syntaxique qu'au niveau de sa structure logique (qui peut s'étendre à celle du discours).

Parmi les activités entreprises ou envisagées, mentionnons l'intégration de la morphologie et de la syntaxe, le calcul de représentations sémantiques à partir d'analyses syntaxiques (catégorielles, minimalistes,...), la résolution d'anaphores, le passage de la structure prédicative de la phrase à la structure logique du discours. Plus précisément, on distingue dans ce thème trois niveaux et techniques :

- La syntaxe est bien évidemment décrite par des grammaires formelles (de mots ou d'arbres, algébriques ou faiblement contextuelles, étendues ou non par l'unification), mais, si l'on vise à extraire de l'analyse syntaxique une représentation utilisable du sens de la phrase, on peut aussi utiliser des systèmes déductifs dans les logiques des ressources, selon la tradition des grammaires catégorielles. La syntaxe doit aussi prendre en compte la structure des mots, qui est généralement décrite au moyen d'automates et de transducteurs.
- La sémantique logique, qui donne la structure prédicative d'une phrase, est pour sa part décrite par la logique d'ordre supérieur représentée dans le lambda calcul typé, comme l'ont fait Church et Montague. On met ainsi en œuvre le principe frégéen de compositionnalité : le sens d'une expression composée est une fonction du sens des parties qui la composent. Lorsqu'on passe de la phrase au discours, ces techniques peuvent s'étendre avec la lambda-DRT.
- Une manière formelle de représenter le rapport des mots au monde ou entre eux est d'utiliser, à la suite de Pustejovsky, des structures de traits qui expriment les rôles que joue un mot par rapport aux autres, et dont les variables incluses dans un lambda terme à la Montague décrivent le rôle qu'occupent les différents aspects d'un mot dans la structure prédicative de la phrase.

La thématique de l'équipe est l'interaction entre ces trois niveaux de représentations d'une phrase et des mots qui la compose, l'automatisation du passage de la phrase à une représenta-

tion de son sens et vice-versa, ainsi que l'étude mathématique des formalismes et transformations utilisés. Comme application particulière on mentionnera notre travail de modélisation de la langue des signes française.

Mots-Clés Linguistique Informatique, Traitement Automatique des Langues, Syntaxe et Sémantique du Langage Naturel, Grammaires Formelles, Logique des Ressources, Logique Linéaire, Calcul de Lambek, Lambda-calcul

Description des activités et principaux résultats

Langue des Signes Française Une application particulière est la formalisation de la syntaxe de la langue des signes française (LSF) et son articulation avec la sémantique. Une allocation de la région Aquitaine (2004-06) finance un doctorat à l'Université Bordeaux-3 (co-encadré par un membre du LABRI) et nous permet l'achat de matériel d'acquisition vidéo pour la réalisation d'expériences linguistiques et la création d'un corpus en LSF, étape préliminaire avant la modélisation formelle. Pour le moment l'étude de de Langhe, Guitteny, Marlet, Portine, Retoré, Voisin s'intéresse à l'ordre naturel des mots dans la phrase simple en langue des signes française. Ils ont montré que cet ordre est souvent Objet Sujet Verbe sans qu'il y ait de topicalisation. Cet ordre est rarissime dans les langues vocales et par conséquent son codage dans les formalismes syntaxiques usuels n'est pas simple [422, 416].

Grammaires Catégorielles Une autre application est le développement de Grail qu'a entrepris Moot. Cet analyseur syntaxique pour les grammaires catégorielles multimodales fonctionne pour l'anglais, le néerlandais et le français. Grail, initialement écrit en SICSTUS Prolog aux Pays-Bas, a été complètement réécrit par Moot en SWI Prolog au LaBRI, en utilisant une représentation plus compacte et des algorithmes de graphes plus efficaces [387], que l'auteur a développés entre-temps. Grail a aussi été utilisé par Moot pour l'étiquetage par des arbres (supertagging) d'un corpus d'un million mots de néerlandais parlé [386]. Les grammaires multimodales ont été aussi utilisées pour décrire une question délicate de syntaxe du français, les pronoms clitics [424], en se servant des modalités pour créer des îlots comme dans les travaux de Moot avec Bernardi (Bolzano) [298].

Sur les grammaires catégorielles simples (de Lambek) Retoré a montré que leurs analyses en réseaux de démonstration peuvent se voir comme des graphes munis d'un couplage parfait tels que les autres arêtes constituent un cographe. Le cographe correspond à la formule démontrée tandis que le couplage parfait décrit les axiomes. Les démonstrations formelles commutatives correspondent exactement aux graphes tels que tout cycle élémentaire alternant les arêtes du couplage et hors du couplage contienne une corde. Cette représentation quotiente les démonstrations formelles par permutation des règles et par les propriétés algébriques des connecteurs logiques (disjonction et conjonction) : associativité et commutativité [330]. Plus récemment ce résultat a été étendu au calcul de Lambek avec Pogodalla (Nancy). Il faut alors considérer un circuit Hamiltonien et demander, outre le critère commutatif, que le couplage parfait soit un bon parenthésage le long du circuit Hamiltonien pour que les graphes correspondent exactement aux preuves en logique non commutative. Cette représentation quotiente les démonstrations formelles par permutation des règles et par l'associativité des connecteurs logiques [400].

Le Nir a aussi montré que dans l'utilisation grammatical du calcul de Lambek associatif ou non, la règle d'introduction de l'implication pouvait être remplacée par des axiomes plus riches et la seule règle d'élimination de l'implication [380, 318, 441].

L'atelier Icharate développé en COQ par Anoun est un système d'analyse semi-automatique

dans les grammaires catégorielles multimodales qui inclut maintenant certaines variantes catégorielles des grammaires minimalistes. Ce genre de système permet aussi de vérifier des propriétés mathématiques des logiques sous-structurelles sur lesquelles se fondent les grammaires catégorielles [412].

Grammaires minimalistes Un des sujets de recherche est le passage de la structure syntaxique à la structure logique pour les grammaires minimalistes. Ce passage est connu et automatisé pour les grammaires catégorielles standard, mais cette correspondance est moins claire si on utilise les grammaires minimalistes de Stabler qui présentent des avantages flagrants : rapports entre analyses de phrases corrélées (affirmation/questions, actif/passif), rapport entre les lexiques de langues distinctes, et analyse polynomiale.

Amblard, Lecomte, Retoré ont montré que, via un codage dans les grammaires catégorielles, on peut cumuler les avantages de ces deux types de grammaires et extraire des analyses minimalistes des représentations sémantiques [429, 417, 340, 423, 401]. Bonato a montré comment ce type d'analyse permet de restreindre le référent des pronoms en interdisant ou en imposant certaines coréférences [349].

Sur ces grammaires, afin de mieux comprendre comment elles se situent et se comparent aux autres formalismes légèrement contextuels, Amblard a montré qu'elles incluaient tous les langages à compteurs [410].

Lexique Génératif Un groupe de travail sur les lexiques sémantiques, en particulier le lexique génératif de Pustejovsky s'est créé début 2005. Il réunit Christian Bassac, Patrick Henry et Renaud Marlet, et a pour objectif la formalisation du lexique génératif ainsi que son implémentation. On pourra alors lever les ambiguïtés d'analyse de Grail (ou d'un autre analyseur) et extraire des représentations sémantiques plus fines.

Projets et perspectives 2005-2008

Maintenant que l'effectif de Signes et son ancrage institutionnel se sont stabilisés, ce projet cherche à développer ses activités. L'un des aspects est la mise en place de cours de Master qui puisse attirer vers la recherche des étudiants. Un autre aspect est la reconnaissance des activités en linguistique informatique du LABRI par la 34e section du CNRS et une collaboration plus étroite avec Toulouse (IRIT, ERSS).

Plus concrètement, les avancées sur les grammaires légèrement contextuelles (TAGs, Minimalisme) et les automates à piles itérées devraient consolider les liens de ce projet avec le reste de l'équipe L3A et du LABRI. Symétriquement, le travail de description et de formalisation de la syntaxe des langues (que ce soit pour la langues des signes ou des langues parlées, Français, Allemand, Catalan, Néerlandais, Sanskrit) devrait renforcer nos liens avec l'ERSS à Bordeaux.

Orthogonalement, les recherches entreprises sur la langue des signes et le développement d'applications sur ce sujet nous orientent vers une collaboration avec l'équipe Image et Son que matérialisent deux propositions de sujet de thèse.

2.4 Publications de l'équipe

2.4.1 Publications d'audience internationale

Revue avec Comité de lecture

- [297] ARNOLD (A), VINCENT (A) et WALUKIEWICZ (I). – Games for synthesis of controllers with partial observation. *Theoretical Computer Science*, vol. 303, n1, 2003, pp. 7–34.
- [298] BERNARDI (R) et MOOT (R). – Generalized quantifiers in declarative and interrogative sentences. *Logic Journal of the IGPL*, vol. 11, n4, 2003.
- [299] BERNET (J), JANIN (D) et WALUKIEWICZ (I). – Permissive strategies : from parity games to safety games. *Theoretical Informatics and Applications (RAIRO)*, vol. 36, 2002, pp. 251–275.
- [300] BONICHON (N), LE SAEC (B) et MOSBAH (M). – Orthogonal drawings based on the stratification of planar graphs. *Discrete Mathematics*, vol. 276, n 1-3, 2004, pp. 43–57.
- [301] BONICHON (N) et MOSBAH (M). – Watermelon uniform random generation with applications. *Theoretical Computer Science*, vol. 307, n2, 2003, pp. 241–256.
- [302] COULBOIS (T), SAPIR (M) et WEIL (P). – A note on the continuous extensions of injective morphisms between free groups to relatively free profinite groups. *Publicacions Matemàtiques*, vol. 47, 2003, pp. 477–487.
- [303] COURCELLE (B) et DUSSAUX (V). – Map genus, forbidden maps and monadic second-order logic. *The Electronic Journal of Combinatorics*, vol. 9, 2002.
- [304] COURCELLE (B) et VANICAT (R). – Query efficient implementations of graphs of bounded clique-width. *Discrete Applied Mathematics*, vol. 131, 2003, pp. 129–150. – Sélectionné par P. Hammer, directeur de la revue parmi l'un des meilleurs articles de 2003 parus dans DAM.
- [305] COURCELLE (B). – The monadic second-order logic of graphs XIV : Uniformly sparse graphs and edge set quantifications. *Tcs*, vol. 299, 2003, pp. 1–36.
- [306] COURCELLE (B). – Clique-width of countable graphs : a compactness property. *Discrete Mathematics*, vol. 276, 2004, pp. 127–148.
- [307] COURCELLE (B). – The monadic second-order logic of graphs XV : On a conjecture by d. seese. *Journal of Applied Logic*, 2005. – à paraître.
- [308] COURCELLE (B) et KNAPIK (T). – The evaluation of first-order substitution is monadic second-order compatible. *Tcs*, vol. 281, 2002, pp. 177–206. – Special issue offered to Maurice Nivat.
- [309] COURCELLE (B) et WEIL (P). – The recognizability of sets of graphs is a robust property. *Tcs*, 2005. – à paraître.
- [310] DEIS (T), MEAKIN (J) et SÉNIZERGUES (G). – Equations in free inverse monoids. *International Journal of Algebra and Computation*, à paraître.
- [311] DREWES (F), KREOWSKI (H.-J) et LAPOIRE (D). – Criteria to disprove context freeness of collage languages. *Tcs*, vol. 290, n3, 2003, pp. 1445–1458.
- [312] DURAND (I). – Autowrite : A tool for term rewrite systems and tree automata. *Electronics Notes in Theoretical Computer Science*, vol. 124, 2005, pp. 29–49.
- [313] DURAND (I) et MIDDELDORP (A). – Decidable call-by-need computations in term rewriting. *Information and Computation*, vol. 196, 2005, pp. 95–126.

- [314] ÉSIK (Z) et WEIL (P). – Algebraic recognizability of regular tree languages. *Theoretical Computer Science*, vol. 340, 2005, pp. 291–321.
- [315] FOUNTAIN (J), PIN (J.-E) et WEIL (P). – Covers for monoids. *Journal of Algebra*, vol. 271, 2004, pp. 529–586.
- [316] JANIN (D) et LENZI (G). – On the logical definability of topologically closed recognizable languages of infinite trees. *Computing and Informatics*, vol. 21, 2002, pp. 185–203.
- [317] JANIN (D) et LENZI (G). – On the relationship between monadic and weak monadic second order on arbitrary trees, with application to the mu-calculus. *Fundamenta Informaticae*, vol. 61, 2004, pp. 247–265.
- [318] LE NIR (Y). – From proof trees in Lambek calculus to Ajdukiewicz Bar-Hillel elimination binary trees. *Journal of Research on Language and Computation*, vol. 1 :3-4, 2003, pp. 181–201.
- [319] LY (O). – Automatic Graph and D0L-Sequences of Finite Graphs. *Journal of Computer and System Sciences*, vol. 67, n3, 2003, pp. 497–545.
- [320] LY (O). – Construction of pseudo-isometries for tree-like non-compact 3-manifolds. *Comptes Rendus De l'Académie Des Sciences*, vol. 337 :7, 2003, pp. 457–460.
- [321] MATIYASEVICH (Y) et SÉNIZERGUES (G). – Decision problems for semi-Thue systems with a few rules. *Theoretical Computer Science*, vol. 330, Issue 1, 2005, pp. 145–169.
- [322] MEAKIN (J) et WEIL (P). – Subgroups of free groups : a contribution to the Hanna Neumann conjecture. *Geometriæ Dedicata*, vol. 94, 2002, pp. 33–43.
- [323] METIVIER (Y), MOSBAH (M), WACRENIER (P) et GRUNER (S). – A distributed algorithm for computing a spanning tree in anonymous t-prime graphs. *Studia Informatica Universalis*, vol. Special Issue Vol.2 HS2, 2002, pp. 141–157.
- [324] MOOT (R) et PUITE (Q). – Proof nets for the multimodal Lambek calculus. *Studia Logica*, vol. 71, n3, 2003, pp. 415–442.
- [325] NEUMANN (J), SZEPIETOWSKI (A) et WALUKIEWICZ (I). – Complexity of weak acceptance conditions in tree automata. *Ipl*, vol. 84, 2002, pp. 181–187.
- [326] NIWIŃSKI (D) et WALUKIEWICZ (I). – A gap property of deterministic tree languages. *Tcs*, vol. 303, n1, 2003, pp. 215–231.
- [327] PIN (J.-E), PINGUET (A) et WEIL (P). – Ordered categories and ordered semigroups. *Communications in Algebra*, vol. 30, 2002, pp. 5651–5675.
- [328] PIN (J.-E) et WEIL (P). – Semidirect products of ordered semigroups. *Communications in Algebra*, vol. 30, 2002, pp. 149–169.
- [329] PIN (J.-E) et WEIL (P). – The wreath product principle for ordered semigroups. *Communications in Algebra*, vol. 30, 2002, pp. 5677–5713.
- [330] RETORÉ (C). – Handsome proof-nets : perfect matchings and cographs. *Theoretical Computer Science*, vol. 294, n3, 2003, pp. 473–488.
- [331] RETORÉ (C) et STABLER (E). – Generative grammar in resource logics. *Research on Language and Computation*, vol. 1, n2, 2004, pp. 3–25.
- [332] SÉNIZERGUES (G). – $L(A)=L(B)$? a simplified decidability proof. *Theoretical Computer Science*, vol. 281, 2002, pp. 555–608.
- [333] SÉNIZERGUES (G). – The bisimulation problem for equational graphs of finite out-degree. *SIAM Journal on Computing*, à paraître.

- [334] THIAGARAJAN (P. S) et WALUKIEWICZ (I). – An expressively complete linear time temporal logic for Mazurkiewicz traces. *Information and Computation*, vol. 179, 2002, pp. 230–249.
- [335] WALUKIEWICZ (I). – Local logics for traces. *Journal of Automata, Languages and Combinatorics*, vol. 7, n2, 2002, pp. 259–290.
- [336] WALUKIEWICZ (I). – Difficult configurations – on the complexity of LTrL. *Formal Methods in System Design*, vol. 26, n1, 2005, pp. 27–43.
- [337] WEIL (P). – Profinite methods in semigroup theory. *International Journal of Algebra and Computation*, vol. 12, 2002, pp. 137–178.
- [338] WEIL (P). – On the logical definability of certain graph and poset languages. *Journal Automata, Languages and Computation*, vol. 9, 2004, pp. 147–165.

Colloques avec Comité de programme et Actes

- [339] AKKAR (M.-L), GOUBIN (L) et LY (O). – About an Automatic Fault Injection Protection System. In : *E-SMART*. – [Http ://www.eurosmart.com/](http://www.eurosmart.com/), Nice 2003.
- [340] AMBLARD (M), LECOMTE (A) et RETORÉ (C). – Syntax and semantics interacting in a minimalist theory. In : *Prospects and Advances in the Syntax/Semantics Interface*, éd. par Duchier (D), pp. 17–22. – Nancy, 2003.
- [341] ANDRONICK (J), CHETALI (B) et LY (O). – Using Coq to Verify Java Card Applet Isolation Properties. In : *TPHOL'2003*, pp. 335–351. – 2003.
- [342] BAUDERON (M) et CARRERE (F). – Decomposing graphs with symmetries. In : *International Conference on Graph Transformation (ICGT 2002)*. pp. 45–59. – Springer-Verlag, 2002.
- [343] BAUDERON (M) et MOSBAH (M). – A unified framework for designing, implementing and visualizing distributed algorithms. In : *Electronic Notes in Theoretical Computer Science*, éd. par Bottoni (P) et Minas (M). – Elsevier, 2003.
- [344] BELLOMI (F) et BONATO (R). – Lexical authorities in an encyclopedic corpus : a case study on wikipedia. In : *Word structure and lexical systems : models and applications*. – Pavia, 2004.
- [345] BERNET (J) et JANIN (D). – Tree automata and distributed games. In : *Fundamentals of Computation Theory (FCT) 2005*. – Springer-Verlag, à paraître.
- [346] BERWANGER (D), GRAEDEL (E) et LENZI (G). – On the variable hierarchy of the modal mu-calculus. In : *CSL'02*, pp. 352–366. – 2002.
- [347] BLUMENSATH (A). – A pumping lemma for higher order pushdown automata. In : *Proceedings of the WASL 2004*. – Auckland, December 2004.
- [348] BOJANCZYK (M) et WALUKIEWICZ (I). – Characterizing EF and EX tree logics. In : *CONCUR'04*, pp. 131–145. – 2004.
- [349] BONATO (R). – Towards inductive semantics for coreference. In : *NASSLLI Student Session*. – 2004.
- [350] BONICHON (N), LE SAEC (B) et MOSBAH (M). – Optimal area algorithm for planar polyline drawings. In : *28th International Workshop, Graph - Theoretic Concepts in Computer Science (WG)*. pp. 35–46. – Springer-Verlag, 2002.
- [351] BONICHON (N), LE SAEC (B) et MOSBAH (M). – Wagner's theorem on realizers. In : *Proceedings of the 29th International Colloquium on Automata, Languages and Programming (ICALP)*. pp. 1043–1053. – Springer-Verlag, 2002.

- [352] BOUQUET (A), SERRE (O) et WALUKIEWICZ (I). – Pushdown games with the unboundedness and regular conditions. *In : FSTTCS'03*. pp. 88–99. – Springer-Verlag, 2003.
- [353] BRUYÈRE (V), CARTON (O) et SÉNIZERGUES (G). – Tree automata and automata on linear orderings. *In : Proceedings WORDS'03*. pp. 222–231. – TUCS General Publication, 2003.
- [354] CHETALI (B) et LY (O). – Formal Verification of the Confidentiality Property in Java Card Technology. *In : E-SMART*. – [Http://www.eurosmart.com/](http://www.eurosmart.com/), Nice 2002.
- [355] COURCELLE (B) et DELHOMMÉ (C). – The modular decomposition of countable graphs : Constructions in monadic second-order logic. *In : CSL'05*. – 2005. à paraître.
- [356] DAWAR (A) et JANIN (D). – On the bisimulation invariant fragment of monadic $s1$ in the finite. *In : FSTTCS*, pp. 224–236. – 2004.
- [357] DERBEL (B) et MOSBAH (M). – Distributing the execution of a distributed algorithm over a network. *In : 7th International Conference on Information Visualization (IV03)*. pp. 485–490. – London, England, 16-18 July 2003.
- [358] DERBEL (B) et MOSBAH (M). – A fully distributed linear time algorithm for cluster network decomposition. *In : the 16th IASTED International Conference on Parallel and Distributed Computing and Systems, PDCS 2004*. pp. 548–553. – MIT, Cambridge, MA, USA, 2004.
- [359] DUONG (V), FERCHAUD (F), GAVOILLE (C) et MOSBAH (M). – Using absorption areas to improve ATFM. *In : ERC Innovative Research Workshop*. – décembre 2003.
- [360] DUONG (V), FERCHAUD (F), GAVOILLE (C) et MOSBAH (M). – A new slot allocation for ATFM. *In : 7th IEEE International Conference on Intelligent Transportation Systems (ITSC)*, pp. 1075–1079. – octobre 2004.
- [361] DUONG (V), FERCHAUD (F), GAVOILLE (C) et MOSBAH (M). – Reducing disturbances by using absorption areas. *In : 1st International Conference on Research in Air Transportation (ICRAT)*, pp. 111–115. – novembre 2004.
- [362] DUONG (V), FERCHAUD (F), GAVOILLE (C) et MOSBAH (M). – Using absorption areas to improve ATFM. *In : 23rd IEEE International Digital Avionics Systems Conference (DASC)*. – octobre 2004.
- [363] DURAND (I). – Autowrite : A tool for checking properties of term rewriting systems. *In : Proceedings of the 13th International Conference on Rewriting Techniques and Applications*. pp. 371–375. – Copenhagen, 2002.
- [364] DURAND (I). – Autowrite : A tool for term rewrite systems and tree automata. *In : Proceedings of the Workshop on Rewriting Strategies*, pp. 5–14. – Aachen, June 2004.
- [365] ÉSIK (Z) et WEIL (P). – On logically defined recognizable tree languages. *In : Proceedings FSTTCS 2003*. pp. 195–207. – Springer-Verlag, 2003.
- [366] FRATANI (S) et SÉNIZERGUES (G). – Iterated pushdown automata and sequences of rational numbers. *In : St Petersburg second Logic Days*. – 2003. Abstract <http://logic.pdmi.ras.ru/2ndDays/index.html>.
- [367] GODARD (E), MÉTIVIER (Y), MOSBAH (M) et SELLAMI (A). – Termination Detection of Distributed Algorithms by Graph Relabelling Systems. *In : Proc. of 1st International Conference on Graph Transformation*. pp. 106–119. – Springer-Verlag, 2002.

- [368] HAMID (B) et MOSBAH (M). – An implementation of a failure detector for local computations in graphs. *In : International Conference on Parallel and Distributed Computing Networks, (PDCN 2005)*. pp. 473–478. – Innsbruck, Austria, 15-17 Février 2005.
- [369] HAMID (B) et MOSBAH (M). – An automatic approach to self-stabilization. *In : the 6th ACIS International Conference on Software Engineering, Artificial Intelligence, Networking, and Parallel/Distributed Computing (SNPD2005)*. – Towson University, Maryland, U.S.A., 23 - 25 Mai 2005. à paraître.
- [370] HAMID (B) et MOSBAH (M). – A formal model for fault-tolerance in distributed systems. *In : THE 24th International Conference on Computer Safety, Reliability and Security*. – Fredrikstad, Norvège, 28 - 30 Septembre 2005. à paraître.
- [371] HAMID (B) et MOSBAH (M). – Visualization of self-stabilizing algorithms. *In : 9th International Conference on Information Visualization (IV05)*. – London, England, 6-8 Juillet 2005. à paraître.
- [372] HUET (G). – Automata mista. *In : Festschrift in Honor of Zohar Manna for his 64th anniversary*, éd. par Dershowitz (N). – Taormina, 2003.
- [373] HUET (G). – Lexicon-directed segmentation and tagging of Sanskrit. *In : XIIth World Sanskrit Conference*. – Helsinki, 2003.
- [374] HUET (G). – Linear contexts and the sharing functor : Techniques for symbolic computation. *In : Thirty Five Years of Automating Mathematics*, éd. par Kamareddine (F). – Kluwer, 2003.
- [375] HUET (G). – Towards computational processing of sanskrit. *In : International Conference On Natural Language Processing*. – Mysore, 2003.
- [376] KNAPIK (T), NIWINSKI (D), URZYCZYŃ (P) et WALUKIEWICZ (I). – Unsafe grammars and panic automata. *In : ICALP'05*. – 2005. à paraître.
- [377] LANGHE (O. D), GUITTENY (P), PORTINE (H) et RETORÉ (C). – A propos des structures OSV en Langue des Signes Française. *In : Journées d'études internationales – La linguistique de la LSF : recherches actuelles*, éd. par Berthonneau (A.-M) et Dal (G). pp. 23–26. – Silex-C.N.R.S., 2003.
- [378] LAPORTE (M), NOVELLI (N), CICCHETTI (R) et LAKHAL (L). – Computing full and iceberg datacubes using partitions. *In : Proceedings in the 13th International Symposium on Methodologies for Intelligent Systems (ISMIS'02)*, pp. 244–254. – 2002.
- [379] LASOTA (S) et WALUKIEWICZ (I). – Alternating timed automata. *In : FOS-SACS'05*. – 2005. à paraître.
- [380] LE NIR (Y). – From NL grammars to AB grammars. *In : Categorical grammars – an efficient tool for natural language processing*. C.N.R.S., pp. 290–304. – Montpellier, 2004.
- [381] LY (O). – Compositional verification : Decidability issues using graph substitutions. *In : MFCS'04*, pp. 537–549. – 2004.
- [382] METIVIER (Y), MOSBAH (M), OSSAMY (R) et SELLAMI (A). – Synchronizers for local computations. *In : International conference on graph transformation*. pp. 271–286. – Springer-Verlag, 2004.
- [383] METIVIER (Y), MOSBAH (M) et SELLAMI (A). – Proving distributed algorithms by graph relabeling systems : Examples of trees in networks with processor identities. *In : Applied Graph Transformations*, pp. 45–57. – Grenoble, 2002.

- [384] MOHALIK (S) et WALUKIEWICZ (I). – Distributed games. *In : FSTTCS'03*, pp. 338–351. – 2003.
- [385] MOORTGAT (M), MOOT (R) et OEHRLE (R). – Teaching tools for logic-based grammar development. *In : To Eutupon*, pp. 1–14. – 2003.
- [386] MOOT (R). – Parsing corpus-induced type-logical grammars. *In : Workshop on Linguistic Corpora and Logic Based Grammar Formalisms*, éd. par Bernardi (R) et Moortgat (M). – 2003.
- [387] MOOT (R). – Graph algorithms for improving type-logical proof search. *In : Categorical grammars – an efficient tool for natural language processing*. C.N.R.S. – Montpellier, June 2004.
- [388] MOSBAH (M) et OSSAMY (R). – Checking global properties for local computations in graphs with applications to invariants testing. *In : Proceedings of the Fifth Mexican International Conference on Computer science, 20-24 September*, éd. par IEEE. pp. 35–42. – Computer Society, 2004.
- [389] MOSBAH (M) et OSSAMY (R). – A programming language for local computation graphs. *In : Eighth Maghrebian Conference on Software Engineering and Artificial Intelligence (MCSEAI'04)*. pp. 137–148. – Centre De Publication Universitaire, 2004.
- [390] MOSBAH (M) et OSSAMY (R). – A programming language for local computations in graphs : Computational completeness. *In : Proceedings of the 5th. Mexican International Conference in Computer Science Colima Mexico 20-24 September*, éd. par IEEE. pp. 12–19. – Computer Society, 2004.
- [391] MOSBAH (M) et SELLAMI (A). – Un environnement général pour étudier et implémenter les algorithmes distribués. *In : Eighth Maghrebian Conference on Software Engineering and Artificial Intelligence (MCSEAI'04)*. pp. 161–172. – Centre De Publication Universitaire, 2004.
- [392] MURAWSKI (A), ONG (L) et WALUKIEWICZ (I). – Idealized algol with ground recursion, and DPDA equivalence. *In : ICALP05*. – 2005. à paraître.
- [393] MURAWSKI (A) et WALUKIEWICZ (I). – Third-order idealized algol with iteration is decidable. *In : FOSSACS'05*, pp. 202–218. – 2005.
- [394] MUSCHOLL (A) et WALUKIEWICZ (I). – An NP-complete fragment of LTL. *In : DLT'04*, pp. 334–344. – 2004.
- [395] RETORÉ (C). – A description of the non-sequential execution of petri nets in partially commutative linear logic. *In : Logic Colloquium 99*, éd. par van Eijck (J), van Oostrom (V) et Visser (A). pp. 152–181. – ASL and A. K. Peters, 2004.
- [396] SÉNIZERGUES (G). – The equivalence problem for t-turn DPDA is co-NP. *In : Proceedings ICALP'03*. pp. 478–489. – Springer-Verlag, 2003.

Conférences invitées

- [397] COURCELLE (B). – Semantical evaluations as monadic second-order compatible structure transformations. *In : FOSSACS'02*, pp. 1–4. – 2002.
- [398] COURCELLE (B). – Recognizable sets of graphs, hypergraphs and relational structures : a survey. *In : DLT'04*, pp. 1–11. – 2004.
- [399] HUET (G). – Zen and the art of symbolic computing : Light and fast applicative algorithms for computational linguistics. *In : Practical Aspects of Declarative Languages (PADL) symposium*. – New Orleans, 2003.

- [400] POGODALLA (S) et RETORÉ (C). – *Handsome Non-Commutative Proof-Nets : perfect matchings, series-parallel orders and Hamiltonian circuits*. – Rapport technique n 5409, INRIA, 2005. Soumis au Journal of Applied Logic.
- [401] RETORÉ (C). – Semantic aspects of minimalist grammars. *In : Algebraic Methods in Language Processing – AMiLP 2003*, éd. par Spoto (F), Scollo (G) et Nijholt (A). pp. 33–38. – University of Twente, 2003.
- [402] SÉNIZERGUES (G). – $L(A)=L(B)$? decidability results from complete formal systems. *In : Proceedings ICALP'02*. pp. 37–37. – Springer-Verlag, 2002. Réception du prix Gödel.
- [403] WALUKIEWICZ (I). – Automata and games for synthesis. *In : AMAST'02*, pp. 15–21. – 2002.
- [404] WALUKIEWICZ (I). – Deciding low levels of tree-automata hierarchy. *In : WoLLIC'02*. – 2002.
- [405] WALUKIEWICZ (I). – Winning strategies and synthesis of controllers (tutorial). *In : CSL*, p. 574. – 2003.
- [406] WALUKIEWICZ (I). – A landscape with games in the background. *In : IEEE LICS'04*, pp. 356–366. – 2004.
- [407] WALUKIEWICZ (I). – How to fix it : using fixpoints in different contexts. *In : LPAR'04*, pp. 184–194. – 2005.
- [408] WEIL (P). – Algebraic recognizability of languages. *In : Proceedings MFCS 2004*. pp. 149–175. – Springer-Verlag, 2004.

Livres d'enseignement et de recherche

- [409] BERTOT (Y) et CASTÉLAN (P). – *Interactive Theorem Proving and Program Development. Coq'Art : The Calculus of Inductive Constructions*. – Springer Verlag, 2004, *Texts in Theoretical Computer Science*.

Posters et autres communications

- [410] AMBLARD (M). – Counting-dependencies languages in minimalist grammars. *In : Logical Aspects of Computational Linguistics 2005*. – 2005.
- [411] AMBLARD (M). – Synchronisation syntaxe sémantique, des grammaires minimalistes catégorielles aux constraint languages for lambda structures. *In : RECITAL05*. – 2005.
- [412] ANOUN (H). – Icharate : Un atelier logique pour les grammaires multimodales. *In : RECITAL04*. – 2004.

Edition d'ouvrages

- [413] BLACHE (P), STABLER (E), BUSQUETS (J) et MOOT (R) (édité par). – *5th International Conference on Logical Aspects of Computational Linguistics, LACL 05*. – Springer-Verlag, 2005, *LNCS/LNAI*, volume 3492.
- [414] MARLET (R) et AMBLARD (M) (édité par). – *Logical Aspects of Computational Linguistics 2005 : Student Session*. – LaBRI-C.N.R.S., 2005.
- [415] RETORÉ (C) et STABLER (E) (édité par). – *Special Issue on Resource Logics and Minimalist Grammars*. – Kluwer, 2004, *Journal of Research on Language and Computation*, volume 2(1).

2.4.2 Publications d'audience nationale

Reuves avec Comité de lecture

- [416] DE LANGHE (O), GUITTENY (P), PORTINE (H) et RETORÉ (C). – A propos des structures OSV en Langue des Signes Française. *Silexicales*, vol. 4, 2004, pp. 115–130.

Colloques avec Comité de programme et Actes

- [417] AMBLARD (M), LECOMTE (A) et RETORÉ (C). – L'interface entre syntaxe et sémantique pour les grammaires minimalistes catégorielles. *In : Journées Scientifiques Sémantique et Modélisation*. – ENS-LSH Lyon, 2004.
- [418] ANOUN (H). – Une bibliothèque coq pour le traitement des langues naturelles. *In : Journées Francophones des Langages Applicatifs*. – Obernai, France, 2004.
- [419] GIMENEZ (E) et LY (O). – *Formal Modelling and Verification of the JavaCard Security Architecture : From Static Checkings to Embedded Applet Execution*. – Rapport technique, <http://www.verificard.org/>, VERIFICARD Project, Marseille 2002.
- [420] HAMID (B) et MOSBAH (M). – Détection de pannes dans un système distribué par échange local de messages. *In : Journées Scientifiques Francophones, JSF 2003*. – Tozeur, 2003.
- [421] JEMILI (I), MOSBAH (M) et BELGHITH (A). – Algorithme distribué à deux sauts pour la clusterisation dans les réseaux ad hoc. *In : Colloque Francophone sur l'Ingénierie des Protocoles (CFIP05)*. – Bordeaux (France), 2005. Papier court.
- [422] LANGHE (O. D), GUITTENY (P), PORTINE (H) et RETORÉ (C). – Vers un traitement informatique de la syntaxe et de la sémantique de la langue des signes française. *In : Workshop on Minority Languages – TALN 2003*. – Batz-Sur-Mer, 2003.
- [423] LECOMTE (A) et RETORÉ (C). – Sémantique dans les grammaires minimalistes. *In : Journées Sémantique et Modélisation*. – C.N.R.S., 2003.
- [424] MOOT (R) et RETORÉ (C). – L'ordre des mots dans les grammaires catégorielles. *In : Pré-Actes des Journées de la syntaxe : ordre des mots dans la phrase française, positions et topologie*, éd. par Gerdes (K) et Muller (C). pp. 56–59. – E.R.S.S.-C.N.R.S., 2004. A paraître dans *Linguisticae Investigationes*.
- [425] MOSBAH (M), SELLAMI (A) et ZEMMARI (A). – Résolution distribuée de conflits dans un réseau par les systèmes de réécritures. *In : Journées Scientifiques Francophones, JSF 2003*. – Tozeur, 2003.
- [426] NOVELLI (N) et MAABOUT (S). – Algorithme efficace de calcul du produit de partitions et ses applications. *In : 19ème conférence Bases de Données Avancées (BDA '03)*, pp. 343–362. – 2003.

Livres d'enseignement et de recherche

- [427] DURAND (I) et STRANDH (R). – *Architecture de l'ordinateur*. – Dunod, 2005, *Sciences sup.*

Chapitres d'ouvrages

- [428] LOPES (S) et NOVELLI (N). – L'inférence des dépendances fonctionnelles. *In : Extraction de motifs dans les bases de données*, éd. par Hermes-Lavoisier, pp. 161–182. – 2004.

2.4.3 Autres publications

Rapports internes et autres publications

- [429] AMBLARD (M). – *Représentation sémantique pour les grammaires minimalistes*. – Rapport technique n5360, INRIA, 2004.
- [430] ANOUN (H), CASTÉРАН (P) et MOOT (R). – *Proof automation for type-logical grammars*. – Rapport technique, European Summer School in Logic, Language and Information, 2004. Notes du cours donné par Pierre Castéran and Richard Moot, disponible sur essli2004.loria.fr.
- [431] BERTOT (Y) et CASTÉРАН (P). – *Examples and exercises on Coq*. – Rapport technique, INRIA and LaBRI, 2004. available on www.labri.fr/~casteran/CoqArt/.
- [432] BLUMENSATH (A) et COURCELLE (B). – *Recognizability and Hypergraph Operations using Local Information*. – Rapport technique, LaBRI, 2005.
- [433] COURCELLE (B) et OUM (S). – *Vertex-minors, monadic second-order logic and a conjecture by Seese*. – Rapport technique, LaBRI, 2004.
- [434] D'AGOSTINO (G) et LENZI (G). – *On modal mu-calculus with explicit interpolants*. – Rapport technique n PP-2002-17, Institute for Logic, Language and Computaion, University of Amsterdam, 2002.
- [435] GIMENEZ (E) et CASTÉРАН (P). – *A Tutorial on [Co-]Inductive types in Coq*. – Rapport technique, INRIA Futurs, Logical project, 2004. Available on coq.inria.fr/doc-eng.html.
- [436] RETORÉ (C). – *Syntaxe et Traitement Automatique des Langues*. – Rr-5459, INRIA, 2005. A paraître dans *La syntaxe au cœur de la grammaire*, PUR, 2005.
- [437] SÉNIZERGUES (G). – Le problème de l'équivalence forte des programmes est décidable. *Bordeaux-1 Recherche*, 2002, no 2, pp. 1–4.

Brevets

- [438] AKKAR (M.-L), GOUBIN (L) et LY (O). – *Procédé de sécurisation d'un dispositif électronique exécutant un algorithme quelconque contre les attaques par introduction d'erreur*. – Rapport technique, European Patent 032906885 (Schlumberger), 2003.

Logiciels

Autowrite, Logiciel de réécriture de termes, implémentant en particulier des stratégies de calcul *par nécessité* (*call by need*), <http://dept-info.labri.fr/~idurand/autowrite/>

Grail, Logiciel d'analyse syntaxique pour les grammaires catégorielles multimodales, développé pour l'anglais, le néerlandais et le français, distribué sous GNU Lesser General Public License, <http://www.labri.fr/Perso/~moot/grail3.html>

Visidia, Logiciel d'implémentation, d'expérimentation et d'animation d'algorithmes distribués

2.4.4 Formation par la recherche

Thèses

- [439] BOJAŃCZYK (M). – *Definable Properties of Tree Languages*. – Thèse de doctorat, Université de Varsovie, 2004.
- [440] BONICHON (N). – *Aspects algorithmiques et combinatoires des réalisateurs des graphes plans maximaux*. – Thèse de doctorat, Université Bordeaux I, 2002.
- [441] LE NIR (Y). – *Structure des analyses syntaxiques catégorielles – application à l'inférence grammaticale*. – Thèse de doctorat, Université Rennes 1, décembre 2003.
- [442] SELLAMI (A). – *Des calculs locaux aux algorithmes distribués*. – Thèse de doctorat, Université Bordeaux I, 2004.

Habilitations

- [443] MOSBAH (M). – *Construction d'algorithmes et Transformation de graphes*. – Habilitation à diriger des recherches, Université Bordeaux I, 2002.
- [444] WALUKIEWICZ (I). – *Logique monadique du second ordre et mu-calcul*. – Habilitation à diriger des recherches, Université Bordeaux I, 2003.

2.5 Collaborations internationales

Les trois thèmes de l'équipe entretiennent un réseau dense de collaborations internationales, manifestées par des contrats européens et des contrats bilatéraux, mais aussi par des échanges de visites, des cotutelles de thèse et des publications jointes.

Les partenariats principaux concernant les thèmes *Logiques et Graphes* et *Automates et Langages* sont structurés autour du projet européen GAMES – avec des liens particulièrement étroits avec Aachen et Varsovie – et d'un réseau franco-indien avec deux instituts de recherche de Chennai. Ces travaux s'appuient aussi sur des collaborations moins structurées (mais donnant lieu cependant à des travaux communs ou des visites) avec les universités de Stuttgart, Princeton, Oldenburg, Mons, Innsbrück, Cambridge, Amagasaki et Aarhus, ainsi qu'avec l'Université de La Réunion pour laquelle, même si elle n'est pas à l'étranger, la distance pose des problèmes pratiques comparables.

Les partenariats privilégiés en *Informatique et Linguistique* sont d'abord aux Pays-Bas (Utrecht), puis parmi les pôles du réseau d'excellence européen COLOGNET, et auprès des Universités de Verona, Haïfa et Ann Arbor.

Enfin, les activités de l'équipe plus proches d'applications en algèbre s'articulent autour de collaborations avec les universités de Barcelone, Szeged, St Pétersbourg et du Nebraska.

2.5.1 Invités

Berwanger, Dietmar, *Université d'Aachen*, février-mars 2005

Blumensath, Achim, *Université d'Aachen*, janvier-décembre 2004

Bojańczyk, Mikołaj, *Université de Varsovie*, octobre-décembre 2002 et janvier-mars 2004

Dawar, Anuj, *Université de Cambridge*, avril et juillet 2003

Delfitto, Denis, *Università di Verona*, 2 semaines en juin 2004

Ésik, Zoltan, *Université de Szeged*, mai 2003

Francez, Nissim, *Technion, Haifa*, 1 semaine en avril 2004

Habel, Annegret, *Université d'Oldenburg*, octobre 2004

Kumar, Narayan, *Chennai Mathematical Institute*, 2 semaines en mai 2002 et 2 semaines en mai 2003

Lodaya, Kamal, *Institute of Mathematical Sciences, Chennai*, 1 mois en septembre-octobre 2002 et 1 mois en mai-juin 2005

Moortgat, Michael, *Universiteit Utrecht*, 1 semaine en mars 2004

Mukund, Madhavan, *Chennai Mathematical Institute*, 2 semaines en mai 2002 et 2 semaines en mai 2003

Ramanujam, R., *Institute of Mathematical Sciences, Chennai*, 1 mois en mai-juin 2005

Phan Trung, Huy, *Institut de mathématiques de Hanoi*, octobre-décembre 2002 et septembre-novembre 2004

Vermaat, Willemijn, *Universiteit Utrecht*, 1 semaine en mars 2004 – 2 semaines en avril 2005

2.5.2 Invitations

Courcelle, Bruno, *Université de La Réunion*, 3 semaines en octobre 2004

Janin, David, *RWTH Aachen*, 1 semaine, juin 2003

Moot, Richard, *OTS Universiteit Utrecht*, juillet et août 2003

Sénizergues, Géraud, *Université de Stuttgart*, août 2004 à juillet 2005

Walukiewicz, Igor, *Université de Aarhus, Danemark*, avril 2002

Walukiewicz, Igor, *National Institute of Advanced Industrial Science and Technology, Amagasaki, Japon*, 3 semaines en novembre 2002

Walukiewicz, Igor, *Université de Varsovie*, avril 2003

Walukiewicz, Igor, *Université de La Réunion*, juin 2004

Weil, Pascal, *Université du Nebraska*, août 2003 à juillet 2004

Weil, Pascal, *Centre de Recerca Matemàtica, Barcelone*, 2 semaines en novembre 2004 et 2 semaines en avril 2005

2.6 Liste des contrats et subventions

2.6.1 Contrats internationaux et européens

Période. 2002-06

Source. Union Européenne, RTN

Partenaires. Aachen, Bordeaux, Edinburgh, Paris-7, Rice University, Uppsala, Varsovie, Vienne

Description. GAMES : étude de la théorie des jeux, et de son utilisation pour la synthèse et la validation.

Période. 2003-05

Source. Union Européenne, Réseau d'excellence

Partenaires. Bordeaux, Nancy, Sarrebrück, Dublin, Haïfa, Bolzano, Poznan, Amsterdam, Londres, Edinburgh

Description. COLOGNET, AREA 6 : applications de la logique computationnelle au traitement des langues naturelles.

Période. 2000-03

Source. Union Européenne, Réseau INTAS

Partenaires. Vienne, Paris-7, Bordeaux, Tel-Aviv, St-Andrews, Manchester, Porto, Ekaterinbourg, Moscou, Saratov, St Petersburg, Vologda.

Description. CGT : étude de la théorie combinatoire et géométrique des groupes et des semigroupes, et de ses interactions avec l'informatique théorique.

Période. 2000-03

Source. CEFIPRA (Centre Franco-Indien de Promotion de la Recherche Avancée)

Partenaires. LIAFA (Paris-7), le LSV (Cachan), Institute of Mathematical Sciences (Chennai), Chennai Mathematical Institute

Description. ACSV : applications de la théorie des automates à la modélisation du parallélisme, méthodes syntaxiques pour la vérification.

Période. 2002

Source. Projet CNRS / ASTV (Viêt-Nam)

Description. Graphes, automates et codes : codage et décodage à l'aide d'outils issus de la théorie des automates et de la théorie des graphes.

Période. 2004-05

Source. Action Intégrée franco-néerlandaise VAN GOGH (MAE et OTS)

Partenaires. Universiteit Utrecht

Description. Grammaires minimalistes et catégorielles pour le traitement de la syntaxe et de la sémantique des langues naturelles.

Période. 2004

Source. Projet CNRS / ASTV (Viêt-Nam)

Description. Applications de la théorie des monoïdes et des groupoïdes pour l'étude et la classification des langages formels.

2.6.2 Contrats avec des entreprises

Période. 2002

Type de financement. Convention CORTECHS

Entreprise. Kimo

Description. Gestion d'une base de données distribuées.

Période. 2002

Type de financement. collaboration conventionnée

Entreprise. MediaSelector

Description. Moteur de recherche intelligent.

Période. 2002-05

Type de financement. Contrat industriel

Entreprise. Eurocontrol

Description. Modélisation du flux du trafic aérien par des graphes.

Période. 2002-03

Type de financement. Contrat ANVAR puis RDTA

Entreprise. GD-Tech

Description. Plateforme de processus de tests industriels.

Période. 2003

Type de financement. Contrat industriel

Entreprise. Schlumberger Systèmes

Description. Sécurité des cartes à puces, compilation.

Période. 2003

Type de financement. Convention CORTECHS

Entreprise. SystemTV

Description. XML, Sécurité et bases de données.

Période. 2003-04

Type de financement. Contrat RDTA

Entreprise. E-critur

Description. Convertisseur d'archives OCR vers XML.

2.6.3 Contrats publics avec des instances nationales

Période. 2001-03

Source. Action Spécifique JEMSTIC (département STIC, CNRS)

Description. Etude de l'interaction entre jeux, logique et automates, et applications à la problématique de la synthèse et de la vérification.

Période. 2002-03

Source. Action Spécifique du Réseau Thématique Pluridisciplinaire *Outils mathématiques pour l'informatique* (département STIC, CNRS)

Partenaires. LaBRI, LIAFA, LSV et VERIMAG

Description. Automates, modèles distribués et temporisés.

Période. 2002-04

Source. ATIP Jeune Equipe (département STIC, CNRS)

Description. Algorithmes distribués sur des structures évolutives.

Période. 2003-06

Source. ACI *Sécurité de l'Information* (MENRT)

Partenaires. LaBRI, LIAFA

Description. VERSYDIS : Vérification de systèmes distribués, développement des méthodes utilisant des modèles sans entrelacement.

2.6.4 Contrats publics avec instances locales ou régionales

Période. 2004 et 2005

Source. Appel d'offre investissement-recherche, Conseil Régional d'Aquitaine

Description. Traitement informatique de la Langue des Signes Française, génération automatique d'énoncés

2.7 Animation de la recherche

2.7.1 Comités de rédaction

Acta Cybernetica
 Computing and Informatics
 Discrete Mathematics and Theoretical Computer Science
 Information and Computation
 Traitement Automatique des Langues (rédacteur en chef)

2.7.2 Comités de programmes de colloques

AFL (Automata and Formal Languages) 2005
 AGT (International Symposium on Applications of Graph Theory) 2003
 AICCSA (International Conference on Computer Systems and Applications) 2006
 BDA (Bases de Données Avancées) 2004 et 2005
 CADE (Conference on Automated Deduction) 2002
 Categorical Grammars 2004
 CONCUR (International Conference on Concurrency Theory) 2002 et 2005
 Défi Fouille de Textes 2005
 Ecole de Printemps d'Informatique Théorique 2004
 EGC (Extraction et Gestion de Connaissances) 2003 et 2004
 ESSLLI (European Summer School in Logic, Language and Information) 2005
 EXPRESS (Workshop on Expressiveness in Concurrency) 2003
 FICS (Fixpoints in Computer Science workshop) 2003 (co-présidence)
 FoSSACS (Foundations of Software Science and Computation Structures) 2003 et 2004
 (présidence)
 FST-TCS (Foundations of Software Technology, Theoretical Computer Science) 2002 et
 2004
 GAMES 2002, 2003 et 2005
 GDV (Games in Design and Verification) 2004
 Human Language Technology 2005
 ICALP (International Colloquium on Automata, Languages and Programming) 2004
 ICGT (International Conference on Graph Transformations) 2001, 2004 et 2005
 IDA (Intelligent Data Analysis) 2005
 INFINITY (International Workshop on Verification of Infinite-State Systems) 2005
 JSF Journées Scientifiques Francophones, Electronique, Télécommunications et Informa-
 tique 2003
 LACL (Logical Aspects of Computational Linguistics), Session étudiante, 2005 (co-présidence)
 LACL (Logical Aspects of Computational Linguistics) 2005
 LCC (Logic and Computational Complexity workshop) 2005
 LICS (IEEE Symposium on Logic in Computer Science) 2005
 Mathematics of Language 2003
 MFCS (International Symposium on Mathematical Foundations of Computer Science)
 2003
 SAC (Annual Symposium on Applied Computing) 2006
 SLP (Workshop on Sign Language Processing), Linguistic Resources European Conference,
 2004
 STACS (Symposium on the Theoretical Aspects of Computer Science) 2003 et 2005
 WG (International Workshop on Graph-Theoretic Concepts in Computer Science) 2005

WoLLIC (Workshop on Logic, Language, Information and Computation) 2004

2.7.3 Organisation de colloques, écoles jeunes chercheurs, ...

Groupe de travail ALP-AMI *Complexité, modèles finis et bases de données*, Arcachon, 2002

Satellite workshop Logic, graph transformations and discrete structures de l'ICGT (*Intern. Conference on Graph Transformations*), Barcelone, 2002

Réunion nationale de l'AS *Automates, systèmes distribués et temporisés* du département STIC du CNRS, Bordeaux, 2002

Workshop *Automata, Concurrency and Logic*, Institute of Mathematical Sciences, Chennai, 2003

JSF Journées Scientifiques Francophones, Tozeur, 2003

Ecole de Printemps d'Informatique Théorique, Luminy, 2004

Satellite workshop Logic, Graph Transformations, Finite and Infinite Structures de l'ICGT (*Intern. Conference on Graph Transformations*), Rome, 2004

GAMES Annual Meeting, Bordeaux, 2004

LACL (Logical Aspects of Computational Linguistics), Bordeaux, 2005

Spring School on Infinite Games and Their Applications, Bonn, 2005

2.7.4 Administration de la Recherche

Responsabilités internationales

Jury du prix de thèse Ackermann décerné par l'EACSL

Jury du prix de thèse E. W. Beth

Comité de pilotage de ETAPS

Comité de pilotage de ICGT

Comité de pilotage du RTN GAMES

Comité de pilotage du STACS

Attaché pour la Science et la Technologie, Consulat de France à Shanghai

Responsabilités nationales

Jury du prix de thèse de l'AFIT, 2002

Jury du concours d'entrée (informatique, 1ère année) à l'ENS Cachan

Responsabilités régionales

Chargé de mission à mi-temps auprès de la Délégation TIC du Conseil Régional

Conseil Scientifique du pôle EITICA (transfert technologique dans les domaines de l'électronique, l'informatique et les technologies de la communication)

Participation à des jurys de thèse extérieurs

20 jurys de thèse, dont 15 comme rapporteurs (et 6 à l'étranger)

Participation à des jurys d'habilitation extérieurs

8 jurys d'habilitation, dont 5 comme rapporteur (et 1 à l'étranger)

Expertise internationale

Fonds FCAR (recherche publique, Québec)

Modélisation, Vérification et Test des systèmes informatiques (MVTsi)

Responsable(s) André ARNOLD

Théorie et pratique de la vérification

Responsable(s) Jean-Michel COUVREUR et Alain GRIFFAULT

*Mots-clés : Vérification, Systèmes infinis, Systèmes hybrides, Logiques comportementales,
Calcul symbolique, Abstraction.*

Tests de systèmes informatisés

Responsable(s) Richard CASTANET

Mots-clés : Test, Protocoles, Robustesse, Interopérabilité, Conformité.

Modélisation du vivant

Responsable(s) David SHERMAN

*Mots-clés : Bioinformatique, Modélisation multi-échelle stochastique, Comportement
cellulaire.*

3.1 Composition de l'équipe

3.1.1 Membres permanents

ARNOLD, André, Pr, Bx 1
CASTANET, Richard, Pr, ENSEIRB
COUVREUR, Jean-Michel, MdC, HDR, Bx 1
DICKY, Anne, MdC, Bx 1
FELIX, Patrick, MdC, Bx 1 (IUT)
GRIFFAULT, Alain, MdC, Bx 1
HERBRETEAU, Frédéric, MdC, ENSEIRB, depuis le 1er septembre 2003
MUSUMBU, Kaninda, MdC, Bx 1
NIKOLSKI, Macha, CR, CNRS
POINT, Gérard, IR, CNRS
SHERMAN, David, MdC, HDR, ENSEIRB
SUTRE, Grégoire, CR, CNRS
VINCENT, Aymeric, MdC, ENSEIRB, depuis le 1er septembre 2004

3.1.2 Doctorants

BERRADA, Ismaël, Doctorant, depuis septembre 2002
BRIAND, Xavier, Doctorant, All. MRT, depuis septembre 2002
CHAUVEAU, Emmanuel, Doctorant, FranceTélécom, Lannion, depuis septembre 2001
FERCHAUD, F., Doctorant, ADERA/EUROCONTROL, depuis septembre 2003
GRIVET, Sébastien, Doctorant, All. MRT, jusqu'en 2004
JEMILI, Imen, Doctorant, bourse franco-tunisienne, depuis septembre 2003
KEHREN, Christophe, Doctorant, DGA, depuis septembre 2002
NGUENA, Omer-Landry, Doctorant, depuis septembre 2003
SAAD KHORCHEF, Farès, Doctorant, depuis septembre 2003
SAGASPE, Laurent, Doctorant, ONERA, depuis septembre 2004
THOUVENIN, Cyrille, Doctorant, CIFRE, depuis septembre 2002
TRAN, The Quang, Doctorant, All. MRT, depuis septembre 2004
YAHIOUN, Karim, Doctorant, CIFRE, depuis septembre 2004

3.1.3 Membres non permanents

ADELAIDE, Michaël, ATER, ENSEIRB, de septembre 2002 à juillet 2004
BREANT, François, PostDoc, Projet Morse, d'octobre 2003 à août 2004
GEHLOT, Shweta, Stage Doc., Contrat EGIDE, de septembre 2004 à mars 2005
KLAI, Kais, PostDoc, Projet Morse, d'octobre 2004 à août 2005
MACKAYA, Marcien, ATER, ENSEIRB, depuis septembre 2003
PAGETTI, Claire, ATER, Bx 4, de septembre 2003 à juillet 2004
SANTOCANALE, Luigi, PostDoc, Bourse Marie Curie, d'avril 2002 à septembre 2003

3.2 Présentation de l'équipe

Les systèmes critiques sont des assemblages complexes de composants logiciels et matériels qui assurent essentiellement des fonctions de contrôle-commande dans divers types de produits allant des avions (commandes de vol) aux produits de grande consommation (par exemple les compteurs électriques à télé-relève) ou dans les processus de production (par exemple la gestion des flux dans les exploitations pétrolières off-shore). Ils se distinguent d'autres applications de l'informatique par le fait que leurs dysfonctionnements ont des conséquences irréversibles : soit la mort de personnes dans les domaines du transport, du nucléaire, de la médecine ; soit la mise en danger économique d'une entreprise, lorsque celle-ci doit faire revenir dans ses ateliers un million de composants du fait d'un défaut de conception.

La conception de tels systèmes doit faire appel à des techniques et des méthodes très rigoureuses pour en garantir le bon fonctionnement, et le monde industriel est de plus en plus attentif aux résultats de la recherche "académique" dans ces domaines. Depuis une vingtaine d'années l'équipe MVTsi du LaBRI se consacre à ce domaine de recherche où elle joue toujours un rôle très original par la conjonction de trois types d'activités :

- une activité théorique consistant à étudier et développer divers types de formalismes permettant de décrire ces systèmes critiques et divers types de logiques afin d'en exprimer les fonctionnalités attendues ;
- une activité algorithmique dont le but est de fournir des outils d'analyse des systèmes critiques, essentiellement des vérificateurs (model-checkers) mais aussi des simulateurs. Elle va de la conception d'algorithmes à leur implémentation efficace en utilisant des représentations symboliques compactes des données à manipuler ;
- une activité plus pragmatique de modélisation dans le cadre de la collaboration et la participation à des projets industriels.

Longtemps limitée aux systèmes à espaces finis d'états, cette activité est en cours d'extension aux systèmes infinis, soit discrets (systèmes à pile, à compteurs), soit continus (systèmes hybrides, stochastiques).

La thématique commune à l'équipe, décrite ci-dessus, se présente de façon différente selon le type de système étudié. Ainsi les activités de recherche sur le *test de systèmes logiciels* se caractérisent par des domaines applicatifs spécifiques et demandent des connaissances particulièrement fines de ces domaines ainsi qu'une spécialisation fort contraignante des techniques générales développées dans l'équipe pour pouvoir traiter efficacement les problèmes qui s'y rattachent. Il en est de même pour les recherches sur les *systèmes biologiques* qui sont susceptibles du même type d'approche que les systèmes physiques informatisés. C'est pourquoi les activités de recherche dans ces deux domaines seront décrites séparément.

3.3 Thèmes de recherche de l'équipe

3.3.1 Théorie et pratique de la vérification

Introduction

Historiquement, ce thème s'est bâti autour des travaux de A. Arnold sur les systèmes de transitions. Le produit synchronisé de systèmes de transitions a été introduit par A. Arnold et M. Nivat en 1982¹ comme un moyen de modéliser les systèmes de processus communicants. Chaque composant d'un système est représenté par un système de transitions qui décrit aussi bien les états dans lesquels ce composant peut se trouver, que la façon dont cet état se modifie lorsque le composant exécute certaines actions ou subit l'influence de certains événements extérieurs. Les interactions entre les différents composants d'un système sont décrites par des vecteurs de synchronisation et le produit synchronisé est lui-même un système de transitions qui représente le comportement global du système. La théorie des automates est donc au centre de nos préoccupations, ainsi que les différentes logiques temporelles, notamment le μ -calcul qui a fait l'objet de nombreux développements (théoriques et algorithmiques) au LaBRI.

Mots-Clés Vérification, Systèmes infinis, Systèmes hybrides, Logiques comportementales, Calcul symbolique, Abstraction.

Description des activités et principaux résultats

Le projet AltaRica Ce projet a débuté en 1996 et consistait à définir un formalisme permettant aussi bien les descriptions fonctionnelles que dysfonctionnelles des systèmes complexes et critiques. En 1997, un partenariat entre le LaBRI, et les sociétés ELF Aquitaine, IXI, Dassault Aviation, Schneider, Renault, Thomson Detexis, et l'IPSN, s'est constitué pour définir le langage ALTARICA, et développer les premiers prototypes. En 2000 la société Dassault Aviation a réalisé un Atelier ALTARICA, basé sur les outils prototypes, et sur un sous-langage d'ALTARICA. Parallèlement, un certain nombre d'outils logiciels ont été développés au LaBRI et dans d'autres laboratoires universitaires avec lesquels l'équipe collabore étroitement (voir ci-dessous).

Depuis 2002 ce langage est intensément utilisé au CERT par l'équipe de Christel Seguin. Il est aussi un des langages de référence du projet européen ISAAC et de l'ACI Persée (voir ci-dessous).

Parallèlement à ces travaux, la communauté ALTARICA organise régulièrement des rencontres : le premier *workshop* a eu lieu en octobre 2002 au CERT-ONERA à Toulouse, le second a eu lieu en octobre 2003 au CIRM à Marseille, et le troisième devrait se dérouler dans la région bordelaise.

Le formalisme AltaRica Ce formalisme, présenté dans la thèse de G. Point² et dont la sémantique formelle est donnée dans³, étend le modèle des produits synchronisés de systèmes de transitions dans trois directions :

¹Comportements de processus. *Colloque AFCEP "Les mathématiques de l'informatique"*, pp 35–68, 1982.

²*Altarica : Contribution à l'unification des méthodes formelles et de la sûreté de fonctionnement*. Thèse de doctorat, Université Bordeaux I, janvier 2000

³A. Arnold, A. Griffault, G. Point, A. Rauzy. The AltaRica formalism for describing concurrent systems. *Fundamenta Informaticae*, 40 :109–124, 2000

- l'interaction des composants ne se décrit pas seulement par des synchronisations d'événements, mais aussi par des échanges instantanés d'informations transportées par des flux,
- le notion de vecteur de synchronisation a été étendue en introduisant les notions de diffusion et de priorité,
- le formalisme permet une construction hiérarchisée des modèles au moyen d'un contrôleur assurant la liaison entre les niveaux inférieurs et supérieurs de la hiérarchie.

Les outils AltaRica Divers outils ont été développés au LaBRI, sans lesquels ALTARICA est inutilisable :

1. Depuis 2001, réalisation de *MecV*, un vérificateur de modèles ALTARICA basé sur une représentation symbolique des états accessibles, par Aymeric Vincent lors de sa thèse soutenue en décembre 2003.
2. Depuis 2001, en collaboration avec l'IRCCyN, définition du langage TIMEDALTARICA, et réalisation d'un compilateur vers UPPAAL, par Claire Pagetti lors de sa thèse soutenue en avril 2004.
3. Depuis l'arrivée de Gérald Point à mi-temps dans l'équipe MVTsi en septembre 2002, développement des outils *altatools* comprenant notamment : *acheck* un vérificateur de modèles ALTARICA basé sur la représentation explicite des états accessibles, des compilateurs vers et depuis le langage *Lustre* (besoin issu du projet européen ESACS), et un prototype de simulateur utile pour l'enseignement.

Synthèse de contrôleurs Les principes de base de la théorie du contrôle des systèmes à événements discrets ont été posés dès le milieu des années 80 par Wonham et ses collaborateurs. Bien que relevant de la théorie des automates finis, cette théorie a été très peu étudiée par les informaticiens, jusqu'à une période récente où les travaux se sont multipliés, en particulier sur le contrôle des systèmes temporisés.

En 2003, A. Arnold, A. Vincent et I. Walukiewicz⁴ ont montré comment la théorie des jeux permettait de résoudre une classe de problèmes de contrôle beaucoup plus vaste que celles considérées jusqu'alors.

La méthode proposée par ces auteurs a été implémentée par G. Point et fait actuellement l'objet d'expérimentation, tandis que X. Briand travaille sur des généralisations de cette méthode à de nouvelles classes de problèmes de contrôle.

Structures de données pour la vérification Dans les années 90, les industries des composants électroniques, dans leur recherche d'outils pour améliorer le niveau de confiance de leurs produits, ont adopté les diagrammes de décisions binaires (BDD) pour traiter des composants de plus en plus complexes. Les BDD sont des structures codant des fonctions booléennes. Ils peuvent être vus comme des graphes dirigés acycliques où les états représentent des choix de valeurs de variables booléennes ; un ordre total sur les variables garantit l'unicité du codage d'une fonction. Dans le cadre de projets industriels, nous avons conçu une nouvelle structure à la BDD, les *Diagrammes de Décisions de Données* (DDD)⁵. Notre objectif était de fournir un outil flexible qui peut être adapté à la vérification de tout type de modèle et qui offre des capacités de traitement similaires aux BDD. À la différence des BDD, les opérations sur nos structures ne sont pas prédéfinies, mais une classe d'opérateurs,

⁴Games for synthesis of controllers with partial observation, *Theoretical Computer Science*, 303 (2003) 7–34.

⁵Couvreur et al. Data Decision diagrams for Petri net analysis, *ICAPTN 2002*.

appelée *homomorphismes*, est introduite pour permettre à un utilisateur de concevoir ses propres opérations. Dans notre modèle, les variables ne sont pas booléennes ; elles prennent leurs valeurs dans des domaines non nécessairement bornés. Une autre caractéristique intéressante est qu'aucun ordre sur les variables n'est présupposé dans la définition. De plus, une variable peut apparaître plusieurs fois dans un même chemin. Cette propriété est très utile quand on manipule des structures dynamiques comme les files. Grâce à la grande flexibilité de la structure, nous avons montré l'aptitude des DDD à traiter des programmes VHDL (dans le cadre du projet CLOVIS, un projet DGA). Les succès de notre première étude ont conduit les DDD à être choisis (dans le projet MORSE, un projet RNTL) comme structures pour la vérification de systèmes décrits en LFP, un langage de prototypage de haut niveau. Nous avons proposé récemment ⁶ et mis en œuvre deux améliorations à la structure originale : (1) adapter la structure à la manipulation de modèles hiérarchiques, (2) ajouter un opérateur local de point fixe pour accélérer les calculs. Le résultat est un gain significatif des temps de vérification de l'ordre de 100 et plus pour certains modèles. Ces méthodes ont été immédiatement adoptées par notre projet et sont un atout essentiel aux succès de notre étude.

LTL La logique temporelle est un langage puissant permettant de décrire des propriétés de sûreté, d'équité et de vivacité de systèmes. Elle est utilisée comme langage de spécification dans des outils tel que SPIN (Holzman, 1993) et SMV (Mac Millan, 1994). Cependant, vérifier qu'un système fini respecte une telle spécification est PSPACE-complet. En pratique, les techniques de vérification sont confrontées à un problème d'explosion combinatoire du nombre d'états du système et de celui de l'automate codant la formule. De nombreuses techniques ont été élaborées pour faire face à ce problème d'explosion. Nous pouvons noter les techniques de vérification à la volée combinées avec des techniques de réduction à base d'*ordre partiel* (SPIN). La représentation symbolique par les diagrammes de décision permet de coder un système et l'automate d'une formule de manière concise et ainsi de repousser les limites de la vérification (SMV). Toutes ces méthodes ont leurs succès sur des systèmes industriels prouvant leur bien-fondé.

En 1999 (Couvreur, 1999), nous avons proposé de nouveaux algorithmes pour résoudre les deux problèmes clefs de la vérification à la volée d'une formule LTL :

- Construire à la demande un automate représentant une formule LTL ;
- Tester à la volée si l'automate résultant du produit synchronisé du système et de l'automate de la propriété est vide.

Dans le cadre de la réalisation de l'outil Spot en collaboration avec le LIP6 de l'université Paris 6, nous avons réalisé de nombreuses expérimentations de ces deux algorithmes. Nous avons montré que malgré sa simplicité, l'algorithmique de construction d'un automate restait très compétitive parmi les nombreuses constructions récentes ⁷ et que l'algorithmique de vérification était ce qui se faisait de mieux actuellement.

En 2003, nous nous sommes intéressés à la vérification de formules LTL sur des systèmes probabilistes. Les algorithmes probabilistes sont conçus dans de nombreux domaines de l'informatique, et plus particulièrement en algorithmique distribuée. En fait, dans ce domaine, il existe des problèmes qui n'admettent pas de solutions algorithmiques déterministes satisfaisantes. Ainsi, le choix de mettre en œuvre des solutions probabilistes devient nécessaire. L'extension et la conception de techniques de vérification efficaces pour les systèmes concurrents probabilistes constituent encore à ce jour un défi. Le principal résultat de notre travail

⁶Bréant et al. Modeling and verifying behavioral aspects, in *Formal methods for embedded distributed systems*, Kluwer Academic, 2004.

⁷Voir : Mémoire d'habilitation de J.-M. Couvreur.

est la conception d'une méthode à base d'automates pour la vérification de formules LTL sur des systèmes probabilistes ⁸. Le point essentiel est que l'algorithme proposé résout le problème avec une complexité optimale. Comme dans le cas non probabiliste, nous synchronisons le système avec l'automate de la négation de la formule. Cependant, nous utilisons une construction particulière des automates de la formule. Cette construction, que nous avons proposée en 2000, produit des automates ayant des propriétés spécifiques. Nous avons exploité ces bonnes propriétés pour éviter la déterminisation de ces automates qui est une étape coûteuse conduisant à une complexité doublement exponentielle en temps.

Ordre partiel Les techniques de dépliage ont montré toute leur efficacité dans le cas des réseaux de Petri pour contrer l'explosion combinatoire qui survient lors de l'exploration de l'espace d'états. Elles consistent d'une part en un parcours intelligent visant à tirer parti de la concurrence intrinsèque au modèle, et d'autre part en une règle de coupure du processus arborescent ainsi construit de façon à obtenir un préfixe fini préservant la propriété considérée (par exemple l'accessibilité).

Nos travaux ont dans un premier temps porté sur l'amélioration de l'efficacité de l'algorithme de dépliage pour certaines classes de modèles ⁹. Nous avons généralisé ces méthodes à des réseaux de Petri intégrant des symétries, et montré comment déplier des produits de réseaux. L'aboutissement de cette étude est la conception d'une méthode efficace pour traiter des systèmes communiquant par rendez-vous et par échanges de messages.

Nos autres travaux ont cherché à étendre les techniques de dépliage aux systèmes infinis. Les systèmes dits "bien structurés" (Abdulla et Jonsson, 1993 ; Finkel et Schnoebelen, 2001) forment une classe de modèles infinis pour lesquels des problèmes classiques de vérification (la couverture, la terminaison, la bornitude, etc.) sont décidables. Cette classe est très générale puisque de nombreux systèmes qui ne sont pas naturellement bien structurés (comme les systèmes à file FIFO) peuvent être transformés en systèmes bien structurés (par exemple, en autorisant les pertes de messages) fournissant ainsi des semi-procédures ¹⁰ pour la décision de ces problèmes. Nous nous sommes donc intéressés à la combinaison de ces deux techniques : le dépliage pour un parcours efficace du graphe d'accessibilité, et la bonne structuration pour arrêter ce parcours dès qu'assez d'information est connue pour déduire la satisfaction ou la non-satisfaction de la propriété considérée. Notre but est la vérification de systèmes hétérogènes. L'un des défis de ce travail a consisté à trouver des techniques de dépliage efficaces pour différentes données : les compteurs et les files, notamment. Nous avons alors produit un cadre général compositionnel : le dépliage d'un produit synchronisé de systèmes est transformé en le produit synchronisé des dépliages de ces systèmes, ce qui préserve la concurrence et donc l'efficacité du dépliage. Enfin, nous déduisons de la bonne structuration de ces systèmes, la bonne structuration du dépliage complet, ce qui nous permet de résoudre les problèmes considérés. Ce travail a fait l'objet du stage de DEA de Tran The Quang, et du début de son doctorat. Un article est actuellement en cours de rédaction.

Vérification des systèmes infinis La vérification formelle d'un système passe par la traduction, dans des modèles mathématiques appropriés, (1) du système (phase de *modélisation*), et (2) des comportements attendus du système (phase de *spécification*). Ces étapes de modélisation et de spécification sont cruciales car la vérification sera réalisée sur ces mo-

⁸Couvreur, Saheb et Sutre. An optimal approach to LTL model checking of probalistic systems, *LPAR 2003*.

⁹Voir le mémoire d'habilitation de J.-M. Couvreur et la thèse de S. Grivet.

¹⁰Au sens où le modèle transformé est une surapproximation du modèle original qui préserve la satisfaction des propriétés, mais non leur insatisfaction.

dèles mathématiques, aussi la traduction doit-elle être fidèle. Le choix des modèles mathématiques est également capital car l'automatisation de la vérification dépend de la puissance d'expression des modèles choisis. D'une manière générale, plus le modèle est expressif, plus la traduction peut être fidèle et moins la vérification peut être automatisée.

Les modèles mathématiques utilisés pour décrire les systèmes sont généralement des *automates étendus* avec des *variables* (par exemple des variables entières, des files, des horloges). Les *localités* de l'automate (aussi parfois appelées états de contrôle) sont en nombre fini. En revanche, les variables peuvent porter sur des *domaines de données* arbitraires. Les transitions de l'automate décrivent l'évolution des variables lors du passage d'une localité à l'autre. Quant aux propriétés, leur expression formelle passe bien souvent par l'utilisation de *logiques*. Les logiques temporelles LTL, CTL et CTL* sont particulièrement bien adaptées à la spécification des systèmes *réactifs*.

Les systèmes les plus simples pourront être modélisés par des automates étendus *finis* (c'est-à-dire avec des domaines de données finis), dont la vérification par *model-checking* est entièrement automatique. Cependant, certains aspects essentiels des systèmes complexes ne peuvent pas être fidèlement modélisés par des automates étendus finis, par exemple : la *communication*, qui est souvent modélisée par des canaux FIFO non bornés (avec ou sans perte), le *temps-réel*, que l'on retrouve généralement sous la forme d'horloges ou de chronomètres à valeurs réelles, la *paramétrisation*, lorsqu'on souhaite raisonner sur des réseaux de processus avec un nombre arbitraire de composants... La modélisation fidèle de systèmes complexes nécessite ainsi l'utilisation d'automates étendus *infinis* (c'est-à-dire avec des domaines de données infinis).

La vérification des automates étendus infinis est indécidable en général. Deux axes principaux de recherche se dégagent des nombreux travaux récents sur l'analyse automatique des systèmes infinis. La première approche cherche à développer des techniques d'analyse incomplètes, mais entièrement automatiques, sur des classes très générales d'automates étendus infinis. L'incomplétude de ces techniques peut provenir (1) de la non-terminaison éventuelle des semi-algorithmes employés (par exemple le calcul exact de l'ensemble d'atteignabilité), ou (2) de la non-exactitude de l'analyse réalisée (par exemple le calcul d'une approximation supérieure de l'ensemble d'atteignabilité). Dans la seconde approche, on cherche à déterminer des classes restreintes (mais les plus expressives possibles) de systèmes infinis pour lesquelles la vérification est décidable (par exemple les *réseaux de Petri*, les *automates à pile*, les *automates à files non fiables*, les *automates temporisés*, ...).

Nos contributions se situent dans ces trois approches, ainsi que dans la mise en œuvre effective de ces techniques.

Calcul symbolique exact Un des principaux problèmes de vérification sur les automates étendus infinis consiste à calculer l'ensemble *post** des états atteignables. En effet, une fois cet ensemble calculé, on peut facilement répondre à de nombreux problèmes de vérification sur l'automate étendu, comme par exemple l'atteignabilité d'un état d'erreur ou encore la finitude de l'ensemble des états atteignables. Le calcul de l'ensemble des états atteignables d'un automate étendu nécessite l'emploi d'une *représentation symbolique*, c'est-à-dire un moyen de manipuler des ensembles infinis d'états. On a également souvent besoin d'*accélérer* le calcul, en collectant en une étape (de complexité raisonnable) un ensemble infini d'états atteignables.

De nombreuses représentations symboliques, munies d'accélération, ont ainsi été développées au cours des dix dernières années, comme par exemple : les *Queue Decision Diagrams* (Boigelot *et al.*, 1997) et les *Constrained Queue Decision Diagrams* (Bouajjani et Habermehl, 1999) pour les automates à files FIFO, les *Simple Regular Expressions* (Abdulla *et*

al., 1998) pour les automates à files FIFO avec perte, les *Periodic Sets* (Boigelot et Wolper, 1994), les *formules de Presburger* (Comon et Jurski, 1998) et les *Number Decision Diagrams* (Wolper et Boigelot, 1995) pour les automates à compteurs, les *polyèdres convexes* (Cousot et Halbwachs, 1978; Halbwachs, 1993) et les *RVA* (Boigelot *et al.*, 1997) pour les systèmes hybrides.

(a)Unification et SLRE. Les techniques de vérification symbolique accélérée développées au cours des dix dernières années sont en fait basées sur des concepts similaires : les hypothèses sur les représentations symboliques sont semblables (clôture effective par union, décidabilité de l'inclusion, etc.) et l'accélération du calcul passe par l'itération de certains circuits. Afin de pouvoir combiner et comparer ces techniques, nous avons formalisé ces concepts en unifiant les travaux existants dans un cadre générique et uniforme de calcul symbolique accéléré ¹¹. Nous avons appliqué notre approche à l'analyse des automates à files FIFO en comparant les représentations symboliques existantes : les QDD et les CQDD. Cette comparaison nous a conduit à proposer une nouvelle représentation symbolique accélérée pour les files FIFO : les *Semilinear Regular Expressions*. Les SLRE, qui sont précisément l'intersection des QDD et des CQDD, permettent une algorithmique plus efficace tout en étant apparemment suffisantes pour représenter les contenus de file de la plupart des protocoles classiques.

(b)Accélération avec les RVA. La vérification des systèmes temporisés et hybrides s'appuie généralement sur des techniques symboliques basées sur des représentations par polyèdres convexes (Halbwachs *et al.*, 1994; Henzinger et Ho, 1994) ou des sous-classes de ceux-ci comme les DBM (Dill, 1989). Ces représentations sont néanmoins d'une expressivité trop faible pour représenter les états accessibles des systèmes temporisés ou hybrides, c'est pourquoi les techniques d'accélération associées en calculent une approximation qui ne permet pas toujours de vérifier la propriété souhaitée (Bouyer, 2003). Les Real Vector Automata (RVA) (Boigelot *et al.*, 1997; Boigelot *et al.*, 2001) permettent de représenter les ensembles descriptibles par les formules de l'arithmétique de Presburger étendue aux réels. Ils constituent donc une alternative aux représentations polyédriques puisqu'ils permettent la représentation d'ensembles périodiques infinis qui apparaissent fréquemment dans les espaces d'états des systèmes temporisés ou hybrides. Nous nous sommes intéressés à la définition d'une technique d'accélération pour les systèmes temporisés et hybrides qui profite de l'expressivité des RVA ¹². Nous avons considéré nos modèles comme des relations convexes polyédriques qui permettent la représentation des relations de transition des automates temporisés et hybrides linéaires, mais aussi celles des automates à compteurs. Ceci permet donc de vérifier des systèmes hétérogènes : compteurs + variables continues. Grâce à notre méthode, nous avons pu calculer exactement l'espace d'états de protocoles de communication temporisés (comme le bit alternant sans canaux de communication) ou de processus dynamiques hybrides (comme le "leaking gas burner").

Abstraction et raffinement La vérification automatique de propriétés de sûreté passe en général par le calcul de l'ensemble des états atteignables de l'automate étendu. Cependant, une *approximation supérieure* de l'ensemble des états atteignables *suffit souvent* pour prouver que le système considéré satisfait la propriété (de sûreté) voulue. En effet, l'idée consiste à faire abstraction de certains détails de l'automate étendu qui ne sont pas pertinents vis-à-vis de la propriété à vérifier, tout en conservant une précision suffisante (éventuellement totale) sur les aspects essentiels pour la satisfaction de la propriété. Dans le cas d'un protocole de communication, par exemple, les données transmises peuvent souvent être ignorées si l'on ne

¹¹Finkel, Iyer et Sutre. Well-abstracted transition systems, *Information and Computation*, 181(2003)1-31.

¹²Boigelot, Herbretreau et Jodogne. Hybrid acceleration using real vector automata, *CAV 2003*.

s'intéresse qu'à la vérification de l'absence de blocage. Pour mettre en œuvre cette approche, il faut donc trouver une abstraction du système qui soit (1) suffisamment simple pour pouvoir être vérifiée automatiquement, et (2) suffisamment précise pour satisfaire la propriété voulue.

Afin d'obtenir de façon quasi-automatique une telle abstraction, Graf et Saïdi ont proposé le mécanisme d'*abstraction par prédicats* (Graf et Saïdi, 1997). L'abstraction est ici obtenue automatiquement (en faisant appel à des procédures de décision) à partir d'un ensemble de *prédicats* fournis par l'utilisateur (un état abstrait étant donné par une valeur de vérité pour chacun des prédicats). Ces prédicats portent sur les localités et sur les variables de l'automate étendu, et doivent essayer de dégager les aspects du système qui garantissent que la propriété voulue est satisfaite. Si l'abstraction ainsi construite ne satisfait pas la propriété, alors de nouveaux prédicats doivent être fournis afin de *raffiner* l'abstraction. Plusieurs travaux ont récemment montré comment obtenir automatiquement ces nouveaux prédicats en exploitant les (faux) contre-exemples fournis par l'étape de model-checking. On obtient ainsi une boucle d'analyse abstraction/model-checking/raffinement qui, lorsqu'elle termine, détermine si le système considéré satisfait la propriété voulue. Cette approche a été notamment appliquée à la vérification de circuits industriels (Clarke *et al.*, 2000) et de programmes C (Ball et Rajamani, 2001).

(a) Abstraction par prédicats paresseuse. En collaboration avec l'équipe de Tom Henzinger (Univ. Californie, Berkeley), nous avons optimisé cette boucle d'analyse en permettant notamment le raffinement local et la construction à la volée de l'abstraction¹³. Nous avons implémenté notre technique dans un prototype d'outil, BLAST, qui vérifie des propriétés de sûreté sur des programmes C. Nous avons utilisé cet outil pour vérifier des propriétés de sûreté simples (par exemple l'absence de blocage) sur des pilotes du noyau Linux et du noyau de Windows NT. BLAST donne des résultats encourageants, et nous a permis de vérifier des pilotes de plus de 50 000 lignes de code. De plus, l'approche que nous avons conçue permet, lorsque la vérification se termine avec succès, de générer une preuve formelle (dans un système de déduction proche de la déduction naturelle) à partir de l'arbre d'atteignabilité symbolique calculé durant la phase de vérification. Cette preuve formelle peut alors servir de certificat de correction du programme¹⁴.

(b) Abstraction par prédicats de systèmes de transitions symboliques. Les méthodes d'abstraction et de raffinement ont été principalement développées pour les automates étendus correspondant à des programmes. Ces techniques étant a priori génériques, nous avons cherché à les étendre à des classes plus générales de systèmes. Nous avons montré comment les appliquer à la classe très générale des *systèmes de transitions symboliques* (Henzinger et Majumdar, 2000), ce qui permet de réutiliser les nombreux travaux sur les calculs symboliques. Nous avons retranscrit dans ce cadre les méthodes de raffinement existant dans la littérature, ce qui nous a permis de les comparer et d'obtenir des résultats de terminaison de la boucle d'analyse abstraction/model-checking/raffinement pour certaines classes de systèmes de transitions symboliques. Ce travail a fait l'objet du mémoire de DEA de Stéphanie Gaudan.

(c) Abstraction de réseaux de régulation génétiques. La modélisation et l'analyse de systèmes biologiques est une thématique de recherche en plein essor. Nous nous sommes intéressés à l'analyse comportementale des réseaux de régulation génétique. La sémantique formelle de ces

¹³Henzinger, Jhala, Majumdar et Sutre. *Lazy Abstraction*, *POPL 2002*.

¹⁴Henzinger et al. "Temporal-Safety proofs for systems code", *CAV 2002*, et "Software verification with BLAST", *SPIN 2003*.

systèmes a fait apparaître une nouvelle classe intéressante de systèmes hybrides : les *systèmes d'attracteurs paramétrés*. Nous avons montré¹⁵ comment effectuer des calculs symboliques d'atteignabilité pour cette classe, à l'aide de contraintes linéaires paramétrées. De plus, notre approche permet de construire des abstractions arbitrairement précises pour les systèmes d'attracteurs paramétrés, en synthétisant automatiquement les contraintes pertinentes sur les paramètres.

Décidabilité pour des classes de systèmes

On sait depuis longtemps que le problème de l'atteignabilité, un des problèmes centraux en vérification, peut devenir très vite indécidable dès qu'on sort du cadre des systèmes finis : même des classes très restreintes d'automates étendus (les machines à deux compteurs) ont la puissance des machines de Turing. Mais il existe également de nombreuses classes décidables d'automates étendus infinis, comme par exemple les *réseaux de Petri*, les *automates à pile*, les *automates à files non fiables*, les *automates temporisés*... D'un point de vue théorique on souhaite affiner la frontière de décidabilité pour les classes (restreintes) d'automates étendus, et pour les classes décidables, on cherche naturellement à calculer une représentation symbolique (manipulable) de $post^*$. Nos travaux sur les automates à file FIFO réactifs et sur les automates à compteurs, se situent dans cette problématique.

(a)Automates à file FIFO réactifs. Les automates à file FIFO réactifs (AFR) permettent la modélisation d'applications réactives avec une vue asynchrone du temps : les tâches ont une durée non nulle, ce qui peut nécessiter de mémoriser des occurrences d'événements qui ne peuvent pas être traités immédiatement. Ils consistent donc en un automate fini étendu avec une file de mémorisation dont la politique FIFO (First In First Fireable Out) est particulièrement bien adaptée au contexte réactif. Les AFR possèdent de bonnes propriétés, en particulier, leurs espaces d'états accessibles sont réguliers et effectivement reconnaissables, et le model-checking de LTL est décidable (Sutre *et al.*, 1999). Cependant, les automates à file FIFO réactifs représentent des systèmes ouverts, et certaines propriétés ne peuvent être vérifiées, comme la bornitude de la file de mémorisation qui nécessite la connaissance de l'environnement. Nous nous sommes donc intéressés au modèle clos des *automates à file FIFO réactifs embarqués (AFR embarqués)*¹⁶. Nous avons montré que les modèles ayant 2 événements et un environnement ultimement périodique sont aussi expressifs que les machines de Turing, et que les modèles ayant un seul événement et un environnement régulier correspondent aux automates à pile. Néanmoins, nous avons montré que si la file est non fiable, alors certaines propriétés deviennent décidables, en particulier, l'ensemble des prédecesseurs d'une configuration est effectivement calculable. Ceci fournit donc une semi-procédure¹⁷ de vérification de propriétés de sûreté pour les AFR embarqués basée sur une approximation de l'espace d'états.

(b)Platitude des systèmes à compteurs. Le calcul symbolique accéléré de l'ensemble $post^*$ des états atteignables s'effectue bien souvent par une classique itération de point fixe, dont la convergence est accélérée par le calcul de la fermeture transitive de certains circuits. Cette approche générique peut être appliquée dès qu'on possède une représentation symbolique avec accélération. Ce calcul ne termine pas en général, mais les outils basés sur ces techniques (LASH, TREX et FAST) se comportent très bien en pratique. Afin de mieux

¹⁵Adelaide et Sutre. Parametric analysis and abstraction of genetic regulatory networks, *BioCONCUR 2004*.

¹⁶Herbreteau, Cassez, Finkel, Roux et Sutre. Verification of embedded reactive FIFO Systems, *LATIN 2002*.

¹⁷Au sens où l'approximation faite ne conserve que la satisfaction des propriétés d'accessibilité, mais pas leur insatisfaction.

comprendre les raisons théoriques de ce succès, il est intéressant d'étudier la terminaison du calcul symbolique accéléré. En collaboration avec Jérôme Leroux, nous avons exhibé une condition nécessaire et suffisante de terminaison, la *platitude*, qui repose sur l'existence d'un schéma semi-linéaire de chemins (i.e. d'un SLRE sur l'alphabet composé des transitions de l'automate étendu) permettant d'atteindre tous les états accessibles. De plus, nous avons montré que la plupart des classes connues d'automates à compteurs pour lesquelles *post** est effectivement semi-linéaire sont en réalité plates, ce qui permet ainsi de calculer *post** par un algorithme générique d'itération de point fixe avec accélération. Ces travaux nous ont également permis d'obtenir de nouveaux résultats de semi-linéarité effective pour des classes d'automates à deux compteurs ¹⁸.

Mise en œuvre

(a)Vérification symbolique des systèmes hétérogènes. Les techniques de vérification symbolique actuelles ne traitent, pour la plupart, que d'automates étendus infinis *homogènes*, c'est-à-dire ne comportant qu'un seul domaine de données infini (par exemple uniquement des compteurs entiers). Or une modélisation fidèle d'un système réel passe bien souvent par l'utilisation simultanée de plusieurs domaines de données infinis, comme par exemple dans le cas des protocoles de communication, qui peuvent comporter des canaux de communication, des compteurs et des aspects temps-réel. Afin de mettre en œuvre l'analyse symbolique d'automates étendus *hétérogènes*, on aimerait pouvoir composer les représentations symboliques homogènes et leurs accélérations au sein d'un même outil. Nous avons amorcé cette étude au sein d'un groupe de travail regroupant des membres du LSV (ENS de Cachan) et du LaBRI. Dans le cadre de ce groupe de travail, un modèle générique d'automates étendus hétérogènes a été défini, ainsi qu'une architecture détaillée d'implémentation en Objective Caml ¹⁹. Cette réflexion a ensuite donné naissance au projet *Persée* de l'ACI *Sécurité Informatique*, qui regroupe le LSV (ENS de Cachan), le LIAFA (Univ. Paris 7) et le LaBRI.

(b)Structures de données symboliques pour la vérification de systèmes infinis. La principale limite des structures de données efficaces utilisées en vérification, les BDD et leurs variantes, est qu'elles ne s'appliquent qu'aux systèmes à nombre fini d'états. Dans le cadre de la vérification de systèmes infinis, nous avons conçu la première structure de données, les *automates partagés* ²⁰, pour manipuler les automates finis basés sur les grands principes des BDD : une représentation canonique forte des automates, l'utilisation d'une table d'unicité et d'un cache de calcul. Nos expérimentations montrent le grand bénéfice des automates partagés quand ils sont utilisés pour l'exploration symbolique de l'espace d'états de systèmes infinis.

Projets et perspectives 2005-2008

Le caractère critique des systèmes enfouis tels que les systèmes de supervision et de production, ou les protocoles de communication nécessite de s'assurer de leur bon fonctionnement. La démarche de vérification formelle consiste alors à élaborer un modèle pertinent du système et des propriétés qu'il doit satisfaire, puis à utiliser une technique de model-checking adaptée.

La phase de modélisation du système nécessite d'extraire de la spécification ses caractéristiques utiles pour le but de vérification poursuivi. Il s'agit principalement des données manipulées par les systèmes enfouis aussi bien en interne, comme la mémoire, qu'en externe en provenance de l'environnement, comme le temps (via des horloges) et les informations

¹⁸Leroux et Sutre. On flatness for 2-dimensional Vector addition systems with states, *CONCUR 2004*.

¹⁹Bardin et al. *Intégration des outils PERSÉE*, Délivrable de l'ACI Persée, 2005.

²⁰Couvreur. A BDD-like implementation of an automata package, *CIAA 2004*

analogiques (mesures, commandes, etc.). Cela conduit à plusieurs modèles, notamment *hybrides* pour combiner les comportements discrets du système et les évolutions continues de son environnement et du temps, *dynamiques* pour tenir compte des manipulations en mémoire (par exemple les files d'attente) ou pour représenter les systèmes ambiants, ou encore *stochastiques* pour modéliser par exemple les dégradations subies par un système. Alors que l'industrie des systèmes embarqués s'est engagée dans une approche par composants des matériels et logiciels critiques, il est aussi important que les modèles reflètent cette structure, notamment pour les aspects *concurrents* qu'il est vital de maîtriser pour que la vérification automatique soit efficace. Nous nous intéressons donc logiquement à l'introduction des caractéristiques hybrides et stochastiques au sein du formalisme ALTARICA.

La spécification des propriétés que doit satisfaire le système nécessite là encore des formalismes suffisamment expressifs. D'une part, la modélisation hybride conduit naturellement à vérifier des contraintes temporelles (notamment le respect de délais), mais aussi des *propriétés quantitatives*, par exemple la durée maximale de fonctionnement d'un système pour une quantité d'énergie donnée. D'autre part, l'analyse des modèles stochastiques requiert de pouvoir exprimer des contraintes telles que l'improbabilité d'un événement redouté. Il est donc nécessaire de disposer de *logiques probabilistes* pour l'analyse de fiabilité des systèmes enfouis. Enfin, l'introduction d'informations de concurrence dans les *logiques d'ordre-partiel* est une voie prometteuse pour combattre les limites connues à la vérification automatique.

Plusieurs modèles et logiques permettant la prise en compte de ces caractéristiques des systèmes enfouis existent, et nous avons contribué à certains d'entre eux. Nous souhaitons d'une part enrichir ces modèles et d'autre part nous intéresser aux problèmes liés aux systèmes hétérogènes qui combinent plusieurs de ces propriétés, comme l'hybride et le stochastique. La vérification formelle de ces modèles pose principalement deux défis : (1) la représentation symbolique d'ensembles infinis (d'états ou de comportements) et (2) l'évaluation efficace de la satisfaction des propriétés par le modèle.

Les projets à moyen terme autour d'ALTARICA La collaboration avec le CERT-ONERA (Toulouse) va se poursuivre et même se renforcer, en particulier dans le domaine de l'utilisation concrète d'ALTARICA pour des applications avioniques, essentiellement par le biais de thèses en co-tutelles.

Les sociétés Dassault Aviation et ARboost Technologies vont certainement continuer le développement de leurs outils respectifs. Les seuls impacts pour l'équipe sont des problèmes d'ingénierie afin de conserver les compatibilités entre les divers outils.

Pour le LaBRI, divers travaux de recherche et développement sont prévus dans les trois prochaines années afin de disposer d'une gamme d'outils complète et expérimentale autour du langage ALTARICA.

ALTARICAFINITEAUTOMATA : L'amélioration des performances de MecV en regardant notamment les BED et la réécriture de formules booléennes est nécessaire afin de pouvoir traiter les gros systèmes qu'étudient nos partenaires industriels. Le développement de compilateurs (tels ceux vers Lustre) est également demandé à travers le projet européen ISAAC (successeur de ESACS).

HYBRIDALTARICA : Le travail de définition du langage HYBRIDALTARICA, commencé cette année par Frédéric Herbreteau dans le cadre de l'AS 155, doit aboutir assez rapidement. Le développement d'un outil prototype est prévu pour l'année 2005-2006.

TIMEDALTARICA : Le travail réalisé par Claire Pagetti va être intégré dans le point suivant.

ALTARICA+ types de données infinies : Le projet Persée (LIAFA, LSV, LaBRI) de l'ACI "Sécurité & Informatique" a choisi le langage ALTARICA comme langage commun pour

pouvoir comparer différentes techniques d'accélération pour l'étude de l'accessibilité dans des systèmes infinis bien spécifiques. Ce choix nous oblige à pouvoir très rapidement décrire de tels systèmes dans notre langage. Le choix retenu est la possibilité d'importer des définitions de types abstraits de données. Ce travail va nécessiter la définition de ce format supplémentaire (objectif à 6 mois), puis la réalisation de nombreux outils. Diverses techniques sont envisagées : la généralisation de celle employée par Claire Pagetti pour la compilation vers UPPAAL, mais également l'interprétation abstraite, technique peu utilisée jusqu'alors dans l'équipe, mais utilisée avec un certain succès par d'autres.

l'atelier ALTARICA : Enfin, un très gros travail d'ingénierie logicielle va commencer cette année. Grâce au recrutement d'Aymeric Vincent, nous conservons le contrôle sur MecV, mais l'objectif est de ne disposer à terme que d'une seule *gamme* d'outils. Il faut donc réfléchir à une architecture logicielle qui soit à la fois conviviale et performante (pour les études de cas et pour nos partenaires), et également ouverte et expérimentale pour les besoins de la recherche.

Synthèse de contrôleurs La théorie de la synthèse de contrôleurs continuera à être développée. Un des objectifs visés est de l'adapter pour qu'elle puisse s'appliquer à des systèmes décrits dans le formalisme ALTARICA et que le contrôleur synthétisé puisse superviser le système en utilisant les outils fournis par le formalisme ALTARICA : assertions et synchronisations d'événements.

Calcul symbolique La vérification de la plupart des propriétés peut se ramener à une analyse d'accessibilité. Pour les modèles que nous considérons, le calcul de l'ensemble des états accessibles présente deux défis : (1) la représentation finie d'ensembles infinis d'états, et (2) le calcul en temps fini de l'ensemble infini des états du système.

Convergence du calcul d'accessibilité En général, il n'existe pas d'algorithme pour ce calcul car les modèles infinis ont pour la plupart l'expressivité des machines de Turing. Nous poursuivons deux voies pour y parvenir. La première s'intéresse à la découverte de classes de modèles pour lesquels nous pouvons prouver la terminaison du calcul d'accessibilité. Nous avons introduit la platitude comme une condition de terminaison pour les systèmes à deux compteurs²¹. Notre objectif consiste à définir la notion de platitude pour d'autres classes de systèmes, notamment les réseaux de Petri et les systèmes à file. À plus long terme, notre but est d'extraire des critères génériques de platitude pour les systèmes hétérogènes.

La seconde approche consiste à définir des techniques d'accélération du calcul d'accessibilité qui lui permettent le plus souvent (i.e. expérimentalement) de converger. Nous avons défini une technique qui permet l'analyse de certains systèmes temporisés et hybrides²². Le calcul de l'ensemble des états accessibles des automates temporisés est un défi important. Nous souhaitons améliorer notre technique afin de calculer l'ensemble des états accessibles des automates temporisés de façon efficace.

Représentations symboliques efficaces Les représentations symboliques actuelles ne permettent pas le passage à l'échelle. Généralement, dès qu'un modèle contient plus de 5 ou 6 variables les ressources nécessaires au calcul d'accessibilité sont insuffisantes. Nous avons montré comment les techniques ayant fait leurs preuves sur les BDD pour les systèmes finis peuvent être appliquées au cas des automates pour la représentation efficace des ensembles d'accessibilité de modèles à variables entières non bornées comme les réseaux de Petri²³.

²¹Leroux et Sutre. On flatness for 2-dimensional Vector addition systems with states, *CONCUR 2004*.

²²Boigelot, Herbreteau et Jodogne. Hybrid acceleration using real vector automata, *CAV 2003*.

²³Couvreur. A BDD-like implementation of an automata package, *CIAA 2004*.

Nous souhaitons étendre ces techniques à la représentation de variables réelles. Ceci permettra d'exploiter leur efficacité pour l'analyse de systèmes temporisés ou hybrides.

Lors du calcul de l'ensemble de états accessibles d'un système, on se frotte bien souvent au problème de l'explosion combinatoire de l'ensemble de ses comportements. Des techniques d'ordre-partiel ont été introduites pour représenter symboliquement des ensembles de comportements d'un modèle. Nous souhaitons poursuivre nos contributions sur le dépliage des systèmes de transitions afin de permettre l'analyse de systèmes de grande taille.

Enfin, la vérification des systèmes enfouis lance de nouveaux défis pour la représentation des ensembles d'états. D'un côté, l'environnement analogique des systèmes requiert une modélisation hybride (continue et discrète), et la prise en compte de lois d'évolution des variables de type exponentielle ou trigonométrique. De l'autre, les systèmes ambiants et les programmes à mémoire dynamique correspondent à des modèles reconfigurables. Il est nécessaire dans ce cas de disposer de représentations symboliques permettant la réécriture. Nous souhaitons donc contribuer à l'analyse de ces modèles en inventant des techniques symboliques adaptées (représentations et procédures de calcul).

Abstraction et raffinement Les fondements théoriques de la vérification par abstraction sont connus depuis une dizaine d'années (Bensalem *et al.*, 1992; Clarke *et al.*, 1992; Dams *et al.*, 1997). Cependant, il aura fallu attendre le début du nouveau millénaire pour que des approches entièrement automatiques soient développées et implantées (Clarke *et al.*, 2000; Ball *et al.*, 2001). Les techniques automatiques de vérification par abstraction forment un axe de recherche privilégié pour le passage à l'échelle des méthodes formelles de validation. Nous avons l'intention de poursuivre nos travaux sur ce thème en concentrant nos efforts sur trois aspects principaux.

Dans un premier temps, nous comptons étendre les techniques d'abstraction par prédicats, qui ont principalement été développées pour la vérification de programmes, à des classes plus générales de systèmes, comme par exemple les *systèmes de transitions symboliques*. En particulier, nous souhaitons intégrer les techniques de calcul symbolique et d'accélération au sein d'une méthode d'abstraction par "prédicats" symboliques, ces derniers étant remplacés par des représentations symboliques. Les défis se situent principalement au niveau du raffinement guidé par contre-exemple : l'inférence de nouveaux "prédicats" suffisamment simples pose de nouvelles questions de calcul symbolique. De plus, les techniques d'accélération devraient nous permettre de faire converger plus rapidement la boucle d'analyse abstraction/model-checking/raffinement, en inférant des "prédicats" permettant d'éliminer en une seule étape une infinité de contre-exemples erronés (ceux correspondant au déroulement d'un circuit). L'étude théorique de la terminaison de cette boucle d'analyse est également fondamentale, car elle nous permettra de mieux cerner les limites de cette approche. Ces travaux se situent dans la continuation du mémoire de DEA de Stéphanie Gaudan.

La méthode d'abstraction par prédicats est limitée par le type de modèles abstraits qu'elle permet de générer : ces modèles sont nécessairement finis. Cela ne pose pas de problème pour la vérification de propriétés de sûreté, mais on sait bien que l'abstraction vers des modèles finis n'est pas suffisante pour la vérification de propriétés plus générales. Nous souhaitons développer de nouvelles techniques de vérification automatique par abstraction afin de repousser ces limites et de permettre notamment la vérification de propriétés de terminaison et de vivacité. Les modèles abstraits devront être légèrement étendus, tout en restant naturellement dans une classe décidable. On pourra par exemple les munir de contraintes d'équité, ou de variables sur des domaines infinis (un compteur, une file, ou des horloges, ...). Les verrous scientifiques se situent au niveau des techniques d'abstraction et de raffinement. Une

première approche consiste à étendre la méthode d'abstraction par prédicats en considérant des prédicats paramétrés, le paramètre étant par exemple un entier qui serait traduit par une variable entière dans le modèle abstrait.

Les méthodes de vérification basées sur des calculs symboliques exacts et celles basées sur le raffinement d'abstractions sont complémentaires. Les premières permettent de calculer, par une itération croissante, l'ensemble $post^*$ des états atteignables. Les secondes sont guidées par la propriété à vérifier, et calculent, par une itération décroissante, une approximation supérieure suffisamment précise de $post^*$. Utilisées séparément, ces techniques ne permettent pas, en général, d'analyser des systèmes complexes de grande taille. Aussi, notre troisième objectif consiste à réduire l'écart entre ces méthodes, en proposant notamment des techniques permettant de maîtriser l'imprécision lors des calculs approchés. Nos premiers travaux dans ce sens cherchent à calculer de façon exacte le plus petit point fixe dans un treillis abstrait, ou en d'autres termes le plus petit invariant inductif représentable par un état abstrait. Le degré de précision est ainsi fixé par le choix du treillis abstrait (par exemple les intervalles, les octogones, les octaèdres, ...), et le calcul de plus petit point fixe est réalisé par des techniques de calculs symboliques. Dans cette approche, les techniques de raffinement pourraient s'appuyer sur les travaux existants dans le raffinement de domaines abstraits.

Systèmes probabilistes infinis La modélisation de systèmes complexes fait intervenir à la fois des aspects probabilistes et des aspects infinis. Elle concerne non seulement l'étude de systèmes informatisés ou automatiques mais aussi le domaine en plein essor des modèles du vivant. Développer des techniques et des outils de vérification et d'analyse pour ces modèles est d'un intérêt fondamental.

Les systèmes probabilistes et les systèmes infinis font l'objet de recherches intensives dans le domaine de la vérification, mais peu de travaux traitent des deux aspects. L'objectif de ce projet est la conception de techniques générales pour la vérification de systèmes infinis probabilistes. Dans un premier temps, il s'agira d'identifier des classes de modèles et de propriétés comportementales sur lesquelles le problème de la vérification est décidable. Un point de départ pourrait être l'étude des systèmes infinis "équivalents" à des systèmes finis, comme les systèmes temporisés. Au delà, il faudra s'intéresser aux systèmes ayant des comportements stationnaires ; cette propriété semble un préambule à la mise en œuvre de méthodes de vérification pour les systèmes infinis. Dans un deuxième temps, il s'agira de développer des algorithmes et des outils de vérification, et de les confronter à des problèmes complexes issus du monde de l'informatique. A cette fin, nous chercherons à étendre nos techniques sur le calcul symbolique et les techniques d'abstraction.

3.3.2 Test de systèmes informatisés

Introduction

L'activité de ce thème est centrée sur la problématique du test et la génération automatique de séquences de tests. Cette activité requiert de nouveaux modèles du fait de la complexité croissante des systèmes et des services. L'optimisation de la phase de test, en particulier avec des techniques de génération automatique de tests à partir d'une spécification intéresse un grand nombre d'industriels. L'arrivée de services temps réel avec qualité de service et de réseaux haut débit implique la nécessité de prendre en compte les contraintes temps réel dans toutes les étapes du cycle d'ingénierie. Les avancées scientifiques récentes du thème de recherche au LaBRI concernent principalement une formalisation du test de conformité et d'interopérabilité, ainsi que la réalisation de deux plate-formes pour le test de systèmes temporisés et de systèmes mobiles. La nécessité pour un grand nombre d'entreprises (exemple : aéronautique, spatial, automobile) de disposer de systèmes matériel/logiciel d'une grande sûreté de fonctionnement nécessite d'intégrer le test dans une démarche qualité du cycle d'ingénierie de systèmes complexes.

Mots-Clés Test, Protocoles, Robustesse, Interopérabilité, Conformité.

Description des activités et principaux résultats

Modèles L'équipe de recherche liée à ce domaine utilise depuis longtemps des modèles à base d'automates qui évoluent du fait de la complexité croissante des systèmes et des services à rendre :

- automates d'états finis pour un niveau simple d'abstraction,
- automates étendus incluant des prédicats et des variables donnant une description plus *riche*,
- automates à entrées-sorties utiles pour faciliter la phase de test dans la mesure où les points d'observation pour le test se situent à l'interface,
- automates temporisés incluant la notion d'horloge et un temps continu,
- p-automates incluant des variables et un temps continu.
- pp-automates (p-automates probabilistes).

Résultats théoriques On s'intéresse à plusieurs types de tests : conformité, interopérabilité, robustesse.

Le test de conformité consiste à vérifier qu'une implantation est conforme à sa spécification. La notion de conformité d'une implantation à une spécification est exprimée à l'aide d'une relation de conformité. Différentes relations de conformité ont été étudiées, e.g , bisimulation, équivalence de trace et préordre, et autres. Vérifier une relation de conformité pour une implantation est particulièrement délicat dans la mesure où les comportements considérés sont généralement infinis dû à la présence de cycles dans les spécifications. De nombreux travaux de recherche s'effectuent dans ce domaine au niveau national (Rennes, Reims, Grenoble, Paris) et international (Canada, Espagne, Allemagne, Grande Bretagne, Pays-Bas, Corée du Sud, Japon).

On propose des méthodes de génération automatique de séquences de test, obtenues à partir de la spécification pour une relation de conformité. Ces séquences de test seront appliquées à l'implantation du protocole. Ce problème de génération automatique est particulièrement important pour la communauté industrielle, car le coût de développement de séquences à la

main est important et source d'erreurs. Cependant, il se heurte à un problème d'explosion combinatoire bien connu.

Deux approches sont utilisées à Bordeaux pour tenir compte de ce problème. La première approche consiste à utiliser une technique de génération par objectifs de tests. Ceci permet de n'avoir qu'un déploiement partiel des comportements. Cette technique est basée sur un type de produit synchronisé entre l'objectif de test, exprimé dans un modèle formel, et la spécification formelle. La deuxième approche opte pour une stratégie de génération permettant de couvrir la spécification ou l'implantation d'une certaine façon (couverture de transitions, d'états, de chemins, de variables,...). La couverture des tests exécutés par le testeur est un concept de mesure de la qualité des tests effectués. Les critères de couverture sont exprimés sous forme de *coloriage de graphes*, ce qui permet de traiter une large classe de critères de couverture d'une manière uniforme. Comme résultat de cette nouvelle approche de couverture nous avons la possibilité de définir des métriques pour mesurer la satisfaction d'un critère. Plusieurs mesures ont été introduites pour exprimer des critères de satisfaction différents à l'exemple de la minimalité d'une couverture. La thèse d'Ismaël Berrada s'inscrit dans cette approche.

Le temps constitue un paramètre important dans le test de protocoles (CSMA-CD, CSMA-CA, protocoles multicast comme PIM ou PGM) et de systèmes répartis. Une définition des automates d'états finis temporisés (modèle proche de celui proposé par Alur et Dill avec un temps dense) est utilisée. Cette prise en compte des aspects temporels a été étendue dans le cadre du projet Calife, labellisé par le RNRT (Francetelecom R&D, CRIL ingénierie, LSV-Cachan, LORIA, INRIA-Rocquencourt, LRI, LaBRI). Le modèle de base retenu est le p-automate. La prise en compte de l'aspect temporel dans la génération de test présente une autre source d'explosion combinatoire, dû au caractère continu du temps. Deux approches ont été utilisées dans le cadre de ce projet.

Une approche est menée sur le lien entre preuve et test. Un système pour la preuve, nommé CClair, est développé avec des logiques d'ordre supérieur et donne la possibilité d'extraire d'une trace de preuve des séquences de test sur des algorithmes de conformité tel que ABR et sur des protocoles à contraintes temporelles tel que PGM. La thèse de Davy Rouillard s'inscrit dans cette approche.

L'autre approche considère un calcul symbolique des états accessibles. Elle repose sur deux opérations définies sur les états symboliques. Le calcul des successeurs (opération `succ()`), d'un état symbolique s , considère les états symboliques accessibles à partir de s . Le calcul des prédécesseurs (opération `pred()`) considère les états symboliques qui ont s comme successeur. Ces deux opérations peuvent être facilement implémentées par les structures de données DBM (Difference Bound Matrix). L'intérêt de cette approche est la possibilité du calcul d'une exécution temporelle (un cas de test) ayant une durée accumulée minimale ou maximale pour atteindre un état. Dans des travaux récents, la minimisation des cas de test a été étudiée et les résultats sont prometteurs. La thèse d'Ismaël Berrada s'inscrit dans cette approche.

Ces approches temporelles ont nécessité la définition d'une méthodologie d'objectifs de test et/ou de coloriage de couverture et le développement d'algorithmes de génération de test basés d'une part sur le graphe de régions temporelles (à la Alur et Dill) et d'autre part sur une approche originale de produit synchronisé entre la spécification formelle et un objectif de test. Cette dernière technique permet de limiter le phénomène d'explosion combinatoire. Il s'agit clairement d'un défi scientifique aussi bien dans la liaison entre preuve et test que dans la recherche de la limitation de l'explosion combinatoire.

Le test d'interopérabilité consiste à vérifier que plusieurs entités sont capables d'interopérer. Le fait que les implantations de ces entités soient conformes à une norme n'implique

pas leur interopérabilité. En effet, le test de conformité ne peut être exhaustif, les normes laissent une certaine liberté aux implémenteurs, les architectures et l'environnement ont une influence sur les aspects temporels. Avec le projet RNTL Averroes, l'étude des systèmes à contraintes temporelles est basée sur un nouveau modèle : les pp-automates (probabilistes). Deux axes de recherche ont été considérés. Le premier axe considère la modélisation des communications (unicast, multicast et broadcast) et le partage des données entre entités. L'introduction du modèle des systèmes communicants, basé sur la définition des ressources communes des entités, et la notion de la topologie de communication qui étend les vecteurs de synchronisation introduits par André Arnold et Maurice Nivat en 1988, permet une modélisation réaliste des communications dans les protocoles. Le deuxième axe de recherche consiste à produire des tests d'interopérabilité et des tests de composants. Le problème de l'explosion combinatoire est dans ce cas encore plus important. On propose de générer des tests en utilisant des techniques de produit synchronisé entre les entités communicantes et un objectif de test. La thèse d'Ismaël Berrada, en cours, fournit une base théorique pour cela. Les protocoles étudiés dans cette approche sont CSMA/CD et CSMA/CA. Dans le cadre du projet Platonis labellisé par le RNRT (Francetelecom R&D, Kaptech, INT, LIMOS (Clermont-Ferrand), LaBRI), on propose une méthodologie et une plate-forme de test pour les systèmes mobiles (WAP, GPRS, UMTS) pour le test d'interfonctionnement d'applications et de services (service de localisation et de suivi d'infrastructures réseaux). Cette méthodologie inclut une approche sur la testabilité et l'observabilité dans une architecture distribuée et une étude sur l'interfonctionnement des différents composants de l'architecture.

Le test de robustesse Ce type de test n'a pas été étudié de façon approfondie. Une des grandes difficultés théoriques est de définir ce qui relève de la conformité et ce qui relève de la robustesse. Que ce soit dans le cadre des systèmes critiques ou embarqués, dans les télécommunications ou dans les systèmes d'information, la tendance actuelle est à l'ouverture et l'interconnexion de composants logiciels. Garantir la sécurité et la fiabilité de tels systèmes nécessite de prendre en compte un environnement d'utilisation variable, voire inattendu, du logiciel. Cette exigence de robustesse du logiciel devient donc primordiale. Par ailleurs, les nouvelles méthodes de conception à base de composants réutilisables doivent prendre en compte l'intégration de tels composants dans des contextes variables, en considérant à la fois les aspects fonctionnels et non fonctionnels, et en particulier leur robustesse. L'IEEE (IEEE Std 610.12-1990) définit la robustesse comme "le degré selon lequel un système, ou un composant, peut fonctionner correctement en présence d'entrées invalides ou de conditions environnementales stressantes". Nous proposons d'élargir cette définition de la robustesse à "la capacité d'un système ou composant à fonctionner de manière acceptable en présence de fautes ou de conditions environnementales stressantes". Cette définition élargie prend en compte la possibilité de dégradation du service, par exemple par passage en mode dégradé. Elle élargit également le type de variation des sollicitations de l'environnement en incluant les interactions erronées ou inopportunes, les fautes internes (exceptions, pannes, fautes de transmission), les variations de profils d'utilisation (réutilisation d'un système ou composant, charge variable), les fautes matérielles ou logicielles. Dans le cadre de l'AS 23 du CNRS, nous avons proposé une méthodologie de test permettant de générer des séquences de test vérifiant la robustesse d'un système. Cette génération est basée sur le graphe de refus de la spécification. Celui-ci est augmenté avec des traces de fautes. Seuls les tests relatifs aux aspects de robustesse sont générés. Ces travaux se poursuivent avec la thèse de Farés Saad Korcheff sur le test de robustesse.

Résultats pratiques Pour le test de conformité, la participation de l'équipe au projet RNRT CALIFE a permis de contribuer à la mise en oeuvre d'une plate-forme pour la modélisation, la vérification, la preuve et la génération de tests. Des protocoles de communication à contraintes temporelles ont servi d'exemples. Cette plate-forme est basée sur un bus logiciel permettant d'interconnecter les outils de modélisation, de vérification, de preuve et de génération de test. Les interfaces sont de type XML.

Pour le test d'interopérabilité, la plate-forme de Calife a été étendue. Les modèles se sont aussi élargis avec les pp-automates. En parallèle, la thèse en cours d'Ismaël Berrada a permis de développer un outil TGSE spécifique au test permettant de modéliser les systèmes, les simuler, les émuler et générer automatiquement des tests. Le modèle de base comprend actuellement la prise en compte du temps, la prise en compte de variables et de prédicats. Dans le cadre du projet Platonis du RNRT, une méthodologie et une architecture de test sont proposées, incluant une étude sur la testabilité et l'observabilité de services applicatifs WML, avec prise en compte de la distribution et des contraintes temporelles (entre les terminaux, téléphones, PDA, et les passerelles et serveurs HTTP). Les architectures testées prennent en compte les protocoles WAP, GPRS et UMTS. La thèse de Marcien Mackaya est centrée sur ces méthodologies et la plate-forme Platonis.

Projets et perspectives 2005-2008

Les perspectives de recherche sont inscrites dans le domaine de l'ingénierie des protocoles de communication, des systèmes distribués et des systèmes réactifs.

De nouveaux modèles sont à prendre en compte, en particulier pour les systèmes temps-réel. On s'oriente vers des approches comparatives de vérification alliant des techniques de preuve, de model-checking et de calcul symbolique.

Les techniques de test (test de conformité, d'interopérabilité et de robustesse) et de génération automatique de séquences (avec calcul de couverture) à partir de la description formelle sont à approfondir afin de tenir compte des aspects temps réel, de la distribution, de la mobilité et de l'aspect modulaire des spécifications. L'ensemble de ces recherches aura aussi bien un aspect pratique avec des implantations de méthodes qu'un aspect théorique avec des études formelles.

Une approche formelle du test, à approfondir, consiste à disposer d'une formalisation/modélisation globale de l'ensemble des éléments intervenant dans le test : type de test (conformité, interopérabilité, embarqué, de composantes), approche de test (test actif et passif), architecture de test (boîte noire, blanche, grise), couverture de test (états, transitions,...). L'intérêt de cette approche est la possibilité de réalisation d'outils *ouverts* supportant différents types, approches, architectures, et critères de couverture de test, ce qui correspond à une attente très importante du milieu industriel. Le modèle des systèmes communicants, déjà introduit, permet une telle uniformisation de ces éléments.

Un certain nombre de projets sont encore en cours, tels que Averroes et le projet européen TAROT. On profitera du cadre de ces projets pour développer des recherches sur les tests de composants. L'impulsion donnée par le pôle de compétitivité ASSE doit aussi être une opportunité pour adapter les méthodes utilisées pour le test de systèmes embarqués.

Un dernier aspect encore peu étudié dans ce thème est la mise en oeuvre des tests géné-

rés dans une architecture donnée (un début d'étude avait été entreprise avec la thèse de Patrice Laurençot). La prise en compte des contraintes temporelles est un problème difficile à régler, ce qui, en conséquence, justifie des recherches fondamentales et appliquées en étroite liaison avec des entreprises, par exemple du milieu aéronautique.

3.3.3 Modélisation du vivant

Introduction

Le point de départ est le constat qu'en biologie, comme en informatique, on rencontre de grandes structures complexes, et qu'une partie importante du défi dans les deux domaines concerne leur manipulation intelligente. Le fait qu'on sait maintenant traiter efficacement de gros systèmes en vérification formelle, est en partie dû à l'adoption d'un modèle élégant avec une sémantique parfaitement comprise. La conception d'outils formels pour modéliser les processus biomoléculaires et pour raisonner sur leur dynamique est une voie de recherche novatrice et prometteuse à laquelle les techniques de simulation et vérification formelle peuvent contribuer efficacement.

Concrètement, nous développons des méthodes de modélisation ainsi que des modèles munis de sémantiques adéquates reflétant des propriétés biologiques sous-jacentes aux comportements observables. Actuellement notre réflexion nous mène vers le choix de la modélisation discrète, stochastique et multi-échelles d'une cellule entière.

L'objectif de ce projet consiste à définir des structures permettant une meilleure compréhension de la biologie cellulaire au travers d'un système dont le fonctionnement se rapproche d'une cellule vivante, et qui ultimement permettrait de mener des expériences *in silico*. Une telle approche nécessite le développement de méthodes de modélisation ainsi que de modèles munis de sémantiques adéquates reflétant des propriétés biologiques sous-jacentes aux comportements observables. La nécessité de tels systèmes, s'inscrivant en droite ligne de méthodes prédictives *in silico*, est très vivement ressentie dans le domaine de la biologie.

En effet, jusqu'à présent la biologie moléculaire s'est principalement concentrée sur l'étude de molécules individuelles, leurs propriétés en tant qu'entités à part ou en tant que complexes au sein d'un système de modélisation très simple. Le défi de la recherche actuelle est de comprendre comment ces composants individuels s'intègrent et forment des systèmes complexes biologiques, ainsi que d'étudier l'évolution dans le temps de ces systèmes. En essayant de comprendre la dynamique de cellules entières, nous devons par exemple, pouvoir prédire les conséquences de modifications introduites dans la cellule et/ou dans son environnement telles que, par exemple, le knock-out de gènes ou la modification de métabolites disponibles dans le milieu.

Mots-Clés Bioinformatique, Modélisation multi-échelle stochastique, Comportement cellulaire

Description des activités et principaux résultats

La question clé est comment concevoir un modèle formel d'un système biologique? Quelles caractéristiques formelles? Composition, modularité, niveaux d'abstraction ou possibilité de traiter des informations incomplètes sont nécessaires pour pouvoir aborder des questions d'intérêt biologique.

Une étude préalable a été effectuée quant à l'adéquation de langages de modélisation au monde du vivant. Dans un premier temps nous avons défini en collaboration avec Michel Aigle et Pascal Durrens un modèle du cycle cellulaire de la levure. Dans un deuxième temps nous l'avons confronté à quelques langages de modélisation existants. Le modèle biologique du cycle cellulaire fait appel à certains concepts en modélisation, notamment il nécessite que le langage puisse exprimer la hiérarchie, des phénomènes stochastiques ainsi que des événements discrets. Parmi les langages testés citons le système de Lindenmayer, UPPAAL, SBML et ALTARICA. Cette étude a permis de faire le choix du langage ALTARICADATFLOW, ainsi

que dresser la liste de modifications à y apporter.

Le langage ALTARICA est un formalisme de description de systèmes complexes (informatisés), qui a été conçu au LaBRI en étroite collaboration avec des partenaires industriels. Ce formalisme dispose d'une syntaxe et d'une sémantique clairement spécifiées, et divers outils d'analyse sont désormais disponibles. Le langage est bien adapté à la modélisation de processus biologiques, car :

1. il permet de décrire un système complexe de façon hiérarchique, sous la forme de composants interagissants. Les interactions peuvent être spécifiées par des synchronisations ou par des échanges de données,

2. la variante "Data Flow" du langage permet une simulation efficace car elle ne comporte plus que des affectation de variables, évitant ainsi le problème de résolution de contraintes qui se pose dans le cas général.

Nous nous appuyons donc sur le langage ALTARICADATAFLOW pour la modélisation de processus biologiques. Dans un premier temps, le langage doit être enrichi par les lois de probabilités afin de pouvoir exprimer des phénomènes biologiques intrinsèquement stochastiques. Pour cela nous nous sommes inspirés de l'expérience d'Antoine Rauzy qui a développé une variante stochastique de ALTARICADATAFLOW dans le cadre de systèmes industriels. Ce travail a donc porté principalement sur une définition de la sémantique appropriée au domaine d'application (cellules vivantes), ainsi que sur le choix adapté des lois de probabilités. Ainsi nous avons défini une sémantique propre à ALTARICADATAFLOW en l'enrichissant de transitions à délais (stochastiques) et en précisant la résolution du non-déterminisme par le mécanisme d'arbitres. Un développement de simulateur d'ALTARICADATAFLOW stochastique a été réalisé dans le cadre du mémoire de Master Recherche de Hayssam Soueidan. L'utilisation de ce simulateur en pratique permet de conclure que l'utilisation de transitions stochastiques assure entre autres une robustesse du système face aux variations de paramètres. Cette validation a été effectuée sur le modèle du cycle cellulaire.

Afin d'explorer et de comprendre les modèles biologiques décrits en ALTARICA, on s'intéresse surtout, dans un premier temps, à la simulation. Du point de vue de la simulation, les modèles biologiques complexes (par exemple de cellules entières) présentent un défi de par leur très grande taille. Pour permettre aux biologistes d'expérimenter *in silico* avec le phénomène qu'ils étudient, on développe des techniques permettant de réduire ces modèles tout en conservant leur sémantique, afin d'obtenir une analyse approchée suffisamment précise.

Nous nous sommes plus particulièrement intéressés à l'analyse comportementale des réseaux de régulation génétique ²⁴. La sémantique formelle de ces systèmes a fait apparaître une nouvelle classe intéressante de systèmes hybrides : les *systèmes d'attracteurs paramétrés*. Nous avons montré comment effectuer des calculs symboliques d'atteignabilité pour cette classe, à l'aide de contraintes linéaires paramétrées. De plus, notre approche permet de construire des abstractions arbitrairement précises pour les systèmes d'attracteurs paramétrés, en synthétisant automatiquement les contraintes pertinentes sur les paramètres.

Projets et perspectives 2005-2008

Le langage ALTARICA de description de systèmes, dans sa version stochastique "Data Flow", est bien adapté à la modélisation de processus biologiques. Il permet de décrire un système complexe de façon hiérarchique, sous la forme de composants interagissants. Cependant, ce langage ne permet pas la création et la destruction dynamiques de composants : l'architecture d'un modèle ALTARICA reste "statique". Cette limitation est gênante pour la

²⁴Adelaide et Sutre. Parametric analysis and abstraction of genetic regulatory networks, *BioCONCUR* 2004.

modélisation de certains phénomènes biologiques, comme par exemple la division cellulaire. Il s'agira donc d'étendre le langage ALTARICA stochastique afin de permettre des architectures "dynamiques".

La perspective majeure de cette recherche consiste à développer des techniques d'analyse (simulation et vérification) de modèles ALTARICA stochastiques à architecture "dynamique", en s'inspirant notamment des travaux existant dans le cas non stochastique.

3.4 Publications de l'équipe

3.4.1 Publications d'audience internationale

Revue avec Comité de lecture

- [445] ARNOLD (A). – Nivat's processes and their synchronization. *Theoretical Computer Science*, vol. 281, n1-2, 2002, pp. 31–36.
- [446] ARNOLD (A), VINCENT (A) et WALUKIEWICZ (I). – Games for synthesis of controllers with partial observation. *Theoretical Computer Science*, vol. 303, n1, juin 2003, pp. 7–34.
- [447] BARRIOT (R), POIX (J), GROPPI (A), BARRÉ (A), GOFFARD (N), SHERMAN (D), DUTOUR (I) et DE DARUVAR (A). – New strategy for the representation and the integration of biomolecular knowledge at a cellular scale. *Nucleic Acids Res.*, vol. 32, 2004, pp. 3581–3589.
- [448] BERRADA (I), CASTANET (R) et FÉLIX (P). – From the feasibility analysis to real-time test generation. *Studia Informatica Universalis*, vol. 3, n2, 2004, pp. 7–34.
- [449] DUJON (B), SHERMAN (D), FISCHER (G), DURRENS (P), CASAREGOLA (S), LAFONTAINE (I), DE MONTIGNY (J), MARCK (C), NEUVÉGLISE (C), TALLA (E), GOFFARD (N), AIGLE (L. F. M), ANTHOUARD (V), BABOUR (A), BARBE (V), BARNAY (S), BLANCHIN (S), BECKERICH (J.-M), BEYNE (E), BLEY-KASTEN (C), BOIRAMÉ (A), BOYER (J), CATTOLICO (L), CONFANIOLERI (F), DE DARUVAR (A), DESPONS (L), FABRE (E), FAIRHEAD (C), FERRY-DUMAZET (H), GROPPI (A), HANTRAYE (F), HENNEQUIN (C), JAUNIAUX (N), JOYET (P), KACHOURI (R), KERREST (A), KOSZUL (R), LEMAIRE (M), LESUR (I), MA (L), MULLER (H), NICAUD (J.-M), NIKOLSKI (M), OZTAS (S), OZIER-KALOGEROPOULOS (O), PELLEZZI (S), POTIER (S), RICHARD (G.-F), STRAUB (M.-L), SULEAU (A), SWENNENE (D), TEKAIA (F), WÉSOŁOWSKI-LOUVEL (M), WESTHOF (E), WIRTH (B), ZENIOU-MEYER (M), ZIVANOVIC (I), BOLOTIN-FUKUHARA (M), THIERRY (A), BOUCHIER (C), CAUDRON (B), SCARPELLI (C), GAILLARDIN (C), WEISSENBAACH (J), WINCKER (P) et SOUCIET. (J.-L). – Genome evolution in yeasts. *Nature*, vol. 430, 2004, pp. 35–44.
- [450] DURRENS (P) et SHERMAN (D). – A systematic nomenclature for chromosomal elements in Hemiascomycete yeasts. *Yeast*, vol. 22, n5, April 2005, pp. 337–42.
- [451] FINKEL (A), IYER (S) et SUTRE (G). – Well-abstracted transition systems : Application to FIFO automata. *Information and Computation*, vol. 181, n1, 2003, pp. 1–31.
- [452] HERMJAKOB (H), MONTECCHI-PALAZZI (L), BADER (G), WOJCIK (J), SALWINSKI (L), CEOL (A), MOORE (S), ORCHARD (S), SARKANS (U), VON MERING (C), ROECHERT (B), POUX (S), JUNG (E), MERSCH (H), KERSEY (P), LAPPE (M), LI (Y), ZENG (R), RANA (D), NIKOLSKI (M), HUSI (H), BRUN (C), SHANKER (K), GRANT (S), SANDER (C), BORK (P), ZHU (W), PANDEY (A), BRAZMA (A), JACQ (B), VIDAL (M), SHERMAN (D), LEGRAIN (P), CESARENI (G), XENARIOS (I), EISENBERG (D), STEIPE (B), HOGUE (C) et APWEILER (R). – The HUPO PSI's molecular interaction format—a community standard for the representation of protein interaction data. *Nature Biotechnol.*, vol. 22, n2, Feb. 2004, pp. 177–83.
- [453] HERMJAKOB (H), MONTECCHI-PALAZZI (L), LEWINGTON (C), MUDALI (S), KERRIEN (S), ORCHARD (S), VINGRON (M), ROECHERT (B), ROEPSTORFF

- (P), VALENCIA (A), MARGALIT (H), ARMSTRONG (J), BAIROCH (A), CESARENI (G), SHERMAN (D) et APWEILER (R). – Intact : an open source molecular interaction database. *Nucleic Acids Res.*, vol. 32, Jan. 2004, pp. D452–5.
- [454] HERTOUGH (B. D), TALLA (E), TEKAIA (F), BEYNE (E), SHERMAN (D), DUJON (B), BARET (P. V) et GOFFEAU (A). – Novel transporters from hemiascomycete yeasts. *J Mol Microbiol Biotechnol*, vol. 6, 2003, pp. 19–28.
- [455] IRAGNE (F), NIKOLSKI (M), MATHIEU (B), AUBER (D) et SHERMAN (D). – Proviz : protein interaction visualization and exploration. *Bioinformatics*, vol. 20, 2004. – Advance Access Publication.
- [456] NIKOLSKI (M), BEYNE (E), DURRENS (P) et SHERMAN (D). – Learning rules for predicting homologues in hemiascomycetous yeasts using génolevures manually-curated alignments. *Yeast*, vol. 20, 2003, p. S345. – Bioinformatics Arena Yeast'2003.
- [457] ORCHARD (S), HERMJAKOB (H), JULIAN (R), RUNTE (K), SHERMAN (D), WOJCIK (J), ZHU (W) et APWEILER (R). – Common interchange standards for proteomics data : Public availability of tools and schema. *Proteomics*, vol. 4, 2004, pp. 490–1.
- [458] SHERMAN (D), DURRENS (P), BEYNE (E), NIKOLSKI (M) et FOR THE GÉNOLEVURES CONSORTIUM (J.-L. S). – Génolevures : comparative genomics and molecular evolution of hemiascomycetous yeasts. *Nucleic Acids Res.*, vol. 32 database issue, 2004, pp. D315–D318.

Colloques avec Comité de programme et Actes

- [459] ADELAIDE (M) et ROUX (O). – A class of decidable parametric hybrid systems. *In : Algebraic Methodology and Software Technology 2002*, éd. par Krichner (H) et Ringeissen (C). pp. 132–146. – Springer, september 9-13 2002.
- [460] ADÉLAÏDE (M) et SUTRE (G). – Parametric analysis and abstraction of genetic regulatory networks. *In : Proc. 2nd Workshop on Concurrent Models in Molecular Biology (BioCONCUR'04), London, UK, Aug. 2004*. – Elsevier, 2004.
- [461] ARNOLD (A) et SANTOCANALE (L). – Ambiguous classes in the games mu-calculus hierarchy. *In : Proc. FOSSACS'2003*. pp. 70–86. – Springer, 2003.
- [462] ARNOLD (A) et SANTOCANALE (L). – Ambiguous classes in the games mu-calculus of tree languages. *In : FICS 2003*. – 2003.
- [463] BARKAOUI (K), COUVREUR (J) et KLAI (K). – On the equivalence between liveness and deadlock-freeness in petri nets. *In : Proc. 26th Int. Conf Applications and Theory of Petri Nets 2005, (ICATPN'2005), Miami, USA*. – Springer-Verlag, 2005.
- [464] BERRADA (I), CASTANET (R) et FELIX (P). – A formal approach for real time test generation. *In : Proc. of Workshop on testing real-time and embedded systems (satellite event of FM)*. – September 2003.
- [465] BERRADA (I), CASTANET (R) et FELIX (P). – Tgse : Un outil générique pour le test. *In : Proc. of CFIP 2005*. – Hermès, Mars 2005.
- [466] BERRADA (I), CASTANET (R) et FÉLIX (P). – Generating interoperability test cases from conformance test case generation tools. *In : FORTE'2004, Madrid, Spain (Short Paper)*. – September 2004.
- [467] BERRADA (I), CASTANET (R) et FÉLIX (P). – Testing communicating systems : a model, a methodology, and a tool. *In : Proc. of TESTCOM 2005*. – Springer, Mai 2005.

- [468] BOIGELOT (B), HERBRETEAU (F) et JODOGNE (S). – Hybrid acceleration using real vector automata. *In : Proc. of the 15th International Conference on Computer Aided Verification (CAV'03)*. pp. 193–205. – Springer-Verlag, juillet 2003.
- [469] CASTANET (R), MACKAYA (M), CAVALLI (A), COMBES (P), LAURENÇOT (P), MEDERREG (A), MONIN (W) et ZAIDI (F). – A multi-service and multi-protocol validation platform - experimentation results. *In : Proc. of TESTCOM 2004*. – Springer, March 2004.
- [470] CASTANET (R) et ROUILLARD (D). – Generate certified test cases by combining theorem proving and reachability analysis. *In : Proc. of TESTCOM 2002*. pp. 249–266. – Kluwer Academic Publishers, March 2002.
- [471] CASTANET (R) et ZARKOUNA (K. B). – Test de robustesse. *In : Proc. of SETIT 2003*. – Mars 2003.
- [472] COUVREUR (J). – A bdd-like implementation of an automata package. *In : Proc. 9th Int. Conf. Implementation and Application of Automata (CIAA'2004) (Poster)*. pp. 310–311. – Springer-Verlag, 2004.
- [473] COUVREUR (J), ENCRENAZ (E), PAVIOT-ADET (E), POITRENAUD (D) et WACRENIER (P.-A). – Data decision diagrams for petri net analysis. *In : ICATPN'02, Adelaide, Australia*. pp. 1–101. – Springer Verlag, 2002.
- [474] COUVREUR (J), SAHEB (N) et SUTRE (G). – An optimal automata approach to LTL model checking of probabilistic systems. *In : Proc. 10th Int. Conf. Logic for Programming, Artificial Intelligence, and Reasoning (LPAR'2003), Almaty, Kazakhstan, Sep. 2003*. pp. 361–375. – Springer, 2003.
- [475] DUJON (B) et THE GÉNOLEVURES CONSORTIUM. – Comparative genomics of hemiascomycetous yeasts : the systematic sequencing of *candida glabrata*, *kluveromyces lactis*, *debaromyces hansenii* and *yarrowia lipolytica*. *In : International Conference in Yeast Genetics and Molecular Biology*. – Göteborg, 2003.
- [476] DURRENS (P), BEYNE (E) et SHERMAN (D). – Génolevures : Comparative genomics in the hemiascomycetes yeasts. *In : Colloque Cartographie Génomique Comparée*. – Lyon, December 2001.
- [477] FELIX (P) et CASTANET (R). – Objectifs de test pour systèmes temporisés. *In : Proc. of CFIP 2003*. – Octobre 2003.
- [478] GRIFFAULT (A) et VINCENT (A). – The mec 5 model-checker. *In : CAV : International Conference on Computer Aided Verification*. pp. 488–491. – Springer, juillet 2004.
- [479] HENZINGER (T), JHALA (R), MAJUMDAR (R), NECULA (G), SUTRE (G) et WEIMER (W). – Temporal-safety proofs for systems code. *In : Proc. 14th Int. Conf. Computer Aided Verification (CAV'2002), Copenhagen, Denmark, July 2002*. pp. 526–538. – Springer, 2002.
- [480] HENZINGER (T), JHALA (R), MAJUMDAR (R) et SUTRE (G). – Lazy abstraction. *In : Proc. 29th ACM Symp. Principles of Programming Languages (POPL'2002), Portland, OR, USA, Jan. 2002*. pp. 58–70. – ACM Press, 2002.
- [481] HENZINGER (T), JHALA (R), MAJUMDAR (R) et SUTRE (G). – Software verification with BLAST. *In : Proc. 10th Int. SPIN Workshop (SPIN'2003), Portland, OR, USA, May 2003*. pp. 235–239. – Springer, 2003. Tool paper.
- [482] HERBRETEAU (F), CASSEZ (F), FINKEL (A), ROUX (O) et SUTRE (G). – Verification of embedded reactive fifo systems. *In : Proc. 5th Latin American Symposium*

- on *Theoretical Informatics (LATIN'2002)*, Cancún, México, Apr. 2002. pp. 400–414. – Springer, 2002.
- [483] LEROUX (J) et SUTRE (G). – On flatness for 2-dimensional vector addition systems with states. In : *Proc. 15th Int. Conf. Concurrency Theory (CONCUR'04)*, London, UK, Aug.-Sep. 2004. pp. 402–416. – Springer, 2004.
- [484] MACKAYA (M) et CASTANET (R). – Modeling and testing. location based application in umts networks. In : *Proc. of IEEE ConTEL 2003*. – June 2003.
- [485] MACKAYA (M), CASTANET (R), CAVALLI (A) et AL. – Une plate-forme de validation multi-protocoles et multi-services - résultats d'expérimentation. In : *Proc. of CFIP 2003*. – Octobre 2003.
- [486] MACKAYA (M), CASTANET (R) et KONE (O). – Modeling location based application in umts networks. In : *Proc. of ACM MSWIM 2002*. – September 2002.
- [487] MUSUMBU (K). – Kallah'game : An application of artificial intelligence. In : *Proc. Int. IPSI-2003 Conference, Sveti Stefan, Montenegro, Nov. 2003*. – 2003.
- [488] MUSUMBU (K). – Simulating security systems based on logigrams. In : *Proc. 19th Int. Conf. Logic Programming (ICLP'2003)*, Mumbai, India, Dec. 2003, pp. 498–500. – 2003.
- [489] MUSUMBU (K). – The detection of dynamic dependencies using a generic abstract interpretation algorithm. In : *International Conference on Computer Science, Software Engineering, Information Technologies*. – ISCA, décembre 2004.
- [490] MUSUMBU (K). – Simulator design for security systems. In : *The 2004 International Conference on Modeling, Simulation and Visualization Methods*. pp. 191–197. – CSREA Press, June 2004.
- [491] MUSUMBU (K). – The semantics of kalah game. In : *4th International Symposium on Information and Communication Technologies*. pp. 191–197. – ACM, janvier 2005.
- [492] NIKOLSKI (M), BEYNE (E), DURRENS (P) et SHERMAN (D). – Learning rules for predicting homologues in Hemiascomycetous yeasts using Génolevures manually-curated alignments. In : *International Conference in Yeast Genetics and Molecular Biology*. – Göteborg, 2003.
- [493] PAGETTI (C), CASSEZ (F) et ROUX (O). – Tarc : A timed hierarchical modelling tool. In : *Workshop on Formal Aspects of Component Software FACS'03*, éd. par Van (H. D) et Liu (Z). – Pisa, Italy, september 2003.
- [494] SAAD KHORCHEF (F) et DELORD (X). – Une méthode de test de robustesse adaptée aux protocoles de communication : application au protocole tcp. In : *CFIP'2005, Bordeaux, France (Papier Court)*. – Mars 2005.

Conférences invitées

- [495] BOZZANO (M), VILLAFIORITA (A), ÅKERLUND (O), BIEBER (P), BOUGNOL (C), BÖDE (E), BRETSCHNEIDER (M), CAVALLO (A), CASTEL (C), CIFALDI (M), CIMATTI (A), GRIFFAULT (A), KEHREN (C), LAWRENCE (B), LÜDTKE (A), METGE (S), PAPADOPOULOS (C), PASSARELLO (R), PEIKENKAMP (T), PERSSON (P), SEGUIN (C), TROTTA (L), VALACCA (L) et ZACCO (G). – Esacs : an integrated methodology for design and safety analysis of complex systems. In : *Proc. ESREL'2003*. – Balkema publisher, 2003.
- [496] CASTANET (R) et CAVALLI (A). – Méthodologies de test. application aux réseaux sans fil. In : *CFIP 2002*. – Mai 2002.

Livres d'enseignement et de recherche

[497] CASTANET (R). – *Ingénierie des Protocoles*. – Hermes, 2005.

Chapitres d'ouvrages

[498] BRÉANT (F), COUVREUR (J), GILLIERS (F), KORDON (F), MOUNIER (I), PAVIOT-ADET (E), POITRENAUD (D), REGEP (D) et SUTRE (G). – Modeling and verifying behavioral aspects. In : *Formal Methods for Embedded Distributed Systems - How to master the complexity*, éd. par Kordon (F) et Lemoine (M), chap. 6, pp. 171–211. – Kluwer Academic, 2004.

[499] CASTANET (R), CHAUMETTE (S), MACKAYA (M) et CONSORTIUM (P). – *PLATONIS : a platform for validation and experimentation of multi-protocols and multi-services*, pp. 217–229. – Hermes Penton Science, 2002.

[500] COUVREUR (J) et POITRENAUD (D). – Dépliage pour la vérification de propriétés temporelles. In : *Vérification et mise en oeuvre des réseaux de Petri - Tome 2*, pp. 127–161. – Hermes Science, 2003.

3.4.2 Publications d'audience nationale

Reuves avec Comité de lecture

[501] CASTANET (R), MACKAYA (M), COMBES (P), MONIN (W), CAVALLI (A), MEDERREG (A), ZAÏDI (F) et LAURENÇOT (P). – Une plate-forme de validation multi-protocoles et multi-services - résultats d'expérimentation. *Annales des Télécommunications*, 2005.

Colloques avec Comité de programme et Actes

[502] GRIFFAULT (A). – Conception et validation d'un protocole avec le modèle AltaRica. In : *AFADL : Approches Formelles dans l'Assistance au Développement de Logiciels*, éd. par Jézéquel (J.-M). pp. 293–307. – IRISA, January 2003.

[503] GRIFFAULT (A), POINT (G) et VINCENT (A). – Vérification formelle des modèles AltaRica. In : *LM : Congrès de maîtrise des risques et de sûreté de fonctionnement ($\lambda\mu 14$)*, éd. par Hermès. – octobre 2004.

[504] MACKAYA (M), CASTANET (R) et A. CAVALLI (A. M). – Test de services basés sur la localisation. In : *GRES'2003*. – Février 2003.

[505] PAGETTI (C). – Une extension temporisée d'AltaRica. In : *Actes du Congrès Modélisation des Systèmes Réactifs, MSR'03*, éd. par Hermès. – Metz, France, 2003.

[506] VINCENT (A). – Synthèse de contrôleurs et stratégies gagnantes dans les jeux de parité. In : *MSR'01 : Congrès Modélisation des Systèmes Réactifs*. pp. 87–98. – Hermès Science, octobre 2001.

Conférences invitées

[507] FELIX (P) et CASTANET (R). – Test de systèmes temporisés : De la conformité à l'interopérabilité. In : *Ecole d'Eté Temps Réel du GDR ARP*. – Septembre 2003.

3.4.3 Autres publications

Rapports internes et autres publications

- [508] ADELAIDE (M). – An introduction to the parametric analysis of hybrid systems. *In : MOdelling and VERification of Paralel processes.* – June 19-23 2002.
- [509] AUBER (D), IRAGNE (F), MATHIEU (B), NIKOLSKI (M) et SHERMAN (D). – Proviz : Protein interaction visualization and exploration tool. – Poster ECCB'2003, Paris, 2003.
- [510] BARDIN (S), HERBRETEAU (F), SIGHIRANEU (M), SUTRE (G) et VINCENT (A). – Intégration des outils Persée (proposition d'architecture). – Délivrable 3.1 de l'ACI Persée, 2005.
- [511] BERRADA (I), CASTANET (R) et FÉLIX (P). – *Techniques de génération de séquences de test temporisées : Manuel, intégration et études de cas.* – Rapport technique, LaBRI, 2001.
- [512] BERRADA (I), CASTANET (R) et FÉLIX (P). – *Techniques de test d'interopérabilité pour les systèmes temporisés.* – Rapport technique, LaBRI, 2003.
- [513] BEYNE (E), DURRENS (P) et SHERMAN (D). – Amino-acid substitution matrices tuned for hemiascomycetous yeasts. – Poster ECCB'2003, Paris, 2003.
- [514] GRIFFAULT (A) et VINCENT (A). – Vérification de modèles AltaRica, octobre 2003. MAJECSTIC : Manifestation des jeunes chercheurs STIC.
- [515] LESUR (I), CHIAVERINI (M), NIKOLSKI (M) et SHERMAN (D). – Malako, a system for type-safe comparison of microarray data from multiple sources. – Poster MGED'6, Aix-en-Provence, 2003.
- [516] POINT (G). – *The Synthesis Toolbox - From modal automata to controller synthesis.* – Technical Report n1342-05, LaBRI, janvier 2005.

`altatools`, Simulation et vérification de modèles ALTARICA. Téléchargeable à `altarica.labri.fr`

`MecV`, Vérification de modèles ALTARICA pour une logique du premier ordre étendue par des points fixes relationnels. Téléchargeable à `altarica.labri.fr`

`synthesis`, Implémentation de la méthode de synthèse de contrôleurs proposée par Arnold-Vincent-Walukiewicz. A demander à `point@labri.fr`

3.4.4 Formation par la recherche

Thèses

- [517] GRIVET (S). – *Vérification de systèmes par des techniques de dépliages.* – Thèse de PhD, Université de Bordeaux I, Bordeaux, France, décembre 2004.
- [518] MACKAYA (M). – *Test des protocoles et services liés à la mobilité.* – Thèse de PhD, Université Bordeaux I, Mars 2005.
- [519] ROUILLARD (D). – *Preuve et test de systèmes critiques à contraintes temporelles. De la conformité à l'interopérabilité.* – Thèse de PhD, Université Bordeaux I, Octobre 2002.
- [520] VINCENT (A). – *Conception et réalisation d'un vérificateur de modèles AltaRica.* – Thèse de PhD, LaBRI, Université Bordeaux 1, décembre 2003.

Habilitations

[521] COUVREUR (J.-M). – *Contribution à l'algorithmique de la vérification*. – Thèse de PhD, Université de Bordeaux I, Bordeaux, France, juillet 2004.

[522] KONE (O). – Contribution au test de protocoles de communication, Décembre 2002.

3.5 Collaborations internationales

La collaboration avec l'Université de Varsovie (en particulier sur les thèmes du mu-calcul, des automates d'arbres et des jeux) est fort ancienne et s'est déroulée à certaines époques dans le cadre de collaborations bilatérales institutionnelles (Programme Polonium). Elle a produit de nombreux travaux communs ayant donné lieu à publication et se poursuit activement.

La collaboration avec l'Université de Liège est plus récente et correspond au développement des travaux sur les systèmes infinis et les systèmes hybrides.

3.5.1 Invitations

Herbreteau, Frédéric, *Université de Liège*, juillet 2002–août 2003

Vincent, Aymeric, *Université de Varsovie*, février 2004–août 2004

3.6 Liste des contrats et subventions

3.6.1 Contrats internationaux et européens

ESACS [Enhanced Safety Assessment for Complex Systems]

Période : 2001-2003

Source : Union Européenne

Partenaires : EADS, SAAB, Alenia, CERT, OFFIS, Prover Technology, IRIST

Financement : 130 000 euros

Description :

ESACS est un projet RTD faisant suite à l'appel européen "Growth 2000" sur le thème "New perspectives in Aeronautics". Il a débuté en février 2001 et s'est terminé fin 2003. Les objectifs étaient multiples : d'abord la définition d'une méthodologie pour l'amélioration des analyses de sûreté lors du développement de systèmes complexes ; ensuite la réalisation d'un environnement supportant la méthodologie et les différents outils utilisés par les participants ; enfin des études de cas pour valider l'approche. Les résultats du projet ont été présentés au colloque ESREL'2003. Dans ce projet, nous fournissions les outils ALTARICA et une expertise sur les techniques formelles.

ISAAC [Improvement of Safety Activities on Aeronautical Complex Systems]

Période : 2004-2007

Source : Union Européenne

Partenaires : Dassault, EADS, SAAB, Alenia, CERT, OFFIS, Prover Tech. IRIST

Description :

ISAAC est un projet européen qui fait suite au projet ESACS précédent. Il a débuté début 2004 et sa durée est de 30 mois. Notre rôle est le même, mais le poids du langage ALTARICA y est bien plus important, grâce notamment aux travaux du CERT-ONERA sur les modèles sûrs d'architectures et à l'utilisation de l'atelier ALTARICA développé par Dassault Aviation.

IDEAS [Interoperability Developments for Enterprise Application and Software]

Période : 2002-2003

Source : Union Européenne

Partenaires : LAP (Bordeaux, responsable du projet : G. Doumeingts), LaBRI, VTT (Finlande), Université de Lisbonne, SAP, BAAN, Computas, Formula, Graisoft, IC Fokus, EADS, Fiat, Intracom.

TAROT [Training And Research On Testing]

Période : 2004-2008

Source : Union Européenne (FP6-2002-Mobility-1)

Partenaires : INT (A. Cavalli, responsable du projet), LaBRI, LAMI-Evry, UCM- Madrid, Université de Göttingen, GMD-Fokus-Allemagne, LSR-IMAG-Grenoble, Brunel University-Grande-Bretagne, KUN-University de Nijmegen, BTU-Cottbus, Université de Tomsk.

Financement : 91 000 euros

3.6.2 Contrats avec des entreprises

AIRBUS - ATOS ORIGIN

Période : 2002

Source : Airbus, ATOS-Origin

Financement : 18 000 euros

Description :

Etude et méthodologie sur le test des futurs réseaux et protocoles mis en place dans l'A380.

EUROCONTROL

Période : 2003

Source : Eurocontrol

Financement : 93 000 euros

Description :

Etude préliminaire de modélisation du trafic aérien.

EUROCONTROL

Période : 2004-2006

Source : Eurocontrol

Financement : 180 000 euros

Description :

Etude préliminaire de modélisation du trafic aérien. Optimisation du trafic aérien par utilisation de modèles de théorie des graphes.

3.6.3 Contrats publics avec des instances nationales

Morse

Période : 2003-2006

Source : RNTL

Partenaires : LaBRI, LIP6 (Paris 6), SAGEM et AONIX

Financement : 90 000 euros

Description :

L'objectif est de fournir une méthodologie et des outils prototypes pour le développement d'applications industrielles certifiables dans le domaine des drones.

Persée

Période : 2003-2006

Source : ACI Sécurité Informatique

Partenaires : LaBRI, LSV (ENS Cachan), LIAFA (Paris 7)

Financement : 45 000 euros

Description :

Techniques symboliques pour la vérification automatique des systèmes critiques hétérogènes.

Calife

Période : 1999-2002

Source : RNRT

Partenaires : LRI, l'INRIA (Rocquencourt et Nancy), LSV-Cachan, LaBRI, Alcatel, CNET, CRIL Ingénierie

Financement : 24 000 euros

Description :

Modélisation, preuve et test de protocoles à contraintes temporelles.

Platonis

Période : 2001-2003

Source : RNRT

Partenaires : INT, Francetelecom R&D, LaBRI, LIMOS, Kaptech Ingénierie

Financement : 111 000 euros

Description :

Plate-forme de test pour les applications, les services et les protocoles liés à la mobilité.

Averroes

Période : 2003-2005

Source : RNTL

Partenaires : LRI, LIX, INRIA (Rocquencourt et Nancy), LSV-Cachan, LaBRI, Alcatel, CNET, CRIL Ingénierie

Financement : 114 000 euros

Description :

Modélisation, preuve et test de protocoles à contraintes temporelles et aspects probabilistes.

AS 23

Période : 2002-2003

Source : STIC-CNRS

Partenaires : LAAS, LaBRI, LRI, IRISA, Vérimag

Financement : 7 500 euros

Description :

Techniques avancées de test de systèmes complexes orientées sur le test de robustesse.

AS 111

Période : 2004

Source : STIC-CNRS

Partenaires : INT, LSR-Grenoble, LAMI-Evry, Francetelecom, CEA, LAAS, LRI.

Financement : 1 500 euros

Description :

Test de composants de télécommunication.

AS 155

Période : 2003-2004

Source : STIC-CNRS

Partenaires : IRCCyN-Nantes, LAAS-Toulouse, LAG-Grenoble, LAGEP-Lyon, LAM-Reims, LISA-Angers, INRIA-POPART, Supélec-Rennes.

Description :

Approches formelles pour l'analyse et la synthèse sûre de contrôle des systèmes dynamiques hybrides.

ATIP

Jeunes chercheurs (2005)

3.6.4 Contrats publics avec instances locales ou régionales

3.7 Animation de la recherche

3.7.1 Comités de rédaction

RAIRO-ITA

3.7.2 Comités de programmes de colloques

RTS2001. Real Time Systems, Paris Mars 2001

PROFES 2001. 3rd International conference on product focused software, Kaiserslautern, Germany, September 2001

CFIP 2002. Colloque francophone sur l'Ingénierie des Protocoles. Montréal (Canada), Mai 2002

RTS 2002. 10ième Conférence internationale sur les systèmes temps réel. Paris, Mars 2002

JDIR 2002. Journées Doctorales Informatique et Réseaux. Toulouse, LAAS, Mars 2002

PROFES 2002. 4th International conference on product focused software. Rovaniemi, Finlande, Décembre 2002

OPODIS 2002. 6th International Conference On Principles Of Distributed Systems. Reims, France, Décembre, 2002

CFIP 2003. Colloque francophone sur l'Ingénierie des Protocoles. Paris, Octobre 2003

WTRTES 2003. Workshop On Testing Real-Time and Embedded Systems Satellite Workshop of FM 2003 Symposium, Pisa, Italy - September, 2003

IWITM 2004. 1st International Workshop on Integration of Testing Methodologies, in cooperation with FORTE 2004, Toledo, Spain, September 2004

CFIP 2005. Colloque francophone sur l'Ingénierie des Protocoles. Bordeaux. Mars-Avril 2005.

NOTERE 2005. Nouvelles technologies de la répartition. Gatineau, Canada, Août 2005.

3.7.3 Organisation de colloques, écoles jeunes chercheurs, ...

CFIP 2005. Colloque Francophone de l'Ingénierie des Protocoles. Bordeaux. Mars Avril 2005.

3.7.4 Administration de la Recherche

Comité du RNRT (Réseau National de Recherche en Télécommunication), commission 4 sur le génie logiciel des télécommunications (1997-2002).

Comité d'évaluation du RNTL (Réseau National des Technologies du Logiciel (2000-2005)).

Présidence du comité d'évaluation CNRS du laboratoire PRISM-Versailles. Février 2005.

Comité d'évaluation CNRS du LAAS-Toulouse. Octobre 2004.

Expertise MENRT (habilitations des diplômés en informatique), 2000 - 2002.

Expertises ANVAR.

Membre du CNU (section 27) depuis 1998.

Rapports pour des HdR ou des thèses extérieures : 30 dont 19 rapports.

Image et Son

Responsable(s) Achille Braquelaire

Analyse et indexation vidéo

Responsable(s) Jenny Benois-Pineau

Mots-clés : analyse du mouvement et de la vidéo – segmentation spatio-temporelles de séquences vidéo et de flux multimédia compressés – indexation cross-média – super-résolution et mosaïques.

Modélisation du son et de la musique

Responsable(s) Myriam Desainte-Catherine

Mots-clés : modèles sonores – analyse et synthèse de son – aide à la composition musicale.

Modélisation et synthèse d'images

Responsable(s) Pascal Guitton

Mots-clés : réalité virtuelle – modélisation et visualisation 3D – interaction 3D – terminaux mobiles.

Structuration et analyse d'image

Responsable(s) Achille Braquelaire

Mots-clés : segmentation et analyse d'image – modèles géométriques et topologiques d'images – représentation, mesure et transformation d'objets discrets – imagerie médicale – reconstruction de surfaces.

4.1 Composition de l'équipe

4.1.1 Membres permanents

BENOIS-PINEAU, Jenny, Pr, Bx 1
BEURTON-AIMAR, Marie, MdC, Bx 2
BLANC, Carole, MdC, Bx 2
BRAQUELAIRE, Achille, Pr, Bx 1
DESAINTE-CATHERINE, Myriam, Pr, ENSEIRB
DESBARATS, Pascal, MdC, Bx 1, depuis 09/2003
DOMENGER, Jean-Philippe, MdC, Bx 1
EYROLLES, Georges, MdC, Bx 1
GONZATO, Jean-Christophe, MdC, Bx 1
GRANIER, Xavier, CR, INRIA
GUEORGUIEVA, Stefka, MdC, Bx 1
GUITTON, Pascal, Pr, Bx 1
LACHAUD, Jacques-Olivier, MdC, IUT Bx 1
LE SAEC, Bertrand, Pr, ENSEIRB
MANSENCAL, Boris, IR, Bx 1
MARCHAND, Sylvain, MdC, Bx 1
POINT, Gérard, IR, Bx 1
SCHLICK, Christophe, Pr, Bx 2
STRANDH, Robert, Pr, Bx 1
THOMAS, Gwenola, MdC, Bx 1
VIALARD, Anne, MdC, Bx 1

4.1.2 Membres non permanents

AGUERRE, Cédric, Doct., CIFRE Philips Medical Systems
ALAYRANGUES, Sylvie, Doct., BDI
BOUBEKEUR, Tamy, Doct., MRT
CARMINATI, Lionel, Doct., BDI
CIEUTAT, Jean-Marc, Doct., FONGECIF, jusqu'à 12/2003
DE DIETRICH, Gabriel, Doct., MRT, jusqu'à 05/2003
DE LA RIVIÈRE, Jean-Baptiste, Doct., MRT
DE VIEILLEVILLE, François, Doct., MRT
DELESSE, Jean-François, Doct., AMN
DENIS DE SENNEVILLE, Baudouin, Doct., CIFRE IGT SA
FRECHOT, Jocelyn, Doct.
GARREAU, Ludovic, Doct., Bourse Agglomération Bayonne-Anglet-Biarritz
HACHET, Martin, Doct., MRT, jusqu'à 12/2003
HANNA, Pierre, Doct., MRT, jusqu'à 12/2003
KABONGO, Luis, Doct., CIFRE Philips Medical Systems
KERAUTRET, Bertrand, Doct., MRT, jusqu'à 12/2004
KRAEMER, Petra, Doct., MRT
LAGRANGE, Matthieu, Doct., CIFRE FT, jusqu'à 12/2004
LALES, Charles, Doct., ACI

LAPEYRE, Jacques-Olivier, Doct., MRT
LEVET, Florian, Doct.
LOUIS, Nicolas, Doct., Région
MANERBA, Francesca, Doct., Bourse Gvt Italien
MOUBA, Joan, Doct., Bourse Gvt Gabonais
NESVABDA, Jan, Doct., Philips Research Pays Bas
PARISY, Olivier, Doct., MRT, jusqu'à 07/04
PERETIE, Guilhem, Doct., CIFRE
POUDEROUX, Joachim, Doct., Région
RASPAUD, Martin, Doct., MRT
REUTER, Patrick, Doct., jusqu'à 12/2003
ROBINE, Matthias, Doct., MRT
SCHMITT, Benjamin, Doct., jusqu'à 12/2002
TATON, Benjamin, Doct., BDI CNRS Région, jusqu'à 09/2004
TOBOR, Irek, Doc., jusqu'à 12/2002

PRIMAUX, Laurent, Vac., Contrat

4.2 Présentation de l'équipe

L'image et le son, qui constituent des éléments essentiels pour la perception et la représentation du monde physique, sont des phénomènes continus qui ont été initialement manipulés au moyen de procédés analogiques, selon l'approche *traitement du signal*. Le développement des ordinateurs a rapidement motivé celui de la numérisation des signaux visuels et sonores et de leur représentation au moyen de données discrètes, accessibles au calcul.

C'est dans ce contexte que s'est progressivement développé au sein de l'informatique un champ d'activité centré sur l'image et le son numériques, et pour lequel on peut identifier plusieurs problématiques :

- la *synthèse* dont le but est de produire un signal — en l'occurrence des images ou des sons — à partir de modèles permettant de spécifier ce signal ;
- l'*analyse* dont le but est d'extraire du signal des informations qualitatives ou quantitatives ;
- le *traitement* qui vise à transformer le signal afin d'en corriger certains défauts — par exemple liés au procédé d'acquisition —, de le simplifier — par exemple en amont de l'analyse —, de l'améliorer — par exemple en aval de la synthèse —, ou encore de produire des effets spéciaux ;
- le *codage* qui vise à structurer et/ou compresser le signal enfin d'en faciliter le stockage, la manipulation ou le transfert.

Si l'analyse et la synthèse d'image et leur application à des domaines comme l'imagerie médicale ou les effets spéciaux sont déjà des activités relativement anciennes, le formidable développement des données multimédia a, plus récemment, fait émerger de nouveaux besoins pour l'image, la vidéo et le son numériques.

Ainsi, l'évolution des performances des machines et des procédés de visualisation qui a permis l'essor de la réalité virtuelle, ou à l'opposé la banalisation de terminaux mobiles communicants aux faibles possibilités d'affichage et de calcul, nécessitent la conception de nouvelles techniques de visualisation et d'interaction. Les besoins croissants en matière de production de bandes son intégrant des éléments musicaux, de la parole et des sons, nécessitent de nouveaux modèles et outils de synthèse et de transformation sonores et musicaux, analogues à ceux conçus trente ans plus tôt pour la synthèse d'image et l'animation. Le développement constant des techniques d'acquisition, de stockage et de consultation — et par conséquent la multiplication des données multimédia disponibles — soulève de nouveaux problèmes, d'une part d'analyse de flux vidéo et d'images fixes, pour faciliter l'extraction et l'authentification d'informations pertinentes ou en améliorer le codage, et d'autre part de synthèse, pour pouvoir disposer de méthodes de rendu adaptées à de gros volumes de données. Enfin, l'essor de l'imagerie volumique, notamment en imagerie médicale, nécessite d'étendre en dimension supérieure les modèles de représentation et les procédés de traitement et d'analyse de l'imagerie 2D.

Les activités de l'équipe *Image et Son* se situent au cœur de ces problématiques en pleine évolution. Elles sont organisées en quatre thèmes :

- analyse et indexation vidéo,
- modélisation du son et de la musique,
- modélisation et synthèse d'images,
- structuration et analyse d'image.

Le thème **Analyse et indexation vidéo** s'inscrit dans la problématique de l'informatique multimédia suivant trois grands axes appartenant à la thématique générale de l'analyse des flux multimédia. Le premier axe concerne la segmentation basée mouvement et luminance/couleur des images animées dans l'objectif d'extraction et de caractérisation dy-

namique des objets d'intérêt afin d'aider à l'interprétation et au codage efficace de la vidéo. Le deuxième axe porte sur la macro-segmentation des documents vidéo en plans et scènes de montage dans l'objectif double de leur indexation et de la navigation. Enfin le dernier point traite de la segmentation cross-média, c'est à dire l'analyse conjointe des flux vidéo et audio dans des contenus audio-visuels numériques. Nous nous intéressons aux outils méthodologiques au niveau pixel et au niveau objet et aux modes d'interaction de ces outils avec les méthodes classiques de segmentation des images en couleur, ceci dans le but d'aboutir à l'extraction des objets d'intérêt dans des séquences vidéo.

Le thème **Modélisation du son et de la musique** concerne la modélisation, l'analyse et la synthèse de son et de musique. La plupart des modèles sonores sont actuellement des modèles hybrides, où la composante bruitée (stochastique) est séparée de la composante sinusoïdale (déterministe). Nous souhaitons définir un modèle spectral hiérarchique unifié, plus souple permettant une représentation des sons à plusieurs niveaux, c'est-à-dire le niveau du son, mais aussi le niveau du contrôle du son, et le niveau des variations du contrôle. Cela permettrait d'agir à tous les niveaux sur le son, du microscopique au macroscopique. Au niveau de la musique, nous cherchons à modéliser diverses formes d'interactions avec des partitions de musique interactive qui pourraient ainsi être spécifiées au moment de la composition du matériau musical. Les résultats de nos recherches trouvent des applications dans la création musicale au sein du SCRIME (Studio de Création et de Recherche en Informatique et Musique Électroacoustique) et dans la pédagogie musicale.

Le thème **Modélisation et synthèse d'images** s'inscrit dans le domaine plus général de la réalité virtuelle (RV) et se consacre aux différentes étapes allant de la création des environnements virtuels (EV) jusqu'à leur visualisation et à leur manipulation. Concrètement, nous travaillons sur des méthodes automatiques de création d'EV à partir de plans numérisés et de photographies, sur différentes méthodes de rendu (par points, non photo-réalistes) et sur l'interaction 3D entre l'utilisateur et l'application. Nos activités ont débuté dans un contexte matériel standard de type station de travail. Puis, nous nous sommes intéressés au cas des systèmes à base de grands écrans (plusieurs dizaines de m²) grâce à l'installation de la salle de RV Hémicyclia. Enfin, depuis deux ans, nous avons abordé un nouveau type d'équipement : les terminaux mobiles communicants (téléphone, PDA) au sein du projet IPARLA (LaBRI - INRIA Futurs). Dans chacun de ces cas, nous essayons de séparer ce qui est spécifique au matériel de ce qui est générique.

Le thème **Structuration et analyse d'image** s'intéresse aux modèles permettant de décrire la structure géométrique et topologique des images discrètes, en dimension deux ou supérieure, et aux applications de ces modèles en analyse d'image. Plus précisément, l'analyse d'image s'appuie sur une étape de segmentation consistant à décomposer une image en différents objets ou composants et sur une étape d'extraction d'informations relatives à ces objets : relations topologiques, mesures et caractéristiques géométriques. Que ce soit pour l'étape de segmentation ou celle de l'analyse de ses résultats, il est nécessaire de disposer de modèles topologiques et/ou géométriques adaptés aux images discrètes pour représenter ces composants images. Nos travaux dans ce contexte portent principalement sur la définition de nouveaux modèles dans le cadre de la topologie et de la géométrie discrètes, sur la conception d'outils et de techniques pour mesurer, reconstruire et lisser des objets discrets, sur le développement de techniques de segmentation basées frontières et basées régions, et sur l'implémentation de logiciels et de bibliothèques mettant en œuvre ces modèles, algorithmes et outils.

Conclusion. Les activités de l'équipe couvrent un large spectre en matière d'image et de son numérique, avec une production scientifique portant sur des problèmes de synthèse, de

traitement, d'analyse, de codage et d'interaction.

La période 2001-2004 a été pour l'équipe celle d'une forte croissance thématique, avec la création du thème « Analyse et indexation vidéo », la montée en puissance du thème « Modélisation du son et de la musique » notamment à travers le SCRIME (Studio de Création et de Recherche en Informatique et en Musique Électro-acoustique), mais également l'extension des domaines de recherche plus « historiques » de l'équipe, avec le démarrage de nouvelles thématiques telles que la réalité virtuelle avec la création d'Hémicyclia (Centre de Compétences en Réalité Virtuelle), le traitement automatique de données issues de systèmes d'acquisition 2D ou 3D, les interfaces hommes-machines, les modèles déformables, imagerie médicale, etc.

Les échanges entre les différentes thématiques de l'équipe sont nombreux et fructueux, et permettent d'aborder des problèmes concrets en bénéficiant d'un large spectre de compétences et de savoir-faire. Cette mise en commun de modèles mathématiques et de techniques algorithmiques dans des domaines de recherche historiquement disjoints constitue l'une des originalités des travaux de l'équipe « Image et Son ». On peut également signaler des collaborations inter-équipe autour de la visualisation d'informations, et de collaborations transdisciplinaire notamment dans le cadre des activités d'interaction homme-machine et d'imagerie médicale.

4.3 Thèmes de recherche de l'équipe

4.3.1 Analyse et Indexation Vidéo

Introduction

Le thème de recherche *Analyse et Indexation Vidéo* (AIV), qui a été initié en septembre 2001, s'inscrit tout naturellement dans l'ensemble des thématiques scientifiques de l'équipe, en ce sens qu'il fait intervenir des notions fondamentales de l'image numérique auxquelles est ajoutée une nouvelle dimension, celle de l'analyse du mouvement. L'activité de recherche de ce thème réside dans l'analyse et l'indexation des contenus vidéos de différentes sources, qu'il s'agisse de journaux télévisés, de documentaires, de vidéos de télésurveillance, etc. L'analyse et l'indexation de la vidéo peuvent s'effectuer à plusieurs niveaux de granularité : soit directement sur les pixels des images successives, soit dans l'espace des descripteurs préalablement extraits de la vidéo. Cette extraction requiert le développement d'outils d'analyse et de caractérisation du mouvement global (celui de la caméra), de segmentation des scènes animées en termes d'objets en mouvement, de détection des ruptures dans le flux temporel correspondant aux frontières des plans de montage, etc... Les travaux entrepris en terme d'analyse vidéo portent également sur la possibilité, à partir d'un flux vidéo de qualité médiocre, d'enclencher des outils d'augmentation de résolution afin d'extraire une image globale de la scène vidéo.

Les méthodes d'analyse s'appuient donc sur les outils méthodologiques classiques tels que la segmentation morphologique des images couleur, les méthodes différentielles d'estimation de flot optique, mais aussi font intervenir les méthodes de décision statistique dans l'espace image ou l'espace des descripteurs. La première cible des méthodes d'analyse est l'indexation vidéo. La quantité croissante des contenus multimédias au format compressé pose des problèmes évidents de gestion des documents, par exemple comment retrouver une séquence vidéo particulière à des fins diverses de réutilisation ou de retransmission. Le projet de recherche AIV s'intéresse aux méthodes de structuration des documents vidéos et exploite les différents niveaux de sa hiérarchie : plan de montage, scène visuelle, objets sémantiques. La caractérisation par la couleur, la forme et le mouvement enrichit la description du document vidéo par de méta données, favorisant ainsi de le retrouver plus aisément.

Par ailleurs, le deuxième objectif de l'analyse vidéo est le « *codage intelligent* » qui s'appuie sur la connaissance d'objets d'intérêt. Ici, on propose des outils pour la segmentation spatio-temporelle adaptée.

Afin de répondre au concept de convergence des média, deux autres thèmes de recherche ont également été initiés en collaboration avec le thème *Modélisation du son et de la musique*, autour de la *cross-modalité*, et avec le projet *Visualisation de l'information* autour de la visualisation de documents vidéo. L'analyse des contenus s'effectue dès lors non seulement sur l'étude du mouvement ou de mesure sur l'image vidéo mais sur la combinaison regroupement scènes visuelles-vidéo et son-vidéo respectivement. Ces deux thématiques font l'objet de projets transversaux au sein même du LaBRI.

Mots-Clés

Analyse du mouvement. Segmentation spatio-temporelle basée-objet. Analyse de la vidéo par apprentissage statistique. Segmentation des séquences vidéo en plans et scènes. Segmentation des flux multimédia compressés. Descripteurs bas-niveau pour l'indexation vidéo. Indexation cross-média. Super-résolution et mosaïques.

Description des activités et principaux résultats

Estimation du mouvement et segmentation spatio-temporelle adaptée

Le mouvement apparent dans le plan image exprimé sous forme de flot optique est un descripteur fondamental qui permet de déduire les caractéristiques du travail de la caméra, et d'extraire et de suivre des objets d'intérêt dans des séquences vidéo. Nous nous sommes intéressés au problème fondamental d'estimation du flot optique par des méthodes différentielles du premier ordre. Nous appuyant sur les approches numériques classiques, nous avons proposé des estimateurs multi-échelle par gradient conjugué et accéléré. L'intégration du modèle pseudo-continu des régions discrètes, *les chemins euclidiens*, développé antérieurement au sein de l'équipe, a permis d'améliorer l'estimation sur des frontières d'objets segmentés dans la vidéo.

De nombreuses normes de codage vidéo répondent à la problématique du codage à bas débit, et sont particulièrement conçues pour assurer une très bonne qualité visuelle malgré le faible coût de codage de l'information, mais à l'heure actuelle l'algorithme de codage normalisé H.264 fournit le meilleur compromis entre qualité de l'image et débit. Cette norme utilise le découpage de l'image à chaque instant de temps en blocs de taille variable et propose un codage optimal au sens débit/distorsion pour toutes les zones de 16x16 éléments dans l'image. Ainsi une telle zone nommée *macro-bloc* peut être encodée par des petits blocs de 4x4 pixels, de plus grands blocs de 8x8 pixels, ou diverses combinaisons de ces derniers. L'image courante dans la zone peut être encodée par rapport à l'image précédente (codage *inter-image*) en exploitant la continuité de la zone au cours du temps ou encore elle peut être encodé en mode *intra-image* sans référence aux autres images de la séquence. Ces calculs du mode optimal de codage sont très lourds et bloquent l'utilisation de la nouvelle norme dans son état actuel pour l'encodage et la transmission en temps réel.

Plutôt que d'effectuer des calculs lourds du mode optimal du codage, le travail que nous avons entrepris propose d'analyser l'image en mouvement et de ne coder que les zones qui appartiennent aux objets en mouvement. Dans les contextes précis de la vidéo surveillance et de la visioconférence, la suite d'images constituant la vidéo est très semblable, et ce malgré le faible nombre d'images à la seconde requis (3 à 6 images par seconde dans notre cas). Ce sont généralement des situations de plans fixes et les différences entre deux images consécutives se font sur des objets mobiles ou sur des changements de luminosité. Le but de ce projet vise à exploiter cette particularité dans le cadre de la détection de mouvement (pour la surveillance) et de la compression vidéo en temps réel (pour le stockage et la diffusion des images). Un des principes fondamentaux de la compression vidéo est de ne coder que les différences entre images successives. Dans le cadre de la télésurveillance, il se pose le problème de « *codage intelligent* », c'est-à-dire qu'il ne s'agit pas de coder simplement un changement de la scène, mais d'être sûr que ce changement correspond bien à l'évolution temporelle d'un objet d'intérêt. Ainsi la détection des objets peut être vue comme un ensemble de deux tâches : la détection du mouvement et la préservation de l'intégrité de l'objet.

Une attention toute particulière a été accordée à la conservation de l'intégrité des objets en mouvement. Un module de détection de mouvement basé sur une analyse statistique des différences entre images successives, permet l'extraction des objets en mouvement de la scène. Un module topologique a été développé pour l'extraction des objets à proprement parler. Tous ces outils opèrent dans le cadre du codage basé blocs et ont été intégré au codec H.264 devenant ainsi optimisé.

Résultats et Logiciels – Le logiciel de référence qui implémente la norme H.264 opère un codage à seulement 0.3 images/sec. Grâce aux outils d'analyse nous avons obtenus un enco-

dage à 3 images/sec, soit une diminution de 90% du temps (en conditions de télésurveillance : plans fixes avec objets mobiles éloignés de la caméra), et à qualité visuelle et objective équivalente. Sur des plans de type visioconférence (personne en gros plan) les performances sont légèrement inférieures puisque l'on atteint en moyenne 2 images/sec. Le débit actuellement atteint sur les images *inter-codées* est équivalent à celui du *Reference Codec Software H.264* : en moyenne moins de 7 Kbits/sec (maximum : 15 Kbits/sec) sur des plans de type télésurveillance, et en moyenne moins de 21 Kbits/sec (maximum : 55 Kbits/sec) sur des plans de type visioconférence.

Apprentissage statistique pour l'analyse vidéo

Dans le contexte de l'analyse des vidéos de surveillance, on s'intéresse à deux problèmes fondamentaux : la détection des objets en mouvement et la reconnaissance et le suivi des objets particuliers dans les séquences vidéos. Ces objets peuvent tout aussi bien être un piéton, une voiture, un visage humain... L'approche retenue implique donc d'être le plus général possible et de ne pas fournir un modèle simpliste qui se limiterait qu'à une seule catégorie d'objets.

Le premier problème que nous avons été amené à résoudre était de localiser la zone de mouvement et à la rendre la plus nette et lisse possible au sens spatio-temporel et ce afin de servir à la recherche d'un ou plusieurs objets d'intérêt. Le problème de détection des zones de mouvement dans la vidéo a été largement étudié dans la littérature mais récemment la modélisation à base de mélange de lois Gaussiennes est devenue très populaire grâce à certains travaux fondamentaux de Grimson et Stauffer notamment. Supposant que l'objet connu est animé d'un mouvement propre différent de celui de la caméra -supposée statique dans notre cas-, ces mêmes travaux montrent de bonnes performances pour des contraintes d'éclairage variables au cours du temps. Considérant donc le cadre du modèle par mélange de lois Gaussiennes, nous avons proposé une méthode originale de détection de mouvement découpé en trois étapes : l'initialisation du mélange Gaussien, la classification de chaque nouveau pixel et la mise à jour du modèle en fonction de la classification. Une régularisation markovienne est également proposée afin d'améliorer le taux de détection tout en ajoutant une cohérence à la fois spatiale et temporelle. De plus contrairement aux travaux précédents, où n'apparaît qu'un simple seuillage, une règle de décision basée sur le maximum de vraisemblance nous permet de déterminer la meilleure distribution Gaussienne parmi le mélange.

Pour ce qui concerne la détection et le suivi d'objet d'intérêt, visages humains dans notre cas, nous avons retenu une autre méthode d'apprentissage : les Support Vector Machine. Le problème de détection de visages est un problème classique de reconnaissance de formes qui peut s'énoncer de la façon suivante : considérant une image d'entrée quelconque, qu'elle soit statique ou extraite d'un flux vidéo, il s'agit de déterminer la présence ou non d'un ou plusieurs visages humains dans cette image. Le travail que nous avons entrepris à pour point de départ les travaux d'Osuna dans lesquels les auteurs proposent un détecteur original à base de machine à vecteurs de support pour classifier les visages en deux classes : visages et non visages. Le but de notre travail concerne donc la possibilité de simplifier au maximum le détecteur -Pré-traitement, fonctions noyaux, nombres de Support Vector- afin de permettre son intégration dans un système temps-réel sans, bien évidemment, perte d'efficacité notoire par rapport aux travaux entrepris par Osuna. De plus nous avons proposé une approche par pyramide multi-résolution de l'image et par propagation des détection entre niveaux de la pyramide multirésolution ce qui permet la localisation des visages dans des prises de vue à différentes échelles et, qui plus est, reste robuste par rapport au faible rotation du visage à son axe de symétrie.

Logiciel – Nous avons développé un ensemble d'outils nécessaires à la télésurveillance per-

mettant de gérer de bout en bout le processus de surveillance de l'acquisition du flux, à sa sortie en fichiers ou à l'écran en passant par une boîte à outil qui permet d'effectuer un ensemble de traitement. Le logiciel en lui-même est découpé en trois catégories de *plug-ins* qui peuvent être intégrés ou pas dans un programme de télésurveillance que nous avons également développé : la première catégorie est constituée de l'ensemble des plug-ins d'entrée : acquisition depuis une caméra vidéo pilotable ou non, fichiers images ou vidéo... La deuxième catégorie de plugin permet des traitements intelligents comme la détection de mouvement, la détection de visage temps réel, la reconnaissance, la détection de piétons. Enfin la dernière catégorie gère l'ensemble des sorties possibles, fichiers image (JPG, PNM, PPM...), vidéo (MPG, MJPEG), écran (Xvideo, SDL, X11...). Le logiciel s'installe et se configure automatiquement sur une distribution linux standard et permet la plupart des opérations en temps réel sur des ordinateurs banalisés de milieu de gamme.

Paradigme dit de « rough indexing »

Dans l'ensemble de nos projets d'indexation, nous nous intéressons aux méthodes de structuration des documents vidéos. Ainsi, nos axes principaux de recherche sont l'identification des plans de montage et des scènes ainsi que l'extraction des objets sémantiques. Du fait de la quantité croissante des contenus multimédias sous forme compressée, notre attention porte sur l'indexation de flux vidéos codés (MPEG 1,2,4 et H.26x). Dans l'ensemble de ces travaux, nous développons de nouveaux outils dans le contexte du paradigme de « *rough indexing* ». Cela signifie que nous traitons des données bruitées à résolution dégradée dans le but d'une compréhension globale du contenu. Ces travaux ont un lien direct avec le nouveau standard de description de contenu multimédia MPEG 7. L'élément commun à tous nos travaux d'indexation est la compréhension et la représentation par un modèle affine du mouvement de la caméra. Un tel modèle nous permet de tester la continuité du mouvement, de classifier le type de mouvement de caméra mais aussi de localiser les zones de mouvement qui ne correspondent pas à celui de la caméra.

Une vidéo au format MPEG est constituée de 3 types d'images : les images I, P et B. Les images I sont codées en *spatiales* ce qui signifie que l'on utilise pas de référence aux autres images pour effectuer le codage. Les images P sont *prédites* par rapport à l'image I ou P précédente, cette prédiction se fait par la compensation du mouvement entre les deux images. Enfin, les images B sont interpolées entre les deux images P (ou I et P) les plus proches. Les images se suivent selon la structure suivante : I B(B) P B(B) P...P B(B) I... avec une image I toute les 12 (ou 15) images. Pour les images P le mouvement est caractérisé par des vecteurs associés à chaque *macro-bloc/bloc* (16x16 pixels) de l'image. Le champ vectoriel ainsi constitué fournit à la fois les informations de mouvement des objets, en avant plan, mais aussi celles du fond visible sur l'image. Il est alors indispensable de savoir distinguer ces deux informations. Pour cela nous cherchons à extraire le mouvement global de chaque image P, généralement celui de la caméra. Ceci nous permettra à la fois de détecter la non-continuité de ce mouvement entre certaines images (changement de plan), et d'extraire les objets en avant plan dans les scènes. Les outils méthodologiques mis en oeuvre sont la modélisation du mouvement global par le modèle affine (de 3 à 6 paramètres) et l'estimation robuste des paramètres de ces modèles. Des techniques de multi-résolution adaptées permettent une estimation adéquate avec le rejet des mesures aberrantes sur les champs de déplacements de macro-blocs issus des flux binaires MPEG, et qui sont généralement très bruités.

Les variations temporelles des paramètres du mouvement global estimé combinées avec le facteur de qualité d'estimation et avec la quantité des blocs codés en mode intra-image permettent de détecter les changements des plans de montage. Pour cela nous avons mis en place un détecteur statistique basé sur la modélisation d'une fonction de ces mesures

par des processus stochastiques suivant les lois normales. Cette statistique est analysée à la volée afin d'en extraire les pics correspondants à des changements de plans. Pour les changements de plan intervenant sur images I un outil spécifique de mesure de similarité a été développé. Ce dernier récolte et interprète toutes les informations de mouvement entre 2 images I successives afin de pouvoir les mettre en correspondance. Cette comparaison se fait alors uniquement sur les zones de l'image I courante qui ont parfaitement suivi le mouvement de caméra depuis la précédente image I.

Ces mêmes outils d'estimation des modèles de mouvements affines à partir des flots optiques bruités sont utilisés dans le contexte de création de nouveaux descripteurs spatio-temporels des plans vidéo, à savoir des mosaïque 1D. Ces travaux s'appuient sur notre recherche antérieure quand à l'estimation du mouvement dans le domaine des transformées de Radon discrète (IRCCyN). Ils ont abouti à une nouvelle approche concernant la classification des plans de montage en hyper-scènes élaborée avec l'IRISA/INRIA de Rennes (équipe TEMICS).

Dans l'ensemble global des tâches d'indexation des contenus vidéos, l'indexation des objets sémantiques est d'une importance primordiale car il est naturel de supposer que la plupart des informations sont contenues dans les objets en avant-plan. L'objectif de cette étude consiste donc à extraire des objets sémantiques dans le cadre du paradigme de *rough indexing*. Ainsi nous utilisons les informations de mouvement associées aux informations de couleur extraites des flux compressés. Nous savons qu'un objet en avant-plan ne suit généralement pas le mouvement de la caméra, alors l'observation des vecteurs de mouvement MPEG2 nous permet déjà d'isoler les régions qui ont des paramètres de mouvement différents de la caméra. Les régions obtenues avec les informations de mouvement sont très bruitées et ne recouvrent pas parfaitement les objets. Pour affiner les régions obtenues nous utilisons les informations de couleur de l'objet, cela nous permet de déterminer si la zone appartient effectivement à l'objet. Sinon il s'agit du fond (en arrière-plan) qui est découvert par l'objet qui se déplace. Ainsi, grâce à cet outil nous disposons du nombre d'objets et de leur forme, nous pouvons ensuite les caractériser par leurs couleurs et leurs mouvements, ce qui apporte un haut niveau de description d'un document vidéo.

Résultats – Nous avons soumis notre détecteur de changements de plans à la campagne internationale TREC VIDEO 2004, qui en plus de la détection des transitions requiert la classification des effets de transition. Nous avons obtenus un rappel et une précision de 72%, ce qui nous classe en neuvième position sur 19 participants à cette tâche. Sans tenir compte de la classification des effets de transition nous obtenons un rappel et une précision de 83%. Le temps de traitement est par trois fois inférieur au temps réel de lecture vidéo.

Logiciel – Dans ce projet nous utilisons un décodeur MPEG1,2 de référence modifié, qui communique directement avec notre outil d'estimation du mouvement global de caméra (GlobalMotionEstimator). Les données ainsi calculées sont alors utilisées soit pour la détection des changements de plans (Indexer), soit pour l'extraction des objets en avant-plan (ForegroundObjectExtraction), soit enfin pour la classification du mouvement de caméra (CameraMotionClassification).

Super-résolution dans le contexte rough indexing

Intuitivement, la méthode la plus simple pour augmenter la résolution d'une image consiste à interpoler -de façon bi-linéaire ou bicubique- tous les pixels d'une image afin de construire une nouvelle image de résolution plus grande. Malheureusement les images numériques souffrent de dégradations dues à l'effet d'aliasing, du bruit, et une simple interpolation ne fournira pas de résultat satisfaisant. Le but des méthodes de super-résolution repose donc sur

un mécanisme de sur-échantillonnage d'une image, c'est-à-dire une augmentation de résolution, accompagné d'une restauration de l'image. Les méthodes de super-résolution utilisent l'information d'une séquence d'images pour sur-échantillonner une image de cette séquence (l'image de référence). Ceci demande le recalage de tous les images de la séquence dans le repère de l'image de référence afin de superposer toutes les images de la séquence. Pour ça, le mouvement de chaque image de la séquence par rapport à l'image de référence doit être estimé et compensé. Afin de restaurer l'image de super-résolution, l'ensemble des paramètres de dégradation doit être déterminé.

Dans le cadre de ce travail, nous proposons un algorithme de super-résolution, basé sur des rétro-projections itératives (Irani, Peleg), pour les mosaïques spatio-temporelles (image panoramique d'une séquence d'images). La nouveauté est que ces mosaïques sont construites à partir des images DC (un pixel dans l'image DC correspond à la moyenne d'un bloc 8×8 dans l'image originale) qui sont extraites directement du flux compressé MPEG. Les images DC sont caractérisées par une résolution très faible et une forte dégradation due à l'effet d'aliasing et de flou.

Afin d'estimer le mouvement entre chaque image de la séquence par rapport à l'image de référence, on estime les champs vectoriels de déplacement à partir du flux compressé. L'estimateur robuste proposé calcule à partir des champs vectoriels le modèle affine à 6 paramètres du mouvement. Cependant cette information de mouvement n'est pas disponible dans tous les types d'images dans le flux compressé. Nous avons donc proposé une méthode d'extrapolation du mouvement dans le cas où l'information serait manquante. Malheureusement lors du mécanisme de super-résolution, si on superpose toutes les images de la séquence sans proposer de méthode de restauration, l'image de super-résolution sera de moins bonne qualité que les images originales du fait qu'elle amplifie les différentes dégradations. Par conséquent, une méthode de restauration couplée à l'algorithme de super-résolution a été proposé.

Pour proposer un algorithme performant de restauration de l'image, il faut au préalable déterminer différents paramètres de la dégradation comme par exemple les flous naturels engendrés par des mouvements de caméra arbitraires ou des effet atmosphériques. Nous avons proposé une méthode de détermination de flou basée sur la réponse du contour dans la direction de mouvement connu à partir du modèle de mouvement.

Indexation cross-média

Toujours dans l'optique de l'indexation des documents vidéo et leur structuration en scènes, nous nous intéressons à la coopération des informations provenant des deux flux : audio et vidéo. Il a donc fallu s'orienter vers une segmentation, non plus en plans de montage, mais en scènes. Cette approche s'appelle l'analyse cross-média des flux vidéos par le contenu.

L'indexation cross-média de la vidéo par le contenu consiste, donc, à fusionner les informations issues de l'analyse des flux audio et vidéo pour proposer un découpage en scènes. Nous avons développé, dans un premier temps, en collaboration avec Philips Research une approche basée sur la détection des silences dans le flux audio et des frontières de changements de plans de montage dans le flux vidéo. Les frontières de scènes sont matérialisées par la correspondance temporelle entre les silences du flux audio et les plans de montage du flux vidéo. La prise de décision s'appuie sur une méthode de classification statistique supervisée basée sur la maximisation du rapport de vraisemblance. La méthode correspond à une classification en deux classes (Frontière de Scène / Non-Frontière de Scène) de l'ensemble des intervalles temporels entre un silence audio et une frontière de plan de montage vidéo. Une étude plus approfondie du comportement du flux audio au voisinage des frontières de scènes a fait ressortir que les silences caractérisaient à environ 66% une frontière de scènes, toutefois les 44% restant sont représentés par d'autres variations du flux audio dont, notam-

ment, les changements de type de bruits : claquement de porte puis bruits de pas avec une frontière de scène entre les deux types de bruits. Ainsi, nous avons adressé le problème de la détection des transitions bruit-bruit et développé une approche statistique semi-aveugle pour les caractériser. Dans la continuité de ces travaux, nous voulions améliorer la segmentation semi-aveugle en utilisant un ensemble optimal de descripteurs de bas-niveau du flux audio. Dans le sens où il s'agit de proposer le plus petit ensemble de descripteurs qui maximise le pourcentage de réussite. Si nous pouvions étiqueter les bruits d'un flux audio, nous pourrions ainsi soumettre un ensemble de descripteurs très spécifiques à la segmentation. Pour cela, nous proposons une classification statistique supervisée et multi-classes des bruits. Avant tout, nous avons dû définir un ensemble de classes et de descripteurs *ad-hoc* pour les bruits les plus couramment rencontrés.

Résultats – L'approche développée pour la caractérisation des frontières de scènes par l'étude de la correspondance temporelle entre les silences audio et les frontières des plans de montage vidéo génère des résultats très prometteurs pour la suite. La détection non-supervisée des changements de types de bruits dans un flux audio continu fournit de bons résultats à condition que la frontière entre les deux types de bruits soit abrupte et non graduelle. La méthode statistique développée pour la classification supervisée des sons bruités a été évaluée sur un grand ensemble de données réelles. Les résultats obtenus sont très satisfaisants et ouvrent la voie à d'autres applications car il s'agit d'un thème assez peu étudié dans la littérature.

Visualisation de documents vidéo

L'indexation des documents vidéo a comme ouverture une vaste problématique de visualisation des données. En s'appuyant sur le modèle structurel des documents vidéo, en plans et en scènes, nous modélisons un document comme un graphe complet. En utilisant les approches issues de l'algorithmique des graphes, nous proposons une structuration en scènes qui peut-être modélisée (collaboration avec le projet transversal *Visualisation*).

Différents critères ont été étudiés à cet effet : descripteurs de couleur issus de la norme MPEG7, algorithme de détection de visages (développé au sein de l'équipe), analyse des motifs périodiques dans l'enchaînement des plans.

Dans un premier temps, nous avons évalué les performances de différents descripteurs de couleur des plans vidéo. La mesure de performance utilisée est la capacité du descripteur à calquer la similitude perçue par l'humain. Ce critère est primordial pour obtenir un groupement des plans similaires en scènes qui soient les plus homogènes possibles. Nous avons donc évalué deux descripteurs de couleur issus de la norme MPEG7 (ColorLayout et DominantColor) et un descripteur de couleur dédié aux plans vidéo (XRay) selon leur mesure d'ANMRR (Average Normalized Modified Retrieval Rate). Un plan vidéo ou une image-clé est choisie comme cible parmi un ensemble de plans vidéo. On compare ensuite la liste des plans classée selon le critère de similitude associé à chaque descripteur avec la liste des plans de l'ensemble triée manuellement du plus similaire au moins similaire (indexation de référence) est. L'ANMRR permet de classer les descripteurs selon leur adéquation avec la perception humaine. Pour caractériser la couleur d'un plan vidéo, nous avons encore cherché à savoir si le fait de ne prendre en compte que quelques frames représentatives du plan vidéo plutôt que la totalité des frames du plan influençait la mesure de similitude de manière importante. Les deux techniques ont également été comparées en termes d'ANMRR.

Le critère colorimétrique retenu permet ensuite de procéder à un groupement des plans les plus similaires en scènes. Cette fragmentation des plans en classes de couleur offre un moyen de navigation non linéaire dans le document vidéo (par opposition au parcours de la vidéo dans l'ordre chronologique). A cet effet, différentes méthodes de fragmentation ont été testées.

Les critères que l'on souhaite maximiser sont l'homogénéité intra-classe et l'hétérogénéité inter-classe. Pour cela, différentes méthodes de fragmentations ont été expérimentées.

La première approche est basée sur des algorithmes de fragmentation utilisés en Datamining et en extraction de connaissances. L'algorithme de fragmentation DBScan est l'une des méthodes qui a été utilisée.

La deuxième approche est basée sur les graphes. L'ensemble des descripteurs de plans vidéo est représenté par l'ensemble des sommets d'un graphe dont les arêtes sont évaluées suivant la mesure de similitude associée à chaque couple de descripteurs. Le graphe construit à partir d'une collection de plans vidéo est donc un graphe complet, évalué. Le graphe complet étant trop dense et sans caractéristiques topologiques significatives, une série de traitements supplémentaires sont nécessaires.

Ainsi, le filtrage de certaines classes d'arêtes du graphe permet de supprimer des relations entre plans jugées inutiles. Par exemple, un seuillage conduisant à la suppression des arêtes de poids fort permet de réduire le nombre d'arêtes du graphe et d'induire des caractéristiques dans le graphe (degré, connexité, paramètres combinatoires) qui sont représentatives du document vidéo sous-jacent.

C'est une approche originale qui permet de tirer partie des nombreux algorithmes dédiés à la manipulation de graphes (arbres couvrants, spanners, calcul de paramètres combinatoires, fragmentation) et à leur représentation graphique (dessins de graphes).

Le groupement des plans similaires en scènes de couleur similaire est un premier niveau d'organisation du document vidéo. Cependant, le faible niveau sémantique de telles scènes ne permet pas de les associer au contenu sémantique du document vidéo. Nous avons donc proposé une définition de scène visant à détecter les dialogues entre deux personnages. Ce type de dialogue est caractérisé par des prises de vues alternées sur le premier et le second interlocuteur. Afin de détecter ces motifs caractéristiques dans l'enchaînement des plans, nous avons intégré deux critères dans la détection des scènes : l'alternance de plans appartenant à deux classes colorimétriques distinctes et la présence de visages dans ces plans. Le second critère servant de filtre afin d'éviter la détection de scènes composées de plans alternés mais n'impliquant pas deux personnages (ex : interaction entre un personnage et un objet).

Pour la détection de plans alternés, nous utilisons une matrice d'affinité entre les plans vidéo basée sur la similitude couleur. Les éléments de la matrice contiennent l'affinité de chaque paire de plans du document vidéo. Différents niveaux de seuillage sont appliqués pour transformer les valeurs d'affinité en 0 et en 1. Cette matrice peut alors être vue comme une image binaire dont les motifs en forme de damier sur la diagonale principale de la matrice caractérisent des plans consécutifs appartenant à deux classes de couleurs distinctes. La détection de visages est ensuite appliquée à ces scènes périodiques.

Résultats – La comparaison des trois descripteurs de couleur et la modélisation de l'ensemble des plans vidéo par un graphe ont fait l'objet de plusieurs publications dans des conférences internationales (CBMI'03, EI'04) et dans un journal (MTAP). La publication concernant la détection des scènes de dialogue visuelle est acceptée dans une conférence internationale (CBMI'05). Les perspectives autour de l'activité de visualisation des documents vidéo concernent la poursuite des travaux d'analyse de la structure des documents vidéo ainsi que la mise en oeuvre de nouvelles interfaces de navigation dans les contenus multimédia indexés basées sur les graphes.

Projets et perspectives 2005-2008

Les axes de recherche que nous avons initiés au sien du LABRI sur l'Analyse et Indexation Vidéo sont à poursuivre. Nous comptons notamment poursuivre la recherche dans des

méthodes d'indexation par apprentissage, en se basant sur des outils d'extraction des descripteurs et des objets déjà développés. Ce thème porteur de la recherche en multimédia est aujourd'hui traité au niveau international sous le volet *Knowledge technologies*. Nous sommes en train de faire les premiers pas pour réduire *l'écart sémantique* entre les résultats de l'extraction des divers descripteurs et objets de la vidéo et l'interprétation de son contenu. Ces méthodes par apprentissage peuvent être déclinées aussi bien pour un média que dans le contexte *cross-média*.

Par ailleurs, les divers problèmes d'analyse et segmentation vidéo sont loin d'être épuisés. Il nous semble prometteur de poursuivre le travail sur la géométrie discrète et l'estimation du mouvement initié au sein de notre équipe Image et Son afin d'aboutir à des schémas de suivi des objets en multi-résolution, ou encore à l'augmentation locale de la résolution sur des objets segmentés. Nous voyons de larges possibilités d'application de ces méthodes aussi bien dans le contexte du multimédia grand public que de l'imagerie médicale.

Un autre axe de recherche qui nous intéresse aujourd'hui est celui de transmodage vidéo. La segmentation vidéo objet et l'ensemble des méthodologies développées dans le paradigme de rough indexing peuvent apporter des solutions intéressantes quant à ce domaine. La reconstitution des informations perdues lors de la transmission de la vidéo via réseaux peut également constituer un axe de futur développement de ces méthodes. Enfin la fusion des informations provenant des outils d'analyse et recherche des meilleurs modes de coopération reste un problème ouvert.

Dans l'immédiat, nous allons mettre en place une campagne d'évaluation nationale des outils d'indexation vidéo ARGOS dans le cadre du programme TECHNOVISION.

4.3.2 Modélisation du son et de la musique

Introduction

D'une manière générale, la recherche en informatique musicale est divisée actuellement en deux domaines assez disjoints : un domaine symbolique qui est lié à la modélisation de pièces musicales, de règles musicales dans le but d'une aide à l'analyse et à la composition ; et un domaine sous-symbolique qui est lié à l'analyse, le traitement et la synthèse de sons. Nous abordons ces deux domaines en tentant de trouver des représentations compatibles pouvant cohabiter dans un même logiciel.

L'objectif principal des recherches du thème Modélisation du Son et de la Musique est de développer des modèles de sons et de partitions permettant diverses interactions en temps réel ou différé. Ces modèles pourraient donner lieu à de nouveaux instruments permettant l'interaction avec des pièces musicales sur support, ces interactions pouvant aller de la spatialisation des sources sonores du morceau jusqu'au jeu improvisé avec un musicien virtuel.

Au niveau symbolique, nous avons surtout travaillé sur des représentations temporelles, explicites ou implicites. Nous étudions un système d'optimisation de contraintes temporelles basées sur les relations de Allen afin d'éditer, sous contraintes, un document sonore hiérarchique.

Concernant la modélisation sonore, nous nous intéressons plus particulièrement aux sons musicaux. Il nous faut donc impérativement considérer la polyphonie (plusieurs sons joués simultanément), ce qui complexifie grandement les méthodes d'analyse sonore à mettre en œuvre. De plus, la composante bruitée des sons musicaux est souvent importante (à la fois quantitativement et qualitativement). Les modèles à définir doivent tenir compte de toutes ces contraintes.

De plus, nous souhaitons que les modèles soient proches de la perception. Il s'agit donc de modéliser les sons au niveau du récepteur (de l'auditeur), et non pas au niveau de la source émettrice (comme c'est le cas pour les modèles physiques par exemple). Ce choix nous amène à travailler en collaboration avec les psychoacousticiens. Les modèles spectraux établissent un lien plus fort entre la perception et les paramètres du modèle, en se basant sur la représentation fréquentielle. Nous nous sommes donc particulièrement intéressés à ce type de modèles.

En plus du modèle, nous souhaitons une chaîne d'analyse / synthèse complète, contrairement à ce qui se fait classiquement en indexation par exemple, où seule l'analyse compte et où la modification des paramètres obtenus en vue d'une resynthèse d'un son transformé est *a priori* exclue.

Nous portons nos efforts sur les algorithmes et les structures de données. Même si des bases solides en traitement du signal ou probabilités sont essentielles, nous pensons arriver à des problèmes plus algorithmiques et combinatoires.

Dans nos recherches, la priorité va à la compréhension (recherche fondamentale) plutôt qu'au traitement (ingénierie). Une fois les mécanismes fondamentaux compris, nous utilisons ces résultats dans de nombreuses applications, souvent implantées sous forme de logiciels libres développés sur systèmes d'exploitation UNIX (Linux, MacOS X). Ces logiciels sont destinés aussi bien aux scientifiques qu'aux musiciens.

Mots-Clés Modèles sonores ; Analyse et synthèse de son ; Représentation temporelles de pièces musicales ; Optimisation de contraintes pour la composition musicale.

Description des activités et principaux résultats

Modèles sinusoïdaux stationnaires

Basés sur les travaux d'Helmholtz et de Fourier, les modèles sinusoïdaux décomposent le son sous forme d'une somme de sinusoïdes (les partiels) dont les paramètres (fréquences et amplitudes) varient lentement dans le temps.

Nous nous sommes dans un premier temps concentrés sur les modèles sinusoïdaux dits stationnaires, qui considèrent qu'à un instant donné les paramètres des partiels sont quasi constants à l'intérieur d'un petit intervalle de temps (la fenêtre d'analyse) [684]. Nous avons proposé une nouvelle méthode d'analyse, tirant parti des dérivées du signal sonore, afin d'améliorer la précision de l'estimation des paramètres des partiels. La comparaison des méthodes d'analyse existantes [617] a montré que cette méthode était une des plus performantes dans le cas des sons musicaux. Nous avons également proposé par le passé une méthode de synthèse rapide, qui recalculait le signal sonore à partir des paramètres du modèle de manière optimale, en utilisant un calcul récursif de la fonction sinus. Nous tenons également compte des résultats de la psychoacoustique afin d'éliminer les partiels inaudibles à la volée pour diminuer la taille du problème et ainsi le temps de calcul.

Modèles sinusoïdaux non stationnaires

Dans le cas des modèles sinusoïdaux non stationnaires, on considère que les paramètres des partiels peuvent évoluer à l'intérieur de la fenêtre d'analyse. Les dérivées des paramètres (fréquences et amplitudes) des partiels ne sont plus nulles, et il convient de trouver une méthode d'analyse pour les estimer, chose proposée dans [631]. La synthèse du signal sonore à partir des paramètres du modèle est alors plus compliquée. Nous avons étendu les modèles polynomiaux de McAulay et Quatieri de l'ordre 3 à l'ordre 5 pour la phase (de l'ordre 2 à l'ordre 4 pour la fréquence, qui est la dérivée de la phase), et de l'ordre 1 à l'ordre 3 pour l'amplitude. La comparaison des différents modèles de phase polynomiaux [592], faite dans le cadre d'une collaboration entre le LaBRI, l'ICP (Institut de la Communication Parlée, Grenoble), le LORIA (Nancy) et l'IRCAM (Paris), a toutefois montré qu'augmenter l'ordre des modèles polynomiaux n'était pas une solution très efficace, et qu'une remise en cause des modèles existants était nécessaire.

Modélisation à long terme

Les paramètres des partiels à chaque trame d'analyse constituent une représentation à court terme. Pour obtenir une représentation à long terme, il s'agit de reconstituer les évolutions dans le temps des paramètres des partiels à partir de la représentation à court terme. Cette phase est communément appelée le suivi de partiels. La représentation sinusoïdale à long terme est essentielle pour de nombreuses applications comme par exemple le codage audio-numérique, la classification des sources sonores, ou les transformations musicales de grande qualité. Cependant, l'existant en matière d'algorithmes de suivi de partiels est assez limité. Nous avons proposé lors de la thèse de Mathieu Lagrange [739], en collaboration avec Jean-Bernard Rault de France Télécom R&D Rennes, de nouveaux algorithmes basés sur l'exploration de trajectoires dans le futur [632], l'extrapolation des trajectoires futures basée sur la prédiction linéaire [630, 633], et l'utilisation du contenu hautes fréquences [634] des évolutions des paramètres des partiels.

La modélisation à très long terme [590] est une perspective de recherche intéressante, notamment dans le cas de la voix parlée. Il s'agit par exemple de coder la totalité des évolutions de chaque partiel au cours d'un segment de parole (une voyelle, voire une phrase entière). Une collaboration en cours avec Laurent Girin de l'ICP laisse entrevoir des perspectives intéressantes pour le codage et la compression des sons de parole.

Modèle sonore statistique

L'essentiel des recherches a été consacré jusqu'à maintenant à la modélisation des parties périodiques du signal. Un son périodique de base, comme par exemple une sinusoïde pure, n'est pourtant pas un son naturel. Ce n'est que depuis quelques années que la partie bruitée du son est prise en considération. La qualité de rendu des sons n'en est que meilleure. Toutefois, la représentation de cette partie repose sur des hypothèses peu restrictives et approximatives, car les sons étudiés restent majoritairement des sons instrumentaux peu bruités. Les travaux présentés dans la thèse de Pierre Hanna [737] ont été menés dans le but de développer un modèle de représentation mathématique des sons dont la partie bruitée est importante, que nous appellerons sons bruités, et de proposer ainsi des contrôles musicaux propres à leurs caractéristiques stochastiques.

Les modèles spectraux existants considèrent les parties bruitées des sons comme des bruits blancs filtrés, sans apporter de justification théorique ou expérimentale. La représentation par du bruit blanc filtré apparaît limitée, par rapport à la richesse et la complexité des sons bruités naturels. Les travaux psychoacoustiques menés par William Hartmann représentent les fondations de notre étude, en mettant en avant la faculté du système auditif à discerner des bandes de bruit de densités spectrales différentes, et en proposant donc une nouvelle dimension perceptive. L'étude de ce nouveau paramètre met en évidence les limites de l'hypothèse précédente. Ainsi, nous avons proposé un nouveau modèle spectral et statistique, appelé CNSS [537], basé sur le modèle du bruit thermique décrivant un bruit comme une somme fixe de sinusoïdes, permettant de prendre en compte de nouvelles propriétés perceptives telles que la densité spectrale [614]. En effet, les paramètres du modèle permettent de contrôler les distributions de fréquences et de phases [613]. Les transformations musicales permises par ce modèle CNSS sont nombreuses : modification de la densité spectrale moyenne, variations localisées de densité spectrale, harmonicité, impulsivité et manipulation de l'échelle temporelle. Cette dernière transformation a été étudiée en détails [615], notamment sur les sons issus de la voix (consonnes, voix chuchotées).

Par ailleurs, le modèle CNSS a été implémenté en tant que modèle de synthèse [610] pour pouvoir expérimenter les paramètres de contrôle musical, et en particulier vérifier les liens entre les paramètres mathématiques du modèle et la perception [611]. Cette implémentation, sous forme de bibliothèque et d'objets jMax, a permis d'effectuer des tests de synthèse et de confirmer son application en temps réel.

Modélisation de sons polyphoniques

Nos recherches se concentrent principalement sur l'étude d'une représentation structurée des sons polyphoniques [739]. De nombreux sons issus d'instruments de musique ou de voix possèdent des modes de vibration qui peuvent chacun être modélisé sous la forme d'une sinusoïde dont les paramètres de fréquence et d'amplitude varient lentement au cours du temps. L'identification de ces partiels à partir d'un enregistrement numérique pose de nombreux problèmes dans le cas polyphonique. La prise en compte de la prédictibilité de l'évolution de ces paramètres ainsi que l'absence théorique de hautes fréquences dans ces mêmes évolutions ont permis de proposer des améliorations conséquentes dans ce domaine (voir plus haut) [634].

L'étape suivante consiste à regrouper les partiels issus de la même entité sonore (note), chaque entité étant perçue par le système auditif humain non plus comme plusieurs sons simples mais comme un unique son complexe. Pour cela, nous devons étudier de nombreux critères tels que l'apparition / disparition simultanée des partiels, la relation d'harmonicité entre les fréquences des partiels et plus généralement l'évolution corrélée des paramètres des partiels. La combinaison pertinente de ces différents critères à l'aide de méthodes de

classification statistiques est actuellement à l'étude. L'objectif est la séparation des divers objets (sources) sonores présents dans la musique.

Modélisation de partitions interactives

Ce sujet de recherche se situe dans le cadre de l'écriture musicale. Contrairement aux outils de composition permettant l'écriture de processus de calcul de la musique, nous abordons la composition comme l'écriture de matériau sonore, correspondant davantage à de l'édition de données plutôt qu'à de la programmation. La prise en compte d'événements interactifs durant le processus de composition dans le cadre de l'écriture est assez limitée dans les outils actuels. Nous souhaitons étudier cette possibilité tout en fournissant au compositeur le moyen de spécifier des propriétés que le résultat sonore devra vérifier.

Nous appelons une "partition interactive", une partition de musique incluant des événements "statiques" et des événements "interactifs", mis en cohérence au moyen d'un ensemble de propriétés.

Nous avons étudié la possibilité d'exprimer les propriétés concernant l'aspect temporel de la spécification en utilisant des relations temporelles liant les événements de la partition. À partir d'une partition interactive entièrement spécifiée, nous proposons des algorithmes permettant d'obtenir une pièce musicale interactive vérifiant les contraintes temporelles spécifiées dans la partition. Ultérieurement, nous étendrons le modèle obtenu par la prise en compte de propriétés concernant d'autres dimensions musicales. Ces résultats ont donné lieu à deux conférences internationales avec actes et un article en revue.

Éditeur de partitions

Gsharp est un projet dont l'objectif est la réalisation d'un éditeur de partitions musicales extensible, interactif et temps-réel. Il est basé sur un algorithme incrémental de calcul de mise en page, permettant à une partition de plusieurs centaines de pages d'être recalculée à la vitesse de frappe de l'utilisateur.

Gsharp est un logiciel *libre*, et distribué avec code source selon la licence GPL (GNU General Public License).

L'expérience de logiciels interactifs similaires (notamment de l'éditeur de texte Emacs) montre la nécessité de l'utilisation d'un langage de très haut niveau pouvant servir à la fois de langage d'implémentation (qui doit pouvoir générer du code efficace) et de réalisations d'extension (scripts). Ces deux objectifs nécessitent une implémentation du langage sous la forme d'un compilateur incrémental. Le langage Common Lisp a été choisi pour toutes ces raisons.

L'utilisation efficace d'un éditeur de partitions musicales nécessite un *modèle* du fonctionnement du logiciel permettant à son utilisateur de prévoir l'effet de ses actions. Le modèle de Gsharp permet à la fois une utilisation efficace de commandes simples, mais également des commandes nécessitant une connaissance sémantique de la musique de la part du logiciel.

Un logiciel comme Gsharp doit être *facile à utiliser*, à savoir que les commandes fréquemment utilisées doivent être facilement accessibles par un utilisateur régulier. Cet objectif est différent de celui des logiciels commerciaux qui sont optimisés pour être vendus, et par conséquent pour un public de débutants.

Afin que le logiciel puisse s'adapter à son utilisateur, en particulier du point de vue de la facilité de déclencher des commandes, Gsharp a été conçu pour être configurable à plusieurs niveaux. Ainsi, les commandes de bases peuvent être associées à des touches différentes selon les souhaits de l'utilisateur, de nouvelles commandes peuvent être construites sous la forme de *macros*, et des fonctionnalités supplémentaires peuvent être chargées dynamiquement dans le logiciel, même en cours d'exécution.

Nous travaillons actuellement sur la quatrième génération du logiciel. Récemment, nous avons fini l'implémentation de l'algorithme incrémental de mise en page d'une manière à permettre des règles de découpage différentes d'être facilement intégrées dans le logiciel. Alors que le travail n'est pas encore terminé, la version actuelle du logiciel permet à un utilisateur futur d'avoir une bonne idée de la présentation finale du logiciel.

Modèle d'édition de partitions musicales

Dans le contexte de la musique contemporaine, nous souhaitons définir un modèle d'édition de partitions musicales adapté au processus incrémental de la création. Faisant suite aux recherches ayant donné lieu au logiciel BOXES validant un modèle d'édition avec optimisation de contraintes temporelles basées sur les relations de Allen, nous cherchons à étendre le champ des contraintes à d'autres dimensions musicales. Nous exhibons ainsi des modèles de contraintes génériques pouvant s'appliquer à des courbes abstraites pouvant représenter l'évolution de divers paramètres musicaux. Les variables du système de contraintes admettent des domaines continus ou bien discrets. Nous travaillons actuellement à la conception d'un algorithme hybride de résolution de ces contraintes, certains domaines étant continus et d'autres finis.

Dolabip : un instrument pour l'éveil musical

L'objectif de ce projet est le développement d'outils pédagogiques pour l'éveil musical en école publique permettant aux jeunes enfants de jouer dans un espace sonore virtuel au moyen de périphériques adaptés à leur motricité. Ces recherches utilisent les modèles sonores développés dans l'équipe image et son car ils sont proches de la perception. Nous avons mené nos premières expérimentations avec un instrument utilisant un joystick et permettant une prise en main rapide et un plaisir immédiat. Cet instrument est utilisé dans plusieurs écoles maternelles.

Les perspectives de ce projet consistent à franchir l'étape de l'exploration sonore pour parvenir à l'exploration musicale, ce qui nécessite la spécification de pièces musicales non terminées, qui pourraient donner lieu à diverses interprétations. Cette perspective rejoint sur le plan formel le travail sur la partition interactive.

Modélisation d'une percussion aérienne

L'objectif de ce projet est de créer et de réaliser un spectacle musical associant des instrumentistes traditionnels (interprètes et improvisateurs) à un méta-instrument électroacoustique. Nous utilisons un dispositif constitué de deux Flocks of Birds, d'un PC permettant l'acquisition des positions en trois dimensions et d'un Macintosh effectuant la synthèse sonore. Nous avons étudié la détection des coups joués par un percussionniste au moyen des deux capteurs. Ce dispositif a été utilisé dans le cadre de plusieurs concerts en France et à l'étranger.

Les perspectives de ce projet consistent à intégrer l'analyse du jeu du musicien et de visualiser la percussion virtuelle avec la position courante des baguettes.

Projets et perspectives 2005-2008

Modèle hiérarchique

Lors de l'analyse d'un son harmonique, seuls quelques paramètres sont extraits par un algorithme de suivi de partiels, en particulier la fréquence et l'amplitude des sinusoides qui composent le son à un moment donné. Mais il existe d'autres paramètres qui pourraient

être extraits par une réanalyse des paramètres estimés lors de la première étape. De manière similaire, il doit être possible d'extraire hiérarchiquement des paramètres du son.

Un modèle hiérarchique sonore [643] permet de modéliser les sons à plusieurs niveaux, c'est-à-dire au niveau du son, mais aussi au niveau du contrôle du son, et au niveau des variations du contrôle. Un modèle hiérarchique permettrait d'agir à tous les niveaux sur le son, du microscopique au macroscopique. La thèse de Martin Raspaud a pour but de proposer un tel modèle pour réaliser un étirement temporel de grande qualité, en agissant sur tous les niveaux de la hiérarchie. Nous avons notamment proposé une technique d'étirement temporel préservant la hauteur, le timbre, et aussi les micro modulations présentes à l'intérieur du son (vibrato, trémolo).

Modèle unifié

La plupart des modèles sonores sont actuellement des modèles hybrides, où la composante bruitée (stochastique) est séparée de la composante sinusoïdale (déterministe). Nous avons montré que la modélisation sinusoïdale, bien adaptée aux sons déterministes, pouvait également à court terme permettre de générer des sons bruités (thèse de Pierre Hanna). Nous souhaitons maintenant définir un modèle spectral unifié, plus souple, où parties déterministe et stochastique seront indissociables. Une collaboration entre Sylvain Marchand et Pierre Hanna est en cours, débutée en 2004 par le co-encadrement d'un master recherche.

Parmi les applications de la modélisation informatique du son musical, citons le tatouage audio-numérique, la séparation de sources, ainsi que la spatialisation sonore.

Tatouage audio-numérique

Le tatouage (*watermarking*) audio-numérique consiste à cacher des informations à l'intérieur des sons, et ce de manière inaudible.

Nous avons montré, en collaboration avec Laurent Girin (ICP), que les modèles sinusoïdaux se prêtaient bien au tatouage des signaux de parole [591]. Cela permettrait de marquer, à des fins de protection de droits d'auteurs par exemple, les dictionnaires utilisés lors de la synthèse vocale.

Nous proposons aussi d'utiliser cette possibilité afin d'inclure dans le son un motif de tatouage suffisamment robuste pour résister aux transformations musicales élémentaires, tout en étant assez souple pour subir également ces transformations en même temps que le son support.

Au final, le but serait, en analysant le son après transformation, et plus précisément la version déformée du motif de tatouage, d'être capable de déterminer les déformations subies par le motif de tatouage initial et d'en déduire une chaîne probable de transformations ayant conduit au son actuel.

Ceci pourrait par exemple permettre d'analyser les pratiques des compositeurs. On peut aussi envisager de retrouver le son d'origine dans le cas de transformations réversibles.

Spatialisation sonore

Nous avons amorcé en 2003 des recherches sur la spatialisation sonore, en nous basant toujours sur une représentation spectrale des sons. Joan Mouba a commencé sa thèse sur le sujet récemment, en 2004.

En partant d'un son stéréo (provenant par exemple d'un CD audio), la technique DUET permet de localiser les sources sonores dans l'espace, en s'inspirant des mécanismes intervenant dans la perception binaurale (avec les deux oreilles). Nous reprenons les résultats récents d'Harald Viste à l'EPFL (École Polytechnique Fédérale de Lausanne, Suisse) sur la modélisation acoustique et psychoacoustique pour la localisation des sons. Il s'agit d'une

modélisation binaurale dérivée des HRTF (*Head-Related Transfer Functions*). La méthode développée permet de localiser les diverses sources sonores, et de les séparer.

Récemment, il a été montré que l'on pouvait aussi interagir avec ces sources durant l'écoute, pour par exemple en changer la position spatiale perçue. Ainsi, si dans un enregistrement rock on entend un chanteur au milieu, une guitare vers la gauche et une batterie à l'avant, on peut vouloir faire glisser la guitare vers la droite, ou supprimer la voix du chanteur (effet karaoké).

Écoute active

Aujourd'hui, l'auditeur est considéré comme un récepteur qui écoute passivement le signal audio fixé sur support (CD, DVD audio, *etc.*). La seule modification qu'il fait généralement subir au son originel est la modification du volume. En observant les pratiques des compositeurs de musique électro-acoustique au sein du SCRIME (Studio de Création et de Recherche en Informatique et Musique Électro-acoustique), nous nous sommes rendus compte qu'il était essentiel de pouvoir interagir en direct avec le son originel sur support, et ce pendant sa diffusion.

Nous proposons de permettre à tout auditeur d'avoir une écoute active, en lui donnant la liberté d'interagir avec le son en temps réel lors de sa diffusion, au lieu de l'écouter de manière habituelle, passive, avec comme seul contrôle des réglages simples comme le volume. Nous proposons d'offrir à l'auditeur la possibilité de changer la localisation spatiale des diverses sources sonores, leurs volumes respectifs, leurs hauteurs, voire d'agir sur le timbre des sons, leurs durées, le rythme de la musique.

Diffusion par le geste

L'objectif de ce projet consiste à substituer les manipulations de curseurs de la table de diffusion par des gestes inspirés de ceux du chef d'orchestre devant les instrumentistes (les enceintes).

Nous utilisons deux capteurs *Flocks of Birds*, un dans chaque main. Ces capteurs transmettent les coordonnées tridimensionnelles ainsi que les rotations toutes les 10 ms. Au moyen de ces capteurs, il faut établir des codes gestuels paramétrés permettant de commander la diffusion, c'est-à-dire de contrôler le volume de chacune des enceintes. Les informations symboliques associées à chaque geste sont utilisées ultérieurement pour commander le volume de chaque enceinte. Le projet mené avec le compositeur Jean-Michel Rivet est dans sa phase finale, et concerne actuellement sa mise en pratique pour déceler les imperfections et le rendre le logiciel fiable.

Le projet MidiDanse

Le MidiDanse est une interface électronique permettant de capter les mouvements d'un musicien/danseur et de les restituer en musique. Le système est composé d'un réseau de capteurs qui détecte les déplacements et les mouvements d'un ou de deux musiciens/danseurs sur une scène. Les informations, ainsi obtenues, sont traitées instantanément par le système et permettent de générer des ordres MIDI transmis ensuite à un ordinateur. L'ordinateur, équipé d'un logiciel musical, pilote un synthétiseur, relié à des enceintes, et permet de diffuser la musique correspondant aux ordres MIDI.

Deux ensembles de sept faisceaux lasers, placés l'un au dessus de l'autre, balayent la scène horizontalement et se croisent en son milieu. Les faisceaux lasers sont constitués par des émetteurs/récepteurs disposés sur des supports verticaux dont la position peut-être modifiée en fonction de l'utilisation. L'espace balayé par les faisceaux couvre une surface circulaire de six mètres de diamètre environ.

Le second prototype du MidiDanse a été utilisé pour la première fois, en situation réelle, lors d'une audition de la classe de composition électroacoustique au conservatoire national d'Angoulême. Il a été très bien accueilli tant par la personne qui l'a utilisé que par le public. Nettement plus fiable et efficace que le prototype précédent, son utilisation est beaucoup plus agréable. La technologie laser pour les faisceaux apporte une grande précision et un attrait visuel très intéressant.

4.3.3 Modélisation et synthèse d'images

Introduction

Notre équipe, qui comprend le projet IPARLA (LaBRI, INRIA futurs), s'intéresse aux problèmes posés par la visualisation et la manipulation de données complexes sur des terminaux mobiles communicants (TMC) comme les assistants personnels ou les téléphones mobiles. Il existe un grand nombre de domaines potentiellement intéressés par ce genre de fonctionnalités depuis les applications de guidage jusqu'aux jeux en passant par la surveillance de procédés industriels ou le pilotage d'expérimentations scientifiques. Les contraintes principales résultent des faibles capacités des TMC (mémoire, processeur, bande passante réseau, taille d'écran) dues à la puissance électrique disponible et à l'encombrement réduit. Pour atteindre ces objectifs, nous travaillons dans plusieurs directions :

- création automatique d'environnements virtuels (terrains, villes) à partir de données numérisées (plans, photos),
- nouvelles approches de structuration et de visualisation de données basées sur des modèles ponctuels et des surfaces implicites (reconstruction, échantillonnage, texturation),
- nouvelles méthodes de rendu non photo-réalistes et à base d'images,
- nouvelles méthodes d'interaction basées principalement sur l'analyse vidéo,
- développement d'une plate-forme logicielle intégrant ces différents résultats et mettant en oeuvre la distribution des informations sur les TMC depuis un client en utilisant un réseau (WiFi ou Bluetooth puis GPRS et UMTS).

Mots-Clés Réalité virtuelle, Visualisation 3D, Interaction 3D, Modélisation, Terminaux mobiles, Modélisation Géométrique, Arithmétique entière

Description des activités et principaux résultats

Rendu

Rendu non-photoréaliste de terrains – Sur les terminaux mobiles la taille mémoire limitée impose de manipuler des modèles géométriques de complexité raisonnable. Dans le cas où l'on souhaite visualiser des objets géométriques complexes une idée peut être d'utiliser une version simplifiée, abstraite de ce modèle. Dans ce contexte, nous avons choisi de représenter des terrains 3D complexes par quelques lignes caractéristiques. Cette méthode est inspirée des techniques de rendu à base de traits issues du domaine du rendu non-photoréaliste. Les lignes caractéristiques d'un terrain 3D sont les lignes de flux et les lignes de niveaux. Afin de rendre un terrain, on peut projeter ces lignes visibles en faisant varier la densité en fonction de l'illumination.

Nous avons mis en oeuvre un outil de visualisation de terrain sur PDA. Nous travaillons en configuration client / serveur. Toute la géométrie du terrain est stockée du côté du serveur. Les lignes caractéristiques (plus légères que le maillage complet) sont transmises au terminal mobile pour le rendu. Nous travaillons à explorer d'autres algorithmes de rendu et aussi à un protocole de communication pour la transmission progressive des données et une mise à jour dépendante du point de vue.

Rendu à Base d'Images – Les terminaux mobiles (TMs) ont des performances de calcul et une mémoire très limitées. Ces caractéristiques techniques font que, sans une évolution du matériel, le *pipeline* de rendu classique restera mal adapté à ce type de plate-formes. Une alternative à ce problème est l'utilisation de techniques de rendu à base d'images (IBR -

Image Based Rendering). Pour rendre des scènes complexes sur TM, nous avons intégré dans une architecture client/serveur des algorithmes issus du domaine de l'IBR. Le serveur calcule des images de référence et les envoie (via un réseau) au TM. Sur le client la navigation dans la scène 3D est réalisée en interpolant le point de vue de l'utilisateur à partir des images de référence fournies par le serveur. En utilisant cette technique la complexité du rendu (sur le TM) dépend uniquement de la résolution des images de référence et non plus de la géométrie de la scène.

Une étude de faisabilité mettant en oeuvre cette approche a été réalisée sur PDA. Cette étude a été complétée par le développement d'un algorithme de placement automatique de caméras (utilisées pour le calcul des images de référence) permettant de réduire le nombre d'artefacts et d'augmenter l'espace 3D visible par un jeu de caméras donné. Ce problème de placement étant clairement difficile à résoudre nous nous sommes limités au cas de scènes urbaines[730].

Visualisation de Surfaces de Points à l'aide de Surfaces de Subdivision – Nous avons créé un algorithme permettant d'obtenir une visualisation de surfaces de points non structurées, en combinant une triangulation 2D locale et une subdivision de surface [704, 562]. L'idée centrale est de produire un ensemble de surfaces polygonales recouvrantes sur le nuage de points, à l'aide d'une triangulation 2D locale de Delaunay, puis de générer une surface de subdivision pour réduire les artefacts des zones de recouvrement. Dans ce cadre, la plate-forme *Osiris* (<http://www.labri.fr/Personnel/boubek/osiris.html>) a été développée pour démontrer la qualité de rendu d'un tel ensemble de surfaces, et la comparer aux techniques concurrentes. Le principal avantage de notre approche est son adéquation avec les systèmes de rendu existant, ce qui permet d'intégrer les surfaces de points aux outils habituels de rendu polygonal. Nous travaillons désormais à une alternative complète aux techniques de rendu par points (usuellement appelées *surface splatting*), en nous basant sur des outils de reconstruction polygonale locale à la volée, pour la modélisation et la synthèse en temps réel de surfaces de points.

Rendu non-photoréaliste d'environnements urbains – Nous avons proposé une méthode de modélisation et de visualisation de bâtiments adaptée à la navigation interactive sur terminaux mobiles dans de gros environnements urbains. Classiquement, un environnement virtuel urbain est composé de bâtiments construits par extrusion d'empreintes au sol. L'apparence des bâtiments résulte du plaquage de textures réalistes (photographies de bâtiments existants, textures procédurales). Dans le cadre d'une application distribuée où le serveur dispose de l'environnement 3D et où le client dispose d'une partie de la scène, il est fondamental d'assurer des transferts d'information efficaces. Le transfert des textures peut-être coûteux. Nous avons exploré les techniques de rendu non-photoréaliste à base de traits comme alternative au rendu texturé. Chaque façade est représentée par une liste de traits représentant des fenêtres et des portes. Cette liste de traits (information vectorielle) remplace les textures classiques. Les traits caractéristiques sont automatiquement extraits de l'image par l'application de filtres détecteurs de contour. Les traits sont ensuite obtenus par vectorisation et post-traitements. Les traits sont ensuite utilisés pour générer une scène 3D (format VRML) composée d'un ensemble de blocs (un par bâtiment). Les traits sont représentés en 3D dans le repère d'une des faces des blocs.

Notre motivation à utiliser du rendu non photoréaliste n'est pas seulement liée à la recherche d'efficacité temporelle, nous pensons aussi que ce type de rendu peut-être satisfaisant d'un point de vue esthétique. En effet, dans les applications de navigation en environnement urbain actuelles, la qualité des textures utilisées n'est pas toujours satisfaisante. L'effet de réalisme escompté par l'utilisation de ces textures n'est pas forcément obtenu et un effet de

monotonie peut-être ressenti par l'utilisateur naviguant dans de telles villes. Nous pensons que l'abstraction fournie par le rendu à base de traits peut-être satisfaisant. Pour le critère esthétique la validation est un point critique. Nous sommes en contact avec des psychologues sur ce point.

Contrôle sur la qualité du rendu – Dans la grande majorité des systèmes de création d'images de synthèse, les calculs y sont effectués dans des espaces de couleur. Ainsi, tout le processus, de l'acquisition (avec par exemple un appareil photo numérique) et de la modélisation, vers l'affichage (comme sur des écrans CRT ou LCD), se déroule dans une même espace de travail, celui de la couleur. Cependant, cela entraîne une approximation dans le calcul de la réflexion d'une source de lumière sur une surface, réduisant notamment la possibilité de préserver l'apparence colorée d'un matériau. Nous avons présenté [722] une nouvelle approche pour le calcul des réflexions dans un espace de couleur, approche qui n'introduit que de faibles changements dans le processus actuel, afin d'y faciliter son intégration. Elle permet aussi un certain contrôle sur le comportement de la réflexion. Enfin, nous avons présenté aussi les possibilités d'application pour cette méthode, comme le calcul de solution d'éclairage global, du ré-éclairage, ou du rendu sur carte graphique.

Modélisation

Reconstruction de Surface – Les technologies récentes d'acquisition 3D fournissent des ensembles de points massifs. Notre but est de développer de nouvelles méthodes de reconstruction de surfaces implicites à partir de grands ensembles de points non organisés. L'idée directrice de notre approche est de diviser le problème en 3 étapes indépendantes : premièrement, partitionner le domaine global de reconstruction en plus petits domaines s'intersectant, ensuite, résoudre le problème de reconstruction dans le sous-domaine en utilisant des bases de fonctions adaptées, et finalement, mélanger les solutions ensemble en utilisant la partition de l'unité.

Nous avons tout d'abord exploré le domaine des *fonctions à base radiale* avec support global pour la seconde étape [662, 741]. Nous avons ensuite proposé deux méthodes utilisant différentes stratégies de partitionnement : la première [676] utilise une décomposition adaptative basée sur un octree et mélange les solutions entre chaque feuille de l'arbre, alors que la seconde utilise un arbre binaire parfait et mélange pour chaque nœud les solutions de ses deux fils [677]. Les deux méthodes offrent de bons résultats quand au temps de reconstruction et l'utilisation mémoire, en respectant le nombre initial de points.

Plus récemment, nous avons développé une alternative aux fonctions à bases radiales durant l'étape de reconstruction locale. Notre nouvelle technique est basée sur les "Enriched Reproducing Kernel Particle Approximation" (ERKPA), qui ont été développées dans le domaine de la mécanique, durant les années 90, pour résoudre certaines équations différentielles partielles sur les structures sans maillage. Une des bonnes propriétés des ERKPA est de permettre l'inclusion d'arêtes vives dans la surface implicite reconstruite.

Modélisation géométrique à base topologique – L'étude et l'utilisation des structures combinatoires topologiques pour la représentation et la manipulation des objets est le cadre théorique d'un grand nombre d'applications aussi bien dans la modélisation géométrique et la CAO que dans la géométrie discrète et l'analyse d'images. Parmi ces structures, une classe importante est la famille des CW-complexes. Un CW-complexe est construit inductivement : à partir d'un ensemble de cellules de dimension 0, puis en rattachant au squelette de dimension i des cellules de dimension $i+1$. Cette construction générale définit des propriétés et des opérations sur les CW-complexes qui rendent possible l'étude générale des différents espaces topologiques sous-jacents et de leurs homologies. Parmi les exemples les plus connus

et utilisées en modélisation géométrique, on peut citer les G-cartes, les ensembles simpliciaux, les ensembles semi-simpliciaux et les ensembles simploldaux.

L'étude des CW-complexes a principalement donné lieu à deux applications [721, 585]. La première concerne le développement d'un modèleur géométrique sur la base de la représentation des ensembles semisimpliciaux et sur l'implémentation des algorithmes de subdivision (split/merge). Les problèmes de la robustesse des calculs géométriques y compris les approches de l'arithmétique à précision finie, arithmétique floue et le calcul géométrique exact, y sont expérimentés. La deuxième, concerne l'utilisation des CW-complexes pour la représentation et la manipulation des objets discrets. Et notamment, la reconstruction de la surface frontière des objets discrets utilisés en imagerie médicale (3D range images) [724].

Texturation Basée Points – Nos techniques de reconstruction ne sont pas limitées à la reconstruction de surfaces implicites, mais s'appliquent aussi à la reconstruction de fonctions continues de dimension arbitraire. Nous avons par exemple utilisé ces méthodes pour définir une nouvelle classe de textures volumique procédurales, générées à partir d'un nuage de points colorés.

Nous avons développé un nouvel environnement pour la *texturation constructive* de la surface de solides 3D arbitraires [661]. Un utilisateur peut texturer la surface en définissant un partitionnement spatial où chaque partition porte un ensemble d'attributs utilisés dans son sous-espace associé. Pour fournir un système interactif, une représentation multirésolution à base de points est utilisée pour le rendu mais aussi pour l'évaluation de la texture. La texture volumique ainsi composée peut finalement être facilement exportée vers les systèmes de rendu prenant pas en compte directement les surfaces de points (CAO, moteurs de rendu photoréaliste, etc).

Échantillonnage adaptatif de surfaces implicites – Nous proposons dans [706, 639] une solution pour adapter le rendu par points différentiels aux surfaces implicites. Le rendu par points différentiels a été initialement développé pour les surfaces paramétriques comme étant un processus d'échantillonnage en deux étapes qui s'appuie très fortement sur les relations de voisinage des échantillons, relations qui n'existent pas naturellement pour les surfaces implicites. Afin de surmonter cette difficulté, nous avons étendu la technique d'échantillonnage par particules développée par Witkin et Heckbert afin de localement prendre en compte les directions principales de courbure de la surface implicite. Le résultat de notre processus est un échantillonnage anisotrope piloté par la courbure où chaque échantillon "commande" un domaine rectangulaire ou ellipsoïdal l'entourant, orienté localement en fonction des directions maximale et minimale de courbures. Comme dans le cas du rendu par points différentiels, ces échantillons peuvent être rendus de manière efficace par l'utilisation d'un shader spécifique d'un GPU programmable.

Nous avons mis cette technique d'échantillonnage dans la section modélisation et non pas dans la section rendu parce que le but final de ces particules anisotropiques n'est pas réellement de visualiser la surface implicite correspondante, mais plutôt de considérer ces ensembles de particules comme un ensemble de paramètres de contrôle qui peuvent être utilisés afin de manipuler cette surface, par exemple dans le cas d'un environnement de sculpture virtuelle.

Vagues océaniques réalistes – Nous avons collaboré avec l'ESTIA (Ecole Supérieure des Technologies Industrielles Avancées) qui développe un simulateur de conduite de navires marins. Simuler de manière physiquement réaliste des vagues en temps réel sur un PC est un gros challenge. Nous avons dû tenir compte de plusieurs facteurs comme le vent, les courants marins, la bathymétrie, les forces appliquées sur le bateau, ... Notre collaboration s'est centrée sur le développement d'un modèle unique de vagues. Ce modèle adapté à notre

problématique utilisant un modèle spectral est basé sur plusieurs théories physiques - mer des vents, vagues approchant les côtes, réfraction et diffraction, vagues capillaires [568, 571].

Cette partie de collaboration s'est terminée par la délivrance d'un doctorat en 2003 [733].

Création automatique de territoires virtuels — Le but de ce projet est de créer, de manière quasi-automatique, un environnement virtuel issu de différents supports : Cartes topographiques scannées, photographies. Les cartes contiennent des données complexes à interpréter et récupérer : le modèle numérique de terrains (MNT) issu des courbes de niveaux, les objets environnementaux (forêts, villes, villages, routes, cours d'eau, ...). La reconstruction du MNT originel consiste en une série d'opérations semi-automatiques utilisant entre-autres analyses d'images, réseaux de neurones et morphologies mathématiques c-à-d sélection des courbes de niveau [670], reconstruction des courbes brisées, affectation des altitudes et enfin l'étape primordiale, l'interpolation du MNT [655]. Nous avons développé un logiciel nommé *AutoMNT* (<http://iparla.labri.fr/software/automnt>) qui s'occupe de transformer nos données de base en un Modèle Numérique de Terrain [707].

Nous sommes en train de développer un algorithme permettant de récupérer automatiquement toutes les informations littérales (nom des lieux, nom des cours d'eau, route, ...) pour les intégrer dans un système d'information géographique complet (SIG). Ces données pourront ainsi être vues sur différentes plate-forme de visualisation : reality centers, PC, interface WEB, PDA et téléphones. De plus, nous travaillons sur l'affichage de ces informations littérales dans une visualisation temps réel 3D. La problématique principale étant d'éviter au maximum des chevauchements de texte au moment de la navigation. Ceci est d'autant plus important pour les écrans de petites tailles.

Interaction 3D

Interaction avec des terminaux mobiles communicants — Les applications pour les terminaux mobiles communicants (TMC) ont longtemps été limitées à de simples calendriers ou éditeurs de notes. Aujourd'hui, les avancées technologiques permettent l'apparition d'applications de plus haut niveau, telles que des applications 3D interactives. Les tâches d'interaction pour ce type d'applications sont différentes des tâches d'interaction pour les applications 2D classiques des TMC. En conséquence, de nouvelles interfaces doivent être imaginées pour favoriser l'interaction de l'utilisateur avec ses données. Dans ce contexte, nous avons développé une interface à base de vidéo qui tire partie de l'interaction bi-manuelle. En déplaçant une cible colorée devant le terminal, l'utilisateur contrôle de manière directe et intuitive 3 degrés de liberté pour la navigation dans de grands documents ou pour la manipulation d'objets 3D [607]. Une évaluation a permis de montrer que cette interface était plus efficace qu'une interface à base de stylets pour une tâche de recherche de cible dans une carte [608]. Cette nouvelle interface a aussi été utilisée pour favoriser l'interaction collaborative avec des environnements virtuels affichés sur grands écrans [606]. Dans un souci d'indépendance du type d'interacteur, un développement, basé sur l'analyse de Fourier du flux vidéo, est en cours. Ceci permettrait de s'affranchir d'une cible colorée spécifique et d'avoir de plus 4 degrés de libertés.

Utilisation d'un laser pour simuler une souris — Nous utilisons 3 caméras vidéo pour suivre un motif laser sur un grand écran où une application est projetée. Des algorithmes d'analyse vidéo extraient les positions 2D du laser sur l'écran. Ces positions sont envoyées à l'application qui les interprète en fonction de ses techniques d'interaction. Ainsi, notre système permet aux utilisateurs d'interagir directement avec l'application grâce à leurs pointeurs laser, sans besoin d'un opérateur extérieur qui manipulerait le modèle.

Reconnaissance de geste en temps-réel – Nous voulons filmer la main ou le corps d'un utilisateur devant un grand écran pour lui permettre d'interagir avec l'application projetée. Alors que la majorité de tels systèmes met en place des approches basées image, nous avons choisi de nous reposer sur une méthode basée modèle. L'algorithme temps réel que nous développons met à jour la pose d'un modèle 3D pour que la projection de sa silhouette converge vers la silhouette en entrée. Nous incluons aussi d'autres caractéristiques telles des limitations de validité sur les angles de flexion. Nous obtenons ainsi des performances temps réel [729, 663, 602]. Les occultations peuvent perturber notre système, mais ajouter des caméras permet de réduire leur influence [697, 573]. Nos améliorations plus récentes portent sur la mise en place d'une collaboration étroite entre quelques méthodes basées image et notre système basé modèle.

CAT : Une interface à 6 degrés de liberté – Les interfaces de visualisation collective (IVC) sont des systèmes à base de projection sur grands écrans qui permettent à plusieurs participants, localisés en un lieu unique, d'être immergés dans un même environnement virtuel (EV). Malgré leur potentiel pour le travail de groupe, ces interfaces sont souvent sous-utilisées car les utilisateurs n'ont pas la possibilité d'interagir facilement et efficacement avec les EVs qu'ils sont en train de visualiser [602]. Nous avons donc développé une interface à 6 degrés de liberté (ddl) adaptée aux contraintes des IVC, d'abord présentée sous le concept d' " Interaction Table " [603, 609]. Ce concept s'est ensuite matérialisé en une véritable interface d'action : le CAT (Control Action Table) [536, 605, 598]. Cette nouvelle interface favorise une interaction non contraignante, rapide et sensitive. Une série d'expérimentations a permis de montrer que les utilisateurs novices n'avaient pas de difficultés d'apprentissage et qu'ils pouvaient réaliser des tâches d'interaction 3D variées avec efficacité.

Étude des facteurs humains impliqués – Notre objectif est d'identifier les facteurs humains qui influencent la performance lors des tâches de manipulation et de locomotion en environnement virtuel afin d'adapter les interfaces aux utilisateurs. Afin de comprendre les comportements et performances des utilisateurs d'environnements de réalité virtuelle (RV) nous avons mené une expérimentation à grande échelle (100 sujets). Pour chaque sujet, nous avons mesuré un ensemble de facteurs cognitifs (capacités spatiales, attentionnelles, motrices...) qui définissent son profil cognitif. Les performances du sujet sont ensuite mesurées lors de deux tâches d'interaction de RV (manipulation et locomotion) face à deux types d'écrans (moniteur et mur géant) et en utilisant deux types d'interacteurs (souris et CAT). Notre objectif est d'expliquer, de prédire les performances de chaque sujet en fonction de son profil cognitif. Après une première analyse statistique des mesures, un résultat intéressant lié à l'influence combinée de la taille de l'écran et des capacités attentionnelles est apparu. La conclusion est que les sujets présentant de bonnes capacités attentionnelles ont des performances similaires sur les petits et grands écrans alors que les sujets montrant des capacités attentionnelles moindres profitent largement des grands écrans. En effet, les sujets à faibles capacités attentionnelles voient leurs performances significativement augmenter (ils surpassent en moyenne les sujets ayant de meilleures capacités attentionnelles) face à un grand écran. On peut ainsi considérer les grands écrans comme des aides cognitives pour certains sujets.

Nous envisageons d'étendre ce genre d'étude au cas des petits écrans (PDA).

Applications

Théâtre interactif – En collaboration avec une compagnie de théâtre, nous avons travaillé sur la possibilité d'utiliser des environnements virtuels affichés au moyen d'interfaces de visualisation collective, pour remplacer les décors de théâtre classiques. Ce projet a donné

lieu à des "performances" dans lesquelles les acteurs et le public ont voyagé dans des mondes virtuels. Nous avons développé des scénarios permettant une grande interactivité entre le public, les acteurs et les EV. Pour permettre cette interactivité, nous avons utilisé une interface à base de pointeurs lasers d'une part, et le CAT d'autre part. Ce travail est décrit plus en détail dans [604].

Geomarketing 3D – Le géomarketing cherche à représenter des informations stratégiques (financières, économiques...) sur des supports géographiques, principalement sur des secteurs urbains. Nous avons développé une suite logicielle permettant la génération automatique d'environnements urbains 3D à partir de données cadastrales. Ces environnements de grandes dimensions sont générés de manière optimisée pour favoriser une navigation et une interaction en temps réel. Des métaphores de représentation de données abstraites en 3 dimensions ont été explorées.

Logiciels

AutoMNT : Logiciel de reconstruction de MNT à partir de cartes numérisées – AutoMNT est un logiciel de reconstruction de Modèles Numériques de Terrains (MNT) à partir de cartes topographiques numérisées. Un MNT est un champ de hauteurs, constitué d'une grille régulière avec une valeur d'élévation (ou altitude) associée à chaque noeud. Les MNTs sont utilisés dans de nombreuses applications comme les systèmes d'information géographique, l'archéologie, les simulateurs pour l'entraînement, les jeux vidéos, etc.

Dans ce logiciel, l'utilisateur peut charger un carte, y appliquer différents filtres dans le but d'extraire des lignes de niveaux, de leurs affecter leur élévation et d'interpoler ces données afin d'obtenir la valeur d'élévation en chacun des points du modèle. Comme ce processus peut être long et fastidieux, AutoMNT contient des algorithmes qui tentent d'automatiser autant que possible cette chaîne de traitements.

AutoMNT est un logiciel fonctionnant sous Microsoft Windows dont le code source n'est, pour l'instant, pas diffusé. Sur le site web associé du logiciel, une version exécutable limitée est cependant téléchargeable librement.

Contribution à OpenSG – Nous avons développé une structure de donnée d'adjacence améliorée afin de permettre la création efficace de bandes triangulées (*triangle strips*) [544]. La structure de donnée peut être créée d'une manière rapide et robuste à partir d'un maillage quelconque, et sa conception optimisant le cache est explicitement conçu afin de répondre aux requêtes fréquentes pendant la création de bandes triangulées en temps constant telle que retrouver un voisin ou encore retrouver un triangle ayant le plus faible degré.

Une implémentation d'un algorithme de création de bandes triangulées [544] a montré une accélération significative comparée aux implémentations existantes, et nous avons intégré cet algorithme dans la nouvelle version 1.2.0 du graphe de scène OpenSource OpenSG (distribution à partir du 18 mars 2003 - OSGNodeGraph - http://www.opensg.org/doc-1.1.0/classosg_1_1NodeGraph.html).

De plus, nous avons développé une méthode de quantification de normales en 2^{3+2n} différentes directions avec un nombre spécifiable n de subdivisions déterminant la précision. Cette méthode est également intégré dans la version 1.2.0 de OpenSG (OSGNormalQuantifier - http://www.opensg.org/doc-1.1.0/classosg_1_1NormalQuantifier.html).

Contribution à HyperFun – Le projet HyperFun (<http://www.hyperfun.org>) est un logiciel libre pour la modélisation, la visualisation et l'animation de formes 3D appelées "Représentation par Fonction" (F-Rep). Le modèle F-Rep peut être considéré comme une généralisation de la *géométrie constructive* (CSG), mais où une plus grande variété de formes

et d'opérateurs sont disponibles. De nombreuses contributions au projet HyperFun projet, principalement basées sur la texturation [546] et la sculpture constructive, ont été développées dans l'équipe au cours de la thèse de Benjamin Schmitt (co-encadrée par Christophe Schlick et Alexander Pasko, de l'Université de Hosei au Japon), qui se termina en janvier 2003 quand il débuta un Post-Doctorat à Digital Media au Japon. La collaboration avec le projet Hyperfun est toujours active; nous prévoyons de développer un système de rendu basé points pour la visualisation interactive des F-Rep.

CoTeX : outil de Texturation Constructive – En se basant sur la *Texturation Constructive* décrite dans [728], le plug-in *CoTeX*, implémentant cette nouvelle méthode de texturation volumique, a été créé pour la plate-forme PointShop3D.

GLUT|ES – GLUT|ES est une adaptation de la bibliothèque GLUT (OpenGL Utility Toolkit, mise au point par Mark Kilgard pour son livre OpenGL "RedBook") pour les ordinateurs embarqués. Celle-ci est pour l'instant compatible avec les systèmes de type Windows Mobile ou Windows PC. GLUT|ES fonctionne donc sur des terminaux de type PocketPC, Smart-Phone et PC. Un port Symbian est prévu. GLUT|ES est basés sur la bibliothèque de rendu 3D dédiée à ces appareils : OpenGL|ES

Projets et perspectives 2005-2008

Multirésolution and Apparence

Pour la transmission de données 3D géométriques ou leur simplification, la multirésolution est une approche naturelle qui a été utilisée dans de très nombreux travaux. Cependant, la plupart de ces solutions sont très liées aux données géométriques et ne peuvent que difficilement être appliquées pour l'apparence, qui est elle une donnée 4D.

Nous travaillons actuellement sur le développement d'une nouvelle représentation de l'interaction lumineuse avec une surface 3D, représentation qui permettra une meilleure préservation de l'apparence d'un objet lors d'une simplification multirésolution. Avec un tel modèle, la richesse d'une visualisation peut être grandement augmentée, même sur un écran de faible résolution. La variation d'éclairage fournira ainsi des détails existant avant la simplification géométrique.

Interface de Sketching

Sur des terminaux mobiles, le stylet est le mode principal d'interaction. Ainsi, pour la modélisation 3D, la métaphore naturelle à utiliser est celle du dessin. Comme il est difficile de faire des rotations 3D précises à l'aide de ce seul stylet, nous développons actuellement une méthode purement 2D à base de technique de "surface à partir d'éclairage". Avec une telle approche, l'utilisateur devra seulement choisir des directions lumineuses représentatives et dessiner l'éclairage résultant sur l'objet.

De même, une première maquette d'une application de sketching de surfaces implicites variationnelles a été développée permettant de créer des surfaces implicites à partir de simples lignes dessinées et de fusionner plusieurs de ces modèles. Nous pensons maintenant intégrer à cette application des outils des modeleurs 3D afin de rendre le sketching plus performant.

Modélisation d'une flamme en temps réel

Modéliser une flamme qui soit juste aussi bien du point de vue physique que du point de vue du rendu est une tâche très gourmande en temps de calcul. Nous nous intéressons ici à modéliser en temps réel une flamme qui soit le plus proche de la réalité. Ce travail se fait en collaboration avec Flavien Bridault, doctorant au sein du Laboratoire d'Informatique

du Littoral. Nous partons de sa technique de modélisation d'une flamme à base de surface NURBS. Le problème avec sa technique, qui fonctionne bien avec une seule flamme, est que la généralisation à un nombre quelconque de flammes (pour modéliser un feu par exemple) n'est pas évidente. Nous avons donc décidé de modéliser une flamme sous forme d'une surface implicite. En effet, il est beaucoup plus facile de fusionner plusieurs surfaces implicites (et donc plusieurs flammes). Ensuite nous triangulons la surface implicite et, en fonction du vent appliqué sur la flamme, nous déplaçons les sommets de la triangulation pour avoir une flamme qui bouge au gré du vent.

Acquisition de sources de lumière

Mesurer précisément le champ proche émis par un luminaire complet présente encore de nombreuses difficultés. Nous avons développé pour cela une nouvelle approche qui projette en 2 temps ce champ proche dans une base pré-définie. D'abord, la source de lumière est filtrée à l'aide d'un filtre optique passe-bas correspondant à la fonction de notre base, afin de garantir, une mesure de haute-précision. Puis, par un processus numérique, ces images ainsi acquises sont combinées pour finaliser la projection. Nous avons aussi mis en place toute la chaîne d'acquisition afin d'expérimenter le processus, et nous étudions actuellement les différentes erreurs possibles.

Rendu et modélisation d'objets massifs

La gestion efficace d'objets très complexes, tant en terme de temps machine que de place mémoire, est clairement un défi pour la synthèse de scènes 3D réalistes en temps réel. Notre contribution dans ce domaine, sera de proposer un ensemble d'algorithmes visant à représenter au mieux l'apparence d'objets très complexes, i.e. composés de plusieurs centaines de millions d'échantillons. Une première technique, dont le but est de proposer un système de visualisation rapide d'objets complexes, a été développée avec succès au printemps 2005. Cette méthode n'est qu'une instance de notre approche générale, qui consiste à identifier les goulots d'étranglements des algorithmes, et à reporter un maximum de traitements sur d'autres parties de l'algorithme, comme par exemple celles pouvant être traités localement en dimension inférieure, ou bien encore celles supportées par une le matériel. Ainsi, un algorithme de visualisation d'objets massifs a pu être développé, en allégeant la phase de reconstitution géométrique 3D, coûteuse en temps de calcul, et peu flexible, puis en utilisant localement une méthode de reconstruction d'image 2D préservant l'essentiel de l'apparence d'une surface au travers de sa distribution de normales associées. Cette méthode s'est également montrée très convaincante pour la reconstitution des propriétés de couleur d'un objet. Notre méthode se prête particulièrement bien aux nuages de points acquis par scanner laser, et permet de traiter à la volée de telles masses de données. Contrairement aux méthodes existantes, notre approche ne nécessite pas de topologie en entrée, souvent très longues à générer, et autorise ainsi le stockage et la transmission d'objets représentés uniquement au travers de leurs propriétés géométriques, sans topologie explicite.

Raffinement et rendu rapide de surfaces

Nous améliorons les PN Triangles — une méthode de raffinement rapide de surface proposée par ATI en 2001 — en enrichissant un maillage basse résolution d'un certain nombre de paramètres scalaires à ses sommets. Ces paramètres permettent de contrôler précisément la génération à la volée de maillages raffinés, en préservant par exemple les arêtes vives, ou bien encore la courbure du maillage d'origine. Nous couplons à cette approche un système générique de génération *matérielle* de maillage, implémentable sur toute carte graphique programmable, et fournissant une unité virtuelle de raffinement de surface, sur le processeur

graphique. Notre approche ne nécessite à aucun moment le stockage en mémoire de l'objet à haute résolution, ni son transfert sur le bus graphique. Outre nos PN Triangles améliorés, cette unité, implémentée par un simple *vertex program* OpenGL SL, permet de mettre en place toutes sortes de techniques de *displacement mapping*, et est particulièrement bien adapté aux textures de hautes fréquences, utilisables pour déformer seulement des surfaces très raffinées.

Reconstruction automatique de la surface d'un océan à partir d'une vidéo

Le but de ce projet de master recherche est de faire un état de l'art sur la reconstruction géométrique à partir d'une séquences images (Shape From Shading, ...) de nous donner une certaine expertise dans ce domaine. A partir d'une vidéo de l'océan (sans déferlement, ni objets), nous souhaitons reconstruire quasi automatiquement la surface en mouvement. Le but est de pouvoir placer des objets virtuels dans cette vidéo et que son déplacement semble aussi réel que possible.

Réalité augmentée sur terminaux mobiles communicants

Actuellement, les différents terminaux mobiles communicants (PDA, BlackBerry, Smartphone, ...) sont quasiment tous équipés de caméras intégrées. Ces périphériques permettent d'entrevoir de nouveaux types de services basés sur la réalité augmentée (c-à-d de rajouter des informations sous forme d'image synthétique sur les images capturées par la caméra).

Par exemple : - Aide à la navigation dans un bâtiment où la caméra filmerait l'environnement réel, et où le système incrusterait des informations comme des directions géographiques, des informations de distance ou de localisation en utilisant un jeu de marqueurs. - Visite de sites historiques où le visiteur pourrait visualiser sur place l'état des structures dans leur état d'origine grâce à des images de synthèse incrustées sur des images du site en l'état actuel.

La démarche proposée pour ce projet est de commencer par une étude des différentes techniques de réalité augmentée existantes sur gros système ou PC, puis de voir comment les appliquer/adapter sur ces terminaux mobiles (peu de mémoire, peu de processeur, petit écran, etc...). Ensuite, il s'agit de concevoir de nouveaux algorithmes de réalité augmentée spécifiques aux périphériques utilisés (et donc de nouveaux services à proposer aux utilisateurs). La plate-forme MAGELLAN (plate-forme Client-Serveur développée à l'INRIA) devra être à priori utilisée pour mener à bien cette étude.

4.3.4 Structuration et analyse d'image

Introduction

L'analyse d'image a pour but d'extraire des informations qualitatives ou quantitatives d'une image ou d'un ensemble d'images. Cela requiert la mise en œuvre et l'enchaînement d'un certain nombre de processus, comme l'identification des différents objets ou composants de l'image encore appelé *segmentation*, le calcul de leurs relations topologiques ou géométriques, l'estimation de mesures globales ou locales, et le cas échéant la caractérisation de ces objets.

La mise en œuvre de ces différentes étapes requiert l'utilisation de modèles topologiques et/ou géométriques, tout d'abord pour décrire puis analyser les composants images issus de segmentation, mais également pour piloter l'étape de segmentation elle-même. Une difficulté importante consiste alors à disposer de modèles et d'outils qui soient à la fois mathématiquement bien fondés et implémentables avec des complexités en temps et en espace raisonnables. C'est d'autant plus crucial que si la discipline s'est construite historiquement sur l'étude des images monochromes 2D, le champ d'application s'est progressivement étendu aux images définies sur des domaines de dimension supérieure et dont les éléments sont eux-même définis sur plusieurs canaux (trois dans le cas des images couleurs usuelles).

Il devient par conséquent rapidement nécessaire de considérer la topologie de subdivisions de l'espace, d'une part dans le cas général, et d'autre part dans le cas du plongement de ces subdivisions dans la grille discrète qui est le support géométrique naturel pour représenter les images. L'étude des propriétés topologiques de grilles discrètes permet alors de caractériser des courbes et des surfaces nécessaires pour définir et représenter les bords des objets discrets des images.

Dans ce contexte, nous nous intéressons à toute la chaîne de l'analyse d'image, c'est-à-dire à l'élaboration de modèles pour la structuration et la représentation topologique et/ou géométrique des images discrètes, à la mise en œuvre d'outils et de méthodes d'analyse reposant sur ces modèles, à leur implémentation efficace, et enfin à leur utilisation pour des applications réelles comme par exemple l'imagerie médicale.

Une part importante de nos travaux en structuration et analyse d'image se situe par conséquent dans le cadre général de la topologie (étude des subdivisions cellulaires, définition et codage de surfaces) et de la géométrie discrète (géométrie des courbes discrètes, lissage de courbes et de surfaces, reconnaissance, parcours, mesures de longueurs, estimation de courbures et de tangentes). Nous travaillons également à la mise en œuvre effective de ces modèles géométriques et topologiques et à leur implémentation efficace, avec comme application importante le développement d'outils et d'applications de segmentation (représentation des images segmentées, modèles déformables continus et discrets). Certains de ces outils commencent à être disponibles sous la forme de logiciel libre.

Récemment, nous avons proposé des solutions pour utiliser des outils de géométrie discrète dans le contexte de la reconstruction de surface. Nous avons également travaillé sur des problématiques de traitement d'image, soit comme prétraitement à l'analyse comme la correction de défauts d'acquisition, soit comme outil de fabrication d'image.

Enfin, en collaborations avec d'autres thématiques de l'équipe, ou avec d'autres équipes du laboratoire, nous nous intéressons aux possibilités d'appliquer nos méthodes et outils à des problématiques comme l'indexation et le traitement de flux vidéo ou la visualisation d'informations.

Mots-Clés

Représentation d'image. Géométrie discrète. Modèles et invariants topologiques. Courbes

et surfaces discrètes. Segmentation. Modèles déformables. Reconstruction de surfaces. Imagerie médicale. Traitement d'image et de flux vidéo. Couleur. Visualisation d'information.

Description des activités et principaux résultats

Modèles et invariants topologiques

Plusieurs modèles ont été proposés pour décrire la topologie de subdivisions dans le contexte de l'analyse d'image. Dans notre équipe, nous travaillons sur des modèles topologiques représentant des subdivisions cellulaires, pas forcément plongées, comme les ordres, les cartes généralisées, les graphes d'incidence ou les complexes cellulaires. Nous cherchons à comparer leur expressivité mais aussi leur exploitabilité dans le cadre de l'analyse d'image. Dans le même ordre d'idée, nous nous intéressons aussi aux invariants topologiques sur des structures cellulaires de dimension quelconque. Plus spécifiquement, nos recherches portent sur des invariants calculables et comparables comme les groupes d'homologie. Ces invariants donnent des éléments pour caractériser la morphologie des constituants mais ils peuvent aussi indiquer les relations spatiales entre objets. En basses dimensions, ces invariants ont un très fort pouvoir de discrimination, ce qui justifie d'autant plus leur étude dans le cadre de l'analyse d'images. Dans ce contexte, nous avons récemment obtenus des résultats sur :

- la comparaison de modèles topologiques génériques de subdivisions et mise en évidence de leur équivalence : ordres et complexes cellulaires, surfaces dans les ordres et cartes généralisées [554, 553] (ces travaux sont conduits en collaboration avec Pascal Lienhardt (SIC, Poitiers) et Xavier Daragon (A2SI/ESIEE, Marne-la-Vallée)) ;
- le calcul d'un invariant topologique sur les complexes simpliciaux et cellulaires : groupes d'homologie et générateurs [651] (ces travaux sont menés en collaboration avec Laurent Fuchs et Samuel Peltier (SIC, Poitiers)).

Géométrie des courbes discrètes

L'étude de la géométrie des courbes discrètes a pour but principal de fournir des outils pour l'extraction d'informations quantitatives des images segmentées. Dans ce contexte, l'étude de la tangente à une courbe discrète, initiée dans le cadre d'une application au lissage de contours extraits d'une image, nous a conduit à travailler d'un point de vue fondamental sur les propriétés géométriques d'une courbe discrète en deux dimensions. Les courbes discrètes étudiées sont des courbes 4-connexes correspondant au contour inter-pixel d'une forme 2D discrétisée dans la grille entière. Étant donnée une forme discrète, nous ne pouvons qu'approcher les caractéristiques géométriques de la forme continue sous-jacente en nous appuyant sur un ensemble d'hypothèses (courbure bornée du contour par exemple). Notre outil d'analyse privilégié est la décomposition des contours discrets en segments de droites discrètes. Nos contributions récentes à cette problématique portent sur deux aspects :

- L'évolution asymptotique des caractéristiques géométriques d'un contour, utile pour connaître le comportement des modèles lorsque la résolution tend vers zéro [574] ; ce travail qui s'appuie sur les propriétés des segments maximaux le long d'un contour discret a été effectué en collaboration avec Fabien Feschet du laboratoire LLAIC1 de Clermont-Ferrand.
- La définition d'un nouvel estimateur de tangente discrète compatible avec les convexités/concavités de la forme initiale, précis à base résolution et asymptotiquement convergent. [629]

Nous finalisons actuellement une étude complète de la convergence asymptotique d'estimateurs géométriques (tangente et courbure) basés sur le calcul de segments maximaux.

Il est également très intéressant de constater que les segments maximaux permettent la caractérisation de la convexité d'une forme discrète.

Surfaces discrètes : représentation, mesures, transformations

Dans la grille cellulaire discrète ou grille de Khalimsky, les bords des objets discrets peuvent être munis d'une topologie de surface. On parle de surface discrète ou *digital surface*. Nous nous intéressons d'abord à ses aspects fondamentaux : définition algébrique des objets et surfaces, algorithmes d'extraction et de suivi, représentation et structures de données pour surfaces discrètes évolutives, reconstruction et analogues continus. Ces travaux sont ensuite enrichis en les croisant avec les recherches menées en géométrie des courbes discrètes. Cela nous a conduit à développer un ensemble d'applications autour des surfaces discrètes : mesures géométriques sur des objets en dimension quelconque, lissage de surfaces discrètes pour la visualisation, modèles déformables discrets pour la segmentation.

Nous avons développé une bibliothèque de représentation d'objets et surfaces discrets en dimension quelconque. Ce travail s'appuie sur une définition algébrique des cellules discrètes et de leur bord qui permet à la fois un codage compact des surfaces discrètes et une implémentation générique nD des algorithmes de parcours de surfaces discrètes. [625]

Nous avons proposé des outils de mesures géométriques sur les surfaces discrètes par croisement de courbes discrètes, notamment pour l'estimation des normales par extraction d'un ensemble de tangentes discrètes 2D, le calcul de l'aire par intégration de normales, et l'estimation de plan tangent pour lissage. [628]

Nous avons donné une formulation discrète des modèles déformables pour lesquels nous avons réalisés des prototypes dans le cas des images 2D. Le principe retenu pour effectuer la segmentation est celui de la minimisation d'énergie *a posteriori*. Nous avons dans ce contexte obtenus des résultats sur la convergence vers les contours actifs continus pour des résolutions croissantes. Ces travaux ont été effectués en collaboration avec le laboratoire LIRIS (ex ERIC) de Lyon, sous la forme d'un projet Jeune chercheur du GdR ISIS.

Parallèlement à ces travaux, nous avons développé une méthode statistique permettant de positionner un plan tangent en tout point d'une surface discrète. Une application de ce résultat est le lissage géométrique de surfaces discrètes selon le modèle des grilles euclidiennes, développé dans l'équipe il y a cinq ans, et qui permet un traitement à la fois automatique et réversible (la discrétisation de la surface lissée redonne exactement la surface discrète de départ) [738, 619].

Représentation d'images segmentées

La segmentation d'image, c'est-à-dire la décomposition d'images en régions satisfaisant certains critères d'homogénéité, est une étape préalable à la structuration et l'analyse. Que ce soit lors de la phase de segmentation proprement dite ou lors de celle de l'analyse du résultat de cette segmentation, il est nécessaire de disposer de modèles permettant de capturer et modifier les propriétés de la décomposition.

Dans le cas des images 2D, la décomposition d'une image en régions induit un graphe des frontières dont la définition, à cause de la nature discrète des images, n'est pas immédiate. Nous avons développé un modèle basé sur des chemins *interpixel* et des cartes combinatoires permettant de représenter explicitement la géométrie et la topologie de la décomposition. Nous avons implémenté un codage efficace de cette structure permettant d'optimiser les opérations géométriques et topologiques usuelles (localisation de la région contenant un point, recherche de régions voisines ou incluses, reconstruction d'une région, découpes, fusions, etc). Les primitives permettant les modifications incrémentales de ce modèle ont été formalisées. Nous avons réalisé implémentation complète du modèle. Outre le développement

proprement dit, un effort important concerne les spécifications de l'interface fonctionnelle de la bibliothèque, avec la définition d'une interface complète et minimale permettant d'exploiter pleinement les possibilités proposées par le modèle théorique. Cette bibliothèque est depuis peu distribuée sous licence GPL sous le nom de GIRL (*General Image Representation Library*) ; elle est disponible à l'adresse www.girl.labri.fr. Cette distribution contient l'intégralité de la bibliothèque associée au modèle d'images 2D segmentées ainsi qu'un exemple d'utilisation basé sur un travail effectué dans le cadre d'un projet EITICA avec la société DIS.

Dans le cas des images discrètes 3D l'intuition fait souvent défaut et il n'est pas généralement pas trivial d'exprimer correctement des propriétés et des algorithmes. Plus encore qu'en 2D une structure géométrique et topologique décrivant le résultat de la segmentation est indispensable, à la fois pour construire cette segmentation et pour en exploiter le résultat [583]. Le modèle précédent de représentation des images 2D a été étendu aux images 3D au moyen de surfaces inter-voxel et de 3-cartes. Il présente les mêmes intérêts : représentation topologique et géométrique explicite de la décomposition et optimisation des opérations usuelles. Des algorithmes de construction, de mise à jour, et d'exploitation des structures de données ont été élaborés et l'utilisation du modèle pour la segmentation volumique est actuellement en cours de développement.

De manière concomitante, un modèle équivalent basé sur les 3 cartes a été développé par Guillaume Damiand (SIC, FRE CNRS 2731, Poitiers). Une différence essentielle avec notre modèle est que ce dernier utilise une définition implicite de la géométrie des éléments topologiques alors que notre modèle utilise une définition explicite de la topologie. Des travaux ont été menés pour fusionner ces deux modèles en un seul afin de tirer partie des avantages des deux représentations. A l'heure actuelle une maquette basée sur la fusion de ces deux modèles est déjà réalisée.

Méthodes pour la segmentation

Nous avons développé une méthode de segmentation par suivi d'organes dans une image 3D permettant de réutiliser les algorithmes de segmentation connus en 2D afin de les appliquer à la segmentation d'organes tubulaires (réseaux bronchiques, vaisseaux sanguins, etc.) sur des examens tomographiques. Nous sommes en mesure d'extraire d'un organe tubulaire, et après une segmentation grossière, une série de coupes placées perpendiculairement à l'axe local de l'organe. Il nous est aussi possible de détecter et gérer les embranchements dans les organes en question et d'appliquer les algorithmes de segmentation 2D [653, 524] de façon incrémentale, en utilisant le résultat de la segmentation de la coupe précédente sur la coupe courante. Cette approche permet d'optimiser des méthodes comme les contours actifs qui procèdent par affinage. Finalement, il est possible d'utiliser les résultats de la segmentation pour recalculer les plans de coupe afin d'augmenter la précision [734, 572].

L'imagerie médicale et biologique fournit des volumes de données représentatifs de structures cellulaires ou anatomiques. Une problématique associée est la segmentation de ces images et la reconstruction des composantes.

Les modèles déformables, qui approchent les composantes par déformations progressives, constituent une approche souple et robuste à ce problème. Ces déformations sont gouvernées par la minimisation d'une fonction coût ou énergie. On considère que le minimum de cette fonction correspond à la meilleure approximation de la forme recherchée dans l'image. C'est un domaine de recherche en plein essor, très étudié au niveau international dans les communautés vision, traitement du signal ou mathématiques appliquées. Sur cet axe de recherche, nous nous intéressons aux modèles hautement déformables continus. Nous cherchons à équilibrer le nombre de paramètres de ces modèles avec la complexité locale des formes à

extraire dans deux buts : limiter la complexité temps/mémoire et conserver l'expressivité de ces modèles.

Nous avons montré que l'utilisation de géométrie riemannienne offre un cadre puissant, à la fois sur le plan théorique et dans le contexte d'applications effectives, pour optimiser les modèles déformables. En particulier, l'utilisation de métriques non-euclidiennes permet d'adapter automatiquement la densité du maillage sur lequel évolue le modèle déformable, en dilatant les zones d'intérêt et en réduisant les zones loin des contours ou bruitées. Cette approche a été validée en 2D et en 3D [698, 673, 626, 743]. L'analyse de la géométrie locale des images 2D et 3D pour quantifier les zones d'intérêt et en déterminer les directions principales (séparation des forts contours et des zones homogènes ou bruitées) peut être effectuée au moyen d'un outil appelé tenseur de structure. On obtient de cette manière une indépendance entre le nombre de paramètres et la résolution image [627, 539].

Reconstruction de surface

La reconstruction d'une surface tridimensionnelle à partir d'une ou de plusieurs images 2D de cette surface est un problème classique en analyse d'image. On se place ici dans le contexte où l'on essaie de reconstruire la profondeur de la surface à partir des informations d'ombres portées (on parle encore de *shape from shading* ou SFS). Une des difficultés de cette problématique est qu'il existe de nombreuses solutions pouvant être associées à l'image ombrée d'une surface.

Les méthodes existantes sont généralement des méthodes numériques globales qui ne peuvent pas être facilement contrôlées localement. Nous avons développé une méthode locale reposant à la fois sur un outil classique de SFS (la carte de réflectance) et sur des techniques de géométrie discrète. En reconstruisant explicitement la surface et en utilisant des caractéristiques géométriques de cette surface reconstruite, il devient possible d'ajouter des contraintes locales permettant de résoudre les ambiguïtés sans introduire de surcoût de calcul [564, 618, 738, 526]. Il en résulte une méthode très souple à utiliser, particulièrement résistante au bruit, que nous avons utilisée pour des applications en archéologie [620], et pouvant également s'appliquer dans un contexte plus large que le contexte habituel du SFS [621]. Cette approche ouvre des perspectives intéressantes en montrant l'intérêt de l'approche géométrie discrète dans le contexte du SFS.

Imagerie biologique et médicale

Les récentes avancées technologiques dans le domaine de l'acquisition d'images biologiques et médicales en 2D, 3D et 4D font apparaître la nécessité de traitements et d'analyse spécifiques. Le besoin principal est actuellement l'extraction d'informations de ces images. La complexité inhérente à ces images fait que les modèles théoriques actuels doivent être adaptés et que de nouveaux modèles doivent être développés. La région Aquitaine dispose de nombreuses plates-formes d'imagerie performantes : CHU, médecine nucléaire, laboratoires de recherche en physique, biologie, neurobiologie, etc. Ce contexte favorable permet une synergie entre les différentes équipes. De telles collaborations ont déjà été mises en place dans le cadre de l'estimation et du recalage de mouvements (IRM), de l'étude de l'arbre bronchique (scanner CT), du recalage d'IRM cérébrales et d'images fonctionnelles (IRM/SPECT), de la représentation graphique de données biologiques (microscope confocal), et de la reconstruction 3D d'images biologiques.

L'équipe a pu appliquer ses compétences en segmentation d'images dans le cadre d'un projet RNTS dénommé « B3S », en collaboration avec AXS Ingénierie. Le thème était le traitement d'images radiographiques. L'objectif est de reconstruire un modèle tridimensionnel de colonne vertébrale à partir de deux radiographies (face et profil). Nous avons proposé

d'utiliser un modèle déformable 3D paramétrique de la colonne pour résoudre ce problème. Le modèle est contraint à s'adapter aux informations de contour de chaque radiographie. Avec une initialisation manuelle raisonnable, la colonne vertébrale est reconstruite avec une meilleure précision dans la plupart des cas.

Dans le cadre de l'imagerie biologique, nous avons proposé des modèles de construction d'objets biologiques [525, 549, 560, 681, 685]. Nous avons également participé à l'élaboration de méthodes d'aide au diagnostic sur la maladie d'Alzheimer (en collaboration avec le Service de Médecine Nucléaire du CHU Pellegrin) [679].

Nous développons une collaboration particulière avec le laboratoire IMF (Imagerie Moléculaire et Fonctionnelle - ERT CNRS / Université Bordeaux 2) en particulier avec l'encadrement de trois thèses de doctorat dans le domaine du traitement et de l'analyse des images IRM. Ces travaux portent sur les aspects suivants :

- traitement du bruit : il s'agit d'étudier les différents facteurs de bruit (bruit électronique, bruit physiologique) qui induisent une erreur sur l'analyse des IRM fonctionnelles dans l'optique de la fusion de ces données avec les potentiels évoqués (EEG/MEG) [552];
- estimation et recalage du mouvement : il s'agit de traiter le problème de mouvements dus aux cycles cardiaque et respiratoire dans l'étude des organes de la cavité abdominale; nous cherchons à analyser ces mouvements en 2D et en 3D dans le but d'améliorer la précision de l'estimation de la thermométrie par PRF [664, 665, 667];
- modélisation de la dose thermique : nous nous intéressons ici aux moyens d'utiliser l'analyse d'image pour rendre les traitements anti-cancéreux par radiothérapie plus sûrs; il s'agit de modéliser la dose thermique apportée par des appareils de radiothérapie (laser, aiguille radio-fréquence, ultrasons focalisés) et de recalibrer cette modélisation en temps-réel sur les images IRM de thermométrie afin de les comparer et de déterminer la zone de nécrose induite tout en permettant la visualisation de zones d'intérêt *sensibles* (vaisseaux sanguins, paroi d'organes, etc.).

Toujours dans le cadre de cette collaboration avec le laboratoire IMF, nous nous intéressons à l'adaptation de l'algorithme des *marching cubes* (Lorenson, 1987) pour obtenir une visualisation topologiquement et volumétriquement correcte de la surface des différents organes [723], et nous participons à un projet européen sur l'étude de la fonction rénale (IMF (Bordeaux, France), LaBRI (Bordeaux, France), CHU Lyon (Lyon, France), Université d'Aarhus (Aarhus, Danemark), GOSH Hospital (Londres, Grande-Bretagne)). Notre apport à ce projet consiste à adapter les techniques de recalage de mouvement que nous avons développées aux séquences spécifiques d'acquisition d'images de cette étude et à utiliser nos compétences en segmentation pour obtenir une mesure précise du cortex rénal.

Traitement et codage avancés des flux vidéo à bas débits

Si dans le passé nous nous sommes explicitement intéressés au développement des méthodes avancées de codage vidéo basé objet, aujourd'hui nous nous intéressons au traitement vidéo en amont et en aval des codecs normalisés. Ces traitements prendront en compte la sensibilité différentielle du système visuel humain pour éliminer des certains détails dans la scène vidéo à coder. En collaboration avec la société VITEC Multimédia, dans le cadre d'une thèse CIFRE, nous avons développé une méthode de filtrage adaptatif en amont de l'encodage vidéo qui préserve la qualité de la vidéo tout en augmentant le taux de compression. Ces traitements s'appuient sur une analyse de l'activité temporelle et spatiale de la scène à encoder ainsi que de la complexité globale de la scène. Les premiers résultats s'avèrent prometteurs et sont actuellement en phase d'évaluation pour s'intégrer dans l'encodeur temps réel développé par la société VITEC.

Modèles de couleur

Nous avons proposé un nouveau modèle pour reproduire de manière plus fiable les effets de transparence en traitement d'image. Ce modèle peut être vu comme la simplification des modèles réflexion/diffusion classique en synthèse d'image au cas particulier du rendu $2D\frac{1}{2}$ (rendu d'objets plats superposables). Nous avons montré qu'une approche spectrale permet d'obtenir des résultats très fiable en comparant les transparences calculés avec celles obtenues avec des filtres réels pour lesquels les caractéristiques spectrales sont connues [527]. Dans le cas d'images usuelles, c'est-à-dire dont les couleurs sont connues dans l'espace RVB, il est possible d'inférer un spectre associé à chaque couleur.

Collaboration au thème Visualisation d'information

Notre équipe collabore activement à ce projet, notamment sur le développement de nouveaux algorithmes de dessin, sur l'étude des paramètres relevants, etc. Nous avons participé au développement d'une heuristique pour mettre en correspondance des parties de sous arbres quasi équivalent. Nous travaillons actuellement à l'adaptation de cette heuristique à la reconnaissance d'objets dans les bases de données vidéo. A partir de la représentation d'une image segmentée, nous extrayons le RAG (Region adjacency Graph) associé. Puis nous adaptons, l'heuristique de quasi-équivalence pour tenir compte des paramètres extrinsèques de la segmentation. En fonction de cette heuristique, nous déterminons la reconnaissance en fonction de la surface relative mise en correspondance.

Un des problèmes majeur en *visualisation d'information* est la fragmentation des données pour extraire des sous-graphes dont les informations sont similaires. Les méthodes de segmentation développées dans le cadre de l'analyse d'image peuvent s'appliquer dans la fragmentation de graphes valués en considérant un sommet valués comme un pixel. Nous sommes entrain d'étudier l'adaptation des techniques de traitement et de segmentation d'images aux problèmes du traitement des graphes valués et de leur fragmentation.

Projets et perspectives 2005-2008

Modèles topologiques

Nos recherches sur les modèles et invariants topologiques ont permis de mieux caractériser les propriétés combinatoires des partitions d'image. Une perspective naturelle est donc leur utilisation dans le contrôle des fusions dans les pyramides d'images. En effet, dans beaucoup d'applications d'analyse d'image de haut niveau, il est important de disposer d'informations topologiques sur les régions abstraites obtenues. On peut citer : la simple connexité des régions, les relations spatiales entre régions comme l'inclusion. Ces travaux seront poursuivis en collaboration avec Pascal Lienhardt et Guillaume Damiand (SIC, Poitiers), qui ont une forte compétence sur les représentations combinatoires des pyramides.

Modèles géométriques

Nous travaillons actuellement sur des estimateurs géométriques nD prenant en compte les avancées récentes obtenues sur les estimateurs $2D$, notamment l'estimation des normales et le repérage de zones caractéristiques comme les arêtes vives. Disposant alors d'estimateurs géométriques robustes, nous pourrions passer à la réalisation effective de modèles déformables discrets en $3D$, dont le comportement approche le comportement des modèles continus. Parallèlement à ces travaux, une collaboration démarre avec Luc Brun et Stéphanie Jehan-Besson (GREYC, Caen) sur l'utilisation de pyramides dans la recherche de la forme discrète optimale et sur les liens avec les modèles déformables implicites type *level-set*.

Représentation d'images nD

Les modèles théoriques pour représenter et manipuler les subdivisions des images de dimension 3 et plus sont maintenant disponibles. Comme ce fut précédemment le cas en dimension 2, il est maintenant nécessaire de les rendre opérationnels en élaborant des structures de données et des algorithmes permettant une mise en œuvre efficace. Cela permettra alors de développer des applications de segmentation volumique et d'analyse d'image nD, avec potentiellement des applications en imagerie médicale, en indexation, en codage, etc. Nous allons poursuivre dans cette direction, notamment dans le cadre de la collaboration en cours avec Guillaume Damiand (SIC, Poitiers).

Segmentation d'image

Plusieurs approches de la segmentation d'images ont été développées parallèlement dans l'équipe : approche région type division-fusion, approche contour type modèle déformable discret. Une thèse démarre sur la coopération possible de ces deux approches pour de la segmentation d'image 3D. Le principe est d'utiliser les structurations région obtenues par division-fusion à différentes échelles (une pyramide), comme espace de déformation multi-échelle du modèle déformable discret. Comme indiqués précédemment, ces travaux seront effectués en collaboration avec Luc Brun (GREYC, Caen). Nous comptons aussi explorer les liens entre modèles déformables discrets et approches variationnelles level-set de la segmentation avec Stéphanie Jehan-Besson (GREYC, Caen).

Projet Girl

Nous disposons désormais d'une implémentation fiable et optimisée du modèles de représentation géométrique et topologique des partitions des images 2D. L'étape suivante va consister à proposer une bibliothèque générique de segmentation de type division/fusion et à réaliser une implémentation homogène des différentes méthodes, puis à faire collaborer, dans un cadre unifié ces méthodes avec diverses méthodes basées frontières, et notamment les modèles déformables discrets développés aux sein de l'équipe. Cela permettra de disposer sous licence GPL d'une plate-forme générale de segmentation d'image 2D, paramétrable, extensible et spécialisable, et pouvant servir de cadre opérationnel à la plupart des applications d'analyse d'image 2D.

Imagerie biologique et médicale

Plusieurs collaborations ont été initialisées avec différents laboratoires de l'université Bordeaux 2 : IMF (ERT CNRS), Laboratoire de Physiopathologie Mitochondriale (INSERM U688), Laboratoire de Physiopathologie des Réseaux Neuronaux Médullaires (INSERM E358), Laboratoire de Résonance Magnétique des Systèmes Biologiques (UMR5536 CNRS/Université Bordeaux 2)). Elles concernent le développement de chaînes de traitement spécifiques pour les différents modes d'acquisition d'images.

Dans le cadre de l'imagerie biologique, nous projetons d'appliquer les algorithmes de segmentation que nous avons développés à la détection de membranes ou des fibres musculaires, et dans celui de l'imagerie médicale, d'explorer de nouvelles méthodes de recalage des images (inter- ou intra-patient, inter- ou intra-modalité), de segmentation et de quantification d'image (aide au diagnostic).

Enfin, une collaboration vient de se mettre en place avec le GRIMAAG (EA 3590 - Groupe de Recherche en Informatique et Mathématiques Appliquées des Antilles-Guyane). Dans un premier temps, un transfert de technologie entre notre équipe et la jeune équipe image du GRIMAAG est en cours et dans un second temps, cette collaboration se poursuivra avec le développement de traitement spécifique dédié à l'aide au diagnostic de la maladie d'Alzheimer

atypique.

Traitement d'image

Nous allons poursuivre le développement du modèle de transparence pour traitement d'images couleurs, notamment en cherchant à améliorer et optimiser le passage du mode RGB au mode spectral dans ce contexte. Nous souhaitons également développer des techniques de traitements d'image basées région, reposant à la fois sur les techniques de segmentation mixte (région/contour) développés dans le cadre du projet *Girl*, et sur des outils de mesure et de lissage développés dans le cadre de nos travaux en géométrie discrète.

Collaboration au thème Visualisation d'information

Une heuristique de quasi-isomorphisme d'arbre a été développée par le thème Visualisation d'information. Étant donné deux arbres, elle permet de détecter les sous-arbres quasi-isomorphes présents dans ces deux arbres. Nous allons utiliser cette heuristique pour détecter les objets communs dans des images. Un autre axe de recherche est la transposition des méthodes de segmentation développées en image à la segmentation de graphe afin de développer de nouvelles de méthodes de fragmentation de graphes.

4.4 Publications de l'équipe

4.4.1 Publications d'audience internationale

Revues avec Comité de lecture

- [523] BENOIS-PINEAU (J) et NICOLAS (H). – A new method for region-based depth ordering in a video sequence : Application to frame interpolation. *Journal of Visual Communication and Image Representation*, vol. 13, n3, September 2002.
- [524] BERGER (P), PEROT (V), DESBARATS (P), TUNON-DE-LARA (J.-M), MARTHAN (R) et LAURENT (F). – Airway thickening in cigarette smokers : Quantitative thin-section ct assessment. *Radiology*, June 2005. – To appear.
- [525] BEURTON-AIMAR (M), LUDINARD (S), KORZENIEWSKI (B), MAZAT (J.-P) et NAZARET (C). – Virtual mitochondria : Metabolic modelling and control. *Journal on Molecular and Cellular Biology*, vol. 29, n1-2, 2002, pp. 227–232.
- [526] BRAQUELAIRE (A) et KERAUTRET (B). – Reconstruction of lambertian surfaces by discrete equal height contours and regions propagation. *Image and Vision Computing*, vol. 23, n2, 2005, pp. 177–189.
- [527] BRAQUELAIRE (A) et STRANDH (R). – A color model for rendering linear passive graphic 2d objects. *Graphical models*, vol. 66, n2, 2004, pp. 71–88.
- [528] BRUN (L.), DOMENGER (J.-P) et MOKHTARI (M). – Incremental modifications of segmented image defined by discrete maps. *Journal of Visual Communication and Image Representation*, no14, 2003, pp. 251–290.
- [529] DE LA RIVIÈRE (J.-B) et GUITTON (P). – Model-based video tracking for gestural interaction. *Springer Virtual Reality*, 2005. – Special Issue on Language, Speech, and Gesture in Virtual Reality.
- [530] DELEST (M), DON (A) et BENOIS-PINEAU (J). – Dag-based visual interfaces for navigation in indexed video content. *Multimedia Tools and Applications*, vol. MTAP 149R1, 2006. – To appear.
- [531] DESAINTE-CATHERINE (M), KURTAG (G), MARCHAND (S), SEMAL (C) et HANNA (P). – Playing With Sounds as Playing Video Games. *ACM Journal : Computers in Entertainment*, vol. 2, n2, April/June 2004, p. 16. – (22 pages).
- [532] GOESELE (M), GRANIER (X), HEIDRICH (W) et SEIDEL (H.-P). – Accurate light source acquisition and rendering. *Transactions on Graphics (Proceedings of ACM SIGGRAPH)*, vol. 22, n3, 2003.
- [533] GRANIER (X) et DRETTAKIS (G). – A final reconstruction framework for an unified global illumination algorithm. *ACM Transaction on Graphics (TOG)*, vol. 23, n2, 2004.
- [534] GRANIER (X) et HEIDRICH (W). – A simple layered rgb brdf model. *Graphical Models*, vol. 65, n4, 2003.
- [535] HACHET (M), GUITTON (P), REUTER (P) et TYNDIUK (F). – The cat for efficient 2d and 3d interaction as an alternative to mouse adaptations. *ACM Transactions on Graphics*, vol. 23, n3, 2004, pp. 731–731.
- [536] HACHET (M), GUITTON (P), REUTER (P) et TYNDIUK (F). – The cat for efficient 2d and 3d interaction as an alternative to mouse adaptations. *Transactions on Graphics (Proceedings of ACM SIGGRAPH)*, vol. 23, n3, 2004. – Invited paper, reprise session of VRST 2003.

- [537] HANNA (P) et DESAINTE-CATHERINE (M). – Cnss model : a statistical and spectral model for representing noisy sounds with short-time sinusoids. *EURASIP Journal on Applied Signal Processing*, 2005.
- [538] KRAEMER (P), BENOIS-PINEAU (J) et DOMENGER (J.-P). – Scene similarity measure for video content segmentation in the framework of rough indexing paradigm. *International Journal of Intelligent Systems, special issue on Intelligent Multimedia Retrieval*, 2005. – To appear.
- [539] LACHAUD (J.-O) et TATON (B). – Deformable model with a complexity independent from image resolution. *Computer Vision and Image Understanding*, 2005. – Accepted. To appear.
- [540] MAHBOUBI (A), BENOIS-PINEAU (J) et BARBA (D). – Tracking of objects in video scenes with time varying content. *EURASIP Journal of Applied Signal Processing (JASP)*, vol. 2002, n6, June 2002, pp. 582–594.
- [541] PASKO. (A), ADZHIEV (V), SCHMITT (B) et SCHLICK (C). – Constructive hypervolume modeling. *Graphical Models Journal, Special issue on volume modeling*, vol. 64, n2, 2002.
- [542] PASKO (A), ADZHIEV (V), SCHMITT (B) et SCHLICK (C). – Constructive hypervolume modeling. *Graphical Models*, vol. 64, n2, 2002. – Special Issue on Volume Modeling.
- [543] REUTER (P), BEHR (J) et ALEXA (M). – An improved adjacency data structure for fast triangle stripping. *Journal of Graphics Tools*, 2005. – To appear.
- [544] REUTER (P), BEHR (J) et ALEXA (M). – An improved adjacency data structure for efficient triangle stripping. *Journal of Graphics Tools*, 2005. – à paraitre.
- [545] RIVIÈRE (J.-B. D. L) et GUITTON (P). – Model-based video tracking for gestural interaction. *Springer Virtual Reality*, 2005. – Special Issue on Language, Speech, and Gesture in Virtual Reality.
- [546] SCHMITT (B), PASKO (A) et SCHLICK (C). – Constructive sculpting of heterogeneous volumetric objects using trivariate b-splines. *The Visual Computer*, vol. 20, n2, 2004.
- [547] TOBOR (I), REUTER (P) et SCHLICK (C). – Scalable implicit surface reconstruction from unorganized point sets using partition of unity radial basis functions. *Graphical Models*, 2005. – To appear.
- [548] TOBOR (I), REUTER (P) et SCHLICK (C). – Scalable implicit surface reconstruction from unorganized point sets using partition of unity radial basis functions. *Graphical Models*, 2005.
- [549] TRINEI (M), VANNIER (J.-P), BEURTON-AIMAR (M) et NORRIS (V). – A hyperstructure approach to mitochondria. *Molecular biology*, vol. 53, n1, 2004, pp. 41–53.
- [550] WEIDENSTEINER (C), KERIOUI (N), QUESSON (B), SENNEVILLE (B. D. D), TRILLAUD (H) et MOONEN (C). – Stability of real-time mr temperature mapping in healthy and diseased human liver. *J. Magn. Reson. Imaging*, vol. 19, n4, Apr 2004, pp. 438–446.

Colloques avec Comité de programme et Actes

- [551] ADZHIEV (V), KARTASHEVA (E), KUNII (T), PASKO (A) et SCHMITT (B). – Cellular-functional modeling of heterogeneous objects. In : *Proc. Of ACM Solid Modeling Symposium 2002, Saarbrücken*. – june 2002.

- [552] AGUERRE (C), DESBARATS (P), DILHARREGUY (B) et MOONEN (C). – 3D animation of cerebral activity using both spatial and temporal fMRI information. *In : IEEE Proceedings of the Fourth International Conference on 3-D Digital Imaging and Modeling*, pp. 103–109. – 2003.
- [553] ALAYRANGUES (S), DARAGON (X), LACHAUD (J.-O) et LIENHARDT (P). – Equivalence between regular n-g-maps and n-surfaces. *In : Proc. Int. Work. Combinatorial Image Analysis (IWCIA'2004), Auckland, New Zealand, December 1-3*, éd. par Klette (R) et Zunic (J). – Elsevier, 2004.
- [554] ALAYRANGUES (S) et LACHAUD (J.-O). – Equivalence between order and cell complex representations. *In : Proc. Computer Vision Winter Workshop (CVWW), Bad Aussee, Austria*, éd. par Wildenauer (H) et Kropatsch (W), pp. 222–233. – feb 2002.
- [555] AUBER (D), DELEST (M), DOMENGER (J.-P), FERRARO (P) et STRANDH (R). – Evat : Environment for visualization and analysis of trees. *In : Proceedings of the IEEE Symposium on Information Visualization*. – Seattle, Washington, October 2003.
- [556] AUBER (D), DELEST (M), FEDOU (J), DOMENGER (J) et DUCHON (P). – New strahler numbers for rooted plane trees. *In : Third Colloquium on Mathematics and Computer Science Algorithms, Trees, Combinatorics and Probabilities*, pp. 203–216. – 2004.
- [557] BENOIS-PINEAU (J), BRAQUELAIRE (A) et ALI-MHAMMAD (A). – Interactive fine object-based segmentation of generic video scenes for object-based indexing. *In : WIAMIS'2003*. pp. 200–203. – in Digital Media Processing for Multimedia Interactive Services, 2003.
- [558] BENOIS-PINEAU (J), DESAINTE-CATHERINE (M) et LOUIS (N). – A method for extraction of audio-visual leitmotif in movies by cross media analysis. *In : EUSIPCO 2002*, pp. 3485–3487. – 2002.
- [559] BENOIS-PINEAU (J) et KHRENNIKOV (A). – Image segmentation in compressed domain by clustering methods with euclidean and p-adic metrics. *In : Invited paper at Special Session on Datamining and Multimedia Systems, DIPMU'2002*, pp. 969–976. – Annecy, France, 2002.
- [560] BEURTON-AIMAR (M), PÉRÈS (S), PARISEY (N) et MAZAT (J.-P). – How to lock a model from a miscomposition of objects? *In : 13th European Conference on Object-Oriented Programming*. – Darmstadt, July 2003.
- [561] BEURTON-AIMAR (M), PÉRÈS (S), PARISEY (N), NAZARET (C) et MAZAT (J.-P). – Modeling biologic networks to use them with heterogeneous treatments. *In : Modelling and Simualtion of biological processes in the context of genomics*. Dieppe. – 12-16 May 2003.
- [562] BOUBEKEUR (T), REUTER (P) et SCHLICK (C). – Visualization of point-based surfaces with locally reconstructed subdivision surfaces. *In : Proceedings of Shape Modeling International (SMI 2005)*. – 2005.
- [563] BRAQUELAIRE (A), DAMIAND (G), DOMENGER (J) et VIDIL (F). – Comparison and convergence of two topological models for 3D image segmentation. *In : Proc. of IAPR International Workshop GbRPR 2003*, éd. par HANCOCK (E) et VENTO (M). pp. 59–70. – Springer, 2003.
- [564] BRAQUELAIRE (A) et KERAUTRET (B). – Reconstruction of discrete surfaces from shading images by propagation of geometric features. *In : Proc of DGCI'03*, éd. par Nyström (I), di Baja (G. S) et Svensson (S). pp. 257–266. – Springer, 2003.

- [565] CARMINATI (L) et BENOIS-PINEAU (J). – Gaussian mixture classification for moving object detection in video surveillance environment. *In : IEEE International Conference on Image Processing (ICIP) 2005*. – Italie (Gènes), 2005.
- [566] CARMINATI (L) et BENOIS-PINEAU (J). – Support vector tracking of human faces with affine motion models. *In : Workshop on Image Analysis for Multimedia Interactive Services (WIAMIS) 2005*. – Suisse (Montreux), 2005.
- [567] CARMINATI (L), BENOIS-PINEAU (J) et GELGON (M). – Human detection and tracking for video surveillance applications in low density environment. *In : SPIE VCIP'2003*, pp. 51–60. – 2003.
- [568] CIEUTAT (J.-M), GONZATO (J.-C) et GUITTON (P). – Navigation training simulation in ocean waves. *In : Sim'Ouest 2002, European Conference on Marine Technology : Industry and Simulation*. – 2002.
- [569] CIEUTAT (J.-M), GONZATO (J.-C) et GUITTON (P). – Wind sea and swell models with ship considerations for maritime training simulator. *In : Proc. of SIMOUEST'2002, Nantes*. – november 2002.
- [570] CIEUTAT (J.-M), GONZATO (J.-C) et GUITTON (P). – A general ocean waves model for ship design. *In : Proceedings of Virtual Concept*. – novembre 2003.
- [571] CIEUTAT (J.-M), GONZATO (J.-C) et GUITTON (P). – A general ocean waves model for ship design. *In : Proceedings of Virtual Concept*. – 2003.
- [572] DE DIETRICH (G) et BRAQUELAIRE (A). – A framework for tubular organs segmentation. *In : Proc. of the 12th WSCG*. WSCG, pp. 41–44. – 2004. ISSN 1213-6972.
- [573] DE LA RIVIÈRE (J.-B) et GUITTON (P). – Image-based analysis for model-based tracking. *In : Proceedings of Mirage*. – 2005.
- [574] DE VIEILLEVILLE (F), LACHAUD (J.-O) et FESCHET (F). – Maximal digital straight segments and convergence of discrete geometric estimators. *In : Proc. 14th Scandinavian Conference on Image Analysis (SCIA'2005), Joensuu, Finland*. – Springer, 2005. To appear.
- [575] DEDIEU (S), GUITTON (P), SCHLICK (C) et REUTER (P). – Reality : an interactive reconstruction tool of 3d objects from photographs. *In : Actes de Vision Modeling and Visualization'2001*, pp. 195–202. – 2001.
- [576] DELEST (M), DON (A) et BENOIS-PINEAU (J). – Graph-based visual interfaces for navigation in indexed video content. *In : Proceedings of Content-Based Multimedia Indexing 2003 (CBMI'03)*, éd. par Gabbouj (M). pp. 49–55. – Rennes, France, Septembre 2003.
- [577] DELEST (M), DON (A) et BENOIS-PINEAU (J). – Intuitive color-based visualization of multimedia content as large graphs. *In : Proceedings of Electronic Imaging 2004*. pp. 65–74. – San Jose, CA, USA, Janvier 2004.
- [578] DEPAULIS (F), COUTURE (N), GARREAU (L) et LEGARDEUR (J). – A reusable methodology based on filters in order to define relevant tangible parts for a tui. *In : Engineering Reality of Virtual Reality 2005*. part of IS&T/SPIE Symposium on Electronic Imaging 2005, San Jose (Californie, USA). – 2005.
- [579] DESAINTE-CATHERINE (M) et BROUSSE (N). – Towards a specification of musical interactive pieces. *In : Proc. of the CIM XIX, Firenze, Italy*. – May 2003.
- [580] DESAINTE-CATHERINE (M) et D'ALLOMBERT (A). – Specification of Temporal Relations Between Interactive Events. *In : Proceedings of the Sound and Music Computing Conference (SMC'04)*. – Paris, France, October 2004.

- [581] DESAINTE-CATHERINE (M) et KURTAG (G). – La pédagogie de l'interactivité. *In : conférence du 10e anniversaire de l'ESCOM (European Society of the Cognitive Sciences of Music), Liège, Belgique.* – 2002.
- [582] DESAINTE-CATHERINE (M) et KURTAG (G). – Spécification de la synthèse sonore appliquée à la pédagogie musicale. *In : Proceedings of the Journées d'Informatique Musicale (JIM'03), Montbéliard, France.* – 2003.
- [583] DESBARATS (P) et DOMENGER (J.-P). – Retrieving and using topological characteristics from 3D discrete images. *In : Proceedings of the 7th Computer Vision Winter Workshop.* pp. 130–139. – PRIP-TR-72, 2002.
- [584] DESBARATS (P) et GUEORGUIEVA (S). – Topological mainframe for numerical representations of objects. *In : Proceedings of the International Conference on Computational Science and its Applications 2003.* pp. 498–507. – Springer, 2003.
- [585] DESBARATS (P) et GUEORGUIEVA (S). – CW complexes : Topological Mainframe for Numerical Representations of Objects. *In : International Conference on Computational Science and its Applications,* pp. 498–507. – 2003.
- [586] DON (A), CARMINATI (L) et BENOIS-PINEAU (J). – Detection of visual dialog scenes in video content based on structural and semantic features. *In : International Workshop on Content-based Multimedia Indexing (CBMI) 2005.* – Létonie (Tampere), 2005.
- [587] DUPUY (W), BENOIS-PINEAU (J) et BARBA (D). – 1-d mosaics as a tool for structuring and navigation in digital video content. *In : VLBV'2003.* – Madrid, Spain, 2003.
- [588] GARREAU (L) et COUTURE (N). – Study of tangible user interface for handling tridimensionnal objects. *In : Proc. Real World User Interfaces, PI'2003, Udine (Italy),* pp. 64–68. – September 2003.
- [589] GARREAU (L), LEGARDEUR (J) et COUTURE (N). – Tangible interface for mechanical CAD parts assembly. *In : proceedings of Virtual Concept 2003, Editors D. Coutellier, X. Fischer, D. Marquis, I. Thouvenin, Biarritz 2003. ISBN 2-9514772-3-6,* pp. 222–227. – novembre 2003.
- [590] GIRIN (L), FIROUZMAND (M) et MARCHAND (S). – Long Term Modeling of Phase Trajectories within the Speech Sinusoidal Model Framework. *In : Proceedings of the INTERSPEECH - 8th International Conference on Spoken Language Processing (ICSLP'04).* – Jeju Island, Korea, October 2004.
- [591] GIRIN (L) et MARCHAND (S). – Watermarking of Speech Signals Using the Sinusoidal Model and Frequency Modulation of the Partial. *In : Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP'04).* Institute of Electrical and Electronics Engineers (IEEE). – Montreal, Quebec, Canada, May 2004.
- [592] GIRIN (L), MARCHAND (S), MARTINO (J. D), ROEBEL (A) et PEETERS (G). – Comparing the Order of a Polynomial Phase Model for the Synthesis of Quasi-Harmonic Audio Signals. *In : Proceedings of the IEEE Workshop on Applications of Signal Processing to Audio and Acoustics (WASPAA'03).* Institute of Electrical and Electronics Engineers (IEEE). – New Paltz, New York, USA, October 2003.
- [593] GONZATO (J.-C) et PONT (B). – A phenomological representation of iridescent butterfly wings. *In : WSCG (Winter School of Computer Graphics).* – 2004.

- [594] GRANIER (X), GOESELE (M), HEIDRICH (W) et SEIDEL (H.-P). – Interactive visualization of complex real-world light sources. *In : Proceedings of Pacific Graphics 2003*. IEEE. – octobre 2003.
- [595] GRIVET (S), AUBER (D), DOMENGER (J) et MELANCON (G). – Bubble tree drawing algorithm. *In : International Conference on Computer Vision and Graphics*, éd. par Verlag (S), pp. 633–641. – 2004.
- [596] HACHET (M) et GUITTON (P). – The interaction table : a new input device designed for interaction in immersive large display environments. *In : Proc. of Eurographics Workshop on Virtual Environments, Barcelone*. – may 2002.
- [597] HACHET (M) et GUITTON (P). – Using virtual reality for "new clowns". *In : Proceedings of International Conference on Virtual Storytelling*. – Springer-Verlag, 2003.
- [598] HACHET (M) et GUITTON (P). – The cat - when mice are not enough. *In : Proceedings of IEEE VR 2004 Workshop : Beyond Glove and Wand Based Interaction*. – 2004.
- [599] HACHET (M), GUITTON (P), REUTER (P) et TYNDIUK (F). – The CAT for efficient 2D and 3D interaction as an alternative to mouse adaptations. *In : Proceedings of ACM Virtual Reality Software and Technology (VRST'03)*, pp. 205–212. – 2003. Best paper award.
- [600] HACHET (M), REUTER (P) et GUITTON (P). – Camera viewpoint control with the Interaction Table. *In : Proceedings of Virtual Reality International Conference (VRIC 2003)*, pp. 41–47. – 2003.
- [601] HACHET (M), RIVIÈRE (J.-B. D. L) et GUITTON (P). – Interaction in large-display vr environments. *In : Proceedings of Virtual Concept*. – novembre 2003.
- [602] HACHET (M), DE LA RIVIÈRE (J.-B) et GUITTON (P). – Interaction in large-display vr environments. *In : Proceedings of Virtual Concept*. – 2003.
- [603] HACHET (M) et GUITTON (P). – The interaction table - a new input device designed for interaction in immersive large display environments. *In : Proceedings of Eighth Eurographics Workshop on Virtual Environments, (EGVE 2002)*. – ACM, 2002.
- [604] HACHET (M) et GUITTON (P). – Using virtual reality for new clowns. *In : LNCS Springer-Verlag 2897, proceedings of International Conference on Virtual Storytelling (ICVS)*. – Springer, 2003.
- [605] HACHET (M), GUITTON (P), REUTER (P) et TYNDIUK (F). – The cat for efficient 2d and 3d interaction as an alternative to mouse adaptations. *In : Proceedings of Virtual Reality Software and Technology, (VRST 2003) BEST PAPER AWARD*. – ACM, 2003. best paper award.
- [606] HACHET (M) et KITAMURA (Y). – 3d interaction with and from handheld computers. *In : Proceedings of IEEE VR 2005 Workshop : New Directions in 3D User Interfaces*. – 2005.
- [607] HACHET (M), POUDEROUX (J) et GUITTON (P). – A camera-based interface for interaction with mobile handheld computers. *In : Proceedings of I3D'05 - Symposium on Interactive 3D Graphics and Games*. – ACM Press, 2005.
- [608] HACHET (M), POUDEROUX (J), GUITTON (P) et GONZATO (J.-C). – Tangimap - a tangible interface for visualization of large documents on handheld computers. *In : Proceedings of Graphics Interface*. – 2005.

- [609] HACHET (M), REUTER (P) et GUITTON (P). – Camera viewpoint control with the interaction table. *In : Proceedings of Virtual Reality International Conference (VRIC 2003)*. – 2003.
- [610] HANNA (P), BEURIVÉ (A) et DESAINTE-CATHERINE (M). – Real-time noise synthesis with control of the spectral density. *In : Proceedings of the Digital Audio Effects (DAFx'02) Conference*. University of the Federal Armed Forces, pp. 151–156. – Hamburg, Germany, September 2002.
- [611] HANNA (P) et DESAINTE-CATHERINE (M). – Adapting the overlap-add method to the synthesis of noise. *In : Proceedings of the Digital Audio Effects (DAFx'02) Conference*. University of the Federal Armed Forces, pp. 101–104. – Hamburg, Germany, September 2002.
- [612] HANNA (P) et DESAINTE-CATHERINE (M). – Detection of sinusoidal components in sounds using statistical analysis of intensity fluctuations. *In : Proceedings of International Computer Music Conference (ICMC'02)*, pp. 100–103. – September 2002.
- [613] HANNA (P) et DESAINTE-CATHERINE (M). – Influence de la densité spectrale sur la synthèse de sons bruités. *In : Proceedings of the Journées d'Informatique Musicale (JIM'02)*, pp. 17–24. – Marseille, France, May 2002.
- [614] HANNA (P) et DESAINTE-CATHERINE (M). – Analysis method to approximate the spectral density of noises. *In : Proceedings of the IEEE Workshop on Applications of Signal Processing to Audio and Acoustics (WASPAA'03)*. Institute of Electrical and Electronics Engineers (IEEE). – New Paltz, New York, USA, October 2003.
- [615] HANNA (P) et DESAINTE-CATHERINE (M). – Time scale modification of noises using a spectral and statistical model. *In : Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP'03)*. Institute of Electrical and Electronics Engineers (IEEE). – Hong Kong, China, 2003.
- [616] HANNA (P), LOUIS (N), DESAINTE-CATHERINE (M) et BENOIS-PINEAU (J). – Audio features for noisy sound segmentation. *In : Proceedings of the 5th International Conference on Music Information Retrieval (ISMIR'04)*. Audiovisual Institute, Universitat Pompeu Fabra, pp. 120–124. – Barcelona, Spain, October 2004.
- [617] KEILER (F) et MARCHAND (S). – Survey on Extraction of Sinusoids in Stationary Sounds. *In : Proceedings of the Digital Audio Effects (DAFx'02) Conference*. University of the Federal Armed Forces, pp. 51–58. – Hamburg, Germany, September 2002.
- [618] KERAUTRET (B). – A robust discrete approach for shape from shading and photometric stereo. *In : Proceedings of the International Conference on Computer Vision and Graphics*, éd. par Kluwer. – Varsawa, Poland, 2004. Distinction of the best student presentation award.
- [619] KERAUTRET (B) et BRAQUELAIRE (A). – A statistical approach for geometric smoothing of discrete surfaces. *In : Proceeding of the International Conference on Discrete Geometry for Computer Imagery*, éd. par Andres (E), Damiand (G) et Lienhardt (P). pp. 405–414. – Springer-Verlag, April 2005.
- [620] KERAUTRET (B), BRAQUELAIRE (A), D'ERRICO (F) et VANHAEREN (M). – Discrete geometry for parameters extraction and shape reconstruction in archeology. *In : Colloque GMPCA, Archeometrie*, p. 164. – 2003.
- [621] KERAUTRET (B), GRANIER (X) et BRAQUELAIRE (A). – Intuitive shape modeling by shading design. *In : Proceedings of the 5th International Symposium on Smart Graphics*. pp. 163–174. – Springer-verlag, 2005.

- [622] KRAEMER (P), BENOIS-PINEAU (J) et DOMENGER (J.-P). – Scene similarity measure for video content segmentation in the framework of rough indexing paradigm. *In : 2nd International Workshop on Adaptive Multimedia Retrieval (AMR 2004)*, pp. 141–155. – August 2004.
- [623] KRAEMER (P), HADAR (O), BENOIS-PINEAU (J) et DOMENGER (J.-P). – Super-resolution mosaicing from mpeg compressed video. *In : International Conference on Image Processing (ICIP)*. – September 2005. To appear.
- [624] KRIVANEK (J), ZARA (J) et BOUATOUCH (K). – Fast depth of field rendering with surface splatting. *In : Proceedings of Computer Graphics International 2003*. – 2003.
- [625] LACHAUD (J.-O). – Coding cells of digital spaces : a framework to write generic digital topology algorithms. *In : Proc. Int. Work. Combinatorial Image Analysis (IWCIA'2003)*, Palermo, Italy, éd. par Del Lungo (A), Di Gesù (V) et Kuba (A). – Elsevier, 2003.
- [626] LACHAUD (J.-O) et TATON (B). – Deformable model with adaptive mesh and automated topology changes. *In : Proc. 4th int. Conf. 3-D Digital Imaging and Modeling (3DIM'2003)*, Banff, Alberta, Canada, éd. par Rioux (M), Boulanger (P) et Godin (G). – IEEE Computer Society Press, 2003. To appear.
- [627] LACHAUD (J.-O) et TATON (B). – Resolution independent deformable model. *In : Proc. 17th int. Conf. on Pattern Recognition (ICPR'2004)*, Cambridge, United Kingdom, 23-26 August. pp. 237–240. – IEEE Computer Society Press, 2004.
- [628] LACHAUD (J.-O) et VIALARD (A). – Geometric measures on arbitrary dimensional digital surfaces. *In : Proc. Int. Conf. Discrete Geometry for Computer Imagery (DGCI'2003)*, Napoli, Italy, éd. par Sanniti di Baja (G), Svensson (S) et Nyström (I). – Springer, 2003.
- [629] LACHAUD (J.-O), VIALARD (A) et DE VIEILLEVILLE (F). – Analysis and comparative evaluation of discrete tangent estimators. *In : Proc. Int. Conf. Discrete Geometry for Computer Imagery (DGCI'2005)*, Poitiers, France, éd. par Andrès (E), Damiand (G) et Lienhardt (P). pp. 140–251. – Springer, 2005.
- [630] LAGRANGE (M), MARCHAND (S), RASPAUD (M) et RAULT (J.-B). – Enhanced Partial Tracking Using Linear Prediction. *In : Proceedings of the Digital Audio Effects (DAFx'03) Conference*. Queen Mary, University of London, pp. 141–146. – London, United Kingdom, September 2003.
- [631] LAGRANGE (M), MARCHAND (S) et RAULT (J.-B). – Sinusoidal Parameter Extraction and Component Selection in a Non Stationary Model. *In : Proceedings of the Digital Audio Effects (DAFx'02) Conference*. University of the Federal Armed Forces, pp. 59–64. – Hamburg, Germany, September 2002.
- [632] LAGRANGE (M), MARCHAND (S) et RAULT (J.-B). – Partial Tracking Based on Forward Path Exploration. *In : Proceedings of the 116-th Convention of the AES*. Audio Engineering Society (AES). – Berlin, Germany, May 2004.
- [633] LAGRANGE (M), MARCHAND (S) et RAULT (J.-B). – Using Linear Prediction to Enhance the Tracking of Partial. *In : Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP'04)*. Institute of Electrical and Electronics Engineers (IEEE). – Montreal, Quebec, Canada, May 2004.
- [634] LAGRANGE (M), MARCHAND (S) et RAULT (J.-B). – Tracking Partial for the Sinusoidal Modeling of Polyphonic Sounds. *In : Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP'05)*. Institute of

- Electrical and Electronics Engineers (IEEE). – Philadelphia, New York, USA, March 2005.
- [635] LALES (C), PARISEY (N), MAZAT (J.-P) et BEURTON-AIMAR (M). – Simulation of mitochondrial metabolism using multi-agents system. *In : 4th International Conf. on Autonomous Agents and Multiagent Systems, W. MAS*BIOMED*. Utrecht, The Netherlands. – 25-29th of July, 2005.
- [636] LEGARDEUR (J), GARREAU (L) et COUTURE (N). – Des interacteurs pour l'assemblage mécanique en CAO. *In : 3rd International Conference : Integrated Design and Production, CPI'2003, Meknes, Maroc*. – 2003.
- [637] LEGARDEUR (J), GARREAU (L) et COUTURE (N). – Experiments to evolve toward a tangible user interface for cad parts assembly. *In : Proc. Electronic Imaging Science and Technology, Stereoscopic Displays and Virtual Reality Systems XI, Editors : Andrew J. Woods, John O. Merrit, Stephen A. Benton, Mark T. Bolas, SPIE, EI'2004 SPIE, San Jose (US)*, pp. 438–445. – 18-22 january 2004.
- [638] LEVET (F), REUTER (P) et SCHLICK (C). – Anisotropic sampling for differential point rendering of implicit surfaces. *Journal of WSCG 2005*, vol. 13, 2005. – To appear.
- [639] LEVET (F), HADIM (J), REUTER (P) et SCHLICK (C). – Anisotropic sampling for differential point rendering of implicit surfaces. *In : WSCG (Winter School of Computer Graphics)*. – 2005.
- [640] LOUIS (N) et HANNA (P). – Statistical classification of natural noisy sounds. *In : CBMI 2005*. – 2005. To appear.
- [641] MANERBA (F), BENOIS-PINEAU (J) et LEONARDI (R). – Extraction of foreground objects from mpeg2 video stream in rough indexing framework. *In : Proc. EI'2004, Storage and Retrieval Methods and Applications for Multimedia 2004*. – San Jose, CA, USA, 2004.
- [642] MANERBA (F), LEONARDI (R) et BENOIS-PINEAU (J). – Real-time extraction of foreground objects in a rough indexing framework. *In : WIAMIS'2005*. – Montreux, SW, 2005.
- [643] MARCHAND (S) et RASPAUD (M). – Enhanced Time-Stretching Using Order-2 Sinusoidal Modeling. *In : Proceedings of the Digital Audio Effects (DAFx'04) Conference*. Federico II University of Naples, pp. 76–82. – Naples, Italy, October 2004. ISBN : 88-901479-0-3.
- [644] MINH (S. T), BENOIS-PINEAU (J), FAZEKAS (K) et GSCHWINDT (A). – Mesh-based error scalable video object codec for variable bandwidth multimedia communications. *In : ICIP'2002*. – Rochester, New York, USA, 2002.
- [645] MINH (S. T), FAZEKAS (K), BENOIS-PINEAU (J) et GSCHWINDT (A). – Full scheme of mpeg4-like codec based on wavelet transform. *In : ICASSP'2002*. – Orlando, Florida, USA, 2002. Summary.
- [646] NESVADBA (J), ERNST (F), PERHAVC (J), BENOIS-PINEAU (J) et PRIMAUX (L). – Comparison of shot boundary detectors. *In : IEEE ICME'2005*. – 2005.
- [647] NESVADBA (J), LOUIS (N), BENOIS-PINEAU (J), DESAINTE-CATHERINE (M) et MIDDELINK (M). – Low-level cross-media statistical approach for semantic partitioning of audio-visual content in a home multimedia environment. *In : IEEE IWSSIP 2004*, pp. 235–238. – 2004.
- [648] NICOLAS (H), MANOURY (A), J.BENOIS-PINEAU, DUPUY (W) et BARBA (D). – Grouping video shots into scenes based on 1d mosaic descriptors. *In : IEEE ICIP'2004*, pp. 667–640. – Singapore, 2004.

- [649] PARISY (O) et SCHLICK (C). – A physically realistic framework for the generation of high level animation controllers. *In : Proc. of SmartGraphics'02, New York.* – june 2002.
- [650] PARISY (O) et SCHLICK (C). – A physically realistic framework for the generation of high level animation controllers. *In : Proceedings of SmartGraphics'02.* – 2002.
- [651] PELTIER (S), ALAYRANGUES (S), FUCHS (L) et J.-O.LACHAUD. – Computation of homology groups and generators. *In : DGCI.* pp. 195–205. – Springer, 2005.
- [652] PELTIER (S), ALAYRANGUES (S), FUCHS (L) et LACHAUD (J.-O). – Computation of homology groups and generators. *In : Proc. Int. Conf. Discrete Geometry for Computer Imagery (DGCI'2005), Poitiers, France*, éd. par Andrès (E), Damiand (G) et Lienhardt (P). pp. 195–205. – Springer, 2005.
- [653] PEROT (V), BERGER (P), DESBARATS (P), BEGUERET (H), MONTAUDON (M) et LAURENT (F). – Chronic obstructive pulmonary disease : measurements of airway lumen and wall areas from high resolution computed tomography data using a laplacian of gaussian algorithm. *In : Euro Radiology.* – Springer, 2002.
- [654] POUDEROUX (J), GONZATO (J.-C), GUITTON (P) et GRANIER (X). – A software for reconstructing 3d-terrains from scanned maps. *In : ACM SIGGRAPH 2004 Sketches and Applications.* – 2004.
- [655] POUDEROUX (J), TOBOR (I), GONZATO (J.-C) et GUITTON (P). – Adaptive hierarchical rbf interpolation for creating smooth digital elevation models. *In : Proceedings of the Twelfth ACM International Symposium on Advances in Geographical Information System 2004.* – ACM Press, 2004.
- [656] PRIMAUX (L) et BENOIS-PINEAU (J). – Optimized real time h.264 encoder for videosurveillance applications. *In : WIAMIS 2004.* – Lisbonne, Portugal, avril 2004.
- [657] PRIMAUX (L), BENOIS-PINEAU (J), KRÄMER (P) et DOMENGER (J.-P). – Shot boundary detection in the framework of rough indexing paradigm. *In : TRECVID Workshop.* – NIST in Gaithersburg, MD, November 2004.
- [658] QUENOT (G. M), MARARU (D), S.AYACHE, CHARHAD (M), BESACIER (L), GUIRONNET (M), PELLERIN (D), GENSEL (J) et CARMINATI (L). – Clips lis lsr labri experiments in trec video retrieval 2004. *In : TREC Video Retrieval 2004.* – 2004.
- [659] REUTER (P), SCHMITT (B), PASKO (A) et SCHLICK (C). – Interactive solid texturing using point-based multiresolution representations. *Journal of WSCG 2004*, vol. 12, n3, 2004, pp. 363–370.
- [660] REUTER (P), TOBOR (I), SCHLICK (C) et DEDIEU (S). – Point-based modelling and rendering using radial basis functions. *In : Proceedings of the 1st international conference on Computer graphics and interactive techniques in Australasia and South East Asia (Graphite 2003)*, pp. 111–118. – 2003. Held in Melbourne, Australia.
- [661] REUTER (P), SCHMITT (B), PASKO (A) et SCHLICK (C). – Interactive solid texturing using point-based multiresolution representations. *In : WSCG (Winter School of Computer Graphics).* – 2004.
- [662] REUTER (P), TOBOR (I), SCHLICK (C) et DEDIEU (S). – Point-based modelling and rendering using radial basis functions. *In : Proceedings of the 1st international conference on Computer graphics and interactive techniques (Graphite 2003).* – 2003.
- [663] RIVIÈRE (J.-B. D. L) et GUITTON (P). – Hand posture recognition in large display vr environments. *In : Proceedings of the 5th International Gesture Workshop.* – Springer, 2003.

- [664] SENNEVILLE (B. D. D), DESBARATS (P), QUESSON (B) et MOONEN (C). – Real-time artefact corrections for quantitative mr temperature mapping. *In : Journal of WSCG*. pp. 87–94. – WSCG, 2003. ISSN 1213-6972.
- [665] SENNEVILLE (B. D. D), DESBARATS (P), SALOMIR (R), QUESSON (B) et MOONEN (C). – Correction of accidental patient motion for on-line mr thermometry. *In : Proceedings of MICCAI 2004, Saint-Malo*. pp. 637–644. – LNCS, 2004.
- [666] SENNEVILLE (B. D. D), QUESSON (B), DESBARATS (P) et MOONEN (C). – 3d motion estimation for on-line mr temperature mapping. *In : Proceedings of IEEE ICIP 2005, Genova*. – IEEE, 2005. To appear.
- [667] SENNEVILLE (B. D. D), QUESSON (B), DESBARATS (P), SALOMIR (R), PALUSSIÈRE (J) et MOONEN (C). – Atlas-based motion correction for on-line mr temperature mapping. *In : Proceedings of IEEE ICIP 2004, Singapore*. pp. 2571–2574. – IEEE, 2004.
- [668] SHANAT (M), FAYOLLE (P.-A), SCHMITT (B), VILBRANDT (T) et HANIWA. – A case study of digital visualization of virtual heritage properties. *In : Proc. of Eurographics UK Conference*, pp. 24–32. – june 2002.
- [669] SHANAT (M), FAYOLLE (P), SCHMITT (B) et VILBRANDT (V). – A case study of digital visualization of virtual heritages properties. *In : Proceedings of Eurographics UK Conference*. – 2002.
- [670] SPINELLO (S) et GUITTON (P). – Contour line recognition from scanned topographic maps. *In : WSCG (Winter School of Computer Graphics)*. – 2004.
- [671] STRANDH (R). – Gsharp, an extensible interactive score editor. *In : Proceedings of the International Lisp Conference*. – San Francisco, California, October 2002.
- [672] STRANDH (R) et MOORE (T). – A free implementation of clim. *In : Proceedings of the International Lisp Conference*. – San Francisco, California, October 2002.
- [673] TATON (B) et LACHAUD (J.-O). – Deformable model with non-euclidean metrics. *In : Proc. 7th European Conference on Computer Vision (ECCV'2002), Copenhagen, Denmark*, éd. par Heyden (A), Sparr (G), Nielsen (M) et Johansen (P). pp. 438–453. – Springer, Berlin, 2002.
- [674] TOBOR (I), REUTER (P) et SCHLICK (C). – Efficient reconstruction of large scattered geometric datasets using the partition of unity and radial basis functions. *Journal of WSCG 2004*, vol. 12, n3, 2004, pp. 467–474.
- [675] TOBOR (I), REUTER (P) et SCHLICK (C). – Multiresolution reconstruction of implicit surfaces with attributes from large unorganized point sets. *In : Shape Modeling International (SMI 2004)*. – 2004.
- [676] TOBOR (I), REUTER (P) et SCHLICK (C). – Efficient reconstruction of large scattered geometric datasets using the partition of unity and radial basis functions. *In : WSCG (Winter School of Computer Graphics)*. – 2004.
- [677] TOBOR (I), REUTER (P) et SCHLICK (C). – Multiresolution reconstruction of implicit surfaces with attributes from large unorganized point sets. *In : Proceedings of Shape Modeling International (SMI 2004)*. – 2004.
- [678] TYNDIUK (F), THOMAS (G), LESPINET-NAJIB (V) et SCHLICK (C). – Impact of large displays on virtual reality task performance. *In : AFRIGRAPH'04, 3rd International Conference on Virtual Reality, Computer Graphics, Visualization and Interaction in Africa*. – 2004.

- [679] VALLET (L), BEURTON-AIMAR (M), DESBARATS (P), FABRIGOULE (C) et ALLARD (M). – Towards computer-aided diagnosis for white matter abnormalities. *In : Proceedings of the International Conference on Computer Vision and Graphics*, éd. par Kluwer. – Varsawa, Poland, 2004.

Conférences invitées

- [680] BEURIVÉ (A) et MARCHAND (S). – Music Composition with Spectral Sounds. *In : Rencontres Mondiales du Logiciel Libre – Libre Software Meeting*. – July 2000. ABUL, URL : <http://www.abul.org>.
- [681] BEURTON-AIMAR (M), PÉRÈS (S), PARISEY (N), BALLEST (P), ZEMMERLINE (A), NAZARET (C) et MAZAT (J.-P). – Modeling biologic networks to use them with heterogeneous treatments. *In : Modélisation et simulation de processus biologiques dans le contexte de la génomique*. – Dieppe (France), Mai 2003.
- [682] BRAQUELAIRE (A). – Representing and segmenting 2D images by means of planar maps with discrete embeddings : From model to applications. *In : Proceedings of IAPR International Workshop GbRPR, Lecture Notes in Computer Science 3434*. pp. 92–121. – Springer, 2005.
- [683] DESAINTE-CATHERINE (M) et BEURIVÉ (A). – Time modeling for musical composition. *In : proceedings of FSKD'02, Singapore*. – 2002.
- [684] MARCHAND (S) et DESAINTE-CATHERINE (M). – Spectral Modeling, Analysis, and Synthesis of Musical Sounds. *Journal of the Acoustical Society of America*, vol. 112, n5, part 2/2, November 2002, pp. 2237–2238. – 144th Meeting of the Acoustical Society of America and First Pan-American / Iberian Meeting on Acoustics, Cancun, Mexico. December 2002.
- [685] MAZAT (J.-P), PÉRÈS (S) et BEURTON-AIMAR (M). – Classification of elementary flux modes in mitochondrial metabolism. *In : Workshop "Metabolic Pathway Analysis"*. Jena. – March 2005.

Chapitres d'ouvrages

- [686] BENOIS-PINEAU (J). – Object-based video indexing. *In : Multimedia Content and Semantic Web*, chap. 6, pp. 163–201. – John Wiley and Sons, ed. g. stamou, s. collias édition, 2005.

Edition d'ouvrages

- [687] BRAQUELAIRE (A), LACHAUD (J.-O) et VIALARD (A) (édité par). – *Tenth international conference on Discrete Geometry for Computer Imagery*. – Springer, 2002. ISBN 3-540-43380-5.
- [688] DOMENGER (J.-P). – Special issue : Discrete geometry for computer imagery (dgc'2002). *Pattern Recognition Letter*, 2003. – Invited Editor.
- [689] LACHAUD (J.-O) et VIALARD (A) (édité par). – *Special issue : discrete topology and geometry for image and object representation*. – Academic Press, may 2003.

4.4.2 Publications d'audience nationale

Reuves avec Comité de lecture

- [690] BENOIS-PINEAU (J), DUPUY (W) et BARBA (D). – Outils de structuration des documents vidéo en vue d'indexation basée sur une approche du signal 1d. *Technique et science informatiques*, vol. 22, n9, 2003, pp. 1167–1200.

Colloques avec Comité de programme et Actes

- [691] ALAYRANGUES (S), DARANGON (X), LACHAUD (J.-O) et LIENHARDT (P). – Equivalence des n-g-cartes fermées sans multi-incidence et des n-surfaces. *In : Journées du Groupe de Travail en Modélisation Géométrique, GTMG 2004.* – 2004.
- [692] BENOIS-PINEAU (J), BRAQUELAIRE (A) et ALI-MHAMMAD (A). – Estimation du mouvement par les méthodes différentielles pour indexation vidéo basée mouvement. *In : Journée d'échanges "Analyse et Indexation Multimédia".* – AS STIC "Indexation Multimédia : Transmodalité et gestion de connaissances", 2002.
- [693] BOUBEKEUR (T), REUTER (P) et SCHLICK (C). – Reconstruction locale et visualisation de nuages de points par surfaces de subdivision. *Actes des 17ièmes Journées de l'Association Française d'Informatique Graphique*, 2004, pp. 147–158. – Prix du meilleur article.
- [694] DESAINTE-CATHERINE (M), KURTAG. (G), GOURDON (B), MARCHAND (S) et SEMAL (C). – Dolabip : Un éveil musical avec l'ordinateur. *In : colloque "Apprendre avec l'ordinateur à l'école".* Université Bordeaux 2. – January 2002.
- [695] LEVET (F), HADIM (J), REUTER (P) et SCHLICK (C). – Echantillonnage anisotrope et rendu par points différentiels pour les surfaces implicites. *Actes des 17ièmes Journées de l'Association Française d'Informatique Graphique*, 2004, pp. 85–94.
- [696] POUDEROUX (J), GONZATO (J.-C) et GUITTON (P). – Création semi-automatique d'un modèle numérique de terrain. *In : Actes des 16ièmes Journées de l'AFIG*, pp. 151–160. – octobre 2003.
- [697] RIVIÈRE (J.-B. D. L) et GUITTON (P). – Silhouettes pour le suivi temps réel basé modèle. *In : Atelier Acquisition du geste humain par vision artificielle et applications, RFIA.* – 2004.
- [698] TATON (B) et LACHAUD (J.-O). – Modèle déformable en métrique non-euclidienne. *In : Actes du 13ème Congrès Francophone de Reconnaissance des Formes et Intelligence Artificielle (RFIA'02).* – Angers, France, jan 2002.
- [699] TOBOR (I), REUTER (P), GRISONI (L) et SCHLICK (C). – Visualisation par surfels. *Actes des 13ièmes Journées de l'Association Française d'Informatique Graphique*, 2000, pp. 193–204.
- [700] TYNDIUK (F), THOMAS (G), SCHLICK (C) et CLAVERIE (B). – Modèles et facteurs humains en ihm : Application à la réalité virtuelle. *In : formels de l'interaction, MFI'03.* – Lille, France, mai 2003.
- [701] TYNDIUK (F), THOMAS (G), SCHLICK (C) et CLAVERIE (B). – Modèles et facteurs humains en ihm. application à la réalité virtuelle. *In : MFI'03 - Modèles Formels de l'Interaction.* – 2003.
- [702] W.DUPUY, J.BENOIS-PINEAU et BARBA (D). – Outils pour l'analyse et l'indexation vidéo basée sur l'approche du signal 1d dans le domaine de la transformée mojette. *In : RFIA'2002.* – Angers, France, 2002.

Conférences invitées

- [703] BENOIS-PINEAU (J). – Analyse des séquences d’images couleur. extraction des objets en mouvement. *In : Conférence invité à l’Ecole d’Hiver sur l’Image Numérique Couleur.* – 2005.
- [704] BOUBEKEUR (T), REUTER (P) et SCHLICK (C). – Reconstruction locale et visualisation de nuages de points par surfaces de subdivision. *In : Actes des 17ièmes Journées de l’Association Française d’Informatique Graphique (AFIG).* – 2004. best paper award.
- [705] DESAINTE-CATHERINE (M). – Logiciels libres du scrime. *In : Résonances, IRCAM, Paris.* – October 2003.
- [706] LEVET (F), HADIM (J), REUTER (P) et SCHLICK (C). – Echantillonnage anisotropique et rendu par points différentiels pour les surfaces implicites. *In : Actes des 17ièmes Journées de l’Association Française d’Informatique Graphique (AFIG).* – 2004.
- [707] POUDEROUX (J), GONZATO (J.-C) et GUITTON (P). – Création semi-automatique d’un modèle numérique de terrain. *In : Actes des 16ièmes Journées de l’Association Française d’Informatique Graphique (AFIG).* – 2003.

Livres d’enseignement et de recherche

- [708] BRAQUELAIRE (A). – *Méthodologie de la programmation en langage C.* – DUNOD, 2005, quatrième édition. Norme C99 - API POSIX. ISBN 2-10-049018-4.
- [709] STRANDH (R) et DURAND (I). – *Architecture de l’ordinateur pour informaticiens.* – MétaModulaire, 2002.
- [710] STRANDH (R) et DURAND (I). – *Initiation à l’informatique.* – MétaModulaire, 2003.

Chapitres d’ouvrages

- [711] BENOIS-PINEAU (J), PATEUX (S) et BARBA (D). – Modèles de représentation d’images vidéo et compression. *In : Compression et codage des images et des vidéos,* chap. V. – Hermès-Lavoisier, 2002.
- [712] HACHET (M). – *Le traité de la réalité virtuelle - 3e édition.* – Presses de l’Ecole des Mines, 2005. En Français.
- [713] MARCHAND (S). – *In : Informatique musicale : du signal au signe musical,* éd. par PACHET (F) et BRIOT (J.-P), chap. La synthèse additive. – Paris, HERMES Sciences, May 2004. ISBN 2-7462-0825-3 (440 pages).
- [714] MARCHAND (S). – *In : Informatique musicale : du signal au signe musical,* éd. par PACHET (F) et BRIOT (J.-P), chap. L’analyse de Fourier. – Paris, HERMES Sciences, May 2004. ISBN 2-7462-0825-3 (440 pages).

4.4.3 Autres publications

Rapports de fin de contrat

- [715] BENOIS-PINEAU (J). – *Contrat de collaboration IRCCyN/Labri. Sous-traitance du projet RNTL DOMUS VIDEUM.* – Rapport technique, LaBRI, 2002. Rapport de la phase 1.

- [716] BENOIS-PINEAU (J), BARBA (D), NICOLAS (H) et MANOURY (A). – *DOMUS Videum, Sous-projet Structuration Vidéo*. – Rapport technique, LaBRI, 2004. Rapport final.
- [717] BENOIS-PINEAU (J), DUPUY (W), BARBA (D), NICOLAS (H) et MANOURY (A). – *DOMUS VIDEUM, Sous-projet Structuration Vidéo*. – Rapport technique, LaBRI, 2003. Rapport de la phase 2.
- [718] DON (A), BAYET (R) et BENOIS-PINEAU (J). – *Incorporation de la norme MPEG4 dans une solution logicielle de production Richmédia*. – Rapport technique, LaBRI, 2004. Rapport d'activités, Convention de recherche SystmTV.
- [719] DON (A) et BENOIS-PINEAU (J). – *Incorporation de la norme MPEG4 dans une solution logicielle de production Richmédia*. – Rapport technique, LaBRI, 2003. Rapport d'activités, Convention de recherche SystmTV.

Rapports internes et autres publications

- [720] DESBARATS (P) et GUEORGUIEVA (S). – *CW complexes : Topological Mainframe for Numerical Representations of Objects*. – Rapport technique nRR-1291-03, LaBRI, Université Bordeaux 1, 2003. 43 p.
- [721] DESBARATS (P) et GUEORGUIEVA (S). – *CW complexes : Topological Mainframe for Numerical Representations of Objects*. – Rapport technique nRR-1291-03, LaBRI, UMR 5800, Université Bordeaux 1, 2003.
- [722] GRANIER (X) et DAMEZ (C). – *Control on Color Reflection Behavior*. – Rapport technique n5227, INRIA, 2004.
- [723] GUEORGUIEVA (S) et DESBARATS (P). – *Reconstruction of topology valid boundary of discrete object from 3D range images*. – Rapport technique nRR-1357-05, LaBRI, Université Bordeaux 1, 2005. 24 p.
- [724] GUEORGUIEVA (S) et DESBARATS (P). – *Reconstruction of topology valid boundary of discrete object from 3D range images*. – Rapport technique nRR-????-05, 24p., LaBRI, UMR 5800, Université Bordeaux 1, 2005.
- [725] KRAEMER (P), BENOIS-PINEAU (J) et DOMENGER (J.-P). – *Construction of Spatio-Temporal Mosaics in the Framework of Rough Indexing Paradigm*. – Rapport technique, LaBRI, June 2004.
- [726] PRIMAUX (L), QUILLET (J) et J.BENOIS-PINEAU. – *Méthodologie et outils logiciel en vue de la compression spatio-temporelle d'images vidéo*. – Rapport technique, LaBRI, June 2005.
- [727] REUTER (P), SCHMITT (B), PASKO (A) et SCHLICK (C). – *Interactive Surface Texturing of Arbitrary 3D Objects by Combining Constructive Texturing and Point-based Rendering*. – Rapport technique, LaBRI, Université Bordeaux 1, 2002.
- [728] REUTER (P), SCHMITT (B), SCHLICK (C) et PASKO (A). – *Interactive Surface Texturing of Arbitrary 3D Objects by Combining Constructive Textures and Point-based Rendering*. – Rapport technique n1284-02, LaBRI, Université Bordeaux 1, 2002.
- [729] RIVIÈRE (J.-B. D. L) et GUITTON (P). – *Hand Posture Recognition in Large Display VR Environments*. – Rapport technique nRR-130003, LaBRI, Université Bordeaux 1,, 2003.
- [730] THOMAS (G), POINT (G) et BOUATOUCH (K). – *A Client-Server Approach to Image-Based Rendering on Mobile Terminals*. – Rapport technique nRR-5447, INRIA, 2005.

- [731] TOBOR (I), REUTER (P) et SCHLICK (C). – *Efficient Reconstruction of Large Scattered Geometric Datasets using the Partition of Unity and Radial Basis Functions*. – Rapport technique n130103, LaBRI, Bordeaux, 2003.

4.4.4 Formation par la recherche

Thèses

- [732] CIEUTAT (J.-M). – *Modélisation physiquement réaliste de sessions de simulation d'entraînement maritime*. – Talence - France, Thèse de PhD, University of Bordeaux 1, décembre 2003.
- [733] CIEUTAT (J.-M). – *Modélisation physiquement réaliste de sessions de simulation d'entraînement maritime*. – Thèse de PhD, Université Bordeaux 1, 2003.
- [734] DIETRICH (G. D). – *Segmentation d'organes tubulaires par suivi de squelette*. – Thèse de PhD, Université Bordeaux 1, Talence, France, may 2003.
- [735] HACHET (M). – *Interaction avec des environnements virtuels affichés au moyen d'interfaces de visualisation collective*. – Talence - France, Thèse de PhD, University of Bordeaux 1, décembre 2003.
- [736] HACHET (M). – *Interaction avec des Environnements Virtuels affichés au moyen d'Interfaces de Visualisation Collective*. – Thèse de PhD, Université Bordeaux 1, 2003.
- [737] HANNA (P). – *Modélisation statistique de sons bruités : étude de la densité spectrale, analyse, transformation musicale et synthèse*. – Thèse de PhD, LaBRI, Université Bordeaux 1, December 2003.
- [738] KERAUTRET (B). – *Reconstruction et lissage de surfaces discrètes*. – Thèse de PhD, Université Bordeaux 1, Talence, France, december 2004.
- [739] LAGRANGE (M). – *Modélisation sinusoïdale des sons polyphoniques*. – Thèse de PhD, LaBRI, Université Bordeaux 1, December 2004.
- [740] REUTER (P). – *Reconstruction et Rendu de Surfaces Implicites à partir de grands ensembles de points non-structurés*. – Talence - France, Thèse de PhD, University of Bordeaux 1, décembre 2003.
- [741] REUTER (P). – *Reconstruction and Rendering of Implicit Surfaces from Large Unorganized Point Sets*. – Thèse de PhD, Université Bordeaux 1, 2003.
- [742] SCHMITT (B). – *Constructive Hypervolume Modelling*. – Talence - France, Thèse de PhD, University of Bordeaux 1, décembre 2002.
- [743] TATON (B). – *Modèle déformable à densité adaptative : application à la segmentation d'images*. – Thèse de PhD, Université Bordeaux 1, Talence, France, 2004.
- [744] TOBOR (I). – *Utilisation des surfels dans le rendu des surfaces 3D*. – Talence - France, Thèse de PhD, University of Bordeaux 1, décembre 2002.
- [745] TOBOR (I). – *Utilisation des surfels dans le rendu des surfaces 3D*. – Thèse de PhD, Université Bordeaux 1, 2002.

4.5 Collaborations internationales

4.5.1 Invités

Burke, Dave, *University of British Columbia*

Ghosh, Abhijeet, *University of British Columbia*

Heidrich, Wolfgang, *University of British Columbia*

Munzner, Tamara, *University of British Columbia*

Trentacoste, Matthew, *University of British Columbia*

Ofer, Hadar, *Ben Gurion University of the Negev.*, Septembre-Octobre 2004

Depalle, Philippe, *McGill University (Montréal, Canada)*

4.5.2 Invitations

Boubekeur, Tamy, *University of British Columbia*

Granier, Xavier, *University of British Columbia*

Hachet, Martin, *University of British Columbia*

Hadim, Julien, *University of British Columbia*

Reuter, Patrick, *University of British Columbia*

Benois-Pineau, Jenny, *Ben Gurion University of the Negev*, Janvier 2004

Benois-Pineau, Jenny, *University of Växjö (Suède)*, Février 2002

4.6 Liste des contrats et subventions

4.6.1 Contrats internationaux et européens

Période. 2004-2007

Source. Equipe associée INRIA-DREI

Partenaire. IMAGER Lab - University of British Columbia, Vancouver, Canada

Description. LIGHT :

Période. 2004-2005

Source. Ministère des relations étrangères, CNRS, INRIA

Partenaire. France, Chine, Corée, Japon, Singapoure, Taiwan

Description. STIC ASIE VIRTUAL REALITY. : Identification et mise en place de collaborations franco-asiatiques

Période. 2002-2003

Source. Philips Research NatLab

Partenaire. LABRI, Philips Research NatLab

Description. INDEXATION CROSS-MEDIA :

4.6.2 Contrats avec des entreprises

Période. 2004-2005

Type de financement. France Télécom

Entreprise. France Télécom

Description. Rendu NPR sur TMC : Modélisation et rendu non photo-réaliste sur terminaux mobiles d'environnements urbains et de territoires

Période. 2001-2004

Type de financement. France Télécom

Entreprise. France Télécom

Description. Analyse sonore et décomposition en entités sonores

Période. 2003-2004

Type de financement. ANVAR, transfert de technologie

Entreprise. Jazz Mutant

Description. Interaction avec la synthèse sonore

4.6.3 Contrats publics avec des instances nationales

Période. 2002-2004

Source. Réseau National des Technologies de la Santé

Partenaire. Surgiview - ENSAM - HU Strasbourg - CHU Bordeaux - CHU Beaujon - Hôp. R. Poincaré - Hôp. R. Debré - AXS ingénierie - Spinevision - Université Lyon 1

Description. B3S - SPINE BALANCE : Développement d'un outil logiciel d'aide au traitement des scolioses à partir d'une paire d'images radio

Période. 2001-2005

Source. Ministère de la culture

Partenaire. Direction de la musique, de la danse, du théâtre et du spectacle

Description. SCRIME : Studio de création et de recherche en informatique et musique électroacoustique.

Période. 2002-2004

Source. projet RNRT (ministère de la recherche)

Partenaire. Surgiview - ENSAM - HU Strasbourg - CHU Bordeaux - CHU Beaujon - Hop. R. Poincaré - Hop. R. Debré - AXS ingénierie - Spinevision - Université Lyon 1

Description. B3S - Spine Balance : développement d'un outil logiciel d'aide au traitement des scolioses à partir d'une paire d'images radio

Période. 2002-2003

Source. AS du dpt. STIC, CNRS

Partenaire. LIS, LIRIS, SIC, ESIEE, LERI, Ecole des mines de Paris

Description. Géométrie discrète pour l'analyse spatio-temporelle d'images.

Période. 2003

Source. AS du dpt. STIC, CNRS

Partenaire. GREYC, INRIA, LIENS, LIF, LIP, LIRIS, LIS-ELESA, LLAIC, LMC-IMAG, LORIA, LSIIT, PSI, SIC

Description. Géométrie discrète et algorithmique.

Période. 2000-2003

Source. projet RNRT (ministère de la recherche)

Partenaire. France Telecom Research and Development, IRISA, Thomson Multimedia, IWI, LaBRI

Description. V2NET : développement d'un prototype d'architecture client-serveur pour la navigation dans des scènes urbaines à travers un réseau (www.telecom.gouv.fr/rnrt/projets/res_d90_ap99.htm).

Période. 2000-2004

Source. projet RNRT (ministère de la recherche)

Partenaire. IRISA, CEA LIST, Ecole des Mines Paris, Image Institut, LRP, LIMSI, ADEPA, Clarté, Dassault, EADS, EDF, Giat, IFP, PSA, Renault, PSA, LaBRI

Description. PERF RV : conception d'une plate-forme pour le développement d'applications de réalité virtuelle dans l'industrie (www.telecom.gouv.fr/rntl/FichesA/Perf-Rv.htm).

Période. 2003-2005

Source. ACI "masse de données", ministère de la recherche.

Partenaire. Grenoble, Nancy, Sophia, IPARLA

Description. SHOW : visualisation de scènes 3D très complexes

Période. 2002-2004

Source. RNTL

Partenaire. Thomson Multimédia, IRISA/INRIA, IRCCyN, SFRS, INA, LABRI

Description. Domus Videum : développement de méthodes de structurations de documents vidéo, pour la navigation dans les vidéos

Période. 2005-2007

Source.

Partenaire. CLIPS-IMAG, Eurecom, INA, LIP6, IRIT, Noveltis, SFRS MENRT, MD, MI TECHNOVISION "ARGOS"

Description. Constitution d'un corpus, indexation de référence, développement d'outils d'indexation vidéo

Période. 2003-2004

Source. CNRS Equiep Projet Pidot

Partenaire. IRIT, CLIPS-IMAG, EURECOM, LIP6, LABRI

Description. Plateforme d'Indexation Distribuée Ouverte Transmédia

Période. 2003-2004

Source. CNRS AS STIC

Partenaire. LPNC,IRCOM-SIC,ENST Paris, LIS,LUSSAC, INSERM, Lyon, LISTIC,LIGIV

Description. Traitement Perceptif d'Images Numériques Couleur

Période. 2003-2004

Source. CNRS AS STIC

Partenaire. I3S, ENST et IRISA TEMIC, LABRI

Description. Compression Scalable et Robuste de Signaux Vidéo

Période. 2001-2003

Source. CNRS AS STIC

Partenaire. IRIT, LABRI, IRCCYN, CIPS-IMAG, LIP6, IRIT, INRIA Rocquencourt

Description. Indexation Multimédia : transmodalité et gestion des connaissances

4.6.4 Contrats publics avec instances locales ou régionales

Période. 2003-2004

Source. Région Aquitaine – EITICA

Partenaire. DIS

Description. Étude, réalisation et validation de composants logiciels pour le traitement et l'analyse d'images 2D haute résolution pour internet

Période. 2003-2004

Source. Région Aquitaine

Partenaire. Compose (INRIA Futurs), Equipe Test (LaBRI), Equipe SOD (LaBRI), IPARLA

Description. Conception d'une plate-forme logicielle pour le développement d'applications multimédia sur des terminaux mobiles

Période. 2002

Source. Pôle Aquitaine Santé

Partenaire. TGS

Description. Plate-forme d'analyse d'image sur Amira.Insertion de composants logiciels d'analyse d'image dans la plate-forme Amira

Période. 2005-2006

Source. Pôle EITICA, ANVAR

Partenaire. BeTomorrow, LABRI

Description. Transmodage video. TRanscription des flux vidéos sportifs naturels encodés en vidéo synthétique pour les services de communication vidéo.

Période. 2002-2004

Source. LABRI, VisualPix

Partenaire. FEDER EITICA

Description. Méthodologie et outils logiciels en vue de la compression spatio-temporelle d'images vidéo

Période. 2002-2004

Source. SystemTV

Partenaire. EITICA

Description. Incorporation de la Norme MPEG4 dans une solution logicielle de production RichMédia

Période. 2002-2004

Source. LABRI

Partenaire. Région Aquitaine

Description. Navigation dans les Bases de Données Vidéo

Période. 2002-2003

Source. LABRI

Partenaire. Université de Bordeaux 1

Description. Collaboration des outils d'analyse de l'image, de la vidéo et du son

Période. 2002-2003

Source. Région Aquitaine

Partenaire.

Description. Expertise scientifique des outils de la compression vidéo

4.7 Animation de la recherche

4.7.1 Comités de rédaction

- IEEE Transactions on Speech and Audio Processing, Perception and Psychophysics
- IEEE Circuits and Systems for Video Technology
- IEEE Transactions on Multimedia
- IEEE Transactions on PAMI
- Pattern Recognition Letters
- Pattern Recognition
- Journal of Applied Signal Processing (JASP)
- Journal of Computer Vision and Image Understanding
- Real-Time Imaging
- Signal Processing : ImageCommunication (EURASIP)
- Computer Graphics Forum
- Traitement du signal
- Numéro spécial “Analyse de document” de la revue Traitement du signal
- Le traité de la réalité virtuelle (3ième édition)
- Membre du comité de programme du numéro spécial sur l’indexation et la recherche par contenu des images et de la vidéo de la revue Techniques et Sciences Informatiques , Hermes, 2002.

4.7.2 Comités de rédaction de colloques

- IEEE ICIP 2003-2005
- IEEE ICASSP 2004
- ACM Multimedia 2004
- PCS 2003
- CBMI 2003-2005
- Wiamis 2004
- CORESDA 2005
- Virtual Conceot 2003
- Web 3D 2003
- Eurographics 2004
- Eurographics Symposium on Virtual Environments 2004
- Graphics Interface 2004
- Shape Modeling International 2004
- Graphics Interface 2005
- Shape Modeling International 2005,
- Colloque international DGCI
- Workshop international IAPR GbR
- Workshop international CVWW (2002)
- Sound and Music Computing
- International Conference on Music Information Retrieval
- International Computer Music Conférence
- Colloque Interdisciplinaire de Musicologie
- AMR (Adaptive Multimedia Retrieval) 2003-2005
- RFIA 2002
- Journées d’Informatique Musicale

4.7.3 Présidence de session de colloques

- Session “Segmentation vidéo” à ICIP 2004
- Session “Video surveillance” à WIAMIS 2004
- Session “Video analysis for indexing” à CBMI 2005
- Session “Analyse synthèse de son” à ICMC 2005

4.7.4 Organisation de colloques, écoles jeunes chercheurs, ...

- Colloque international DGCI 2002
- Réunion nationale de l’AS STIC géométrie discrète
- Organisatrice de la session spéciale "Intelligent Systems for Video Processing" du congrès international IPMU’2002, Juillet 2002, Annecy, France
- Organisatrice des journées nationales Recherche/Industrie Analyse et Indexation Multimédia
- Colloque CBMI (Content-Based Multimedia Indexing) 2003 et 2005
- Journées d’Informatique Musicale
- Colloque international DGCI

4.7.5 Administration de la Recherche

- Expert-membre de la Commission Multimédia SDTETIC MEN et MER depuis 1998.
- Expert pour les programmes FP4, FP5, FP6/IST de la Commission Européenne en "Multimédia".
- Expert-rapporteur du Groupe de Conseil Technologique (Technical Advisory Group) du Programme Européen MEDIAPLUS EC DG Inf.So and Media 2001-2005.
- Conseil d’administration de l’AFIG
- Présidence de l’AFIM (Association Française d’informatique Musicale)
- Membre du comité de pilotage du RTP 7 "Synthèse d’image et réalité virtuelle"
- coresponsabilité de l’AS "Rendu temps réel"
- Responsabilité de l’AS STIC Géométrie Discrète (2002-2003)
- Co-responsable de l’AS Indexation Multimédia : Transmodalité et gestion des connaissances, 2001-2002.
- Membre de l’AS Compression Scalable et robuste des Signaux vidéo 2003 - 2004.
- Membre de l’AS Traitements perceptifs d’images numériques couleur 2003 - 2005, RTP 25.
- Membre de l’équipe projet "PIDOT : plate-forme d’indexation distribuée cross-média", 2003 - 2004.
- Membre du comité de pilotage du Pôle Aquitaine Santé de 2001-2004
- Présidence du comité scientifique et technique ICI (12ième contrat de plan état région Poitou-Charentes), 2001- 2004
- Membre du comité scientifique GDR Algorithmique, Langages et Programmation
- Membre du comité scientifique ACI GRID
- Membre du comité scientifique GRID 5000
- Membre du comité scientifique INRIA Futurs
- Membre du comité scientifique Projet Visitor (action Marie Curie)
- Membre du comité d’Orientation Scientifique et Technologique de l’INRIA
- Membre du jury de thèse SPECIF
- Expertises pour les ACI (GRID, Jeunes chercheurs, Masse de données)
- Expertises pour les régions Bretagne, Pays de Loire et Rhône - Alpes.

- Expertise pour le fond de la recherche nature et technologie du Québec
- Expertise pour le ministère italien de l'éducation nationale
- Membre du comité d'évaluation du LORIA (Nancy)
- Membre du comité d'évaluation de la fédération de Recherche en "Physique et image de la ville de Nantes
- Expertises de contrats CIFFRE
- Expertises de projets RIAM
- Participation à 30 jurys de thèse (dont 22 en tant que rapporteur)
- Participation à 3 jurys d'habilitation à diriger des recherches (dont 2 en tant que rapporteur)

Informatique Parallèle et Distribué (PARADIS)

Responsable(s) Jean ROMAN

Comet

Responsable(s) Francine KRIEF

*Mots-clés : Autonomie dans les réseaux, Qualité de service et sécurité des communications,
Transport des flux multimédia, Réseaux B3G*

Phoenix

Responsable(s) Charles CONSEL

*Mots-clés : Services de communication, Analyse et transformation de programmes,
Systèmes d'exploitation et réseaux*

Runtime

Responsable(s) Raymond NAMYST

Mots-clés : Support exécutif, Calcul haute performance, Grappe de calcul, Grille de calcul

ScAlApplix

Responsable(s) Jean ROMAN

*Mots-clés : Algorithmique et calcul haute performance à grande échelle, Applications
complexes de grande taille*

Sod

Responsable(s) Serge CHAUMETTE

*Mots-clés : Systèmes et objets distribués, Plate-formes et environnements, Communication
et mobilité*

5.1 Composition de l'équipe

5.1.1 Membres permanents

AHMED, Toufik, MdC, ENSEIRB, depuis septembre 2004
AUMAGE, Olivier, CR, INRIA, depuis octobre 2003
BARON, Josy, Assistante, INRIA, depuis février 2003
BEAUMONT, Olivier, MdC, ENSEIRB
CHAUMETTE, Serge, Pr, Bx1, depuis septembre 2003
CONSEL, Charles, Pr, ENSEIRB
COULAUD, Olivier, DR, INRIA
COUNILH, Marie Christine, MdC, Bx1
DELORD, Xavier, MdC, ENSEIRB, depuis septembre 2004
DENIS, Alexandre, CR, INRIA, depuis octobre 2004
FURMENTO, Nathalie, IR, CNRS, depuis février 2005
GUERMOUCHE, Abdou, MdC, Bx1, depuis septembre 2005
HÉNON, Pascal, CR, INRIA, depuis janvier 2003
KRIEF, Francine, Pr, ENSEIRB, depuis septembre 2004
LARUE, Brigitte, Assistante, INRIA, depuis juillet 2003
LÉPINE, Jean Michel, MdC, Bx1
NAMYST, Raymond, Pr, Bx1, depuis septembre 2002
PELLEGRINI, François, MdC, ENSEIRB, en détachement INRIA de septembre 2002 à septembre 2004
RAMET, Pierre, MdC, Bx1, en délégation INRIA de septembre 2003 à septembre 2005
RÉVEILLÈRE, Laurent, Mdc, ENSEIRB, depuis septembre 2002
ROMAN, Jean, Pr, ENSEIRB, en délégation partielle INRIA depuis septembre 2004
RUBI, Franck, MdC, Bx1
VAUQUELIN, Bernard, Pr, Bx1
WACRENIER, Pierre-André, MdC, Bx1

5.1.2 Membre non permanents

ANCIAUX, Guillaume, Doct, INRIA-CEA, depuis septembre 2003
ATALLAH, Eve, Doct, Région Limoges, depuis septembre 2004
BHATIA, Sapan, Doct, INRIA-Région Aquitaine, depuis septembre 2003
BENMAMMAR, Badr, Doct (CDD), Bx1, depuis septembre 2004
BURGY, Laurent, Doct, Région Aquitaine, depuis octobre 2004
CASTEIGTS, Arnaud, Doct, MENR, depuis septembre 2004
CHEVALIER, Cédric, Doct, BDI CNRS-Région Aquitaine, depuis septembre 2004
CIROU, Bertrand, Doct puis ATER, MENR, jusqu'à août 2005
DANJEAN, Vincent, Doct, MENR, jusqu'à août 2004
DJAMA, Ismail, Doct, MENR, depuis septembre 2005
ESNARD, Aurélien, Doct et ATER, MENR, depuis septembre 2001
FORTIN, Pierre, Doct et ATER, BDI CNRS-Région Aquitaine, depuis septembre 2002
GAVAID, Usman, Doct, France Telecom, depuis octobre 2005
GRANGE, Pascal, Doct puis ATER, MENR, depuis septembre 2000
KARRAY, Achraf, Doct, co-tutelle Univ. Sfax-Tunisie, depuis janvier 2005

LATRY, Fabien, Doct, INRIA, depuis octobre 2004
LATU, Guillaume, Doct puis ATER, MENR, jusqu'à août 2003
LE MEUR, Anne-Françoise, Doct puis ATER, MENR, jusqu'à août 2003
MBAREK, Nader, Doct (CDD), Bx1, depuis septembre 2004
MERCIER, Guillaume, Doct, MENR, jusqu'à février 2005
MINARD, Mathieu, Doct, Thomson, depuis octobre 2002
MUBASHAR, Mushtaq, Doct, bourse Pakistan, depuis octobre 2005
PALIX, Nicolas, Doct, INRIA, depuis octobre 2004
PÉRACHE, Marc, Doct, CEA, depuis septembre 2003
RICHART, Nicolas, Doct, INRIA-Région Aquitaine, depuis octobre 2005
SAUVERON, Damien, Doct puis ATER à Limoges, CIFRE, jusqu'à décembre 2004
TESSIER, Gaël, Doct, Région Aquitaine, depuis septembre 2003
THIBAUT, Samuel, Doct, MENR, depuis septembre 2004
VIGNERAS, Pierre, Doct puis ATER, MENR, jusqu'à décembre 2004

CHARDRON, Jean-Noël, Ingénieur stagiaire, CNAM, depuis octobre 2005
DANJEAN, Vincent, IR, CEA, de septembre 2004 à août 2005
DUSSÈRE, Michaël, Ingénieur expert, INRIA, jusqu'à octobre 2005
LECAS, Dimitri, Ingénieur associé, INRIA, jusqu'à décembre 2004

5.2 Présentation de l'équipe

Les travaux de l'équipe PARADIS concernent de manière complémentaire diverses problématiques de l'informatique parallèle et distribuée à grande échelle. La structure de cette équipe a beaucoup évolué ces dernières années, soit de par l'arrivée de nouvelles composantes comme Comet et Runtime, soit de par l'évolution forte de composantes plus anciennes comme Sod, et surtout AliENor et Compose au titre des projets ScAIApplix et Phoenix qui sont communs avec l'INRIA Futurs.

Les membres de PARADIS font leurs recherches plus particulièrement dans trois domaines qui sont les réseaux, les objets distribués, et le calcul hautes performances. L'équipe est structurée en cinq parties dont les activités sont introduites ci-après.

Les activités de recherche du thème Comet, qui a démarré au LaBRI en 2004, portent sur la négociation, le contrôle, la configuration et l'adaptation de la qualité de service et de la sécurité de flux de différents types (données, audio, vidéo, streaming) pour des terminaux fixes et mobiles sur les réseaux IP de nouvelle génération.

Les principaux axes de recherche couvrent des aspects comme la négociation et le déploiement automatique de services sécurisés avec contraintes de qualité de service dans des environnements fixes et mobiles, la gestion autonome de réseaux et de services, les assistants de négociation, et le transport adaptatif des flux audiovisuels.

Ces différents thèmes de recherche convergent vers le *context-aware management* et sont menés en grande partie dans le cadre de contrats de recherche.

L'activité de Phoenix, qui est en cours d'expertise pour devenir un projet commun LaBRI-INRIA, couvre les aspects fondamentaux de la création de service de communication, la définition de nouveaux *langages dédiés à un domaine* (ou DSLs pour Domain Specific Languages) en utilisant la technologie des langages de programmation pour permettre la spécification de services robustes, l'étude des couches supportant les services de communication pour améliorer la flexibilité et les performances, et enfin l'application à des domaines concrets pour valider ces approches.

Les domaines traités actuellement sont la téléphonie sur IP, les serveurs de streaming multimédia (application d'un modèle client serveur programmable), les serveurs à hautes performances, et la spécialisation de programmes.

Le projet commun LaBRI-INRIA Runtime, créé en 2004, s'inscrit dans le cadre du calcul parallèle à hautes performances. Il s'intéresse à l'étude et à la conception des principes, ainsi qu'à la mise en œuvre et à l'évaluation des mécanismes qui seront au cœur des supports exécutifs parallèles de demain. Ces supports devront permettre l'exploitation efficace des configurations hétérogènes et hiérarchiques que sont les grappes de grande taille, les grappes de grappes et les grilles. Le projet de recherche s'articule autour des trois principaux défis qui sont la maîtrise des configurations hétérogènes de grande taille, l'exploitation des technologies nouvelles de manière optimale, et l'amélioration de l'intégration des environnements et des applications.

L'objectif du projet commun LaBRI-MAB-INRIA ScAIApplix créé en novembre 2002 est la mise en œuvre de compétences scientifiques complémentaires pour une recherche pluridisciplinaire dans les domaines de l'informatique hautes performances et de la modélisation numérique, et ce dans le but d'analyser et de résoudre efficacement des problèmes de calcul scientifique provenant d'*applications complexes nécessitant des puissances de calcul et ma-*

nipulant des tailles de données téraflopiques. La résolution effective de cette gamme de problèmes est un véritable défi qui nécessite clairement une approche pluridisciplinaire. Cette synergie de compétences scientifiques concernent plus particulièrement en mathématiques appliquées le domaine des schémas numériques, et en informatique les domaines de l'algorithme parallèle et du savoir-faire de la mise en œuvre de codes hautes performances sur diverses plates-formes de calcul.

Il s'agit donc, dans le cadre de cette synergie scientifique, de *contribuer à toutes les étapes de la chaîne qui va de la conception de schémas performants à la mise en œuvre optimisée et efficace de codes de simulation pour des applications complexes et/ou de très grande taille.* Cette chaîne contient aussi la partie interprétation des résultats ; ceci se fera en pilotant les codes de calcul hautes performances parallèles et/ou distribués par une application de type réalité virtuelle, faisant ainsi de la simulation numérique un outil efficace pour interagir avec la modélisation via une analyse de sensibilité interactive des paramètres de cette simulation.

Les activités du thème Systèmes et objets distribués du LaBRI ont pour but de contribuer à rendre possible et efficace l'utilisation et l'administration des équipements de calcul ou fonctionnels connectés à un réseau fixe ou mobile. Cela passe naturellement par la fourniture d'une plate-forme logicielle (APIs, communication asynchrone, migration, ...) et d'un environnement validé qui offrent un niveau élevé d'abstraction tout en permettant un niveau de contrôle relativement fin des couches sous-jacentes. Cette activité de recherche repose sur la technologie des objets distribués, paradigme qui permet d'aborder les problématiques ci-dessus de façon adéquate. Le thème SOD s'appuie sur les compétences suivantes qu'il a développées en son sein : technologies objets et modèles associés ; intergiciels, machines virtuelles ; sécurité dans un cadre distribué (PKIs, cartes à puce, ...) ; technologies de la mobilité et modèles associés.

5.3 Thèmes de recherche de l' quipe

5.3.1 Comet

Introduction

Les activit s de recherche du th me COMET (COntext-aware ManagEment and neT-working) portent sur la n gociation, le contr le, la configuration et l'adaptation de la QoS (Qualit  de Service) et de la s curit  de flux de diff rents types (donn es, audio, vid o, streaming) pour des terminaux fixes et mobiles sur les r seaux IP de nouvelle g n ration.

Ces travaux sont orient s volontairement vers le « context-aware management » qui permet une gestion plus rapide et personnalis e, que ce soit pour la configuration, la s curit , les performances ou la composition de services.

Mots-Cl s Gestion autonome, Qualit  de Service, Adaptation des flux multim dia, Mobilit , R seaux B3G, S curit .

Description des activit s et principaux r sultats

Les principaux axes de recherche couvrent les aspects suivants :

- N gociation et d ploiement automatique de services s curis s avec contraintes de qualit  de service dans des environnements fixes et mobiles ;
- Gestion autonome de r seaux et services ;
- Interface de n gociation ;
- Transport adaptatif des flux audiovisuels.

Ces travaux sont men s en grande partie dans le cadre de projets de recherche.

La n gociation et le d ploiement automatique de services Cet axe concerne plus pr cis ment :

- La n gociation et le d ploiement automatique de services s curis s avec contrainte de QoS (1) ;
- Le maintien de la qualit  de service dans des environnements mobiles (2).

S curit  et Qualit  de Service (1) Offrir la s curit  de la communication et la qualit  de service exig e par l'application font parti des d fis majeurs que la gestion de r seaux et services doit relever pour permettre   l'utilisateur d' tre connect  en continu de partout et   haut d bit. Une m thode pour offrir   un utilisateur un service de s curit  quand ses communications transitent par des r seaux administr s par des op rateurs ind pendants est d'int grer ce probl me   celui plus g n ral de la n gociation de SLS (Service Level Specification) dans un cadre inter-domaine en utilisant un protocole de n gociation. Ceci permet une politique de s curit  n goci e et configurable en temps r el entre les diff rents acteurs ainsi qu'une s curit  ajust e   la sensibilit  des informations et   la menace. Concernant la configuration automatique de services, tr s peu de travaux traitent, aujourd'hui, du couplage des politiques de qualit  de service et de s curit , qui sont en g n ral impl ment es s par ment.

L'objectif est de sp cifier et d'impl menter un environnement de n gociation et de configuration dynamique de services s curis s avec contraintes de QoS. Il s'agit de couvrir l'aspect de n gociation, d'adaptation et de contr le de la QoS et de la s curit  dans la cha ne de distribution des services, afin de supporter une QoS et une s curit  de bout en bout. Le but est double, d'une part, d finir des strat gies de QoS et de s curit  sur chaque entit  individuelle de la cha ne, et d'autre part, harmoniser leurs fonctionnalit s sur les r seaux h t rog nes.

En particulier, il faudra :

- Proposer et gérer un profil de qualité de service et de sécurité ;
- Définir une garantie de QoS de bout en bout, impliquant l'interopérabilité, l'uniformité et la fiabilité entre différentes ressources ;
- Définir le(s) protocole(s) associé(s) et les mécanismes requis ;
- Étudier la faisabilité et l'évaluation des performances des solutions d'architecture.

Ces travaux rentrent dans le cadre du projet RNRT SWAN (Self-aWare mAnagement). Les principaux résultats sont les suivants.

- Définition d'un SLS générique : Dans toute négociation dynamique, la signalisation joue un rôle capital et il est important de connaître les éléments sur lesquels porte la signalisation. Nous avons proposé un SLS générique recouvrant les différents paramètres nécessaires pour définir un SLA en vue de tenir compte des besoins utilisateurs (fixes ou mobiles) pour réaliser une signalisation. Il s'agit outre les paramètres de qualité de service, des paramètres liés à la mobilité tels que des indications sur les changements de localisation prévus ou les priorités lors d'un handover et des paramètres de sécurité.
- Etude de l'impact de la sécurité sur la qualité de service dans un environnement IPSec
- Spécification d'une architecture et des protocoles associés pour la négociation, le contrôle et la configuration dynamique de services dans les réseaux IP offrant des garanties de QoS.

Ces travaux ont été menés dans le cadre du projet RNRT IP-SIG (SIGnalisation générique du monde IP). Ce projet s'attaque au problème fondamental de l'hétérogénéité des produits des équipementiers et devrait permettre de bâtir des réseaux inter constructeurs et inter opérateurs. De plus, ce projet s'attaque à la dynamique de l'ensemble lorsque les comportements du réseau changent ou lorsque le client est mobile.

Mobilité et Qualité de service (2) La gestion de la qualité de service se complexifie lorsque les terminaux se déplacent. Les garanties de QoS accordées aux flux à destination ou en provenance d'un mobile peuvent être remises en cause suite à un changement dynamique de la topologie du réseau et des ressources, à des déconnexions soudaines et fréquentes, à des délais très variables, à des interférences. Les tâches de gestion sont donc profondément influencées par la mobilité de l'utilisateur. Ces tâches incluent la gestion du handover, le contrôle de flux, l'allocation des ressources, le contrôle de congestion, le contrôle d'admission et l'approvisionnement de la QoS.

Notre objectif est d'améliorer la technologie des réseaux sans fil de façon à répondre le mieux possible aux besoins de qualité de service des utilisateurs et à maintenir les performances de la connexion lors de la mobilité de l'utilisateur. Notre démarche consiste à :

- Améliorer les protocoles et algorithmes existants. Nous mettons l'accent sur les méthodes pro-actives. Ainsi la prévision des futures localisations du terminal et la connaissance du profil utilisateur et applicatif peuvent être utilisées pour l'allocation des ressources nécessaires.
- Optimiser les mécanismes de handover. Nous proposons d'intervenir sur les mécanismes d'optimisation des handovers de façon à diminuer le temps d'interruption de communication qui peut s'avérer trop long pour certaines applications, comme par exemple les applications temps réel (voix sur IP) ou des applications à moyenne contrainte de temps (streaming).
- Optimiser l'utilisation de la bande passante. Un réseau radio est limité par sa bande passante. Aussi, l'introduction d'agents mobiles peut permettre de réduire le trafic sur le lien radio.

Ces travaux ont débuté avec le projet RNRT IP-SIG.

Les principaux résultats sont les suivants.

- Spécification d'un profil de mobilité;
- Modélisation de la prédiction de la mobilité de l'utilisateur;
- Spécification d'un environnement de signalisation permettant la réservation de ressources à l'avance et minimisant la dégradation du service lors du handover horizontal et vertical. Dans ce cadre, le protocole QoS NSLP proposé par le working group NSIS (Next Steps in Signaling) de l'IETF a été étendu. Ce groupe de travail est tout particulièrement concerné puisqu'il a défini un cadre pour la signalisation relative à la qualité de service ainsi que des protocoles de signalisation spécifiques.

La gestion autonome La gestion évolue vers plus d'autonomie et deviendra sensible au contexte. Par autonomie, on entend habituellement la capacité des processus de gestion et de l'infrastructure sous-jacente à se déployer, s'organiser et opérer sans aide extérieure. Le rôle de l'administrateur se limite au guidage de ces processus en leur fixant des objectifs et des paramètres opérationnels. Offrir cette autonomie à la gestion est crucial pour répondre aux besoins croissants de qualité de service et de mobilité des utilisateurs.

Trois axes principaux de recherche sont considérés :

- L'apport des politiques (1);
- L'adaptation automatique du réseau au contexte (2);
- L'adaptation automatique des services au contexte (3).

Les politiques (1) L'approche déclarative de la gestion par politique permet à l'opérateur d'établir des objectifs de service et des politiques qui sont par la suite interprétés par les ressources réseau. Ainsi, la décision sur l'allocation et la configuration des ressources peut être prise localement de façon autonome. De plus, la modification dynamique de ces politiques peut fournir un moyen souple de contrôler le comportement du réseau lorsque l'environnement change. Pour adapter les politiques, les processus de contrôle utiliseront des informations de contexte, c'est-à-dire des données relatives à la localisation, l'identité, le profil des utilisateurs et des services, les ressources et services disponibles, etc.

L'objectif est de concevoir une architecture de supervision de réseau générique favorisant au maximum l'automatisation des activités et l'autonomie des composants entrant dans la livraison des services à l'aide des politiques. La configuration se fera de façon totalement automatique. Dans ce cadre, le contrat de service doit, tout d'abord, être traduit automatiquement en règles de politique. Celles-ci doivent, ensuite, être raffinées jusqu'à devenir compréhensibles par les équipements à configurer. Il est également important de vérifier que l'introduction d'une nouvelle règle de politique ne rentre pas en conflit avec les politiques existantes.

Un outil de traduction automatique du contrat de service en règles de politique a été défini et est en cours de développement dans le cadre du projet RNRT SWAN.

Gestion autonome de réseaux (2) Nos travaux concernent principalement :

- Les réseaux d'opérateurs (a);
- Les réseaux de capteurs sans fil (b).

Les réseaux d'opérateurs (a) : La nécessité d'automatiser l'activité de supervision est aujourd'hui un élément crucial des réseaux d'opérateurs en raison de la demande croissante de qualité de service par les usagers, de la complexité croissante des fonctions de gestion qui permettent sa mise en œuvre et son respect, de la difficulté à trouver des personnes très fortement qualifiées pour maîtriser cette complexité et du besoin de maîtrise des coûts liés à cette activité.

Dans ce contexte, nos travaux portent principalement sur l'ingénierie de trafic des réseaux GMPLS. L'objectif est de proposer une infrastructure de gestion à base de politique permettant d'optimiser les ressources du réseau d'opérateur. Pour cela, les LSP (Label Switched Path) sont établis en fonction des contrats de service.

Une architecture de gestion par politique des réseaux GMPLS a été spécifiée et est en cours de validation à l'aide de l'outil de simulation Opnet.

Les réseaux de capteurs sans fil (b) : Ils forment une nouvelle génération de réseaux aux propriétés qui n'entrent pas dans le cadre des architectures classiques. Un champ de recherche important est ouvert aujourd'hui pour rendre ces réseaux autonomes. Un réseau de capteurs se définit comme un ensemble de capteurs connectés entre eux, chaque capteur étant muni d'un émetteur-récepteur. La miniaturisation des capteurs pose des problèmes de communication et de ressource d'énergie. Il faut que le capteur rassemble l'information requise et l'émette quand c'est nécessaire. De plus, le processeur du capteur ne doit pas être utilisé trop intensivement afin de consommer le moins d'énergie possible. Enfin, pour assurer un bon débit, la portée des émetteurs-récepteurs est nécessairement faible. La mise en place d'un réseau de capteurs pose donc des problèmes de routage.

Nos recherches dans ce domaine portent actuellement sur l'optimisation de la consommation d'énergie. Elles s'articulent autour de la classification sous contraintes de données spatio-temporelles.

Un algorithme de routage adaptatif basé sur le clustering a été proposé. Ce dernier s'appuie sur des techniques d'apprentissage connexionniste et plus exactement sur les cartes auto-organisatrices de Kohonen auxquelles de nouvelles règles pour le choix des représentants des clusters ont été ajoutées.

La gestion autonome de services web (3) Aujourd'hui, la technologie des Web Services (WS) est considérée comme une technologie prometteuse pour faire collaborer des services via le Web. Elle permet également la composition de services dans le cadre de l'Internet ambiant. Les WS reposent sur une architecture hétérogène distribuée. Dans le cadre, notamment, de transactions impliquant plusieurs entreprises, la question de savoir qui va prendre la responsabilité de superviser le déroulement des processus, gérer les flux de données, facturer l'utilisation d'un service, gérer la sécurité, les fautes et les erreurs, s'occuper de la maintenance des applications se pose. De plus, Les aspects de durée de vie courte, d'évolution rapide mais aussi de frontières entre organisations impliquent que chaque WS ne pourra avoir qu'une visibilité réduite des WS avec lesquels il interagit. Tous ces aspects montrent que la gestion doit pouvoir s'adapter rapidement aux changements tout comme pour l'Internet ambiant.

Notre objectif est de proposer des solutions pour une administration de haut-niveau des WS. Cela vise également à étudier des fonctionnalités nouvelles telles que la cohérence de politiques distribuées, et la négociation distribuée de SLS avec recherche de compromis. Nous adopterons une architecture décentralisée, dans laquelle les échanges se font par invocation de services Web. L'utilisation d'appels de services Web inclus dans des documents permet de mettre à jour dynamiquement les données et de contrôler finement la périodicité de ces mises à jour et la pérennité des informations. Les systèmes participants peuvent échanger des données intensionnelles, c'est à dire des données dont certaines parties sont fournies par leur définition en terme d'appels de services Web. Ces travaux seront effectués dans le cadre du projet RNRT SWAN.

Les principaux résultats concernent la spécification des paramètres de SLA/SLS pour les WS.

Interface de négociation Avec les réseaux B3G (beyond 3G), l'équipement terminal de l'utilisateur aura plusieurs technologies d'accès sans fil à sa disposition et souhaitera être connecté au mieux, n'importe où, n'importe quand et avec n'importe quel réseau d'accès. Pour cela, les différentes technologies sans fil qui seront présentes dans les réseaux B3G doivent coexister de manière à ce que la meilleure technologie puisse être retenue en fonction du profil de l'utilisateur, du type de réseau d'accès, de terminal et d'application demandé.

Notre objectif est de spécifier et d'implémenter une interface de négociation coté terminal qui choisira dynamiquement un service pour le compte de l'utilisateur de façon à ce que celui-ci soit connecté au mieux à tout moment. Cette interface, connaissant les préférences de l'utilisateur (services utilisés, QoS demandée, coût que l'utilisateur est prêt à payer pour utiliser certains services, etc), doit :

- Identifier, voire anticiper, les choix de l'utilisateur en matière de qualité de service, de sécurité et de mobilité afin de pouvoir négocier les paramètres techniques avec le fournisseur de service et ceci de façon totalement transparente pour l'utilisateur ;
- Sélectionner le meilleur fournisseur de services disponible ;
- Négocier sur la base de micro-flots, ce qui permettra de n'offrir à l'utilisateur que le service dont il a besoin ;
- Assurer un contrôle dynamique des paramètres de QoS. Ce contrôle permettra de déclencher une demande de renégociation des paramètres de SLS si l'écart entre la satisfaction de l'utilisateur et le service offert est trop important dans un sens ou un autre ;
- Vérifier que l'utilisateur a bien accès à un service avec une qualité conforme à ce qui a été négocié.

Un assistant de négociation, coté terminal utilisateur, a été spécifié. Celui-ci a pour principale tâche, l'identification des besoins de l'utilisateur afin de négocier dynamiquement, pour le compte de ce dernier, les SLA/SLS avec le/les fournisseurs de service. Un protocole de négociation de SLS, conforme à l'environnement NSIS, a également été proposé.

Nous envisageons de séparer l'interface utilisateur du terminal en l'intégrant à la JavaCard, les utilisateurs pouvant utiliser d'autres terminaux lors de leurs déplacements.

Transport adaptatif des flux audiovisuels Les réseaux de nouvelle génération deviennent multiservices, convergents, supportant des utilisateurs fixes, mobiles et sans-fil à travers des infrastructures et des technologies d'accès hétérogènes. Il est clair qu'un nombre important de réseaux de communications, tels que le WLAN (802.11a/g, HIPERLAN), le WMAN (802.16) et les réseaux cellulaires 2G/3G (GSM, GPRS, UMTS), peuvent être considérés comme candidats pour la distribution des services audiovisuels dans les réseaux de nouvelle génération. Cependant, les réseaux de télédiffusions numériques (DVB), satellite (DVB-S), terrestre (DVB-T) et pour terminaux mobiles (DVB-H) peuvent être améliorés pour le transport des services interactifs multimédia basé sur IP. L'intégration des technologies bidirectionnelles combinée à la maturité du système DVB (S/T/H) offre un support parfait pour développer le transport des flux audiovisuels sur ces réseaux.

Les travaux du thème COMET sur le transport des flux audiovisuels visent à (1) définir une architecture de transport des flux audiovisuels sur les réseaux IP avec gestion de qualité de service (QoS) (2) proposer une architecture adressant la convergence des services multimédia dans les réseaux de nouvelle génération fixes et mobiles.

Les résultats obtenus s'articulent autour de :

- Un modèle de classification automatique d'objets audiovisuels intégré à la couche système MPEG-4 ;
- Un protocole robuste de fragmentation et d'encapsulation de flux vidéo sur RTP avec contrôle inégal d'erreurs (FEC, Unequal FEC) et de débits (TCP-Friendly, TFRC) ;

- Un algorithme de marquage dynamique de paquets vidéo IP pour routeurs d'accès Diff-serv asservi aux variations de bande passante et aux caractéristiques des flux multimédia ;
- Une architecture de passerelle de signalisation supportant les protocoles de contrôle de sessions IETF SIP (Session Initiation Protocol) et OSI MPEG-4 DMIF (Delivery Multimedia Integrated Framework). Cette passerelle multimédia de signalisation permet d'assurer la continuité du service, la résolution d'adresse et la négociation des capacités des terminaux pour favoriser une plus grande interopérabilité des services multimédia MPEG-4 avec des terminaux IP fixes et mobiles.

Les travaux en cours du projet COMET pour le transport des flux audiovisuels ont pour but :

- La définition d'une solution complète capable d'optimiser les ressources utilisées dans les réseaux pour satisfaire les besoins d'un grand nombre d'utilisateurs fixes et mobiles ;
- L'intégration des composants et des normes pour faciliter l'harmonisation des technologies, la création, la protection, la distribution et la consommation des flux audiovisuels ;
- La garantie de QoS de bout en bout, impliquant l'interopérabilité, l'uniformité et la fiabilité entre différentes ressources.

Projets et perspectives 2005-2008

Nos principaux objectifs sont les suivants :

- Spécification et implémentation d'un environnement de négociation et de configuration dynamique de services sécurisés avec contraintes de qualité de service ;
- Amélioration de la technologie des réseaux sans fil de façon à répondre le mieux possible aux besoins de qualité de service des utilisateurs et à maintenir les performances de la connexion lors de la mobilité de l'utilisateur ;
- Conception d'une architecture de supervision de réseau générique favorisant au maximum l'automatisation des activités et l'autonomie des composants entrant dans la livraison des services ;
- Proposition de solutions pour une administration de haut-niveau des WS ;
- Spécification et implémentation d'une interface intelligente pour les réseaux B3G ;
- Définition d'une architecture de transport des flux audiovisuels sur les réseaux IP avec gestion de qualité de service et proposition d'une architecture adressant la convergence des services multimédia dans les réseaux de nouvelle génération fixes et mobiles.

Ces différents thèmes de recherche convergent vers la gestion autonome et les réseaux B3G. Un projet fédérateur autour de l'autonomie dans les réseaux de future génération devrait permettre, à terme, d'intégrer de façon homogène les différents résultats théoriques ou d'implantation.

5.3.2 Phoenix

Introduction

Les rapides avancées dans le domaine des communications multimédia, associées à la convergence entre les télécommunications et les réseaux informatiques, ouvre un large spectre de nouvelles fonctionnalités. La création de services devient un moyen fondamental d'offrir ces changements aux utilisateurs finaux.

Cette situation a trois conséquences principales : (1) la création de services apparaît comme un domaine dans lequel le développement logiciel est intensif; (2) étant donné que les services de communication sont considérés comme obligatoirement fiables, la création intensive de services se doit de préserver la *robustesse*; (3) la nature multimédia grandissante des services nécessite de *hautes performances* de la part des services et des couches sous-jacentes. Le groupe Phoenix a pour but de concevoir des principes, des techniques et des outils pour le développement de *services de communication*. Pour satisfaire les exigences de ce domaine, le spectre de nos recherches comprend les éléments clés qui sous-tendent les services de communication : l'infrastructure permettant de déployer des services (*e.g.* plateforme de signalisation, protocoles de transport, et description de sessions), l'architecture logicielle supportant les services (*e.g.* le modèle client-serveur, les interfaces de programmation, et la notion de logique du service), et enfin, les terminaux permettant la communication (*e.g.* fonctionnalités des terminaux et systèmes embarqués).

Notre approche couvre trois aspects clés du domaine des services de communication : (1) la définition de nouveaux *langages dédiés à un domaine* (ou DSLs pour Domain Specific Languages), utilisant la technologie des langages de programmation pour permettre la spécification de services robustes ; (2) l'étude des couches supportant les services de communication pour améliorer la flexibilité et les performances ; (3) l'application à des domaines concrets pour valider notre approche.

Mots-Clés Services de communication, Téléphonie, Conception de langage, Analyse et Transformation de programme, Spécialisation, Compilation, Systèmes d'exploitation, Réseau, Analyse de domaine.

Description des activités et principaux résultats

Services de téléphonie De nos jours, la téléphonie évolue très rapidement. Après avoir convergé avec les réseaux informatiques et multimédia, elle offre maintenant un large spectre de nouvelles fonctionnalités. En tête de cette évolution, on trouve le protocole Session Initiation Protocol (SIP) destiné à la Voix sur IP (VoIP) et la téléphonie mobile de troisième génération. SIP fournit un support à la mobilité des utilisateurs et à la notification de présence. Il permet également de gérer divers types de dialogues multimédia incluant les conversations audio/vidéo, les jeux et les messageries instantanées. La plate-forme de signalisation SIP permet d'interfacer des services de téléphonie avec d'autres applications (courrier électronique, Web...).

L'étendue grandissante du domaine de la téléphonie demande au développeur des services de téléphonie une profonde connaissance des règles et contraintes inhérentes à la téléphonie, ainsi qu'une grande maîtrise du protocole SIP et des protocoles connexes. Programmer une plateforme SIP nécessite de plus une expertise en systèmes distribués, réseaux, et de maîtriser une interface de programmation SIP souvent vaste et complexe. Ces exigences font du développement de services de téléphonie un véritable défi.

Pour programmer des services, deux stratégies sont possibles. La première stratégie, tradi-

tionnelle, consiste à interfacier la plate-forme avec un langage généraliste tel que Java. Etant donnée l'expressivité d'un tel langage, seul un administrateur du réseau (ou un programmeur accrédité) peut bénéficier d'un accès aussi large à la plate-forme sans compromettre son intégrité ; une telle situation limite donc considérablement l'introduction de nouveaux services. La deuxième stratégie vise à restreindre le langage de programmation de services, pour permettre à un utilisateur quelconque de définir son propre service, sans compromettre son intégrité.

La communauté SIP tente de promouvoir cette deuxième stratégie en énonçant un véritable cahier des charges auquel doit répondre un langage conçu pour permettre à un utilisateur quelconque de programmer des services de téléphonie. Ce cahier des charges est consigné informellement dans le "Request For Comments" (RFC) 2824 ; il détaille notamment les exigences de sécurité qui doivent être remplies par un tel langage.

Une tentative de langage restreint pour la programmation de services de téléphonie a été développé à l'Université de Columbia. Toutefois, l'examen de ce langage montre qu'il a été exagérément bridé. D'autre part, aucun travail formel n'a démontré que ce langage était conforme au RFC 2824.

Notre objectif vise à formaliser le cahier des charges contenu dans le RFC 2824 en utilisant un formalisme approprié. D'autre part, nous avons conçu un langage de programmation dédié au développement de services de téléphonie. Ce langage, nommé SPL (*Session Processing Language*), est à la fois expressif, pour permettre une forte interaction avec diverses applications (Web, bases de données...), tout en étant conforme au RFC 2824. Pour remplir nos objectifs, nous avons défini formellement la sémantique de SPL et commencé à vérifier la conformité de sa sémantique par rapport à notre cahier des charges. Ce travail est basé sur nos travaux sur le développement de langages dédiés. Nous avons également conçu une machine virtuelle SIP qui fournit une interface haut-niveau dédiée au développement de services.

Une première implémentation de SPL est déployée sur la plateforme de signalisation SER. La machine virtuelle est en cours de développement pour une plateforme Java.

Un modèle client-serveur programmable Nous concevons et développons une nouvelle architecture logicielle permettant aux programmeurs de développer rapidement et avec concision des services de communications multimédia robustes et efficaces. Cette approche repose sur deux composants principaux pour une famille de services donnée : un langage de description de services et un serveur programmable.

Langage de description de services Pour les aspects réseaux et télécommunications, notre projet consiste à développer un langage de description de services permettant à des développeurs de définir de nouveaux services en tenant compte des besoins et préférences d'utilisateurs et des ressources disponibles. Pour ce faire, nous proposons une approche qui vise à modéliser une large famille de services avec un langage de programmation dédié au domaine visé. Un service pourra ainsi être défini avec concision en utilisant à la fois des constructions syntaxiques et une sémantique appropriées. A la différence de langages généralistes, tels que C ou Java, notre langage de description sera conçu de telle sorte que des propriétés de sûreté et de sécurité pourront être plus facilement vérifiées sur les services.

Serveur programmable/extensible Le modèle client-serveur est une architecture communément utilisée pour mettre en oeuvre une famille de services dans les réseaux et les télécommunications. Un serveur implémente un ensemble de services. Des clients, connectés au serveur par le réseau, envoient des requêtes pour accéder aux services. Quand le serveur

a traité une requête, il envoie une réponse au client correspondant. Bien qu'éprouvé, le modèle client-serveur est grandement limité par le fait qu'il est insensible aux contraintes du client : caractéristiques du terminal, bande passante disponible, fiabilité de la connexion, mode de facturation, préférences de l'utilisateur... Cette limitation est fondamentale lorsque l'on considère les terminaux mobiles communiquants qui, par nature, ont des contraintes fortes d'utilisation. En outre, lorsque des applications graphiques sont ciblées, la prise en compte des contraintes du client devient impérative. Avant de déployer un service, notre nouveau serveur invoquera une phase de vérification, pour s'assurer de leur sûreté et leur sécurité, puis une phase de compilation, pour générer une implémentation efficace.

Application aux services de courrier électroniques. Nous avons modifié un serveur IMAP (Internet Message Access Protocol) pour le rendre extensible. Cette extensibilité permet de définir des vues sur des boîtes aux lettres distantes. L'objectif d'une vue est d'adapter l'accès à une boîte aux lettres à différentes contraintes telles que les caractéristiques du terminal et de la bande passante, et aussi aux préférences de l'utilisateur. Une vue peut ainsi filtrer les messages nouveaux pour minimiser le sommaire. En fonction de critères définis par l'utilisateur, chaque message se voit affecté une catégorie. En fonction de son importance, chaque catégorie correspond à une présentation plus ou moins condensée. Les services sont définis grâce à un langage, nommé PEMS, développé par le groupe Phoenix.

Application aux services de *streaming* multimédia Le traitement de flux multimédia est un domaine en rapide évolution qui nécessite un développement logiciel intensif et exige des performances élevées. Développer une application de traitement de flux implique souvent de la programmation de bas niveau, une gestion critique de la mémoire, et un ordonnancement très précis des différentes étapes de traitement.

Nous avons développé un langage dédié, nommé Spidle, qui permet de décrire des applications de traitement de flux multimédia. Le langage Spidle offre des constructions déclaratives de haut niveau, contrairement aux langages généralistes, sans toutefois compromettre la performance des applications. Il améliore également leur robustesse en permettant la vérification de différentes propriétés spécifiques au domaine.

Application aux serveurs à haute performance La spécialisation de programmes est maintenant disponible pour des langages aussi répandus que C et Java. De plus, elle a été utilisée avec succès pour des applications de taille réelle dans des domaines comme les systèmes d'exploitation, les réseaux ou le graphique. Toutefois, la lourdeur du processus de spécialisation le confine à des machines à haute performance. Cette situation réduit drastiquement l'éventail des applications de la spécialisation. En particulier, elle empêche la spécialisation de cibler les systèmes embarqués de petite taille alors que leurs besoins correspondent parfaitement aux bénéfices de la spécialisation, c'est-à-dire, l'amélioration en temps et la réduction en taille d'un programme.

Notre contribution a consisté à développer une approche pour lever cette limitation. Elle permet à des programmes d'être spécialisés à distance durant leur exécution. Ainsi, un système embarqué peut requérir la spécialisation de code en continu, et ce, au fur et à mesure que les opportunités de spécialisation deviennent disponibles. Pour ce faire, nous avons développé une interface de programmation et une infrastructure basée sur un modèle client-serveur.

Notre approche a été implémentée et validée sur une étude de cas : la pile TCP/IP de noyau Linux. Les expérimentations ont été conduites sur trois plate-formes différentes : le Pentium III (600 MHz), l'ARM SA1100 (200 MHz) sur COMPAQ iPAQ et le 486 (40 MHz). Les versions spécialisées de la pile de protocoles ont un temps d'exécution amélioré

d'environ 25% et une taille jusqu'à 20 fois inférieure au programme original. De plus, dans notre première mise en oeuvre, nous avons limité la latence totale pour requérir et obtenir un code spécialisé à environ une demi-seconde.

Optimisation de couches systèmes Récemment nous avons optimisé avec succès la pile TCP/IP de Unix. En outre, nous avons développé une infrastructure permettant d'effectuer la spécialisation de programmes à distance. Cette innovation rend la spécialisation de programmes accessible aux systèmes embarqués à faible ressources.

Spécialiseur de programmes C Nous avons conçu et développé un évaluateur partiel pour des programmes C, nommé Tempo. Une innovation importante apportée par ce système est qu'il permet la spécialisation de programmes à la compilation et à l'exécution. Diverses analyses dont le but est de préparer la phase de spécialisation ont été conçues pour ce système. Etant donné la richesse du langage C et le fait qu'il ait été peu étudié dans le contexte de l'évaluation partielle, le développement de ces analyses a constitué une partie importante de notre travail. Les principales analyses de programmes sont les suivantes : analyse d'alias, analyse d'effets de bord, analyse de temps de liaison et analyse d'actions (transformations de programmes). Tempo a été utilisé avec succès pour la spécialisation de code système, scientifique et graphique.

Déclaration de spécialisation Nos travaux ont également eu pour but l'intégration de la spécialisation de programmes dans le processus de développement logiciel. Ainsi, nous avons travaillé sur une approche déclarative qui vise à intégrer la notion de spécialisation de programmes dans le processus de développement logiciel. Cette approche repose sur la création de *modules de spécialisation*. Un module de spécialisation prend la forme de déclarations qui spécifient les scénarios de spécialisation d'un composant logiciel, c'est-à-dire les opportunités de spécialisation de ce programme. Ces déclarations sont écrites par le programmeur lors du développement et sont distinctes du code. Non seulement ces déclarations documentent les opportunités de spécialisation mais elles correspondent également aux conditions à remplir afin de garantir que toutes les opportunités de spécialisation sont bien prises en compte pendant le processus de spécialisation.

Nous avons développé un compilateur pour le langage de déclaration des modules de spécialisation. La compilation de modules de spécialisation permet d'extraire des informations qui servent à :

- Configurer automatiquement notre spécialiseur Tempo.
- Visualiser graphiquement comment les scénarios de spécialisation de plusieurs programmes se composent. Grâce à cette représentation, le programmeur possède une vue globale des déclarations à tous les moments du développement.
- Effectuer des vérifications relatives à la cohérence des scénarios de spécialisation vis à vis des résultats des analyses du spécialiseur. Ces vérifications sont primordiales pour garantir que le spécialiseur génère un programme dont le degré de spécialisation satisfait les déclarations du programmeur.

Nous avons illustré notre approche avec le développement d'encodeurs de données pour les codes correcteurs. Ceci nous a permis de montrer qu'il est possible de spécialiser efficacement et rapidement un ensemble de programmes génériques, si leur développement est couplé avec notre approche déclarative.

Distinctions De nombreux travaux ont fait l'objet de publications dans de grandes conférences internationales. L'article intitulé *A programmable client-server model : Robust exten-*

sibility via DSLs, par C.Consel et L.Réveillère, publié dans *Proceedings of the 18th International Conference on Automated Software Engineering (ASE)* en 2003, a été sélectionné pour le prix du meilleur article.

L'article intitulé *Automatic specialization of protocol stacks in OS kernels*, par S.Bhatia, C.Consel, A.-F. Le Meur et C. Pu, publié dans *Proceedings of the 22th Annual IEEE Conference on Local Computer Networks* a obtenu le prix du meilleur article.

Enfin, en 2002, l'*ASF* (Association ACM SIGOPS de France) a décerné à L. Réveillère le deuxième prix *ASF* de la recherche en système, qui récompense la meilleure thèse francophone en système, pour son travail intitulé : *Approche langage au développement de pilotes de périphériques robustes*.

Projets et perspectives 2005-2008

Les services de téléphonie Cette activité s'articule autour des thèmes suivants.

- **SPL sous forme graphique.** Au-delà de la représentation textuelle, une version graphique de SPL est en cours de développement pour rendre la programmation de services accessibles à des non-informaticiens. Cet environnement de développement graphique permettra à n'importe quel utilisateur de définir son propre service de téléphonie par un simple assemblage d'icônes représentant des fonctions élémentaires telle qu'une redirection d'appel ou une messagerie vocale.
- **Passage à l'échelle.** Un déploiement à grande échelle constitue l'une des prochaines étapes de l'expérimentation. Une plateforme de téléphonie (appelée TelIP) est en cours de déploiement au sein de l'École Nationale Supérieure d'Électronique d'Informatique et de Radiocommunications de Bordeaux (ENSEIRB). Il s'agit ensuite d'augmenter de manière conséquente le nombre d'utilisateurs de TelIP pour vérifier le passage à l'échelle de la solution proposée par l'équipe Phoenix. Des services réels sont en cours de déploiement sur cette plateforme. Ainsi, un système de gestion de file d'attente pour les appels d'un secrétariat sera prochainement opérationnel pour le département Télécommunications de l'École.
- **Ouverture.** Par ailleurs, la plate-forme sera ouverte à certains groupes d'utilisateurs, tels que des étudiants par exemple, pour le déploiement de leurs propres services écrit en SPL. Une telle ouverture engendre de nombreux défis en termes de fiabilité, de sécurité et de performance. Relever ces défis avec succès permettra de valider l'approche SPL.
- **Au-delà de la téléphonie.** L'équipe Phoenix étudie l'extension du langage SPL et de son serveur d'applications pour inclure d'autres types de communication entre des entités SIP. Ainsi, les systèmes d'alertes, de présence, de géo-localisation et de messagerie instantanée sont actuellement étudiés. Les perspectives de services combinant ces différentes fonctionnalités sont très prometteuses. Des applications dans des domaines tel que la médecine et la télésurveillance sont envisagées.
- **Déployer les services de manière sûre.** Nous travaillons également à d'autres approches permettant de garantir l'exécution des services. Pour nos expérimentations, nous nous servons du langage SPL et des plateformes de téléphonie comme sujet d'étude

Langage de programmation à base de capacités. Pour résoudre le problème de robustesse inhérent à l'ouverture fournie par la programmation de services, de nombreuses approches ont été explorées sans toutefois susciter un consensus. La dichotomie (client/administrateur) des développeurs est en pratique trop grossière. Par contraste, on peut remarquer que les politiques de gestion de ressources dans un système d'exploitation ne sont pas dichotomiques. En fait, l'accès aux ressources d'un système est défini par les capacités de

l'utilisateur. La même approche devrait être applicable à un langage de programmation dans la mesure où ce langage peut être vu comme une interface à des ressources. Dans le cas de SPL, on peut aisément imaginer qu'il y a une variété de politiques de contrôle de ressources à appliquer à un service suivant les attributions, et donc les capacités de celui qui déploie ce service. Ce contrôle fin des ressources devrait s'appliquer à l'utilisation du réseau, de services Web, d'opérations de signalisation, etc. Actuellement, nous travaillons à développer une approche permettant de compiler la description d'un service en fonction des capacités de celui qui le déploiera. Cette approche ne requiert qu'un seul langage de programmation ; l'accès aux ressources que permet ce langage est défini par les capacités accompagnant le service à compiler. Ce travail passe par la définition d'un modèle de capacités pour les ressources d'une plate-forme de signalisation SIP, la définition d'une analyse de programmes paramétrée par des capacités et permettant de vérifier statiquement qu'un service utilise les ressources que lui autorisent ses capacités. Enfin, la génération de code pourra également être paramétrée par des capacités pour choisir l'implémentation de certaines opérations en fonction de leur coût.

Approche *Goal-Directed* au développement de services. Par ailleurs, il convient de mettre en relation les capacités de l'utilisateur du service et l'utilisation des ressources faite par le service. Toutefois, une question reste encore en suspens : quelle quantité de ressources est en train d'être consommée par un service ? Cette question demeure importante ; en effet, bien qu'un service soit vérifié avec succès et utilise seulement des ressources autorisées, son exécution peut néanmoins dégrader, voire effondrer, les performances de la plateforme. Comme dans un système d'exploitation, une plateforme ne possède qu'une quantité finie de ressources. Avant d'accepter l'exécution d'un nouveau service, elle doit s'assurer qu'elle a les moyens de le faire. En d'autres termes, une plateforme doit mettre en place un procédé de contrôle d'admission au niveau des services. Comme dans un système d'exploitation, le contrôle d'admission ne peut être purement statique mais doit également inclure des vérifications dynamiques lors de l'accès aux ressources. Les services sont très sensibles à ce problème étant donné que l'absence d'une ressource conduit à une interruption du service, et qu'une pénurie de ressources entraîne une dégradation des performances de la plateforme. Certains domaines de services sont particulièrement sensibles à cette problématique. Par exemple, dans la téléphonie sur IP, le traitement d'un appel est soumis à de fortes contraintes temps-réel. Si une plateforme ne possède pas de procédé efficace de contrôle d'admission, une pénurie de ressources peut entraîner un mauvais traitement des appels. Dans le pire des cas, il peut même arriver que des appels soient tout simplement perdus, ce qui n'est pas acceptable pour l'utilisateur. De même, il arrive que les plateformes de diffusion de flux multimédia subissent une dégradation de leur qualité de service. Il est important de noter ici que nous ne cherchons pas à résoudre le problème de la qualité de service (QoS) au niveau réseau (comme dans DiffServ et RSVP, par exemple), mais que nous nous concentrons sur le niveau applicatif. Le contrôle des ressources utilisées par les services dans une plateforme ouverte présente de fortes similitudes avec le déploiement de composants logiciels dans les intergiciels et les systèmes distribués. D'une manière équivalente, déployer trop de composants peut globalement dégrader le système. Une étude préliminaire a conduit à établir un état de l'art sur le contrôle des ressources ; certaines notions ont été clarifiées, comme la notion de ressources et de QoS, et une étude portant sur les solutions existantes autour du contrôle de ressources a été réalisée. Le travail suivant vise à développer des principes de conception pour permettre aux DSLs d'assurer que le but de l'application sera atteint, de manière dégradé ou non, même dans les cas extrêmes d'exécution (pénurie de ressource, contraintes temporelles...). Dans le cas de la téléphonie sur IP, cela consiste à assurer que tout appel émis se terminera

par une action de signalisation (réception de l'appel, refus d'émission...). Cette approche vise donc à prévoir les "erreurs" imprévisibles du service, et à fournir des alternatives afin d'assurer la terminaison correcte de l'application, vis-à-vis de son but original. Le principe est de proposer un procédé complet de gestion (statique/dynamique), couplé à la notion de ressources pour le service, et ceci à de multiples phases (e.g., compilation, déploiement, invocation, et exécution). À partir de ces critères, de nouveaux concepts de programmation devront être proposés pour permettre à des développeurs de services d'exprimer une dégradation des ressources. Cela devrait permettre à une plateforme de diminuer la quantité de ressources exigées par un service et au service de pouvoir s'exécuter correctement.

5.3.3 Runtime

Runtime est un projet LaBRI-INRIA Futurs créé en avril 2004.

Introduction

Le projet de recherche Runtime s'inscrit dans le cadre du calcul parallèle haute performance. Il consiste à contribuer à l'étude et à l'élaboration des supports d'exécution sur lesquels s'appuieront les environnements de programmation et les applications de calcul intensif de demain.

Par support exécutif (*runtime*), nous entendons la couche logicielle destinée à apporter aux applications parallèles les fonctionnalités qui ne sont fournies ni par le système d'exploitation et ni par les pilotes de périphériques. Dans le domaine du parallélisme, et plus précisément de l'exploitation de machines parallèles distribuées, la complexité de mise en œuvre des environnements de programmation, des bibliothèques numériques spécialisées, des compilateurs ou encore des machines virtuelles fait que l'utilisation de supports exécutifs de bas niveau, apportant à la fois des propriétés de portabilité et d'efficacité, est devenue incontournable.

Notons au passage que la frontière entre un support exécutif et le système d'exploitation sous-jacent est floue car le support exécutif intègre parfois des extensions/modifications spécifiques du système d'exploitation, telles que des modifications de l'ordonnancement des processus, par exemple. En revanche, il se distingue des couches logicielles de plus haut niveau par le fait qu'il constitue réellement une extension logique du système d'exploitation.

Mots-Clés Support exécutif, Calcul haute performance, Grappe de calcul, Grille de calcul, Système d'exploitation

Contexte matériel : des grappes et grilles d'ordinateurs pour le calcul scientifique

Nos travaux visent les configurations matérielles de moyenne gamme (centre de calculs industriels ou de recherche) et de haut de gamme (telles celles appartenant au TOP 500 des configurations les plus performantes au monde). De fait, l'architecture de ces grandes plateformes de calculs parallèles a considérablement évolué ces quatre dernières années : les ordinateurs massivement parallèles (70% du TOP 500 en 2000) ont été supplantés par les grappes de station (près de 60% du TOP 500 en 2004). Ces grappes sont des ensembles souvent homogènes de nœuds interconnectés par des réseaux courte distance de type SAN (*System Area Network*). Les nœuds d'une grappe vont du simple PC du commerce aux machines multiprocesseurs symétriques dotés de 2 à 4 processeurs. Les technologies d'interconnexion utilisées sont souvent au sommet de l'état de l'art et permettent d'obtenir des performances de tout premier plan : latence de l'ordre de la microseconde et débits de plusieurs centaines de Mo/s! Constituées de 10 à 100 nœuds hier, de 100 à 2000 nœuds aujourd'hui, et vraisemblablement de plusieurs dizaines de milliers de nœuds demain, ces machines ne disposent pas encore de logiciels – et en particulier de supports exécutifs – capables d'en exploiter aisément toute la puissance. Cet état de fait s'explique par des raisons technologiques (évolution très rapide des technologies utilisées), méthodologiques (choix du modèle d'exécution à adopter) et fonctionnelles (augmentation des contraintes applicatives). Notons que cette difficulté d'exploiter les architectures de type "grappe" s'accroît encore aujourd'hui avec l'émergence de configurations encore plus complexes :

- À l'instar de la toute prochaine configuration de calcul du CEA, les grappes ont une structure hiérarchique de plus en plus profonde en interconnectant par un réseau rapide des nœuds multiprocesseurs qui peuvent eux-mêmes être des configurations hiérarchiques où l'accès mémoire n'est pas uniforme (*Non Uniform Memory Architectures*,

machine de quatre à quelques dizaines de processeurs) et où les ressources de calcul ne sont pas symétriques comme les processeurs multi-core SMT (puces rassemblant plusieurs processeurs chacun étant capable d'affecter des ressources de calculs à plusieurs flots d'instructions en parallèle). La difficulté est évidemment de répercuter ces caractéristiques au niveau du modèle d'exécution sans pour autant compromettre la portabilité des applications sur des grappes plus "conventionnelles".

- Qui plus est, la présence de multiples grappes dans les laboratoires et centres de calcul, qui diffèrent potentiellement par leur réseau d'interconnexion et par la nature des nœuds, conduit souvent à les interconnecter par des liaisons plus ou moins performantes pour former ce que l'on appelle des "grappes de grappes" fortement couplées, des fédérations de grappes (plusieurs grappes réparties sur plusieurs sites), et des grilles de calcul (interconnexion sur plusieurs sites et domaines d'administrations différents de plusieurs grappes, avec souvent hétérogénéité des architectures). Le passage d'une grappe à une configuration à plus large échelle complexifie la gestion des communications. Aux problématiques d'exploitation du réseau haute performance local à chaque grappe s'ajoutent les problématiques d'exploitation des réseaux longue distance et des interactions entre les deux types de réseau.
- On observe également une tendance nette à l'augmentation importante du nombre de nœuds composant les grappes. Ces grappes de grande taille posent des problèmes aux supports exécutifs qui n'apparaissent pas (ou très peu) sur les grappes de petites tailles, notamment en matière de consommation des ressources de communication. Beaucoup de bibliothèques de communication de très bas niveau, par exemple, ne permettent pas physiquement d'établir des connexions point-à-point entre tous les nœuds d'une configuration au delà de quelques dizaines de nœuds. Ainsi, il faut imaginer et mettre en place des systèmes de communication intégrant le fait qu'un nœud ne peut communiquer très efficacement qu'avec un voisinage restreint.
- Finalement, les évolutions technologiques récentes des processeurs, des réseaux, ou encore des systèmes d'entrées/sorties imposent souvent l'utilisation de nouvelles techniques pour les exploiter pleinement. Par exemple, l'exploitation d'un réseau à capacité d'adressage comme SCI s'effectue de manière radicalement différente d'un réseau à transmission de messages tel que Myrinet. La technologie d'interconnexion Infiniband, quant à elle, amène à repenser les interfaces d'entrées/sorties actuelles. Pour parvenir à intégrer un tel éventail de technologies et favoriser la réactivité des applications, les supports exécutifs doivent être conçus dès le départ dans cette optique.

C'est donc dans ce contexte que s'inscrivent les travaux de recherche que nous menons, et qui visent à concevoir une nouvelle génération de supports exécutifs capables d'exploiter efficacement les architectures parallèles de type "grappe" sur lesquels pourront s'appuyer les environnements à valeur ajoutée abstraite spécifique de programmation et d'exécution pour les applications parallèles de calcul intensif et du calcul scientifique à grande échelle (i.e. grid computing).

Contexte logiciel : les supports exécutifs pour le calcul parallèle Aujourd'hui, les supports exécutifs sont omniprésents dès lors que l'on s'intéresse à la mise en œuvre d'applications/environnements parallèles complexes. Les compilateurs, par exemple, génèrent du code qui sollicite de plus en plus les fonctionnalités d'un support exécutif sous-jacent (e.g. compilateur HPF Adaptor, compilateur de bytecode Java d'Hyperion), pour des raisons de portabilité, de simplicité du code généré, mais aussi parce que certains traitements complexes ne peuvent être effectués qu'à l'exécution (ramasse-miettes, équilibrage de charge).

Pendant longtemps, la majorité des supports exécutifs parallèles ont principalement consisté

en une glue logicielle “évoluée” entre des implémentations de bibliothèques standard, telles que MPI pour la gestion des communications et POSIX-threads pour la gestion du multithreading, par exemple. On peut notamment mentionner les travaux autour d'Athapascan (ID-IMAG, France), Chant (NASA, USA) ou encore PM² (thèse de R. Namyst au LIFL). Même si de telles approches sont encore répandues aujourd'hui, elles souffrent de nombreuses limitations liées à des incompatibilités “fonctionnelles” entre les différents composants (performances dégradées) et parfois même à des incompatibilités d'implantation (ex : problèmes de réentrance du code)

Plus près de nous, plusieurs travaux, comme Nexus (Argonne, USA), Panda (université de Vrije, Pays-Bas) et PM², ont montré qu'une meilleure approche consiste à concevoir des supports exécutifs réalisant une intégration maîtrisée des différents aspects que sont les communications, les entrées/sorties et le multithreading. Pour s'approcher d'une solution optimale, ces supports exécutifs s'appuient le plus souvent sur des bibliothèques de très bas niveau, comme par exemple BIP (LIP, France), GM (Myricom, USA), FM (université d'Illinois, USA) ou LFC (université de Vrije) pour l'exploitation du réseau Myrinet, de façon à conserver un contrôle fin sur le matériel. C'est une des raisons pour lesquelles leur conception demeure délicate.

Ainsi, de nombreux supports exécutifs “maison” ont vu le jour pour répondre aux besoins d'environnements spécifiques (e.g. Athapascan0 pour l'environnement Athapascan1 (ID-IMAG), Panda pour le compilateur Orca (université de Vrije), PM pour l'environnement SCORE (université de Tsukuba, Japon), PM² pour les régulateurs de charge par migration de processus légers), entraînant du même coup une certaine duplication des efforts de développement sur des architectures pourtant souvent identiques.

C'est pourquoi de récents efforts ont été entrepris pour tenter de définir des “*micro-supports exécutifs*” (par analogie avec les micro-noyaux) qui définiraient en quelque sorte des supports minimaux génériques sur lesquels on pourrait bâtir un large éventail de supports exécutifs de plus haut niveau. Un exemple de tel micro-support est μ PM² (habilitation de R. Namyst au LIP), qui réalise une intégration des communications et du multithreading sans imposer de modèle d'exécution particulier, sur des grappes de multiprocesseurs. De tels travaux ont effectivement permis une meilleure réutilisation des supports exécutifs au sein de différents environnements, et une plate-forme telle que μ PM², par exemple, a été utilisée avec succès comme base pour l'implémentation d'une machine virtuelle Java distribuée (collaboration LIP - UNH, USA), d'un courtier d'objets Corba (LIFL, France), d'un environnement multi-applications pour la grille (IRISA, France) et même d'une version multi-protocoles de la bibliothèque MPICH au sein de notre équipe.

Aujourd'hui, même si plusieurs problèmes restent encore ouverts (optimisation des communications, réactivité des applications aux E/S), on peut tout de même dire que l'on dispose de supports exécutifs capables d'exploiter efficacement les grappes homogènes de petite taille. En revanche, il reste beaucoup de travaux à mener pour parvenir à maîtriser les configurations de grande taille, hiérarchiques et potentiellement hétérogènes (i.e. grappes de grappes), qui posent une multitude de nouveaux problèmes tels que le routage haute performance des messages en contexte hétérogène, la gestion de configurations dynamiques (tolérance aux pannes). En matière de grappes de grappes hétérogènes, il faut noter l'existence de travaux de recherche intéressants visant à concevoir des bibliothèques de communication capables d'interconnecter virtuellement tous les nœuds deux à deux (e.g. MPICH-G2, PACX-MPI). Toutefois, ces travaux se focalisent pour l'instant sur des configurations à grande échelle (le protocole TCP/IP est utilisé sur les liaisons inter-grappes, les grappes étant supposées géographiquement dispersées) et l'approche est donc inadaptée aux configurations dotées de liens inter-grappes performants.

Par ailleurs, même si les quelques *success stories* évoquées précédemment démontrent que les supports exécutifs actuels apportent réellement une plus-value importante en termes de portabilité et de performances aux environnements parallèles complexes, il n'en demeure pas moins que beaucoup de progrès restent à faire dans le domaine de l'exploitation optimale des fonctionnalités des supports par les couches supérieures, qui les utilisent encore comme de simples "boîtes noires". Plus précisément, nous pensons que l'expertise acquise par un concepteur de support exécutif doit pouvoir être formalisée puis transférée de manière systématique (analyse de code, spécialisation) aux couches supérieures. À notre connaissance, il n'existe pas encore de tels travaux dans le cadre des supports exécutifs parallèles.

Axes de Recherche Le projet Runtime s'intéresse à l'étude et la conception des principes, ainsi qu'à la mise en œuvre et à l'évaluation des mécanismes qui seront au cœur des **supports exécutifs parallèles** de demain. Plus précisément, il s'agit de définir, d'implanter et de valider une famille de supports exécutifs la plus générique possible constituant une base flexible et performante pour la construction d'environnements/applications dans le domaine du calcul parallèle intensif. Ceux-ci devront permettre l'exploitation efficace des machines parallèles que sont les grappes de grande taille, hétérogènes et hiérarchiques.

Le projet de recherche s'articule autour de trois *challenges* principaux :

Maîtriser les configurations hétérogènes, de grande taille. Il s'agit de proposer de nouveaux modèles, principes et mécanismes permettant d'organiser les communications (en particulier le routage haute performance en contexte hétérogène), l'ordonnancement des threads et les entrées/sorties sur ce type d'architectures, de manière à la fois portable et performante. Il s'agit également de réfléchir à l'introduction des nécessaires propriétés de dynamique, de tolérance aux pannes et de passage à l'échelle au sein de cette nouvelle génération de supports exécutifs, en minimisant l'impact négatif sur les performances des applications.

Exploiter les technologies nouvelles de manière optimale. Il s'agit de continuer une activité de suivi technologique pointue (réseaux, processeurs) de façon à bien rester en phase avec les contraintes réelles imposées par les machines de "production" et surtout à comprendre comment exploiter ces nouvelles technologies de manière performante (ex : prise en charge de nouveaux paradigmes dans les modèles de communication) Sur ce point, il est clair que nous devons poursuivre les travaux amorcés sur l'expressivité des interfaces proposées qui permettent de séparer les contraintes applicatives des optimisations effectuées par le support exécutif. En particulier, nous sommes sur le point de démarrer deux études : l'une sur la technologie Infiniband et les optimisations possibles des communications, l'autre sur le guidage de l'ordonnancement de threads sur multi-processeurs SMT.

Améliorer l'intégration avec les environnements et les applications. Il s'agit d'explorer la frontière entre les supports exécutifs et les environnements de plus haut niveau afin d'étudier la mise en place d'optimisations encore plus performantes. Différentes voies sont à suivre : proposition d'extensions des interfaces des environnements qui encapsulent trop souvent l'interface du support sous-jacent, exploitation d'information fournies par un analyseur de programme dans le but d'améliorer les heuristiques utilisées au sein des supports, raffinement du code applicatif par un spécialiste à l'aide d'informations remontées par le support au moment du déploiement, etc.

Description des activités et principaux résultats

Nos travaux s'articulent autour de 3 axes qui sont l'ordonnancement de processus sur les architectures multiprocesseurs, la communications sur réseaux rapides et les supports d'exécution pour grilles de calcul. Ces travaux sont intégrée à notre suite logicielle PM2 diffusée librement et utilisée par des chercheurs avec qui nous menons des collaboration contractualisée CEA/DAM avec qui collaborons au sein des ACI Grid (Grid 5000, RMI, ALTA) et Masse de Données (GridExplorer), par des industriel (Alcatel, Bull) ou pour l'enseignement (Université de Bordeaux, Calais, Nice, INSA Lyon).

Ordonnancement de processus sur multiprocesseurs La bibliothèque Marcel synthétise nos travaux sur l'ordonnancement des processus. Marcel offre toute la souplesse et l'efficacité des processus légers et est spécialisable selon le type d'architecture visée. Ainsi, sur une machine multiprocesseur symétrique, Marcel adoptera le modèle des processus légers hybrides (ordonnancement à deux niveaux) qui est le plus à même pour exploiter pleinement ce type d'architecture.

Cependant l'utilisation conjointe de processus légers et de bibliothèques de communication en environnement hautes performances fait apparaître des difficultés. Ces deux composants ne se marient pas très bien, même s'ils sont performants indépendamment l'un de l'autre. Des propriétés comme la *réactivité* des applications peuvent grandement souffrir de l'ignorance réciproque de ces deux composants. Dans sa thèse soutenue en décembre 2004, Vincent Danjean a conçu pour Marcel un *serveur d'événements* répondant, partiellement, à ce problème de réactivité. Il permet à l'application (ou aux bibliothèques extérieures) de proposer un ensemble de méthodes de détection d'événements. Ce serveur choisit alors la plus adaptée au contexte, en fonction du type de bibliothèque de processus légers et des fonctionnalités du système disponibles. Cette sélection est transparente pour l'application à qui il est fourni une *interface uniforme* pour attendre des événements. Intégrer le serveur à la bibliothèque de processus légers a plusieurs avantages : le choix de la méthode de détection adaptée est facilité par la connaissance des propriétés de la bibliothèque de processus légers elle-même. En cas de scrutation, l'ordonnanceur est capable de garantir une fréquence régulière indépendante du nombre de processus légers prêts et sans nécessiter de changements de contexte inutiles. Enfin, les requêtes de même type peuvent éventuellement être agrégées et réduire ainsi le coût de la détection d'événements. Cette réponse n'est cependant que partielle car elle ne permet pas toujours de minimiser le temps de réaction aux événements matériels : seule l'utilisation du mécanisme interruption matériel permet d'obtenir les meilleures performances, il s'agirait donc d'utiliser des appels systèmes bloquants. Mais ces derniers, utilisés sans précaution, paralyseront l'application entière au lieu du seul processus léger incriminé. Pour résoudre cette difficulté, nous avons implémenté le modèle des *Scheduler Activations* d'Anderson après l'avoir adapté et amélioré dans notre contexte d'utilisation (calcul hautes performances) sous Linux.

Dans le même esprit et dans le cadre d'une thèse réalisée en co-tutelle avec le CEA, Marc Pérache étudie l'intégration au sein des ordonnanceurs de mécanismes de communication entre threads. Il s'agit d'implémenter efficacement et de manière *portable*, les opérations de communication (barrières de synchronisation, opérations de réduction, communications point à point, ...) entre threads sur des architectures multiprocesseurs hiérarchiques. L'ordonnanceur détient en effet toutes les informations lui permettant d'optimiser les transferts de données, les outils de synchronisation à utiliser ou encore les processeurs inactifs utilisables pour déporter certains traitements. Les performances obtenues [876] sur les opérations élémentaires ainsi que sur des codes de calcul scientifique montrent la supériorité de l'approche

sur des solutions logicielles classiques mises en œuvre en dehors de l'ordonnanceur.

Depuis 2004, dans le cadre de son mémoire de DEA et de sa thèse, Samuel Thibault adapte la bibliothèque Marcel aux serveurs de calcul à architecture très hiérarchisée tels les machines NUMA basées sur des processeurs multi-cores. Pour cela, nous avons défini une interface permettant à l'application de décrire de façon structurée les relations d'affinités qu'entretiennent ses différents flots de calculs. Par exemple, cette interface permet de regrouper dans des classes d'équivalence imbriquées les processus qui travaillent sur les mêmes données. L'ordonnanceur de la bibliothèque Marcel utilise alors cette information pour répartir le calcul sur la machine hiérarchique afin de pouvoir tirer parti des effets de cache tout en limitant le coût du facteur NUMA. Ce travail est réalisé en collaboration avec le CEA/DAM et la société Bull. Un prototype a été développé [878] et a été évalué au CEA par Marc Pérache [876] afin de comparer l'efficacité de différentes stratégies d'ordonnement de l'exécution d'une application de calcul de conduction de chaleur. Les résultats obtenus sont très prometteurs.

Enfin, pour pouvoir suivre le déroulement exact et précis des applications, en particulier pour connaître l'ordonnement exact des processus légers sur les processeurs physiques de la machine, nous avons conçu et implémenté des mécanismes de traces [818, 857] peu intrusif. Grâce à la récolte d'une trace noyau et d'une trace applicative, nous sommes en mesure de reconstituer l'exécution précise des applications.

Communications sur réseaux rapides La bibliothèque de communication Madeleine (thèse d'Olivier Aumage au LIP) et l'environnement de programmation MPICH-Madeleine rassemblent nos travaux sur l'axe *communications sur réseaux rapides*.

Madeleine, socle de nos travaux en matière de communication, pilote des technologies réseau aussi diverses que Quadrics, Myrinet/GM, Myrinet/MX, SCI ou encore VIA avec des performances extrêmement proches des limites matérielles et tout en satisfaisant aux exigences de portabilité propres aux supports exécutifs [773]. L'interface de Madeleine est fondée sur une construction incrémentale des messages. Chaque segment de message possède ses propres contraintes de transmission et de réception. Son originalité est de permettre une sélection dynamique d'optimisations de transmission des données en fonction de ces contraintes, exprimées par l'application. La sélection proprement dite des optimisations met en œuvre des heuristiques intégrant les caractéristiques du réseau sous-jacent ainsi que la taille et les contraintes de chacun des segments traités depuis le début d'un message.

L'environnement de programmation MPICH-Madeleine est une implémentation haute performance du standard MPI (*Message Passing Interface*) reposant sur Marcel et Madeleine. Cet environnement cible en particulier les configurations nécessitant l'utilisation concertée de plusieurs protocoles telles les configurations hiérarchiques ou hétérogènes comme les grappes de grappes. Dans sa thèse soutenue en décembre 2004, Guillaume Mercier a repensé la façon dont la progression des communications est orchestrée dans les implémentations actuelles de MPI et ce en introduisant un moteur multithreadé de progression des communications. Dans ce cadre, une fois la communication initiée, sa progression s'effectue indépendamment du calcul. Cette technique apporte non seulement de substantiels gains de réactivité (d'autant plus que l'on bénéficie du serveur d'événements de Marcel) mais aussi permet aux applications multithreadées d'utiliser le standard MPI (notre implémentation est dite "MPI_THREAD_MULTIPLE"). Enfin, l'utilisation de Madeleine comme interface de communication de bas-niveau générique offre un haut niveau de performances et les services indispensables pour leur conservation dans des environnements complexes comme ceux des grappes de grappes hétérogènes. Ainsi les performances obtenues avec notre implémentation sont similaires ou parfois même meilleures que celles obtenues par des implémentations MPI op-

timisées pour une technologie réseau particulière comme l'a montrée une étude comparative réalisée avec l'équipe CoC-Grid du Professeur Rehm de l'université de Chemnitz.

Mentionnons que Guillaume Mercier a été recruté en tant que chercheur contractuel en post-doctorat au laboratoire états-unien Argonne (au sein de l'équipe qui a élaboré MPICH) où il travaille actuellement à la transposition dans MPICH2 de certains des mécanismes développés au cours de sa thèse.

Support d'exécution pour grilles de calcul L'axe *support d'exécution pour grilles de calcul* est incarné essentiellement par les travaux sur la plate-forme de communication PadicoTM, dont la conception a démarré pendant les travaux de thèse d'Alexandre Denis à l'IRISA et est poursuivie dans le projet Runtime. La plate-forme de communication PadicoTM permet à une large palette d'intergiciels — MPI, CORBA, Java RMI, Web Services, DSM, etc. — d'être déployés sur les réseaux des grilles de calcul. PadicoTM repose sur Madeleine pour les communications intra-grappe, et sur des méthodes spécifiques aux réseaux longue distance pour les communications inter-grappe et inter-site.

L'originalité de l'approche de PadicoTM réside dans sa généralité par rapport aux intergiciels. Alors que la plupart des travaux autour des communications réseaux sur les grilles s'articulent autour d'un intergiciel donné (MPI pour MPICH-G2, modèle spécifique pour Ibis, PVM pour Harness) qui impose un modèle de programmation, PadicoTM n'est pas lié à un intergiciel particulier. La gestion des communications est faite de façon *générique*, et l'accès au réseau est offert au travers d'interfaces standard permettant d'utiliser de nombreux intergiciels différents. Ainsi, PadicoTM propose par exemple une implémentation CORBA sur les mêmes technologies réseau et aussi performante que MPI. Le programmeur d'application peut alors choisir l'intergiciel le plus adapté à son modèle de programmation, indépendamment des contraintes de portabilité des intergiciels sur les différents réseaux.

Pour parvenir au niveau de flexibilité nécessaire pour s'adapter aux réseaux très variés des grilles de calcul, PadicoTM est construit selon les principes des *composants logiciels*, assemblables à la carte selon les besoins et dynamiquement. Pour chaque intergiciel supporté et chaque réseau pris en charge par PadicoTM, il existe un assemblage de composants de communication permettant à l'intergiciel donné d'utiliser le réseau choisi. De plus, cette architecture permet d'ajouter facilement le support de nouveaux réseaux ou de nouvelles méthodes de communications, sans devoir remettre en question ce qui existe déjà.

Par ailleurs, suite à des séjours post-doctoraux à l'université Vrije, Olivier Aumage et Alexandre Denis collaborent au projet Ibis d'Henri Bal. Ce projet a pour objectif la production d'une plateforme de programmation pour grille et, ce, en s'appuyant sur le paradigme "*compile once, run everywhere*" du langage Java. Ainsi des applications peuvent être exécutées sans recompilation malgré l'extrême hétérogénéité propre aux grilles. Leurs travaux concernant le support des communications et de la connectivité en contexte de grille ont débouché sur la production du logiciel NetIbis [821, 772]. Le logiciel NetIbis constitue le sous-système réseau de la plateforme Ibis et prend en charge l'ensemble des communications, qu'elles s'effectuent sur un réseau local rapide tel que Myrinet ou sur des interconnexions longues distances tels que les épines dorsales des grilles.

Projets et perspectives 2005-2008

Ordonnancement de processus sur multiprocesseurs Afin d'optimiser l'exploitation des machines à architecture très hiérarchisée, nous allons développer et expérimenter des algorithmes de distribution adaptatifs, répartissant la charge de calcul de manière appropriée selon les informations d'affinités entre flots de calculs, et selon leurs poids respectifs (qui

peuvent évoluer au cours du temps). Afin d'améliorer plus encore la qualité de la répartition, il sera également intéressant de fournir aux applications des primitives d'allocation mémoire plus expressives que celles fournies usuellement par les environnements d'exécution. L'ordonnanceur de la bibliothèque Marcel pourra alors décider d'éventuelles migrations de données lorsqu'une nouvelle répartition des flots de calculs est établie. Il sera enfin envisageable d'intégrer nos algorithmes de distribution adaptatifs au noyau Linux pour répartir de manière adaptée l'ensemble des processus et leurs threads d'un système.

Communications sur réseaux rapides La nature des applications, les caractéristiques des plateformes de calculs, et plus généralement les usages, évoluent :

- Les applications deviennent modulaires, composites, chacun des fragments applicatifs étant un consommateur de ressources réseau en puissance; quelque soit la technologie de communication utilisée, le matériel n'est pas en mesure d'isoler de telles quantités de flux communicants.
- Les plateformes intègrent un nombre croissant d'unités de calcul. Il n'est plus combinatoirement envisageable de connecter directement chaque unité à toutes les autres.

En conséquence de ces observations, les communications doivent désormais être appréhendées de façon globale, c'est-à-dire en considérant les flux entrants et sortants dans leur ensemble, plutôt qu'isolément dans le contexte restreint de leur connexion point-à-point propre. Il s'agit donc à présent de concevoir Madeleine, non plus comme une interface générique d'accès aux technologies réseau mais bien comme un ordonnanceur/optimizeur général de flux de communication. L'optimisation des communications nécessitera entre autres d'élaborer des stratégies de réagencement dynamique des transferts, en s'appuyant sur des annotations ou méta-données applicatives, ou encore sur une vue globale des communications en instance obtenue à l'aide d'outils d'analyse de code, par exemple.

L'environnement MPICH-Madeleine jouit maintenant d'une bonne visibilité internationale, comme l'atteste les publications étrangères du domaine y faisant référence. Notre implémentation de MPI est désormais reconnue comme une des solutions majeures pour l'exploitation des configurations de type "grappes de grappes", voire des grilles, même si cela n'était pas notre objectif premier. À ce titre des contacts ont été établis avec les équipes de Yutaka Ishikawa (université de Tsukuba, Japon) et de William D. Gropp (Laboratoire Argonne, USA). Dans la perspective de doper la diffusion de MPICH-Madeleine, il s'agit tout d'abord de consolider cet environnement tout en continuant d'améliorer ses performances en optimisant certains schémas de communication. Par exemple, on peut envoyer directement les types utilisateurs représentant des données non contiguës en mémoire au-dessus de Madeleine au lieu d'utiliser la méthode initiale de MPI qui regroupe les données dans un nouveau buffer afin de les envoyer comme une zone de mémoire contiguë. Un prototype montrant de bonnes performances dans le cas de données découpées en quelques blocs de grande taille a été développé par Nathalie Furmento et a fait l'objet d'un rapport interne de recherche.

Support d'exécution pour grilles de calcul Les configurations évoluant vers une plus grande taille, un plus grand nombre de sites, et une plus grande dynamique, il reste encore de nombreux défis à relever :

- gérer les communications en présence de filtrage réseau, de pare-feux et de réseaux privés non-routés, qui sont mis en œuvre de plus en plus fréquemment dans les grappes nouvellement mises en service, pour répondre aux nouvelles exigences de sécurité sur internet. D'une manière plus générale, prendre en compte les aspects liés à la sécurité résultant de l'utilisation de réseaux non-dédiés.

- étudier les interactions entre les réseaux longue distance et les réseaux haute performance, et notamment réduire l'impact de la scrutation de ces premiers sur la performance de ces derniers.
- choisir automatiquement les méthodes de communication à utiliser en fonction d'observations, plutôt que de se reposer sur une configuration donnée par l'utilisateur qui est de moins en moins en mesure de donner ces informations à mesure que les configurations s'étendent sur des sites plus nombreux.
- adapter dynamiquement les méthodes de communications en fonction des changements des conditions de trafic réseau, de réservation de ressources ou de configuration, sans intervention de l'utilisateur et sans arrêter l'application.

Un objectif à moyen terme du projet serait de parvenir à un intergiciel "unique" pour la grille, capable d'exploiter de manière performante l'épine dorsale du réseau ainsi que les configurations situées aux extrémités. Il s'agit de réaliser une véritable mutualisation de certaines couches logiques de Madeleine et de PadicoTM, plutôt qu'un empilement des bibliothèques. Plus précisément, les objectifs secondaires sont de parvenir à découpler l'optimisation des paquets et la réalisation des transmissions sur la topologie réseau, ainsi que de garantir une adaptation aisée à de nouvelles technologies (Quadrics, InfiniBand, etc.). L'enjeu est donc de maîtriser toute la chaîne des optimisations des transferts, et ce de "bout en bout", entre n'importe quelles machines sur la grille, et pour n'importe quel environnement cible (MPI, Corba, etc.)

5.3.4 ScAlApplix

Introduction

SCALAPPLIX (Schémas et Algorithmes Hautes Performances pour les Applications Scientifiques Complexes) est un projet commun du Laboratoire Bordelais de Recherche en Informatique (LaBRI, UMR 5800), du Laboratoire de Mathématiques Appliquées de Bordeaux (MAB, UMR 5466) et de l'INRIA FUTURS. Ce projet a été créé le 1^{er} novembre 2002 (<http://www.labri.fr/scalapplix>).

L'objectif de SCALAPPLIX est la mise en œuvre de compétences scientifiques complémentaires pour une recherche pluridisciplinaire dans les domaines de l'informatique hautes performances et de la modélisation numérique, et ce dans le but d'analyser et de résoudre efficacement des problèmes de calcul scientifique provenant d'*applications complexes nécessitant des puissances de calcul et manipulant des tailles de données téraflopiques*. La résolution effective de cette gamme de problèmes est un véritable défi qui nécessite clairement une approche pluridisciplinaire. Cette synergie de compétences scientifiques concernent plus particulièrement en mathématiques appliquées le domaine des schémas numériques, et en informatique les domaines de l'algorithmique parallèle et du savoir-faire de la mise en œuvre de codes hautes performances sur diverses plates-formes de calcul.

Il s'agit donc, dans le cadre de cette synergie scientifique, de *contribuer à toutes les étapes de la chaîne qui va de la conception de schémas performants à la mise en œuvre optimisée et efficace de codes de simulation pour des applications complexes et/ou de très grande taille*. Cette chaîne contient aussi la partie interprétation des résultats ; ceci se fera en pilotant les codes de calcul hautes performances parallèles et/ou distribués par une application de type réalité virtuelle, faisant ainsi de la simulation numérique un outil efficace pour interagir avec la modélisation via une analyse de sensibilité interactive des paramètres de cette simulation.

Tout naturellement, ce projet contient un important volet de transfert technologique et de valorisation dans le cadre de collaborations avec des industriels ayant à traiter des applications grandes consommatrices de puissance de calcul et manipulant des grandes masses de données.

Dans ce rapport d'activités, nous nous restreindrons à décrire la partie de nos travaux plus spécifiquement rattachée aux thématiques du LaBRI. En ce qui concerne le MAB, les membres permanents qui font partie de SCALAPPLIX sont Rémi Abgrall (Professeur et IUF), Boniface Nkonga (MCF et HDR), Christophe Berthon (MCF) et Mario Ricchiuto (CR2 INRIA) ; les membres non permanents sont Mikaël Papin (Postdoctorant), Benjamin Braconnier (Doctorant CEA) et Cédric Tavé (Doctorant CNES). Nous donnons aussi une description des résultats obtenus en collaboration avec M.C. Counilh et B. Cirou au titre de l'ancienne équipe ALiENor dans le domaine des modes d'expression pour le parallélisme.

Mots-Clés Calcul hautes performances, Algorithmique parallèle, Régulation de charge et ordonnancement, Algèbre linéaire creuse, Grappes et grilles de calcul, Pilotage de simulations numériques, Applications irrégulières de grande taille multi-physiques et multi-échelles, Mécanique des fluides, Dynamique moléculaire, Dynamique des populations.

Description des activités et principaux résultats

Algorithmique et ordonnancement pour les plates-formes à grande échelle

Introduction et méthodologie Ces dernières années, de nombreux travaux ont porté sur la conception d'algorithmes efficaces pour les plates-formes distribuées à grande échelle. On peut classer ces plates-formes en fonction de leur stabilité, aussi bien au niveau des ressources (de calcul et de communication) que de la topologie. A une extrémité, on trouve les grilles de calcul dédiées, pour lesquelles la topologie et les performances des ressources sont fixées. A l'autre extrémité, on trouve les systèmes pair à pair, pour lesquels la topologie (à cause des connexions et déconnexions fréquentes) et les ressources (par l'utilisation de liens internet et de machines non dédiées) varient très rapidement. L'hétérogénéité est une caractéristique essentielle de ces plates-formes, aussi bien au niveau des ressources de calcul que des ressources de communication.

L'essentiel de notre travail a consisté à prendre en compte cette hétérogénéité, pour la conception d'algorithmes d'ordonnement de calculs et de communications collectives. La littérature en ordonnancement regorge de résultats de NP-Complétude et d'inapproximabilité, même dans le cas de ressources homogènes et de modèles de communication peu réalistes (multi-port), et la prise en compte de l'hétérogénéité et de modèles de communication plus réalistes complique encore les problèmes ! Toutefois, la nature des plates-formes, en particulier les latences importantes et le temps de déploiement des applications, réduit leur utilisation à des problèmes réguliers qui s'expriment généralement sous la forme d'un grand nombre de tâches indépendantes ou d'un grand message à échanger (qui peut être découpé). En s'inspirant des travaux de Bertsimas et Gamarnik sur le routage de paquets, nous avons changé la métrique usuelle en ordonnancement, qui mesure le temps de complétion de la dernière tâche, pour la maximisation du débit, qui mesure le nombre d'opérations effectuées par unité de temps.

Ordonnement de calculs L'utilisation de la maximisation du débit a permis de concevoir de nombreux algorithmes polynomiaux sous un modèle de communication et de recouvrement calcul/communication réaliste. L'idée pour concevoir des algorithmes efficaces du point de vue de la maximisation du débit est de se concentrer sur les opérations "en régime permanent". On cherche la solution sous la forme d'une somme pondérée de schémas d'allocation et on écrit les contraintes pour chaque ressource (de calcul ou de communication). Par exemple, si on considère l'allocation d'un grand nombre de tâches indépendantes initialement générée par un processeur maître, un schéma d'allocation correspond à un chemin dans le graphe de la plate-forme entre le maître et l'esclave sur lequel la tâche sera traitée. Les contraintes de ressources expriment qu'en temps unitaire, le temps de calcul sur chaque esclave, le temps de communication sur chaque lien et l'occupation de chaque port d'entrée ou de sortie de chaque machine, sont tous inférieurs à 1. Cette écriture des problèmes conduit généralement à la solution d'un programme linéaire en nombre rationnels, qui peut donc être résolu en temps polynomial. Une phase de reconstruction est alors nécessaire, pour passer de la solution rationnelle à un ordonnancement effectif des calculs et des communications.

Cette technique a été appliquée avec succès à la distribution d'un grand nombre de tâches indépendantes identiques sur des plates-formes quelconques. Des généralisations à certains types de graphes de tâches (on suppose toujours disposer d'un grand nombre de tâches indépendantes, mais chaque tâche peut elle-même être exécutée en parallèle) ont également été proposées. Dans le cas où la profondeur de dépendance du graphe est bornée, nous avons montré que ce dernier problème peut être résolu en temps polynomial. Néanmoins, pour un graphe de tâches quelconque, la recherche du débit maximal reste un problème NP-Complet.

Récemment, nous avons formalisé l'ensemble de ces travaux sous un cadre unifié, en utilisant la méthode des ellipsoïdes pour résoudre en temps polynomial des programmes linéaires dont le nombre de contraintes peut être exponentiel. Nous avons également développé un

cadre général pour passer de la solution du programme linéaire à un ordonnancement des calculs et des communications. A défaut de fournir des algorithmes utilisables "en pratique" (du fait de l'utilisation de la méthode des ellipsoïdes), ces travaux permettent de déterminer très rapidement la complexité des problèmes d'ordonnancement en régime permanent.

Ordonnancement de communications collectives Plus récemment, nous nous sommes intéressés à la conception d'algorithmes de diffusion efficace. L'idée ici est de considérer un message de grande taille (pour appliquer les techniques de maximisation de débit) qui doit être diffusé à partir d'une source unique à l'ensemble des nœuds d'une plate-forme (broadcast) ou à un sous ensemble de ces nœuds (multicast). La principale difficulté supplémentaire associée à ces problèmes provient du fait que le même message doit être envoyé à l'ensemble des nœuds participant, ce qui oblige à considérer des schémas d'allocation plus complexes.

Nous avons montré que le problème de la diffusion totale (broadcast) peut être résolu en temps polynomial. Par contre, le problème de la maximisation du débit pour le calcul de préfixes parallèles et la diffusion partielle (multicast) restent NP Complets.

Algorithmes et solveurs hautes performances

Factorisations complètes La résolution de très grands systèmes linéaires creux $Ax = b$ est une brique de base algorithmique fondamentale pour les applications scientifiques de calcul intensif. C'est souvent l'étape la plus consommatrice aussi bien en temps CPU qu'en espace mémoire. La taille des systèmes utilisés pour les grandes applications complexes fait que le calcul hautes performances est de fait incontournable (on doit résoudre aujourd'hui des systèmes de *plusieurs millions ou dizaines de millions d'inconnues* pour des problèmes 3D).

Dans nos travaux, nous avons contribué doré et déjà à l'analyse, la conception et le développement d'outils logiciels parallèles concernant les méthodes *directes* de résolution qui sont bien connues pour être très robustes numériquement même pour des problèmes très mal conditionnés comme ceux provenant du domaine de la mécanique des structures. L'algorithmique mise en place est très fine et nécessite essentiellement des techniques de mathématiques discrètes issues de la théorie des graphes. Ainsi, nous avons travaillé sur toutes les étapes de post-traitement des données, et sur les solveurs parallèles eux-mêmes.

Le premier objectif de cet axe de recherche a été tout d'abord de continuer autant que faire se peut à augmenter la qualité et les fonctionnalités des solveurs parallèles utilisant des méthodes directes. Ainsi, le solveur PASTIX bien que déjà très performant a été rendu encore plus performant en poursuivant l'amélioration de la qualité du prétraitement des données et des calculs ainsi qu'en améliorant sa scalabilité mémoire. Ces travaux ont concerné les domaines suivants :

- *le partitionnement de graphes irréguliers de très grandes tailles*. Il s'est agi ici d'étudier comment obtenir un partitionneur *parallèle* ayant une très bonne efficacité et permettant d'obtenir au moins la même qualité de séparation que les meilleurs logiciels séquentiels actuels.
- *le partitionnement de maillages d'éléments finis*. L'utilisation de la structure native d'un maillage d'éléments finis plutôt que de celle du graphe d'adjacence permet une réduction significative du coût mémoire. De plus, en utilisant des parcours efficaces dans la structure du maillage et en étendant les algorithmes de base du partitionneur, le surcoût en temps reste très acceptable en particulier pour les maillages 3D de grande taille.

- *une gestion plus optimisée de l'occupation mémoire.* Les méthodes directes sont intrinsèquement coûteuses en mémoire, et ce malgré le choix de renumérotations performantes permettant de bien conserver le creux des matrices au cours de la factorisation. Nous avons travaillé dans un premier temps sur des techniques d'aggrégation partielle sous contrainte mémoire.
- *l'extension aux clusters de nœuds SMP.* Le solveur PASTIX a résolu avec succès des systèmes 3D de plus d'un million d'inconnues et des systèmes ayant jusqu'à 25 millions d'inconnues pour des problèmes $2\frac{1}{2}$ D tous issus d'applications industrielles du CEA/CESTA. Ces exécutions ont atteint sur un grand nombre de processeurs (256 processeurs sur le SP3 du CINES) près de 50 % de la puissance de crête du calculateur. Une algorithmique spécifique a été mise en place pour utiliser au mieux les clusters de nœuds SMP et pour préserver une scalabilité mémoire raisonnable, en conservant dans un premier temps un modèle de programmation MPI. Néanmoins, cette version généraliste de PASTIX n'exploite pas une propriété importante de ces architectures qui est la possibilité d'utiliser la mémoire partagée comme un espace d'adressage unique pour tous les processeurs d'un même nœud via l'utilisation de threads. Devant l'importance que prennent les architectures à base de nœuds SMP (allant jusqu'à 64 processeurs par nœud SMP), nous avons travaillé sur une version dédiée de PASTIX pour de telles plates-formes de calcul. Ainsi nous nous sommes orientés vers une programmation hybride Thread/MPI du solveur qui a pour principe de partager les données d'un même nœud SMP via l'utilisation de threads POSIX. Toutes les communications entre les processeurs d'un même nœud sont remplacées par des ajouts locaux dans la matrice globale distribuée par nœud et non par processeur comme c'était le cas jusqu'à présent. Outre l'avantage d'un système de d'échange beaucoup plus performant, cette version offre l'avantage algorithmique très important qui est celle de la scalabilité mémoire.
- *l'ajout de nouvelles fonctionnalités dans PASTIX.* Nous avons étendu les fonctionnalités du solveur pour pouvoir traiter des matrices non symétriques mais à structure symétrique; dans ce cas, nous nous autorisons à faire du pivotage numérique statique ce qui nécessite après résolution une phase de raffinement itératif.
Nous envisageons également d'améliorer les techniques de distribution et d'ordonnement du solveur PASTIX en séparant plus clairement le parallélisme de tâches et le parallélisme de données.

Factorisations incomplètes par blocs Le second objectif principal de cet axe de recherche est de travailler, pour passer la limite en taille des solveurs parallèles précédents, sur des solveurs par blocs (pour profiter toujours de l'efficacité des primitives BLAS) *hybrides* basés sur un couplage fin entre méthodes directes et méthodes itératives. A partir de numérotations adaptées de type dissections emboîtées, nous calculons en parallèle un préconditionneur de type Cholesky incomplet par blocs qui sera ensuite intégré dans des versions parallèles par blocs d'algorithmes de type gradient conjugué ou GMRES. Le but de ces travaux est donc d'utiliser au mieux la technologie algorithmique parallèle par bloc qui est utilisée pour les solveurs directs haute performance pour développer des préconditionneurs parallèles de type Cholesky incomplet de manière adaptative et paramétrée, et qui soient robustes numériquement, suffisamment économes en mémoire et très performants en temps CPU (puissance tirée par processeur importante, bonne scalabilité). L'idée centrale de cette première approche consiste donc à trouver un compromis entre une diminution importante de la taille mémoire pour stocker le facteur L de la méthode directe, et une conservation d'une certaine dose de remplissage (les blocs conservés seront toujours considérés comme pleins) pour exploiter suffisamment les effets superscalaires dans les calculs BLAS3 du précon-

ditionneur et atteindre globalement de bonnes performances en temps.

Le premier point crucial est donc le choix de la partition initiale pour la factorisation symbolique complète. Cela ne peut pas être directement la partition supernodale issue de la technique de renumérotation que nous utilisons car il faut être capable de creuser les blocs diagonaux associés à cette partition. Nous utilisons donc une partition raffinée de notre partition supernodale initiale en définissant un paramètre de blocage qui définira ainsi le bloc élémentaire plein de notre préconditionneur. C'est donc ce paramètre important qui fixera le compromis décrit plus haut.

Le deuxième point crucial concerne le calcul de différents critères pour ensuite supprimer des blocs de la structure calculée sur notre partition raffinée. On peut calculer par exemple pour chaque bloc le level-fill durant l'élimination par bloc, le nombre de fois où il est modifié, le nombre de fois où il participe à une contribution modifiant un autre bloc.

Cette étude a été initiée dans le cadre d'un Projet NSF/INRIA mettant en commun les compétences de spécialistes américains (Yousef Saad/Minneapolis, Randall Bramley/Indiana, Esmond Ng/Berkeley, Maria Sosenkina/Minneapolis, John Gilbert/Xerocs) et français qui sont parmi les membres participant à ce projet.

Factorisation incomplète multi-niveaux Depuis quelques années, différentes techniques de factorisation incomplète LU ont été développées dans le but de combiner la généralité des préconditionneurs ILU avec la très bonne scalabilité des méthodes multi-grilles. La caractéristique principale de ces techniques est de réordonner le système pour pouvoir extraire du parallélisme de manière *naturelle*.

Dans ce contexte, nous avons généralisé les notions de "faces" et "d'arêtes" utilisées dans les décompositions de type "Wire-Basket" à des graphes irréguliers quelconques. Nous avons ainsi défini une *décomposition hiérarchique de l'interface* entre des sous-domaines obtenus par partitionnement du graphe d'adjacence de la matrice creuse du système à résoudre. A partir de cette décomposition, nous avons défini une factorisation ILU par niveaux successifs. La factorisation obtenue par ce principe peut être vue comme un calcul récursif du complément de Schur par classe (ou grille) de connecteurs.

Les algorithmes de décomposition hiérarchique, de factorisation multiniveaux et différentes méthodes itératives ont été implémentés dans une bibliothèque qui a pour nom PHIDAL (Parallel Hierarchical Interface Decomposition Algorithms).

Pilotage de simulations numériques distribuées par la visualisation

Introduction L'accroissement des puissances de calcul et de visualisation a entraîné de nouveaux besoins chez les modélisateurs, les développeurs, comme la visualisation "en ligne", l'interaction avec l'application depuis la visualisation. La chaîne de traitement classiquement utilisée en mode batch à savoir modélisation, simulation puis post-traitement se trouve modifiée. Maintenant l'utilisateur souhaite un retour immédiat, visuel ou sonore, en dépit de la complexité des équations qui régissent le modèle ou du très grand nombre de données à manipuler. Cette visualisation interactive doit être adaptée au besoin de l'utilisateur. Cette approche permet à un utilisateur de superviser en temps réel ses simulations numériques et d'interagir directement avec elles. Les environnements de pilotage et de visualisation sont aujourd'hui au carrefour d'axes de recherche porteurs comme la réalité virtuelle, le Grid computing et la simulation numérique interactive.

Un environnement de pilotage peut-être défini comme une infrastructure de communication couplant une simulation avec une interface utilisateur distante, appelée le système de pilotage. La plupart des environnements, comme le plus connu CUMULVS, sont basés sur

une instrumentation du code source de l'application. Cette approche autorise un pilotage à grain fin et fournit de bonne performance à l'exécution. Concernant la plate-forme de communication, et lorsque l'on considère des applications parallèles et distribuées, les points suivants sont importants : le transfert de données hétérogènes, les protocoles de communication et la redistribution des données. Les applications scientifiques que l'on vise sont des applications de grandes tailles parallèles ou distribuées pouvant coupler différents codes spécialisés. On trouve ces applications dans tous les domaines scientifiques comme par exemple en aéronautique (interaction écoulement-structure), en biologie (modélisation du coeur), en climatologie (interactions océan-atmosphère-glace), matériaux (propagation d'une fissure).

L'objectif principal de cet axe est de proposer un modèle de pilotage, et de développer un environnement pour permettre le pilotage de codes parallèles disséminés sur le réseau, tout en offrant une interaction forte entre eux et les utilisateurs. Cet environnement est constitué d'une interface pour les codes de simulation permettant l'extraction en temps réel des données du calcul ainsi que leur modification, d'une plateforme de communication adaptée et enfin d'applications de visualisation et d'interaction. Pour aborder ce couplage entre un code de visualisation et des codes de simulation parallèles, nous nous engageons dans une démarche générique sans se focaliser sur une application particulière. Ce projet a été initialisé dans le cadre d'une ACI-GRID, regroupant des spécialistes de l'interaction mais aussi des applicatifs de différents domaines comme la mécanique des fluides, la chimie moléculaire et l'environnement.

Résultats Pour réaliser un tel environnement de pilotage (EPSN), nous nous intéressons aux problèmes suivants. Le premier concerne la modélisation des simulations parallèles ou distribuées. Il s'agit de définir un modèle représentant le flot d'exécution de la simulation pour agir sur les données de manière sûre, cohérente et efficace. Le deuxième aspect, concerne la redistribution des données (pouvant être des tableaux, des maillages ou des particules) entre la simulation et la visualisation parallèle.

Notre approche est basée sur une modélisation de la simulation et une annotation des sources au travers de l'API. Cette modélisation est réalisée à l'aide d'un graphe de tâches reproduisant le flot d'exécution de l'application. Le premier modèle a été étendu aux applications parallèles de type SPMD. Ce modèle découpe l'application en différentes tâches hiérarchiques ; on obtient ainsi un graphe dirigé acyclique. Les tâches modélisent des sections de codes, des structures itératives ou conditionnelles. Cette modélisation permet pour les applications SPMD d'avoir un système de datation précis de l'évolution des processus (caractère acyclique) et une robustesse face à des différences d'exécutions mineures (tâches hiérarchiques). Le modèle est décrit en XML et est utilisé dans la simulation par l'API. Le pilotage proposé est dirigé par le client, les requêtes de pilotage peuvent s'exécuter sur un point de pilotage ou une plage. L'utilisation de plages d'accès pour les requêtes de transferts de données ou pour le déclenchement d'actions permet de recouvrir les communications par le calcul de la simulation. Les requêtes des clients sont planifiées pour être exécutées à un moment précis de la simulation et cette exécution s'effectue de manière indépendante sur chaque processeur. Cette approche nous permet de garantir la cohérence des données et d'être efficace dans le traitement des requêtes en minimisant le surcoût.

Le deuxième point clé d'un tel environnement concerne la redistribution des données entre la simulation parallèle ou distribuée et la visualisation parallèle. Nous avons développé une bibliothèque de redistribution permettant d'échanger des données entre deux codes parallèles. Cette bibliothèque se compose de deux modules. Le premier module RedSYM décrit de manière générique et symbolique les données, à savoir le type, la distribution des données entre les processeurs et sur un processeur l'organisation en mémoire de ces données. Ayant

cette description, nous avons développé un algorithme permettant de construire en parallèle la matrice de communication de la distribution. Les données supportées sont des données régulières comme les tableaux, et des données irrégulières comme les particules et les maillages non structurés. Le deuxième module RedCORBA est la couche de communication réalisant le transfert des données au-dessus de l'intergiciel CORBA. Cette bibliothèque a été conçue comme un plug-in et s'interface avec EPSN et les objets parallèles PaCO++ développés au sein du projet PARIS de l'IRISA.

La plate-forme logicielle EPSN permet de piloter des applications parallèles existantes (écrites en C, C++ et Fortran). Elle permet de coupler la simulation avec une interface utilisateur fournissant de la visualisation à la demande et des interactions de pilotage. Elle est dynamique et permet à plusieurs clients de se connecter à une ou plusieurs applications. L'API permet de développer des clients de pilotage indépendamment d'une simulation particulière par exemple en QT VTK,...

Cet ensemble de travaux ont fait l'objet de la thèse d'Aurélien Esnard.

Schémas et algorithmes pour la chimie informatique

Introduction En raison de l'augmentation de la puissance des ordinateurs, de nouvelles applications telles que l'étude des chemins de réaction, le calcul d'énergie libre, les simulations de macromolécules, la propagation de fissure, sont régulièrement réalisées par les chimistes. Ces calculs simulent des systèmes allant jusqu'à plusieurs dizaines de milliers d'atomes voire quelques millions, et pendant de grandes échelles de temps, par exemple plusieurs nanosecondes. Pour obtenir des résultats réalistes, ces simulations doivent prendre en compte l'environnement de ces gros systèmes, par exemple de l'eau. Cela conduit à des systèmes pouvant aller jusqu'à environ plusieurs centaines de milliers d'atomes voire beaucoup plus en physique du solide. En outre, la simulation de l'agrégation des protéines, qui est critique pour des biologistes étudiant des virus, exige des modèles allant jusqu'à un million d'atomes, avec un temps de simulation pouvant aller jusqu'à la microseconde. Pour pouvoir simuler ces systèmes avec des temps raisonnables, les simulations atomistiques doivent être accélérées de plusieurs ordres de grandeur. Pour obtenir cette vitesse, les algorithmes numériques et parallèles utilisés aujourd'hui doivent être améliorés.

Nous nous intéressons à quelques aspects de ces problèmes. D'abord, nous essayons d'améliorer les modèles et les algorithmes. Pour cela, nous diminuons la complexité des algorithmes classiques en introduisant de nouvelles approximations dans les algorithmes et dans le modèle (c'est le cas pour des méthodes à croissance linéaire comme la méthode *divide-and-conquer*), et en proposant de nouveaux algorithmes. En second lieu, nous développons des méthodes multi-échelles pour diminuer le nombre d'atomes qui sont considérés au niveau le plus fin (électronique ou atomistique). L'idée est d'utiliser un modèle macroscopique (milieu continu) pour prendre en compte les effets électrostatiques de l'environnement sur le système, ou pour prendre en compte des conditions aux limites dans les cristaux. La difficulté ici réside dans la construction d'un schéma efficace qui couple les différentes échelles sans perdre les propriétés des schémas de chacun des modèles. L'algorithmique du couplage est importante pour atteindre l'efficacité désirée. Nous pouvons aussi récrire nos algorithmes sous forme de calculs par blocs afin d'employer des bibliothèques performantes comme les BLAS, ou pour mettre en place des stratégies d'équilibrage plus efficaces.

Résultats En ce qui concerne la méthode *Divide and conquer*, nous nous intéressons ici à déterminer l'état fondamental d'une molécule biologique par un modèle quantique semi empirique. Cet état est obtenu comme la solution d'un système aux valeurs propres non

linéaire; pour diminuer la complexité de la méthode, on va découper la molécule en différents morceaux et sur chaque morceau (ou sous domaine), on va résoudre séparément le problème aux valeurs propres, puis on recollera les solutions locales en classant les valeurs propres. Nous avons intégré cette méthode dans le code semi empirique QC++ que nous développons. La méthode *divide-and-conquer* a été parallélisée en utilisant la programmation par échanges de message. Les principales caractéristiques de la parallélisation sont les suivantes. Tout d'abord, plusieurs sous domaines ont été placés sur chaque processeur, ce qui nous permettra d'appliquer des méthodes d'équilibrage de charge. En second lieu, des communications asynchrones sont utilisées pour recouvrir les communications par des calculs. Deux schémas de communication sont utilisés : un anneau lorsque l'information doit être connue par tous les sous domaines et en point à point pour les relations de voisinages. Troisièmement, un partitionneur de molécules basé sur la bibliothèque SCOTCH et appelé KIMIKA a été développé. Plusieurs types de stratégies de découpage (par atome ou par fragment) ont été évaluées, et ont montré que lorsque le nombre de sous domaines augmente, le déséquilibre de charge peut être important. De nouveaux algorithmes de découpage sont à développer en tenant compte du recouvrement des sous domaines dans les fonctions de coût. Ces développements sont validés au travers de l'ACI SIRE.

L'étude de la propagation d'une fissure dans un verre de silice en couplant une méthode de dynamique moléculaire et une méthode d'élasticité a débuté en collaboration avec le CEA Ile-de-France en décembre 2003. La difficulté réside dans la transmission des informations entre un modèle discret et un modèle continu. Par exemple, on doit éviter que les ondes issues du milieu atomique rebondissent sur le milieu continu. Différents algorithmes ont été analysés et validés sur un modèle unidimensionnel et une analyse de la dispersion a été faite pour trouver les fréquences qui ne sont pas affectées par le couplage. Une interface logicielle générique permettant de coupler un code d'élasticité et un code de dynamique moléculaire est en cours de terminaison. Cette interface nous permet d'étudier le couplage dans les dimensions supérieures. Les aspects liés au calcul parallèle sont également considérés. Nous travaillons sur des simulations à grandes échelles issues de modèles tridimensionnels pour des matériaux réels. Ces simulations nécessitent de prendre en compte plusieurs millions d'atomes et doivent donc être fortement parallèles.

Dans la plupart des applications de calcul scientifique considérées de nos jours comme de grands challenges, par exemple les systèmes biologiques, l'astrophysique ou l'électromagnétisme, la résolution rapide d'un problème à N corps est cruciale. L'introduction des méthodes hiérarchiques, comme la méthode des multipôles rapides, et basées sur une structure en octree, réduisent la quantité de calcul requise pour simuler ces systèmes pour une précision donnée. Nous avons amélioré la vitesse de cette méthode en utilisant les bibliothèques BLAS. Une première amélioration basée sur des calculs par blocs dans les interactions entre particules pour évaluer le champ proche a été réalisée. Pour le champ lointain, une reformulation du problème en terme de produit matriciel a permis l'utilisation du niveau 3 de cette bibliothèque BLAS. En améliorant les structures de données, et en utilisant cette nouvelle formulation, nous avons amélioré très significativement les temps d'exécution de la méthode. Notre version a été comparée avec succès à des améliorations basées sur l'utilisation de FFTs et de matrices de rotation.

Algorithmique hautes performances en dynamique des populations

Introduction En dynamique des populations, les systèmes peuvent avoir des comportements complexes et être difficiles à analyser d'un point de vue purement mathématique. Le but de ce travail interdisciplinaire a été de développer des outils numériques pour les modèles de dynamique des populations concernant les systèmes complexes et hétérogènes de type hôte-parasite. Quelques hétérogénéités typiquement prises en compte sont la localisation spatiale, l'âge ou la capacité à recruter des macroparasites pour l'hôte, l'âge des macroparasites. Nos objectifs sont : comprendre l'impact de la structure d'une population hôte sur la dynamique de la population parasite, développer des simulations numériques précises à l'aide du parallélisme, modéliser des méthodes de prophylaxie. Pour beaucoup de systèmes hôte-parasite, différentes échelles de temps entre la population hôte (e.g. une période d'une année) et le virus (e.g. un hôte infecté meure en quelques semaines) requiert un petit pas de temps. Les schémas numériques du modèle épidémiologique non linéaire résultant sont complexes à réaliser en environnement spatialement hétérogène. Des résultats numériques fiables deviennent difficiles à obtenir quand la taille du domaine spatial croît. Par ailleurs, de nombreux paramètres (facteurs biologiques et environnementaux) sont pris en compte pour comparer les résultats de simulations avec les observations des études sur le terrain. Ainsi un simulateur réaliste a un coût de calcul important et une parallélisation est nécessaire.

Les modèles individus-centrés se montrent de plus en plus utiles pour décrire les systèmes biologiques. Les interactions entre les individus sont simples et locales, pourtant elles peuvent donner lieu à des déroulements complexes au niveau global. Le principe est de répliquer plusieurs fois la simulation afin d'obtenir des résultats statistiquement significatifs. L'approche des modèles individus-centrés contraste avec une approche modélisant la population de manière plus agrégée, en ce sens qu'elle fournit des mécanismes de bas niveau pour gérer les interactions de la population. Les simulations stochastiques reproduisent des processus élémentaires et aboutissent souvent à des complexités calculatoires prohibitives; nous avons donc encore besoin d'algorithmique parallèle.

Dans nos développements de modèles stochastiques et déterministes, les processus biologiques sont combinés pour atteindre un bon niveau de réalisme. Pour les systèmes hôte-parasite, cela fait une grande différence avec les modèles purement mathématiques, pour lesquels les résultats numériques peuvent difficilement être comparés aux observations. Les simulations numériques parallèles imitent certaines des dynamiques observées sur le terrain, et fournissent un outil utilisable pour valider les modèles. Ce travail a nécessité un effort de collaboration interdisciplinaire entre la dynamique des populations, les mathématiques et l'informatique.

Résultats Les ectoparasites des poissons interagissent continuellement avec leur population hôte. Un modèle décrivant les stratégies démographiques de telles populations de poissons et de parasites a été développé pour le système *diplectanum aequans*-bar. Ce modèle est principalement déterministe avec quelques aspects stochastiques. L'agrégation des macroparasites sur les hôtes n'est pas attribuée au modèle mais survient ou non en fonction de la dynamique de la population parasite. Les simulations discrètes, déterministes, précédentes avaient des temps d'exécution trop longs. On a implémenté un simulateur haute performance s'exécutant sur machines parallèles (IBM SP3, Regatta et SGI Origin 3800, machines parallèles du CINES – Montpellier) et fournissant des calculs plus précis. L'étude algorithmique et l'analyse des performances ont établi l'efficacité et la scalabilité (vérifiée jusqu'à 448 processeurs) de l'algorithme parallèle. Le simulateur parallèle effectue des calculs plus précis que la version séquentielle. L'efficacité parallèle atteint 77 % avec 448 processeurs pour une simulation complète. Une étude des accès mémoire et de l'utilisation des caches a conduit

à une implémentation atteignant 60 % de la puissance de crête dans le noyau de calcul sur l'IBM SP3 et l'Origin 3800. Une simulation complète de 1.45 PFLOP a été réalisée en seulement deux heures sur 256 processeurs (IBM SP3). Une analyse complète des simulations numériques nous a permis de calibrer le modèle afin d'obtenir un comportement qualitatif réaliste. Puis, une validation exhaustive du modèle a été réalisée avec P. Silan (UPS CNRS 2561, Guyane). Les résultats numériques des simulations montrent l'effet de la sur-dispersion, de la mortalité du parasite et de l'hôte sur la distribution du parasite, et la régulation de l'hôte (intervenant par cycles).

Par la suite, un modèle individu-centré a été développé. Nous avons utilisé un algorithme de Monte Carlo pour ce simulateur stochastique pour lequel nous avons décrit trois niveaux différents de parallélisme. Les résultats et performances de ce programme hybride MPI/OpenMP ont été étudiés sur un cluster de nœuds SMP jusqu'à 256 processeurs. La qualité des résultats des deux simulateurs parallèles a été comparée. L'amélioration du modèle a permis de comprendre plus en détails les processus intervenant dans le système biologique réel. Même si la méthode diffère de celle déterministe, les résultats sont qualitativement similaires pour des jeux de données identiques.

Une coopération faisant intervenir une biologiste (Agnès Calonnec - INRA UMR Santé Végétale 1065 - Villenave d'Ornon) et un doctorant en informatique (Gaël Tessier) a débuté en 2003. A l'aide de méthodes numériques et de techniques parallèles, nous nous sommes intéressés à la modélisation de la propagation de l'*oïdium*, une maladie de la vigne. La prédiction correcte des épidémies de ce type de parasite nécessite un simulateur réaliste, et pourrait avoir un impact industriel. Un modèle architectural des ceps de vigne est utilisé pour deux raisons : l'étude de la croissance des ceps et l'influence de la structure du couvert végétal sur la dispersion de l'*oïdium*. Dans ce modèle, on considère un grand nombre d'éléments infectieux et plusieurs paramètres environnementaux spatialement hétérogènes. En effet, la dispersion de l'*oïdium* est un phénomène multi-échelles se produisant à la fois au sein des ceps et le long et à travers des rangs du vignoble. Une version initiale d'un simulateur parallèle a été développée. Une caractérisation de l'algorithme implémenté est en cours ; nous évaluons en particulier le coût des communications et le déséquilibre de charge. Les premiers résultats indiquent une bonne scalabilité jusqu'à 128 processeurs pour des parcelles de taille réaliste.

Programmation des grappes de SMP

Les travaux décrits ici ont débuté dans le cadre des activités de l'équipe ALiENor sur les modes d'expression et environnements de programmation pour le parallélisme. Dans ce cadre, nous proposons un mode d'expression facilitant l'écriture des applications parallèles, notamment irrégulières, et l'environnement associé (appelé PRFX) permettant d'exploiter les caractéristiques prévisibles de ces applications (caractéristiques calculables à partir de données structurantes) et l'hétérogénéité des machines telles que les grappes de SMP. Le mode d'expression, basé sur le langage C et des appels à la bibliothèque PRFX, permet un découpage de l'application en tâches et ne nécessite ni synchronisation ni communication explicite entre elles. PRFX fournit un espace d'adressage global facilitant la structuration des données. Les caractéristiques des tâches, dont la description des données (paramètres) auxquelles elles accèdent et leur coût, sont décrites par le programmeur. La mise en œuvre repose sur trois éléments essentiels : une iso-mémoire assurant la validité des pointeurs en mémoire distribuée ; un fonctionnement de type inspection/exécution ; une modélisation de l'exécution du programme par un DAG de tâches. Les communications et synchronisations entre les tâches sont capturées dans ce DAG et déduites de l'analyse des accès aux données

réalisée lors de l'inspection statique consistant en une pré-exécution partielle du programme. Ensuite, un ordonnanceur séquentiel applique une heuristique d'ordonnement intégrant les optimisations (par exemple, spécification d'un placement initial) décrites par le programmeur. Enfin, un exécuteur parallèle exécute le DAG de tâches en respectant l'ordonnement des tâches et des communications précédemment obtenu. Il exploite la performance des communications unilatérales pour les communications inter-noeuds et tire profit de la mémoire partagée d'un noeud avec des threads. Cette approche a été validée par l'implémentation des algorithmes de résolution de l'équation de Laplace, de factorisation LU de matrices pleines et de factorisation de matrices creuses par la méthode de Cholesky et par leur expérimentation en vraie grandeur sur des grappes de SMP IBM NH2 et p690+. Les principales perspectives de ce travail sont de permettre la prise en compte de programmes quasi-prévisibles par un résolveur de dépendances à la volée permettant d'affiner les dépendances de données entre tâches qui ont été sur-estimées lors de l'inspection par manque d'information, et de permettre l'équilibrage dynamique de la charge par migration des tâches en iso-mémoire. L'ensemble de ce travail a constitué le travail de thèse de B. Cirou.

Projets et perspectives 2005-2008

Algorithmique et ordonnancement pour les plates-formes à grande échelle Comme nous l'avons vu, le passage de la minimisation du temps de complétion à la maximisation du débit a montré sa pertinence et son efficacité. En effet, il est possible de construire des algorithmes polynomiaux pour résoudre de nombreux problèmes d'ordonnement, pour des plates-formes hétérogènes et sous des modèles de communication plus réalistes que ceux utilisés généralement en ordonnancement. Néanmoins de nombreux problèmes restent ouverts :

- Tout d'abord, il reste à vérifier qu'il est possible d'instancier en pratique le modèle de communication que nous avons proposé. En particulier, il est important de développer des logiciels réalisant la découverte automatique de la topologie des plates-formes et des performances (bande passante, latence, vitesse de calcul) des ressources. Des prototypes existent (ENV ou AINEM), mais ils sont encore beaucoup trop lents pour être utilisés en pratique.
- Ensuite, toutes les solutions que nous avons proposées s'appuient sur une connaissance centralisée de la topologie de la plate-forme pour pouvoir calculer l'ordonnement. Cette solution pose clairement un problème de passage à l'échelle, si on veut considérer des plates-formes très grandes. Il est donc nécessaire de construire des algorithmes d'ordonnement décentralisés (l'utilisation des multi-flots et du "network coding" peut être une solution à ces problèmes).
- De plus, de nombreux paramètres architecturaux n'ont pas été pris en compte dans notre modèle (latence des liens, taille des buffers de communication, taille de la mémoire de noeuds participant...). Ces paramètres sont clairement également très hétérogènes dans le cadre des plates-formes visées et doivent être considérés (l'utilisation des techniques de contrôle peut être une solution).
- Enfin, nous avons vu que la dynamique des capacités des ressources de communication et de calcul, ainsi que de la topologie, est une autre caractéristique essentielle des systèmes distribués à grande échelle. On doit donc construire des algorithmes décentralisés et susceptibles de s'adapter aux variations de performance et de topologie du système. Ce point peut être également abordé en utilisant les techniques de flots et le "network coding", mais il suppose une adaptation des modèles utilisés. En particulier, dans le cas des réseaux pair-à-pair, il est important de considérer à la fois le réseau logique (qui assure la robustesse et la tolérance aux pannes) et le réseau physique sous-jacent.

Algorithmes, applications et solveurs hautes performances Pour la partie algèbre linéaire haute performance, il conviendra tout d'abord de valider le prototype de partitionneur parallèle actuellement en cours de conception et de mise en œuvre (thèse de C. Chevalier). On travaillera aussi sur l'intégration de fonctionnalités de *repartitionnement* permettant de traiter certains problèmes plus dynamiques à frontières variables.

Pour ce qui est des solveurs directs, un premier objectif sera de réaliser une plate-forme générique *d'Out-Of-Core* distribuée; cela permettra de repousser encore les limites de la consommation mémoire. Pour les solveurs hybrides directs/itératifs, nous travaillerons sur la qualité des partitions pour les calculs par blocs en essayant de combiner critères numériques et critères combinatoires. Une importante collaboration avec les collègues de mathématiques appliquées du projet permettra l'intégration complète de nos solveurs dans un logiciel de simulation d'écoulements multifluides multiphasiques qui s'appelle FLUIBOX. Il est prévu aussi que ces travaux soient impliqués dans le projet ITER, dans un contrat cadre de collaboration avec EDF au titre de l'INRIA, et dans le Pôle de Compétitivité Aquitaine – Midi Pyrénées AESE (Aéronautique, Espace et Systèmes Embarqués) récemment labellisé.

Concernant nos travaux algorithmiques et applicatifs (multipôles – thèse de P. Fortin, propagation de l'oïdium dans les parcelles de vigne – thèse de G. Tessier, propagation de fissures dans la silice – thèse de G. Anciaux), l'objectif sera de finaliser les implémentations pour pouvoir valider en qualité de précision et en performance les solutions proposées et cela pour des cas en vraie grandeur.

Pilotage par la visualisation de simulations numériques distribuées Les perspectives de ce thème s'articulent autour de quatre axes : évaluation de la plate-forme, amélioration de la bibliothèque de redistribution, extension de notre modèle aux applications distribuées, étude et proposition d'un modèle pour des interactions provenant de la visualisation.

Le premier axe concerne l'évaluation des capacités de notre plate-forme de pilotage sur des cas et des environnements réalistes (simulations parallèles de grandes tailles, environnement d'exécution sécurisé, ordonnancement, déploiement, etc.). La plate-forme Grid5000 nous permettra de tester nos différentes approches.

Dans le deuxième axe, nous souhaitons améliorer les fonctionnalités, les performances de la bibliothèque de redistribution notamment pour des données issues de maillages irréguliers. Deux des applications du projet, la propagation de fissures et la mécanique des fluides utilisent ce type de discrétisation. Il s'agit de prendre en compte le graphe de partitionnement du maillage donné par un partitionneur, par exemple celui de la bibliothèque SCOTCH, pour améliorer l'algorithme de redistribution.

Dans le cadre des applications distribuées (couplage de codes parallèles), nous devons étendre notre modèle. La difficulté liée à notre approche provient premièrement du fait que l'on n'a plus un seul flot d'exécution mais autant de flots que de codes; il faut donc pouvoir décrire de façon cohérente et coordonnée les différents flots d'exécution (parallèles) constituant l'application. Deuxièmement, les données dans une application couplée peuvent être de différents types : local à un ou plusieurs codes, distribuées ou dupliquées dans les codes. La caractérisation de ces données est liée au type de couplage dans le modèle que l'on simule; par exemple dans un couplage frontière, les données sont propres à chacun des codes tandis que dans un couplage multi-échelle et volumique certaines sont dupliquées et d'autres sont locales à un code.

Le dernier axe concerne le pilotage proprement dit de la simulation à travers la visualisation. Ici, les actions seront effectuées depuis la visualisation comme par exemple le raffinement de la discrétisation, les modifications de paramètres, ...

5.3.5 Sod

Introduction

Avec la multiplication des équipements connectés au réseau une demande fondamentale de l'utilisateur au sens large, i.e l'utilisateur final, le programmeur ou l'administrateur, est de pouvoir utiliser ses applications de façon simple et efficace sur la meilleure configuration disponible avec un maximum de sécurité. Selon son niveau de compétence et l'importance que revêt son application, il souhaitera aussi contrôler plus ou moins finement le support sous-jacent – pour réaliser par exemple un réglage de performance ou collecter des traces d'exécution pour effectuer la mise au point –. Il faut donc disposer de couches logicielles qui offrent un niveau élevé d'abstraction tout en permettant un niveau de contrôle et de mesure relativement fin. Ceci est de plus en plus pertinent, puisqu'avec des technologies comme Bluetooth ou Wi-Fi, le réseau devient présent partout et acquière une dimension dynamique difficile à maîtriser. La multiplication des processeurs connectés, la mobilité associée et la multiplication des technologies de communication sous-jacentes font exploser les possibilités mais aussi la complexité des systèmes résultants, ce qui a pour effet d'accroître les besoins et donc les attentes des utilisateurs, des développeurs et des administrateurs.

Mots-Clés Parallélisme, Mémoire distribuée, Environnements, Mise au point, Langages, Sémantique, Modèles, Validation, Systèmes distribués, Technologie objet, Java, CORBA, Mobilité, Non filaire, Wi-Fi, Téléphonie mobile, Nomadique, Bluetooth, Cartes à puce, Java Card.

Description des activités et principaux résultats

Phase initiale du projet La plate-forme Jem, *Experimentation environMent for Java* a été le développement initiateur du projet *Systèmes et Objets Distribués*. Cette action a pour une grande partie été menée dans le cadre de la thèse de Asier Ugarte, qui a soutenu en janvier 2001. Elle a consisté à développer une plate-forme distribuée basée sur Java, RMI et CORBA, i.e. sur des technologies de type objets distribués. Jem fournit une plate-forme offrant un accès homogène aux ressources hétérogènes et distribuées d'un réseau et permet bien entendu d'exploiter ces ressources. Il est simple de contribuer au système pour y intégrer de nouvelles ressources, donc de l'étendre.

Structure du projet et principales actions de recherche Le développement de Jem nous a permis d'acquérir une forte compétence dans le domaine des technologies des objets mobiles et distribués, des langages de type Java et des modèles associés. Nous avons aussi largement été confrontés aux problématiques associées aux supports d'exécution. Ceci nous a conduit à structurer nos activités autour d'un certain nombre d'actions de recherche qui toutes concourent à notre projet global qui vise à faciliter l'exploitation – au sens large du terme – des ressources connectées au réseau. Ce sont pour l'essentiel :

1. **le pilotage de codes numériques** Dans le cadre d'une approche de type grille, les plates-formes Jem et DHPCCM (architecture Java/CORBA conçue avec l'IFP pour le pilotage de codes numériques) que nous avons développées peuvent être utilisées pour deux fonctions principales : encapsuler des codes de sorte à ce qu'ils offrent une interface homogène; réaliser la couche de pilotage d'une grille de calcul. Ce sont ces aspects que nous avons mis en avant dans le cadre de notre participation à l'ACI GRID 2002 (projet EPSN) et au projet ScAlApplix commun avec l'UR INRIA Futurs.

2. **l'aide à la distribution automatique de codes Java multithreadés** Ces travaux font actuellement l'objet de la thèse de Pascal Grange. Plus précisément ses travaux portent sur la distribution automatique d'une application objets en réalisant une analyse de code. Une des difficulté et originalité majeure de ses travaux est qu'ils se placent dans un cadre *multithread*.
3. **le développement, la modélisation et la validation de supports d'exécution distribués** Les plates-formes réalisées (Jem, DHPCCM et Mandala) présentent la caractéristique d'offrir un accès homogène aux ressources hétérogènes disséminées sur un réseau ; ces travaux, et ceux sur Mandala en particulier, se sont déroulés dans le cadre de la thèse de Pierre Vignéras soutenue en décembre 2004.
4. **la validation formelle et pratique de systèmes SERMA** Technologies et le LaBRI collaborent dans le domaine des cartes à puce. Cette relation a été amorcée dans le cadre d'un projet intitulé *Sécurité Java Card*, portant sur l'analyse de la sécurité des cartes à puce Java Card et la validation des applications embarquées sur ces cartes. Ce projet, labélisé Société de l'Information par le Secrétariat d'Etat à l'Industrie, a été en partie mené dans le cadre de la thèse CIFRE de Damien Sauveron soutenue en décembre 2004, et avec le concours d'Iban Hatchondo, ingénieur recruté sur ce projet. Cette action a été financée par le ministère de l'Industrie.

Ces travaux sont menés dans le cadre des structures nationales indiquées ci-dessous, certaines n'existant d'ailleurs plus au moment de la rédaction de ce document. Tout d'abord les activités de SOD se déroulent pour partie au sein de l'action transversale PRO, *Parallélisme, Répartition et Objets*, commune aux GDR ALP (Algorithmique, Langages et Programmation) et ARP (Architecture, Réseaux et systèmes, Parallélisme). Ils s'intègrent aussi dans l'action transversale *Grappes* du GDR ARP qui regroupe des équipes " dont les travaux coordonnés permettront de développer l'utilisation des stations de travail interconnectées par des réseaux à hautes performances (grappes) ". Nos activités sont aussi référencées par le *Java Special Interest Group* d'*EuroTools*, pendant européen du *Java Grande Forum* américain. Nous participons au projet EPSN de l'ACI GRID 2002 et à l'Action de Recherche ScAIApplications de l'UR INRIA Futurs. Enfin, nous participons au RTP CNRS *Systèmes Embarqués* et à son Action Spécifique *Sécurité Logicielle*.

Projets et perspectives 2005-2008

La ligne directrice de nos travaux demeure centrée sur la conception et le développement d'outils et d'environnements d'exécution permettant de répondre aux problématiques actuelles nées de la multiplication des équipements connectés. Plus précisément, nous développons et nous faisons évoluer cette thématique dans la direction de la sécurisation des calculs sur la grille et de la gestion avancée de la mobilité et de la nomadique, ce dernier point rendant d'ailleurs la sécurité encore plus critique. En particulier, nous exploitons la compétence acquise dans le domaine connecté pour étudier plus avant ces systèmes dans lesquels les équipements présents et les moyens de communication disponibles peuvent évoluer à tout moment (un PDA qui s'éteint ou s'allume, un terminal Wi-Fi dont la communication est temporairement coupée par un tunnel, etc.). Nous utilisons aussi la compétence relative à la sécurité que nous avons acquise dans nos travaux sur les cartes à puce.

Pour cela nous déployons deux plates-formes :

- la première est à base de cartes à puce Java. Même si on ne peut en attendre aucune efficacité, elle nous permet de valider les aspects fonctionnels et sécuritaires. En effet on peut charger dans une carte à puce Java des données et du code dont on peut être certain qu'ils resteront totalement confidentiels. Les travaux sur cette thématique sont

en cours dans le cadre de la thèse d'Achraf Karray, en co-tutelle avec l'Université de Sfax en Tunisie. Cette plate-forme est déjà déployée. Elle a donné lieu à deux publications. Au delà des tests relatifs au calcul sécurisé sur la grille, cette plate-forme nous offre un environnement de simulation de la mobilité. En effet, elle permet aussi de simuler des terminaux mobiles par arrachement et re-insertion de cartes. Elle est exploitée dans le projet *Mise en œuvre et sécurisation d'un environnement distribué pour la modélisation et la simulation de structures en biologie cellulaire* financé par l'Université Bordeaux 1 dans le cadre du BQR (Bonus Qualité Recherche). Ce projet est réalisé par quatre équipes du LaBRI, dont SOD. SOD intervient ici sur les aspects sécurité à base de cartes à puces. Ces travaux sont supportés en particulier par GemPlus, IBM, et SunMicrosystems.

- la seconde est à base de vrais terminaux mobiles communicants (téléphones, PDAs et PCs) équipés de connexions Wi-Fi et/ou Bluetooth. Nos travaux sont soutenus par des industriels, par la région Aquitaine et l'Université Bordeaux 1. Cette plate-forme est en particulier exploitée dans un projet transversal au LaBRI *Plate-forme pour le développement d'applications multimédia fiables sur terminaux mobiles communicants*. Ce projet est porté par quatre équipes du LaBRI, dont SOD, et fait par ailleurs l'objet d'une demande de soutien FEDER (Europe).

Enfin, ces activités trouvent aussi un point d'ancrage avec un autre projet auquel nous participons. Il s'agit d'un projet transversal au Laboratoire : le projet Visidia, *Visualisation de structures complexes évolutives et d'algorithmes distribués sur ces structures*. Ce projet est porté par les équipes CombAlgo, L3A et SOD. Ces travaux ont pour objectif de traiter la chaîne complète d'écriture d'algorithmes distribués. On aborde la conception, la visualisation de l'exécution et l'exécution effective. C'est sur ce dernier point que nous intervenons, en offrant nos compétences en termes de supports pour effectivement exécuter ces algorithmes. De plus, les environnements de type PDA, mobiles, et plus généralement non filaires que nous étudions actuellement lèvent de nouvelles problématiques. En effet, le fait que des équipements puissent apparaître ou disparaître du réseau à tout moment pose le problème du dynamisme du graphe, alors que jusqu'à présent on raisonnait à graphe constant, i.e. statique. Des travaux sont en cours sur cet aspect dans le cadre de la thèse d'Arnaud Casteigts qui a débuté en septembre 2004. Il se font en étroite collaboration avec Mohamed Mosbah, responsable du projet Visidia décrit ci-dessus.

Dans une phase ultérieure nous pourrions envisager un passage à l'échelle et une multiplication du facteur d'hétérogénéité de nos plates-formes, aussi bien en terme de sécurité (passage à de gros processeurs) que de mobilité. Les aspects passage à l'échelle et application à la sécurisation d'une grille plus classique font l'objet de la thèse d'Eve Atallah, en co-tutelle avec le laboratoire LMSI de l'Université de Limoges.

L'expérience que nous gagnons ainsi nous permettra de proposer des évolutions en termes de langages, de méthodologies de programmation et d'outils qui pourraient faciliter le développement et la validation d'applications embarquées et/ou mobiles basées sur le paradigme objet. En effet La plupart des problématiques effectives rencontrées sont liées ou s'expriment relativement aux spécifications du langage. Souvent, c'est parce que le langage ou le code intermédiaire sous-jacent peut trop ou trop peu que la réalisation effective ou la validation formelle de certains mécanismes ou applications nous échappe.

Tout ceci participe à notre objectif qui est de contribuer à la conception de supports d'exécution, de bibliothèques et de formalismes associés qui permettront d'intégrer au mieux le triptyque modélisation / implémentation / validation dans le monde des logiciels de manipulation des équipements connectés et de la nomadique.

5.4 Publications de l'équipe

5.4.1 Publications d'audience internationale

Revue avec Comité de lecture

- [746] AHMED (T), ASGARI (A. H), MEHAOUA (A), BORCOCI (E), BERTI-ÉQUILLE (L) et GEORGIOS (K). – End-to-end quality of service provisioning through an integrated management system for multimedia content delivery. *Computer Communication Journal SI on Emerging Middleware for Next Generation Networks*, winter 2005.
- [747] AHMED (T), MEHAOUA (A), BOUTABA (R) et IRAQI (Y). – Cognitive video streaming over next generation networks : A cross-layer approach. *IEEE Journal on Selected Areas in Communications (J-SAC), Special issue on Intelligent Services and Applications in Next Generation Networks*, vol. 23, Feb. 2005.
- [748] AUMAGE (O), BOUGÉ (L), DENIS (A), EYRAUD (L), MÉHAUT (J.-F), MERCIER (G), NAMYST (R) et PRYLLI (L). – A portable and efficient communication library for high-performance cluster computing (extended version). *Cluster Computing*, vol. 5, n1, janvier 2002, pp. 43–54. – Special Issue : Selected Papers from the IEEE Cluster 2000 Conference.
- [749] AUMAGE (O), BOUGÉ (L), MÉHAUT (J.-F) et NAMYST (R). – Madeleine II : A portable and efficient communication library for high-performance cluster computing. *Parallel Computing*, vol. 28, n4, avril 2002, pp. 607–626.
- [750] BANINO (C), BEAUMONT (O), CARTER (L), FERRANTE (J), LEGRAND (A) et ROBERT (Y). – Scheduling strategies for master-slave tasking on heterogeneous processor platforms. *IEEE Trans. Parallel Distributed Systems*, vol. 15, n4, 2004, pp. 319–330.
- [751] BEAUMONT (O), BOUDET (V), RASTELLO (F) et ROBERT (Y). – Partitioning a square into rectangles : NP-completeness and approximation algorithms. *Algorithmica*, vol. 34, 2002, pp. 217–239.
- [752] BEAUMONT (O), CASANOVA (H), LEGRAND (A), ROBERT (Y) et YANG (Y). – Scheduling divisible loads on star and tree networks : results and open problems. *IEEE Trans. Parallel Distributed Systems*, vol. 16, n3, 2005, pp. 207–218.
- [753] BEAUMONT (O), LEGRAND (A), MARCHAL (L) et ROBERT (Y). – Scheduling strategies for mixed data and task parallelism on heterogeneous clusters. *Parallel Processing Letters*, vol. 13, n2, 2003.
- [754] BEAUMONT (O), LEGRAND (A), MARCHAL (L) et ROBERT (Y). – Pipelining broadcasts on heterogeneous platforms. *IEEE Trans. Parallel Distributed Systems*, vol. 16, n4, 2005, pp. 300–313.
- [755] BEAUMONT (O), LEGRAND (A), MARCHAL (L) et ROBERT (Y). – Steady-state scheduling on heterogeneous clusters. *Int. J. of Foundations of Computer Science*, 2005. – to appear.
- [756] BEAUMONT (O), LEGRAND (A), RASTELLO (F) et ROBERT (Y). – Dense linear algebra kernels on heterogeneous platforms : Redistribution issues. *Parallel Computing*, vol. 28, 2002, pp. 155–185.
- [757] BEAUMONT (O), LEGRAND (A) et ROBERT (Y). – Static scheduling strategies for heterogeneous systems. *Computing and Informatics*, vol. 21, 2002, pp. 413–430.

- [758] BEAUMONT (O), LEGRAND (A) et ROBERT (Y). – The master-slave paradigm with heterogeneous processors. *IEEE Trans. Parallel Distributed Systems*, vol. 14, n9, 2003, pp. 897–908.
- [759] BEAUMONT (O), LEGRAND (A) et ROBERT (Y). – Scheduling divisible workloads on heterogeneous platforms. *Parallel Computing*, vol. 29, 2003, pp. 1121–1152.
- [760] CONSEL (C), LAWALL (J) et LE MEUR (A.-F). – A tour of Tempo : A program specializer for the C language. *Science of Computer Programming*, 2004.
- [761] COULAUD (O), DUSSÈRE (M), HÉNON (P), LEFEBVRE (E) et ROMAN (J). – Optimization of a kinetic laser-plasma interaction code for large parallel systems. *Parallel Computing*, vol. 29, n9, 2003, pp. 1175–1190.
- [762] HÉNON (P), RAMET (P) et ROMAN (J). – PaStiX : A High-Performance Parallel Direct Solver for Sparse Symmetric Definite Systems. *Parallel Computing*, vol. 28, n2, janvier 2002, pp. 301–321.
- [763] HUOT (F), GHIZZO (A), BERTRAND (P), SONNENDRUCKER (E) et COULAUD (O). – Instability of the time splitting scheme for the one-dimensional and relativistic vlasov-maxwell system. *Journal of Computational Physics*, vol. 185, n2, 2003, pp. 512–531.
- [764] LANGLAIS (M), LATU (G), ROMAN (J) et SILAN (P). – Performance analysis and qualitative results of an efficient parallel stochastic simulator for a marine host-parasite system. *Concurrency & Computation : Practice and Experience*, vol. 15, n11-12, 2003, pp. 1133–1150.
- [765] LE MEUR (A.-F), LAWALL (J) et CONSEL (C). – Specialization scenarios : A pragmatic approach to declaring program specialization. *Higher-Order and Symbolic Computation*, vol. 17, n1, 2004, pp. 47–92.
- [766] SCHULTZ (U), LAWALL (J) et CONSEL (C). – Automatic program specialization for Java. *ACM Transactions on Programming Languages and Systems*, vol. 25, n4, 2003, pp. 452–499.

Colloques avec Comité de programme et Actes

- [767] ÅBERG (R. A), LAWALL (J. L), SÜDHOLT (M), MULLER (G) et LE MEUR (A.-F). – On the automatic evolution of an OS kernel using temporal logic and AOP. *In : Proceedings of the 18th IEEE International Conference on Automated Software Engineering (ASE 2003)*. pp. 196–204. – Montreal, Canada, octobre 2003.
- [768] AHMED (T), ASGARI (A), BORCOCI (E), ECCLES (S), KORMENTZAS (G), KOURTIS (A), MEHAOUA (A) et XILOURIS (G). – Enthroned core networking elements for end-to-end QoS provision over heterogeneous settings. *In : IST Mobile & Wireless Communications Summit*. – Jun 2005.
- [769] ATALLAH (E), CHAUMETTE (S), DARRIGADE (F), KARRAY (A) et SAUVERON (D). – A Grid of Java Cards to Deal with Security Demanding Applications Domains. *In : Proceedings of the 6th International Conference of e-Smart05*. – Nice, France, September 21-23 2005.
- [770] AUMAGE (O). – Heterogeneous multi-cluster networking with the Madeleine III communication library. *In : Proc. 16th Intl. Parallel and Distributed Processing Symposium, 11th Heterogeneous Computing Workshop (HCW 2002)*. Held in conjunction with IPDPS 2002. – Fort Lauderdale, avril 2002. 12 pages. Extended proceedings in electronic form only.

- [771] AUMAGE (O), BAHY (J. M), CONTASSOT-VIVIER (S), COUTURIER (R), DENIS (A), NAMYST (R), PAPAURÉ (G), PEREZ (C) et SAUGET (M). – Alta : Asynchronous loss tolerant algorithms for grid computing. *In : 3rd International workshop on Parallel Matrix Algorithms and Applications (PMAA'04)*. CIRM. – Marseille, France, octobre 2004.
- [772] AUMAGE (O), HOFMAN (R) et BAL (H). – Netibis : An efficient and dynamic communication system for heterogeneous grids. *In : Proc. of the Cluster Computing and Grid 2005 Conference (CCGrid 2005)*. ACM/IEEE. – Cardiff, UK, mai 2005. 8 pages.
- [773] AUMAGE (O) et MERCIER (G). – MPICH/MadIII : a Cluster of Clusters Enabled MPI Implementation. *In : Proc. 3rd IEEE/ACM International Symposium on Cluster Computing and the Grid (CCGrid 2003)*. IEEE, pp. 26–36. – Tokyo, mai 2003.
- [774] BANINO (C), BEAUMONT (O), LEGRAND (A) et ROBERT (Y). – Scheduling strategies for master-slave tasking on heterogeneous processor grids. *In : PARA'02 : International Conference on Applied Parallel Computing*. pp. 423–432. – Springer Verlag, 2002.
- [775] BEAUMONT (O), BOUDET (V) et ROBERT (Y). – The iso-level scheduling heuristic for heterogeneous processors. *In : PDP'2002, 10th Euromicro Workshop on Parallel, Distributed and Network-based Processing*. – IEEE Computer Society Press, 2002.
- [776] BEAUMONT (O), BOUDET (V) et ROBERT (Y). – A realistic model and an efficient heuristic for scheduling with heterogeneous processors. *In : HCW'2002, the 11th Heterogeneous Computing Workshop*. – IEEE Computer Society Press, 2002.
- [777] BEAUMONT (O), CARTER (L), FERRANTE (J), LEGRAND (A) et ROBERT (Y). – Bandwidth-centric allocation of independent tasks on heterogeneous platforms. *In : International Parallel and Distributed Processing Symposium IPDPS'2002*. – IEEE Computer Society Press, 2002.
- [778] BEAUMONT (O), DAOUDI (E), MAILLARD (N), MANNEBACK (P) et ROCH (J.-L). – Tradeoff to minimize extra-computations and stopping criterion tests for parallel iterative schemes. *In : PMAA'04 Parallel Matrix Algorithms and Applications*. – CIRM, Marseille, 2004.
- [779] BEAUMONT (O), LEGRAND (A), MARCHAL (L) et ROBERT (Y). – Assessing the impact and limits of steady-state scheduling for mixed task and data parallelism on heterogeneous platforms. *In : HeteroPar'2004 : International Conference on Heterogeneous Computing, jointly published with ISPDC'2004 : International Symposium on Parallel and Distributed Computing*. – IEEE Computer Society Press, 2004.
- [780] BEAUMONT (O), LEGRAND (A), MARCHAL (L) et ROBERT (Y). – Complexity results and heuristics for pipelined multicast operations on heterogeneous platforms. *In : 2004 International Conference on Parallel Processing (ICPP'2004)*. pp. 267–274. – IEEE Computer Society Press, 2004.
- [781] BEAUMONT (O), LEGRAND (A), MARCHAL (L) et ROBERT (Y). – Pipelining broadcasts on heterogeneous platforms. *In : IPDPS'2004*. – IEEE Computer Society Press, 2004.
- [782] BEAUMONT (O), LEGRAND (A), MARCHAL (L) et ROBERT (Y). – Steady-state scheduling on heterogeneous clusters : why and how ? *In : 6th Workshop on Advances in Parallel and Distributed Computational Models APDCM 2004*. – IEEE Computer Society Press, 2004.

- [783] BEAUMONT (O), LEGRAND (A), MARCHAL (L) et ROBERT (Y). – Independent and divisible task scheduling on heterogeneous star-shaped platforms with limited memory. *In : 13th Euromicro Conference on Parallel, Distributed and Network-based Processing*. – IEEE Computer Society Press, 2005.
- [784] BEAUMONT (O), LEGRAND (A) et ROBERT (Y). – A polynomial time algorithm for allocating independent tasks on heterogeneous fork-graphs. *In : ISCIS'02, 17th International Symposium on Computer and Information Sciences*. – CRC Press, 2002.
- [785] BEAUMONT (O), LEGRAND (A) et ROBERT (Y). – Static scheduling strategies for dense linear algebra kernels on heterogeneous clusters. *In : Parallel Matrix Algorithms and Applications*. – Université de Neuchâtel, 2002.
- [786] BEAUMONT (O), LEGRAND (A) et ROBERT (Y). – Static scheduling strategies for heterogeneous systems. *In : ISCIS XVII, Seventeenth International Symposium On Computer and Information Sciences*. – CRC Press, 2002.
- [787] BEAUMONT (O), LEGRAND (A) et ROBERT (Y). – Optimal algorithms for scheduling divisible workloads on heterogeneous systems. *In : HCW'2003, the 12th Heterogeneous Computing Workshop*. – IEEE Computer Society Press, 2003.
- [788] BEAUMONT (O), LEGRAND (A) et ROBERT (Y). – Scheduling strategies for mixed data and task parallelism on heterogeneous clusters and grids. *In : PDP'2003, 11th Euromicro Workshop on Parallel, Distributed and Network-based Processing*. pp. 209–216. – IEEE Computer Society Press, 2003.
- [789] BEAUMONT (O), MARCHAL (L) et ROBERT (Y). – Broadcast trees for heterogeneous platforms. *In : International Parallel and Distributed Processing Symposium IPDPS'2005*. – IEEE Computer Society Press, 2005.
- [790] BEAUMONT (O), RAMET (P) et ROMAN (J). – Asymptotically optimal algorithm for laplace task graphs on heterogeneous platforms. *In : Proceedings of Fifth International Conference on Parallel Processing and Applied Mathematics, Workshop HeteroPar, Czestochowa, Pologne, LNCS 3019*, pp. 880–887. – septembre 2003.
- [791] BENMAMMAR (B) et KRIEF (F). – Advance resource reservation in a wman environment based on the QoS NSLP signaling application and the CTP protocol. *In : IFIP Open Conference on Metropolitan Area Networks Architecture, protocols, control, and management*, pp. 253–272. – HCMV, Vietnam, 11-13 April 2005.
- [792] BENMAMMAR (B) et KRIEF (F). – An advanced resource reservation protocol in wireless networks based on user mobility profile. *In : The Fifth IEEE Workshop on Applications and Services in Wireless Networks (ASWN 2005)*. – Paris, France, 29 June - 1 july 2005.
- [793] BENMAMMAR (B) et KRIEF (F). – MQoS NSLP : A mobility profile management based approach for advance resource reservation in a mobile environment. *In : The 7th IEEE/IFIP International Conference on Mobile and Wireless Communications Networks (MWCN 2005)*. – Marrakech, Morocco, 19-21 September 2005.
- [794] BHATIA (S) et CONSEL (C). – Implementing high-performance in-kernel network services with WYKIWYG. *In : ECOOP Workshop on Programming Languages and Operating Systems*. – Oslo, Norway, juin 2004.
- [795] BHATIA (S), CONSEL (C), LE MEUR (A.-F) et PU (C). – Automatic specialization of protocol stacks in OS kernels. *In : Proceedings of the 29th Annual IEEE Conference on Local Computer Networks*. – Tampa, Florida, novembre 2004. Awarded best paper.

- [796] BHATIA (S), CONSEL (C) et PU (C). – Remote customization of systems code for embedded devices. *In : Proceedings of the 4th ACM International Conference on Embedded Software.* – Pisa, Italy, septembre 2004.
- [797] BURGY (L), CAILLOT (L), CONSEL (C), LATRY (F) et RÉVEILLÈRE (L). – A comparative study of SIP programming interfaces. *In : Proceedings of the ninth International Conference on Intelligence in service delivery Networks (ICIN 2004).* – Bordeaux, France, octobre 2004.
- [798] CALONNEC (A), LATU (G), NAULIN (J.-M), ROMAN (J) et TESSIER (G). – Parallel simulation of the propagation of powdery mildew in a vineyard. *In : Proceedings of Euro-Par'05, Lisbonne, Portugal.* – Springer Verlag, août 2005.
- [799] CHAUMETTE (S), , GRANGE (P) et VIGNERAS (P). – JToe : a Java API for object exchange. *In : Parallel Computing 2003.* – Dresden, Germany, September 2-5, 2003.
- [800] CHAUMETTE (S) et GRANGE (P). – Parallelizing multithreaded Java programs : a criterion and its pi-calculus foundation. *In : FMPPTA '2002, 7th International Workshop on Formal Methods for Parallel Programming : Theory and Applications.* – IEEE, 2002. Held in conjunction with the 16th International Parallel & Distributed Processing Symposium (IPDPS'2002), Marriot Marina, Fort Lauderdale, FL., USA. April 15-19, 2002.
- [801] CHAUMETTE (S), GRANGE (P), KARRAY (A), SAUVERON (D) et VIGNÉRAS (P). – Secure distributed computing on a Java Card Grid. *In : Workshop on Java for Parallel and Distributed Computing, held in conjunction with IPDPS 2005.* – April 2005. Denver, Co., USA, April 2005.
- [802] CHAUMETTE (S), GRANGE (P), MÉTROT (B) et VIGNÉRAS (P). – JToe : a Java API for object exchange. *In : PARCO,* pp. 135–142. – 2003.
- [803] CHAUMETTE (S), GRANGE (P), SAUVERON (D) et VIGNÉRAS (P). – Computing with Java Cards. *In : CCCT'2003 and 9th ISAS'03.* – August 2003. Orlando, FL., USA.
- [804] CHAUMETTE (S), HATCHONDO (I) et SAUVERON (D). – JCAT : an environment for attack and test on Java Card. *In : CCCT'2003 and 9th ISAS'03,* pp. 270–275. – August 2003. Orlando, FL., USA.
- [805] CHAUMETTE (S) et SAUVERON (D). – An efficient and simple way to test the security of java cards. *In : WOSIS 2005, 3rd International Workshop on Security in Information Systems.* – April 2005. Miami, FL., USA, April 2005.
- [806] CHAUMETTE (S) et VIGNERAS (P). – Extensible and Customizable Just-In-Time Security (JITS) Management of Client-Server Communication in Java. *In : Parallel Computing 2003.* – Dresden, Germany, September 2-5, 2003.
- [807] CHAUMETTE (S) et VIGNERAS (P). – A framework for seamlessly making object oriented applications distributed. *In : Parallel Computing 2003.* – Dresden, Germany, September 2-5, 2003.
- [808] CHAUMETTE (S) et VIGNÉRAS (P). – Behavior model of mobile agent systems. *In : FCS'05 - The 2005 International Conference on Foundations of Computer Science.* – Las Vegas, USA, june, 27–30 2005.
- [809] CIROU (B), COUNILH (M. C) et ROMAN (J). – PRFX : a runtime library for high performance programming on clusters of SMP nodes. *In : PARALLEL COMPUTING : Software Technology, Algorithms, Architectures & Applications.* pp. 189–200. – Elsevier, septembre 2003.

- [810] CIROU (B), COUNILH (M. C) et ROMAN (J). – Programming irregular scientific algorithms with static properties on clusters of SMP nodes. *In : HPSEC Proceedings.* – IEEE Computer Society Press, juin 2005.
- [811] CONSEL (C). – Generative programming from a domain-specific language viewpoint. *In : Unconventional Programming Paradigms.* – Mont Saint Michel, France, septembre 2004.
- [812] CONSEL (C), HAMDI (H), RÉVEILLÈRE (L), SINGARAVELU (L), YU (H) et PU (C). – Spidle : A DSL approach to specifying streaming application. *In : Second International Conference on Generative Programming and Component Engineering.* – Erfurt, Germany, septembre 2003.
- [813] CONSEL (C) et RÉVEILLÈRE (L). – A programmable client-server model : Robust extensibility via DSLs. *In : Proceedings of the 18th IEEE International Conference on Automated Software Engineering (ASE 2003).* pp. 70–79. – Montréal, Canada, novembre 2003.
- [814] COULAUD (O), DUSSÈRE (M), HÉNON (P) et ROMAN (J). – Optimisation of a kinetic laser-plasma interaction code for massively parallel systems. *In : Proceedings of PMAA'2002, Neuchâtel, Suisse.* – novembre 2002.
- [815] COULAUD (O), DUSSÈRE (M) et ESNARD (A). – A time-coherent model for the steering of parallel simulations. *In : Euro-Par 2004 Parallel Processing.* pp. 90–97. – Springer Verlag, 2004.
- [816] COULAUD (O), DUSSÈRE (M) et ESNARD (A). – Toward a computational steering environment based on corba. *In : Parallel Computing : Software Technology, Algorithms, Architectures & Applications,* éd. par Joubert (G), Nagel (W), Peters (F) et Walter (W). pp. 151–158. – Elsevier, 2004.
- [817] DANJEAN (V) et NAMYST (R). – Controlling Kernel Scheduling from User Space : an Approach to Enhancing Applications' Reactivity to I/O Events. *In : Proceedings of the 2003 International Conference on High Performance Computing (HiPC '03).* – Hyderabad, India, décembre 2003.
- [818] DANJEAN (V), NAMYST (R) et WACRENIER (P.-A). – An efficient multi-level trace toolkit for multi-threaded applications. *In : EuroPar.* – Lisbonne, septembre 2005. To appear.
- [819] DEHNI (L), BENNANI (Y) et KRIEF (F). – Low energy adaptative connectionist clustering for wireless sensor networks. *In : The 2nd IEEE/IFIP International Workshop on Mobility Aware Technologies and Applications (MATA 2005).* – Montréal, Canada, 17-19 October 2005.
- [820] DEHNI (L), KRIEF (F) et BENNANI (Y). – Power control and clustering in wireless sensor networks. *In : Fourth Annual Mediterranean Ad Hoc Networking Workshop (MedHocNet'2005).* – Porquerolles Island, France, 21-24 June 2005.
- [821] DENIS (A), AUMAGE (O), HOFMAN (R), VERSTOEP (K), KIELMANN (T) et BAL (H). – Wide-area communication for grids : An integrated solution to connectivity, performance and security problems. *In : Proc. of the Thirteenth IEEE International Symposium on High-Performance Distributed Computing (HPDC'13).* IEEE. – Honolulu, Hawaii, juin 2004.
- [822] HÉNON (P), NKONGA (B), RAMET (P) et ROMAN (J). – Using of the high performance sparse solver pastix for the complex multiscale 3d simulations performed by the fluidbox fluid mechanics software. *In : Proceedings of PMAA'2004, Marseille, France.* – octobre 2004.

- [823] HÉNON (P), PELLEGRINI (F), RAMET (P) et ROMAN (J). – An efficient hybrid mpi/thread implementation on a network of smp nodes for the parallel sparse direct solver pastix : ordering / scheduling / memory management / out-of-core issues, and application to preconditioning. *In : Sparse Days and Grid Computing, Saint Giron, France.* – juin 2003.
- [824] HÉNON (P), PELLEGRINI (F), RAMET (P) et ROMAN (J). – Towards high performance hybrid direct-iterative solvers for large sparse systems. *In : International SIAM Conference On Preconditioning Techniques For Large Sparse Matrix Problems In Scientific And Industrial Applications, Napa Valley, USA.* – octobre 2003.
- [825] HÉNON (P), PELLEGRINI (F), RAMET (P) et ROMAN (J). – Blocking issues for an efficient parallel block ilu preconditioner. *In : International SIAM Conference On Preconditioning Techniques For Large Sparse Matrix Problems In Scientific And Industrial Applications, Atlanta, USA.* – mai 2005.
- [826] HÉNON (P), PELLEGRINI (F), RAMET (P), ROMAN (J) et SAAD (Y). – Applying parallel direct solver skills to build robust and highly performant preconditioners. *In : PARA'04, Workshop on state-of-the-art in scientific computing, Copenhagen, Denmark.* – juin 2004.
- [827] HÉNON (P), PELLEGRINI (F), RAMET (P), ROMAN (J) et SAAD (Y). – High performance complete and incomplete factorizations for very large sparse systems by using Scotch and PaStiX softwares. *In : Eleventh SIAM Conference on Parallel Processing for Scientific Computing, San Francisco, California, USA.* – février 2004.
- [828] HÉNON (P), RAMET (P) et ROMAN (J). – Parallel factorization of very large sparse SPD systems on a network of SMP nodes. *In : Proceedings of PMAA'2002, Neuchâtel, Suisse.* – novembre 2002.
- [829] HÉNON (P), RAMET (P) et ROMAN (J). – Efficient algorithms for direct resolution of large sparse system on clusters of SMP nodes. *In : SIAM Conference on Applied Linear Algebra, Williamsburg, Virginie, USA.* – juillet 2003.
- [830] HÉNON (P), RAMET (P) et ROMAN (J). – A blockwise algorithm for parallel incomplete cholesky factorization. *In : Proceedings of PMAA'2004, Marseille, France.* – octobre 2004.
- [831] HÉNON (P), RAMET (P) et ROMAN (J). – On using an hybrid MPI-Thread programming for the implementation of a parallel sparse direct solver on a network of SMP nodes. *In : Proceedings of Sixth International Conference on Parallel Processing and Applied Mathematics, Workshop HPC Linear Algebra, Poznan, Pologne.* – septembre 2005.
- [832] HÉNON (P) et SAAD (Y). – A parallel ilu factorization based on a hierarchical interface decomposition algorithm. *In : International SIAM Conference On Preconditioning Techniques For Large Sparse Matrix Problems In Scientific And Industrial Applications, Napa Valley, USA.* – octobre 2003.
- [833] HÉNON (P) et SAAD (Y). – Phidal : A parallel hierarchical interface decomposition algorithm for solving sparse linear systems. *In : Sparse Days and Grid Computing, Saint Giron, France.* – juin 2003.
- [834] JRAD (Z), BENMAMMAR (B), CORREA (J), KRIEF (F) et MBAREK (N). – A userassistant for QoS negotiation in a dynamic environment using agent technology. *In : Second IFIP International Conference on Wireless and Optical Communications Networks WOCN'05.* – Dubai, UAE, 6-8 March 2005.

- [835] KRIEF (F) et BOUTHINON (D). – A learning and intentional local policy decision point for dynamic QoS provisioning. *In : Network Control and Engineering for QoS, Security and Mobility (Net-con'2004)*. pp. 277–288. – Palma de Mallorca, Spain, 2-5 November 2004.
- [836] LE MEUR (A.-F), CONSEL (C) et ESCRIG (B). – An environment for building customizable software components. *In : IFIP/ACM Conference on Component Deployment*, pp. 1–14. – Berlin, Germany, juin 2002.
- [837] LE MEUR (A.-F), LAWALL (J) et CONSEL (C). – Towards bridging the gap between programming languages and partial evaluation. *In : ACM SIGPLAN Workshop on Partial Evaluation and Semantics-Based Program Manipulation*. pp. 9–18. – Portland, OR, USA, janvier 2002.
- [838] MBAREK (N) et KRIEF (F). – SLN NSLP : A service level negotiation signaling protocol in the NSIS environment. *In : 12th IEEE International Conference on Telecommunications (ICT'05)*. – Capetown, South Africa, 3-6 May 2005.
- [839] MOUNGLA (H) et KRIEF (F). – Conflict detection and resolution in qos policy based management. *In : The 16th International Symposium on Personal Indoor and Mobile Radio Communications*. – Berlin, Germany, 11-15 September 2005.
- [840] MOUNGLA (H) et KRIEF (F). – A probabilistic heuristic for conflict detection in diffserv policy based management networks. *In : The 2nd IEEE/IFIP International Workshop on Mobility Aware Technologies and Applications (MATA 2005)*. – Montréal, Canada, 17-19 October 2005.
- [841] PELLEGRINI (F) et GOUDIN (D). – Using the native mesh partitioning capabilities of Scotch 4.0 in a parallel industrial electromagnetics code. *In : Eleventh SIAM Conference on Parallel Processing for Scientific Computing, San Francisco, California, USA*. – février 2004.
- [842] SAMAAAN (N), BENMAMMAR (B), KRIEF (F) et KARMOUCH (A). – An advanced resource reservation protocol in wireless integrated networks based on hosts mobility prediction. *In : IEEE Canada, 18th Annual Canadian Conference on Electrical and Computer Engineering (CCECE05)*. – Saskatchewan, Canada, 1-4 May 2005.

Chapitres d'ouvrages

- [843] BEAUMONT (O), BOUDET (V), LEGRAND (A), RASTELLO (F) et ROBERT (Y). – Static data allocation and load balancing techniques for heterogeneous systems. *In : Annual Review of Scalable Computing*, éd. par Yuen (C), chap. 1, pp. 1–37. – World Scientific, 2002.
- [844] CASTANET (R), CHAUMETTE (S), MACKAYA (M) et THE PLATONIS CONSORTIUM. – Applications and Services in Wireless Networks. chap. PLATONIS : A platform for Validation and Experimentation of Multi-protocols and multi-services, pp. 217–229. – HERMES PENTON SCIENCE, 2003.
- [845] CONSEL (C). – Domain-specific program generation ; international seminar, dagstuhl castle. chap. From A Program Family To A Domain-Specific Language, pp. 19–29. – Springer-Verlag, 2004.
- [846] CONSEL (C) et RÉVEILLÈRE (L). – Domain-specific program generation ; international seminar, dagstuhl castle. chap. A DSL Paradigm for Domains of Services : A Study of Communication Services, pp. 165 – 179. – Springer-Verlag, 2004.

Posters et autres communications

- [847] CHAUMETTE (S) et GRANGE (P). – Optimizing the execution of a distributed object oriented application by combining static and dynamic information. – 2003. Nice, France.
- [848] CHAUMETTE (S) et SAUVERON (D). – The smart cards grid project. – November 2003. Paris, France.
- [849] CHAUMETTE (S) et SAUVERON (D). – Modèles de mémoire en java card. introduction du concept de pré-persistance. *In : MAJECSTIC04 (MANifestation des JEunes Chercheurs STIC)*. – Octobre 2004. Calais, France.
- [850] HÉNON (P), RAMET (P) et ROMAN (J). – A Parallel Direct Solver for Very Large Sparse SPD Systems. *In : Poster session, SuperComputing'2002, Baltimore, USA*. – novembre 2002.
- [851] HÉNON (P), RAMET (P) et ROMAN (J). – A Parallel Direct Solver for Very Large Sparse SPD Systems. *In : Poster session (with Grid-TLSE ACI project), IPDPS'2003, Nice, France*. – avril 2003.
- [852] HÉNON (P), RAMET (P) et ROMAN (J). – Parallel Complete and Incomplete Blockwise Factorisations for Very Large Sparse Systems. *In : Poster session, SuperComputing'2004, Pittsburgh, USA*. – novembre 2004.

Edition d'ouvrages

- [853] CAROMEL (D), CHAUMETTE (S), FOX (G) et GRAHAM (P) (édité par). – *Workshop on Java for Parallel and Distributed Computing*. – Fort Lauderdale, Fl., USA, April 2002. Publié par l'IEEE et tenu conjointement avec le 14th IPPS/SPDP'2002.
- [854] CAROMEL (D), CHAUMETTE (S), FOX (G) et GRAHAM (P) (édité par). – *Workshop on Java for Parallel and Distributed Computing*. – Nice, France, April 2003. Publié par l'IEEE et tenu conjointement avec le 15th IPPS/SPDP'2003.
- [855] CAROMEL (D), CHAUMETTE (S), FOX (G) et GRAHAM (P) (édité par). – *Workshop on Java for Parallel and Distributed Computing*. – Sante Fe, NM., USA, April 2004. Publié par l'IEEE et tenu conjointement avec le 16th IPPS/SPDP'2004.
- [856] CAROMEL (D), CHAUMETTE (S), FOX (G) et GRAHAM (P) (édité par). – *Workshop on Java for Parallel and Distributed Computing*. – Denver, Co., USA, April 2005. Publié par l'IEEE et tenu conjointement avec le 17th IPPS/SPDP'2005.

5.4.2 Publications d'audience nationale**Reuves avec Comité de lecture**

- [857] DANJEAN (V) et WACRENIER (P.-A). – Mécanismes de traces efficaces pour programmes multithreadés. *Technique et science informatiques*, 2005. – To appear.

Colloques avec Comité de programme et Actes

- [858] BEAUMONT (O), BOUDET (V), DESPREZ (F), RAMET (P), ROMAN (J) et TRAVERS (C). – Modélisation de pipelines hétérogènes. *In : GRID'2002, Aussois, France*. – décembre 2002.

- [859] BEAUMONT (O), LEGRAND (A) et ROBERT (Y). – Ordonnancement en régime permanent pour plateformes hétérogènes. *In : GRID'2002, Actes de l'école thématique sur la globalisation des ressources informatiques et des données*. pp. 325–334. – INRIA Lorraine, 2002.
- [860] BENMAMMAR (B). – Les réseaux sans fil et la nouvelle signalisation IP. *In : Réseaux sans fil (Journées scientifiques DNAC'2004)*. – Egypte, 11-18 décembre 2004.
- [861] BENMAMMAR (B) et KRIEF (F). – La mobilité dans la future génération de protocoles de signalisation du monde IP. *In : 6èmes Journées Doctorales Informatique et Réseau*, pp. 7–15. – Lannion, 2-4 novembre 2004.
- [862] BENMAMMAR (B) et KRIEF (F). – Gestion dynamique du handover horizontal et vertical basée sur le profil de mobilité de l'utilisateur. *In : Colloque Gestion de REseaux et de Services*, pp. 159–170. – Luchon, du 28 février au 2 mars 2005.
- [863] CIROU (B). – PRFX : une plate-forme pour les applications parallèles utilisant les grappes de nœuds SMP. *In : Actes de RenPar'15*. pp. 101–108. – INRIA, octobre 2003.
- [864] DANJEAN (V). – Mécanismes de traces efficaces pour programmes multithreadés. *In : Actes des Rencontres francophones du parallélisme (RenPar 15)*. – La Colle sur Loup (France), octobre 2003.
- [865] DEHNI (L), BENNANI (Y) et KRIEF (F). – Une approche neuronale adaptative de routage minimisant la consommation d'énergie dans les réseaux de capteurs. *In : Colloque Gestion de REseaux et de Services*, pp. 115–127. – Luchon, du 28 février au 2 mars 2005.
- [866] DUSSÈRE (M) et ESNARD (A). – Vers le pilotage des simulations numériques sur la grille. *In : 15ème Rencontres francophones du parallélisme*, pp. 196–203. – 2003.
- [867] ESNARD (A). – Modèle pour la redistribution de données complexes. *In : 16èmes Rencontres francophones du parallélisme, RenPar'16*. – 2005.
- [868] GOUDIN (D), HÉNON (P), MANDALLENNA (M), MER (K), PELLEGRINI (F), RAMET (P), ROMAN (J) et PESQUÉ (J.-J). – Outils numériques parallèles pour la résolution de très grands problèmes d'électromagnétisme. *In : Séminaire sur l'Algorithmique Numérique Appliquée aux Problèmes Industriels, Calais, France*. – mai 2003.
- [869] HÉNON (P) et RAMET (P). – Optimisation de l'occupation mémoire pour un solveur parallèle creux direct hautes performances de type supernodal. *In : ACTES RenPar'2002, Hamamet, Tunisie*. – avril 2002.
- [870] KRIEF (F). – La gestion autonome et contextuelle des réseaux IP. *In : 6èmes Journées Doctorales Informatique et Réseau*. – Lannion, 2-4 novembre 2004.
- [871] KRIEF (F). – Les réseaux de capteurs et l'économie d'énergie. *In : Réseaux sans fil (Journées scientifiques DNAC'2004)*. – Egypte, 11-18 décembre 2004.
- [872] KRIEF (F). – L'autonomie dans les réseaux. *In : Réseaux IP : Synthèse et perspective (Journées scientifiques DNAC'2005)*. – Berouth, Liban, 25-27 avril 2005.
- [873] LATU (G). – Simulation stochastique parallèle d'un système hôte-parasite par programmation hybride MPI+OpenMP. *In : RenPar'14, 14ème Rencontres francophones du parallélisme, des architectures et des systèmes (Hammamet)*, pp. 83–90. – 2002.
- [874] MBAREK (N) et KRIEF (F). – Négociation de niveau de service dans un environnement NSIS : SLN NSLP. *In : 18ème Congrès DNAC : Contrôle, Maîtrise et Autonomie des Réseaux*, pp. 140–152. – Paris, du 29 novembre au 1er décembre 2004.

- [875] NGUYEN (T), PUJOLLE (G) et KRIEF (F). – Archises – nouvelle architecture pour la création de services à valeur ajoutée. *In : 6èmes Journées Doctorales Informatique et Réseau*, pp. 93–102. – Lannion, 2-4 novembre 2004.
- [876] PÉRACHE (M). – Nouveaux mécanisme au sein des ordonnanceurs de threads pour une implantation efficace des communication collective sur machines multiprocesseurs. *In : Actes des Rencontres Francophones du Parallélisme (RenPar'16)*, pp. 231–236. – Le Croisic (France), mars 2005.
- [877] TESSIER (G). – Simulation parallèle de la propagation de l'oïdium dans une parcelle de vigne. *In : 16ièmes Rencontres francophones du parallélisme, RenPar'16*. – 2005.
- [878] THIBAUT (S). – Un ordonnanceur flexible pour machines multiprocesseurs hiérarchiques. *In : Actes des Rencontres Francophones du Parallélisme (RenPar'16)*, pp. 77–88. – Le Croisic (France), mars 2005.

Chapitres d'ouvrages

- [879] AUMAGE (O), BOUGÉ (L), EYRAUD (L) et NAMYST (R). – Communications efficaces au sein d'une interconnexion hétérogène de grappes : Exemple de mise en oeuvre dans la bibliothèque Madeleine. *In : Calcul réparti à grande échelle*, éd. par Françoise Baude (I.-U. d. N.-I). – Hermès Science Paris, 2002. ISBN 2-7462-0472-X.
- [880] BEAUMONT (O), BOUDET (V), DUTOT (P), ROBERT (Y) et TRYSTRAM (D). – “fondements théoriques pour la conception d'algorithmes efficaces de gestion de ressources”. *In : Informatique répartie : architecture, parallélisme et système*. – Hermès Publications, 2005.
- [881] BENMAMMAR (B). – Agents et mobiles de 3ème et 4ème générations. *In : Traité IC2, Intelligence dans les réseaux*. – Hermès Science, 2005.
- [882] KRIEF (F). – Application des agents dans la gestion par politique. *In : Traité IC2, Intelligence dans les réseaux*, pp. 103–136. – Hermès Science, 2005.
- [883] MORIN (C), DENIS (A), NAMYST (R), AUMAGE (O) et LOTTIAUX (R). – Des réseaux de calculateurs aux grilles de calcul. *In : Encyclopédie de l'informatique*. – Vuibert, 2005. À paraître.

Edition d'ouvrages

- [884] SILAN (P), CALTRAN (H) et LATU (G). – *Taxonomy, ecology and evolution of meta-zoan parasites*. – Presses Universitaires de Perpignan, 2003. C. Combes & J. Jourdané editors.

5.4.3 Autres publications

Rapports de fin de contrat

- [885] BENMAMMAR (B), JRAD (Z), KRIEF (F) et MBAREK (N). – *Dynamique de l'environnement*. – Rapport technique, Projet RNRT IP-SIG, 2005.
- [886] CHRISTY (S), RAMET (P) et ROMAN (J). – *Développement de la phase d'assemblage de la chaîne EMILIO pour un solveur parallèle 2D*. – Rapport technique, C.E.A. / C.E.S.T.A, 2003. Rapport Final.
- [887] GOUREMAN (A), RAMET (P) et ROMAN (J). – *Développement de la phase d'assemblage de la chaîne EMILIO (distribution du maillage et multi-threading)*. – Rapport technique, C.E.A. / C.E.S.T.A, 2004. Rapport Final.

- [888] HÉNON (P), LECAS (D), RAMET (P) et ROMAN (J). – *Amélioration et Extension du Solveur Direct Parallèle pour Grandes Matrices Creuses du CESTA*. – Rapport technique, C.E.A. / C.E.S.T.A, 2003. Rapport Final.
- [889] HÉNON (P), PELLEGRINI (F), RAMET (P) et ROMAN (J). – *Etude sur l'applicabilité de méthodes itératives nouvelles aux problèmes du CESTA*. – Rapport technique, C.E.A. / C.E.S.T.A, 2004. Rapport final.

Rapports internes et autres publications

- [890] CHAUMETTE (S), GRANGE (P), KARRAY (A), SAUVERON (D) et VIGNÉRAS (P). – *Secure distributed computing on a Java Card Grid. Position paper*. – Rapport technique n1331-04, LaBRI, Université Bordeaux 1, 2004.
- [891] CHAUMETTE (S), GRANGE (P), MÉTROT (B) et VIGNÉRAS (P). – *Implementing a High Performance Object Transfer Mechanism over JikesRVM*. – Rapport de recherche n1324-04, Laboratoire Bordelais de Recherche en Informatique (LaBRI), 2004.
- [892] CHAUMETTE (S), LACOUSTILLE (J.-P) et SAUVERON (D). – *Projet Sécurité Java Card, Rapport d'avancement, Phase 2, Novembre 2001 à Septembre 2002*. – Rapport technique, LaBRI, 2002.
- [893] CHAUMETTE (S) et SAUVERON (D). – *New security problems raised by open multi-application smart cards*. – Rapport technique n1332-04, LaBRI, Université Bordeaux 1, 2004.
- [894] HATCHONDO (I) et SAUVERON (D). – The JCatools website, <http://sourceforge.net/projects/jcatools/>.
- [895] THIBAULT (S). – *Un ordonnanceur flexible pour machines multiprocesseurs hiérarchisées*. – Rapport de stage DEA, École Normale Supérieure de Lyon, juillet 2004.

5.4.4 Production de logiciels

Langage Spidle, Développement d'applications de streaming, architecture client-serveur programmable, personnalisation de serveurs en fonction des utilisateurs.

Langage SPL, Conception et développement d'une machine virtuelle SIP, architecture logicielle pour le déploiement de services de téléphonie.

Tempo, Plateforme de spécialisation automatique pour divers systèmes Unix, développement d'un évaluateur partiel pour les programmes écrits en C.

Madeleine, Bibliothèque de communication multi-grappes qui implémente efficacement un concept de canal de communication pouvant être soit physique (configurations homogènes) soit virtuel (configurations hétérogènes).

Marcel, Bibliothèque de threads utilisateurs s'adaptant à la plateforme visée (architectures mono-processeur, SMP, NUMA) et intégrant un serveur d'événements améliorant le temps de réaction des applications aux entrées/sorties.

LinuxActivations, Extension du noyau Linux permettant de contrôler l'ordonnancement des threads utilisateurs lorsqu'ils effectuent des opérations d'entrées/sorties bloquantes dans

le noyau.

MPICH/MADIII, Implémentation multithreadée et multi-protocole du standard de communication par passage de messages MPI.

PadicoTM, Déploiement d'intergiciels sur les réseaux des grilles de calcul.

PM2, Support exécutif générique de bas niveau qui intègre la gestion des threads, la gestion des communications et une gestion mémoire distribuée élémentaire.

EPSN, Plateforme couplée de pilotage interactif et de visualisation pour des simulations numériques parallèles et distribuées.

PaStiX, Bibliothèque de solveurs parallèles haute performance basés sur des factorisations complètes et incomplètes pour des très grands systèmes linéaires creux issus de problèmes 3D.

PHIDAL, Bibliothèque de solveurs parallèles itératifs basés sur des décompositions hiérarchiques et des factorisations multiniveaux.

Scotch, Bibliothèque pour le partitionnement et le placement de graphes et de maillages irréguliers de grande taille.

Jem, Plateforme de manipulation de ressources connectées à un réseau fixe.

JToe, Bibliothèque Java de transfert efficace d'objets.

Mandala, Plateforme pour le développement d'applications Java distribuées intégrant la notion de conteneur et d'appel de méthodes distantes asynchrones.

5.4.5 Formation par la recherche

Thèses

- [896] CIROU (B). – *Expression d'algorithmes scientifiques et gestion performante de leur parallélisme avec PRFX pour les grappes de SMP*. – PhD thesis, Université Bordeaux 1, juin 2005.
- [897] DANJEAN (V). – *Contribution à l'élaboration d'ordonnanceurs de processus légers performants et portables pour architectures multiprocesseurs*. – 46, allée d'Italie, 69364 Lyon cedex 07, France, Thèse de doctorat, spécialité informatique, École normale supérieure de Lyon, décembre 2004.
- [898] LATU (G). – *Algorithmique parallèle et calcul haute performance dédiés à la simulation d'un système hôte-macroparasite*. – PhD thesis, Université Bordeaux 1, décembre 2002.
- [899] LE MEUR (A.-F). – *Approche Déclarative à la spécialisation de programmes C*. – Thèse de doctorat, Université de Rennes 1, France, décembre 2002.
- [900] LEGRAND (A). – *Algorithmique parallèle hétérogène et techniques d'ordonnancement : approches statiques et dynamiques*. – Thèse de PhD, École Normale Supérieure de Lyon, décembre 2003.

- [901] MERCIER (G). – *Communications á hautes performances portables en environnements hiérarchiques, hétérogènes et dynamiques.* – Labri, Bordeaux, France, Thèse de doctorat, Université de Bordeaux 1, décembre 2004.
- [902] SAUVERON (D). – *Étude et réalisation d'un environnement d'expérimentation et de modélisation pour la technologie Java Card. Application à la sécurité.* – Thèse de PhD, LaBRI, Université Bordeaux 1, 2004.
- [903] VIGNÉRAS (P). – *Vers une programmation locale et distribuée unifiée au travers de l'utilisation de conteneurs actifs et de références asynchrones.* – Thèse de PhD, LaBRI, Université Bordeaux 1, 2004.

Habilitations

- [904] BEAUMONT (O). – *Nouvelles méthodes pour l'ordonnancement sur plates-formes hétérogènes.* – Thèse de PhD, Habilitation à diriger des recherches de l'Université de Bordeaux 1, décembre 2004.

5.5 Collaborations internationales

5.5.1 Invités

Feng, Wu-Chan, *Portland State University - USA*, 7 jours en 2004

Pu, Calton, *Georgia Institute of Technology - USA*, 7 jours en 2004

Russell, Robert D., *New Hampshire University - USA*, 3 jours en 2004

Yousef, Saad, *University of Minnesota - USA*, 15 jours en 2002 et en 2005

5.5.2 Invitations

Consel, Charles, *Microsoft Research - USA*, 2 jours en 2003

Consel, Charles, *Georgia Institute of Technology - USA*, 25 jours en 2003 et en 2005

Consel, Charles, *Yale University - USA*, 3 jours en 2004

Consel, Charles, *Oregon Graduate Institute - USA*, 15 jours en 2004

Réveillère, Laurent, *University of Berkley - USA*, 2 jours en 2002

Réveillère, Laurent, *Microsoft Research - USA*, 2 jours en 2002

Réveillère, Laurent, *Carnegie Mellon University - USA*, 2 jours en 2002

Aumage, Olivier, *University of New Hampshire - USA*, 3 jours en 2005

Aumage, Olivier, *Argonne National Laboratory - USA*, 3 jours en 2005

Brunet, Elisabeth, *University of New Hampshire - USA*, 2 mois en 2004

Mercier, Guillaume, *Argonne National Laboratory - USA*, 15 jours en 2004

Namyst, Raymond, *University of New Hampshire - USA*, 3 jours en 2005

Namyst, Raymond, *Argonne National Laboratory - USA*, 3 jours en 2005

Pagès, Jonathan, *University of New Hampshire - USA*, 7 semaines en 2005

Thibault, Samuel, *University of New Hampshire - USA*, 3 mois en 2003

Beaumont, Olivier, *University of San Diego - USA*, 10 jours en 2003 et en 2005

Hénon, Pascal, *University of Minnesota - USA*, 15 jours en 2003 et en 2005

Ramet, Pierre, *University of Minnesota - USA*, 15 jours en 2003 et en 2005

Roman, Jean, *University of Minnesota - USA*, 15 jours en 2003 et en 2005

5.6 Liste des contrats et subventions

5.6.1 Contrats internationaux et européens

Amigo – Informatique ubiquitaire pour le réseau domestique

Période : 2004-2007

Source : Projet Européen – *Integrated Project* – IP6

Financement : 190Keuros (INRIA)

Description :

Le projet Amigo a pour objectif de rendre utilisable le réseau domestique en développant un intergiciel inter-opérable, standardisé et ouvert. Cet intergiciel permettra la configuration dynamique et automatique de périphériques et de services. Ce projet regroupe les leaders dans le domaine de l'électronique domestique, des télécommunications et du logiciels (Philips, France Telecom, Telefonica, Microsoft Europe...), ainsi que de grands laboratoires de recherche (Inria, Franhofer, VTT...).

Notre objectif dans ce projet est d'utiliser l'approche langages dédiés pour faciliter la création de services dans l'infrastructure proposé par Amigo. Nous comptons développer des aspects méthodologiques et des outils pour assister le développement de langages dédiés. Nous illustrerons ce travail avec deux familles de services : la téléphonie et les flux multimédias.

Serveurs de stockage distribués sur réseaux hautes performances

Période : 2005

Source : Action de collaboration internationale INRIA avec l'Université du New Hampshire - USA

Financement : 5Keuros (INRIA)

Description :

L'objectif de cette collaboration est de concevoir, développer et évaluer de nouvelles versions de iSCSI spécialement adaptées aux réseaux rapides. Ce protocole permet d'adresser des requêtes directement à un contrôleur disque distant via le protocole IP, et permet ainsi de concevoir de nouvelles techniques de stockage distribué dont les accès aux données seraient entièrement pilotés par les clients. Ces travaux ouvriront la voie à la conception de nouveaux serveurs de fichiers distribués pour grappes d'ordinateurs hautes performances, en remplacement de NFS, dont les limites sont largement dépassées sur de telles architectures.

Préconditionneurs parallèles robustes

Période : 2002-2004

Source : Projet NSF – INRIA

Financement : 5Keuros / 15Keuros (IRISA)

Description :

Ce projet a consisté en la conception et la mise en œuvre hautes performances de solveurs hybrides utilisant des techniques issues des méthodes directes et des méthodes itératives pour très grands systèmes linéaires creux provenant d'applications industrielles en vraie grandeur. Les partenaires français étaient le projet ALADIN de l'IRISA, le projet Graal de l'INRIA Rhône-Alpes et l'ENSEEIH IRT ; les partenaires américains étaient des chercheurs des Universités de Minneapolis, d'Indiana, de Berkeley et de Santa Barbara.

5.6.2 Contrats avec des entreprises

Plate-forme programmable de service de communications sur Windows

Période : 2003-2004

Source : Microsoft

Financement : 20.8Keuros (ENSEIRB)

Description :

Nous avons construit une plate-forme de communications, basé sur une modèle client-serveur, permettant de programmer le serveur pour s'adapter à la diversité des client qui peuvent interagir avec lui. Cette adaptation prend la forme de programmes écrits dans un langage restreint. Ce programme permet d'adapter les transactions entre client et serveur pour prendre en compte des éléments tels que la bande passante, les contraintes matériels du client et les préférences de l'utilisateur. Un tel serveur a été baptisé *serveur programmable*.

Le but de ce projet a été de porter cette plate-forme dans un environnement Windows.

Service .NET Programmable

Période : 2003-2004

Source : Microsoft

Financement : 20Keuros (ENSEIRB)

Description :

Dans ce projet, nous avons utilisé la programmabilité introduite au niveau du serveur, pour rendre des services Web adaptables à différents clients. Cette étude s'est effectuée dans le cadre de .NET.

Routeur parallèle tolérant aux pannes

Période : 2002-2003

Source : Alcatel

Financement : 61Keuros / 122Keuros (INRIA Rhône-Alpes)

Description :

Ce contrat fit l'objet d'une collaboration entre les projets INRIA PARIS d'une part (Yvon Jegou et Christine Morin), et la société Alcatel d'autre part. Le projet s'est articulé autour d'un bus logiciel haute performance pour la mise en œuvre de routeurs parallèles sur des architectures de type grappe. L'essentiel de notre travail a consisté à concevoir une extension de Madeleine tolérante aux pannes dans les grappes, c'est-à-dire qui permette à une application de continuer à s'exécuter même en présence de défaillances graves, ainsi que d'accepter la ré-intégration de nœuds "à chaud".

Bibliothèque de multithreading sur machines CC-NUMA

Période : 2004-2005

Source : CEA/DAM Ile-de-France

Financement : 65Keuros (INRIA)

Description :

L'objectif de ce contrat est d'adapter de façon performante notre bibliothèque hybride de threads Marcel aux calculateurs hiérarchiques de type NUMA, tels les serveurs BULL Novascale basés sur le processeur Itanium2, briques de base de la configuration majeure de calcul du CEA. Outre son portage, il s'agit aussi de rendre notre bibliothèque compatible binaires avec les threads POSIX et de développer un environnement de recueil et d'analyse des performances des applications multithreadées. Vincent Danjean a été recruté en tant

qu'ingénieur de recherche (CDD d'une durée d'un an) par le CEA/DAM pour effectuer ce travail.

Parallélisation du code CALDER d'interaction Laser/Plasma

Période : 2002

Source : CEA/DAM Ile-de-France

Financement : 51Keuros (LaBRI)

Description :

Ce contrat a eu pour objectif d'optimiser des techniques d'algorithmique parallèle pour mettre en œuvre des méthodes PIC hautes performances pour la simulation en vraie grandeur de l'interaction Laser/Plasma. Le code CALDER issu de ces travaux est en production sur la machine TERA de la DAM.

Parallélisation du code CODDEX de dynamique des dislocations

Période : 2003

Source : CEA/DAM Ile-de-France

Financement : 32.8Keuros (LaBRI)

Description :

Ce contrat a eu pour objectif d'optimiser des techniques d'algorithmique parallèle pour mettre en œuvre des méthodes dites "sans maillage" hautes performances pour la simulation en vraie grandeur de dynamique des dislocations de la matière sous l'effet de lasers haute énergie. Le code CODDEX issu de ces travaux est en production sur la machine TERA de la DAM.

Simulation de la propagation de la fissuration de la silice due au laser

Période : 2004-2006

Source : CEA/DAM Ile-de-France

Financement : 89.7Keuros (INRIA)

Description :

Ce contrat a pour objectif la modélisation de la propagation de fissures dans la silice dues à l'effet laser (problème de l'endommagement des lentilles du Laser Méga Joule lié aux hautes énergies des faisceaux) en utilisant un couplage multiéchelle "dynamique moléculaire + éléments finis". Cela sera ensuite réalisé par un couplage fin de codes parallèles hautes performances simulant chacun une physique.

Parallélisation du code MIRO

Période : 2002

Source : CEA/CESTA

Financement : 22.7Keuros (LaBRI)

Description :

Ce contrat a eu pour objectif d'optimiser des techniques d'algorithmique parallèle pour mettre en œuvre des méthodes de transformées de Fourier hautes performances pour la simulation en vraie grandeur de la propagation des faisceaux laser dans la chaîne optique du Laser Méga Joule. Le code MIRO parallèle issu de ces travaux est en production sur la machine TERA de la DAM.

Chaîne logicielle EMILIO – Solveur direct hautes performances**Période** : 2002-2004**Source** : CEA/CESTA**Financement** : 90.1Keuros (LaBRI et INRIA)**Description** :

Ce contrat a eu pour objectif de transférer au CEA le résultat de nos recherche sur les solveurs directs parallèles pour résoudre de très grands systèmes linéaires creux provenant d'applications de grande taille (on traite ainsi des systèmes de plusieurs dizaines de millions d'inconnues pour des problèmes 3D). La bibliothèque parallèle EMILIO issue de ces travaux est en production sur la machine TERA de la DAM. Ce transfert a donné lieu aussi à des prestations de consulting auprès des ingénieurs du CEA.

Solveur hybrides hautes performances**Période** : 2002-2004**Source** : CEA/CESTA**Financement** : 27.3Keuros (LaBRI)**Description** :

Ce contrat a pour objectif de montrer la faisabilité de nouveaux solveurs utilisant conjointement des techniques issues des méthodes directes et des techniques issues des méthodes itératives. Le but est de traiter des problèmes réputés comme numériquement difficiles de très grandes tailles (plusieurs centaines de millions d'inconnues, voir un milliard d'inconnues) en faisant sauter le verrou de l'occupation mémoire qui est réhibitoire pour les méthodes directes dans ce cadre.

Téléphonie et nomadique**Période** : juillet à août 2003**Source** : Sinpag**Financement** : 6.2Keuros (LaBRI)**Description** :

L'objectif de ce projet était la sécurisation d'échange de contenus et la gestion des droits (au sens propriété intellectuelle) dans un contexte de téléphonie mobile.

5.6.3 Contrats publics avec des instances nationales**SWAN – Self-aWare mAnagement****Période** : 2004-2006**Source** : RNRT**Financement** : 105.8Keuros (LaBRI)**Description** :

Le projet SWAN (Self aWare mANagement) se propose de développer et d'expérimenter des méthodes de "gestion autonome". Ce projet a pour objectifs de dégager des problématiques de gestion autonome communes à la supervision de réseaux et à l'administration des web services, de contribuer au développement d'outils mathématiques pour la gestion autonome et d'expérimenter ces méthodes au sein de deux plates-formes, l'une concernant l'auto-configuration d'un équipement, l'autre déployant des web services interconnectés.

Composition et raffinement de systèmes sûrs**Période** : 2003-2005**Source** : ACI sécurité

Financement : 28.4Keuros / 142Keuros (INRIA)

Description :

Ce projet est une collaboration de chercheurs en systèmes et en méthodes formelles. Il vise à étudier des méthodes et des outils pour le développement de services systèmes sûrs.

Notre contribution se concentre sur le développement de services de téléphonie sûrs grâce à l'introduction d'un langage dédié à ce domaine. Nous espérons pouvoir formaliser et vérifier des propriétés spécifiques aux services de téléphonie.

Noyau d'infrastructure répartie adaptable

Période : 2000-2003

Source : RNRT

Financement : 166Keuros (INRIA)

Description :

Cette action de recherche s'inscrit dans le prolongement de notre collaboration avec le CNET sur l'adaptation de systèmes réflexifs au moyen de langages dédiés. Le CNET, le projet SOR de l'INRIA-Rocquencourt et le LIP6 sont nos partenaires au sein de cette action.

RMI

Période : 2002-2003

Source : ACI Grid

Financement : 36Keuros (LaBRI)

Description :

L'objectif de ce projet était de promouvoir un modèle de programmation pour les grilles de calcul combinant à la fois des modèles du calcul parallèle et du calcul distribué. Ce modèle s'appuie sur le concept d'objet distribué et de composant logiciel pour la programmation répartie. Le projet visait à concevoir et expérimenter une infrastructure logicielle de communication haute performance permettant à la fois la communication efficace entre objets ou composants ainsi que la programmation parallèle. Il en a résulté la plate-forme distribuée à objets multi-applications PadicoTM capable d'exploiter de multiples technologies réseaux de manière transparente. Cette plate-forme s'appuie sur une extension des bibliothèques Marcel et Madeleine développées au sein de notre groupe.

Protocoles hautes performances, Stockage à grande échelle, Indexation distribuée

Période : 2003-2005

Source : ACI Masses de données

Financement : 10.5Keuros / 1876Keuros (LaBRI)

Description :

Le projet Data Grid Explorer, dirigé par Frank Cappello du LRI a pour but de construire un "grand appareil" pour émuler la grille et les systèmes pair à pair. Cet émulateur est basé sur un grand cluster (plus de 1000 processeurs), une base de données de mesures expérimentales et un ensemble d'outils dédiés à l'expérimentation et à l'analyse d'outils. Notre objectif est de fournir des outils logiciels permettant d'émuler efficacement sur une grappe de machines multiprocesseurs une grille de PC. Pour cela, nous étendons la bibliothèque d'ordonnancement mixte Marcel et la bibliothèque de communication Madeleine. Il s'agit aussi de définir des outils pour la prise de traces et de mesures précises sous ces environnements.

Méthodologie de programmation des grilles, Action d'animation**Période** : 2003-2004**Source** : AS CNRS STP8**Financement** : 5Keuros (LaBRI)**Description** :

Ce projet d'action spécifique a permis la conduite d'une réflexion sur les directions de recherche qui se dessinent aujourd'hui en matière de programmation/exploitation des grilles de calcul et de déterminer les axes important à privilégier.

ALTA : Asynchronous Loss Tolerant Algorithms**Période** : 2003-2004**Source** : ARC INRIA et ACI Grid**Financement** : 36Keuros (IRISA)**Description** :

Il s'agissait d'étudier l'impact des protocoles de communication à tolérance de perte ajustable. Pour ce faire nous avons implémenté une interface de programmation de niveau utilisateur offrant des primitives contrôlant quelles parties de données peuvent être perdues dans un message et à quelle fréquence. Cette approche a été validée sur des méthodes de résolution de type IACA (Itérations Asynchrones - Communications Asynchrones) développées au LIFC (Besançon) où toutes les unités de calcul font leurs itérations sans se préoccuper de l'avancement des autres (sans attendre la réception des données fraîchement calculées) et envoient les données utiles aux autres unités de façon asynchrone.

EPSN – Environnement pour le Pilotage de Simulations Numériques distribuées**Période** : 2002-2005**Source** : ACI GRID**Financement** : 300Keuros (INRIA)**Description** :

L'objectif du projet EPSN est d'analyser, de concevoir et de développer une plate-forme logicielle permettant de piloter et de diriger une application numérique distribuée par la visualisation. Le but est de coupler les potentialités de la réalité virtuelle avec des schémas de résolution numérique hautes performances pour visualiser les résultats obtenus et interagir avec la simulation.

RedGrid – Redistribution de données sur la grille**Période** : 2003-2004**Source** : ARC INRIA**Financement** : 5Keuros / 40Keuros (INRIA)**Description** :

L'objectif de l'action RedGrid est d'étudier et de développer une bibliothèque de redistribution de données de grande taille dans un environnement CORBA pour la grille. Le domaine d'application est le pilotage de simulations numériques distribuées (projet EPSN). Ceci s'est fait en collaboration avec les projets INRIA Graal, PARIS et Algorille.

TLSE – Test pour grands systèmes d'équations**Période** : 2002-2005**Source** : ACI GRID

Financement : 10Keuros / 300Keuros (ENSEEIH-IRIT)

Description :

Le but du projet GRID-TLSE est de constituer un site d'expertise dédié au traitement des matrices creuses et de rendre transparent l'accès et l'analyse comparative des algorithmes. Les développements et les recherches impliquées dans ce projet se situent à trois niveaux : une infrastructure efficace pour le calcul sur une grille ; une interface web pour le calcul sur la grille permettant, à partir d'un simple navigateur Web, d'enchaîner une suite de requêtes de calcul ; des développements pour le site d'expertise. Ceci se fait en collaboration avec le CERFACS, l'ENSEEIH-IRIT, le projet INRIA Graal et des partenaires industriels (CEA, CNES, EADS, IFP).

GRID2 – Action d'animation scientifique

Période : 2002-2004

Source : ACI GRID

Financement : 5Keuros / 15Keuros (LaBRI)

Description :

GRID2 est un projet de l'ACI GRID du ministère de la recherche. Ce groupe a été mis en place pour assurer la formation de jeunes chercheurs, la diffusion d'information et plus généralement les rencontres entre les chercheurs impliqués dans des travaux sur le GRID en France. GRID2 a organisé au moins une fois par trimestre des réunions thématiques portant sur des domaines scientifiques et techniques pointus visant à la diffusion des connaissances entre les chercheurs ; des écoles ont été organisées pour faciliter l'intégration des jeunes chercheurs dans cette communauté en leur donnant les éléments d'une culture commune. Enfin des conférences ont également été organisées afin d'assurer une dissémination rapide des résultats des travaux dans la communauté scientifique.

Simulation informatique des systèmes enzymatiques

Période : 2004-2006

Source : ACI IMPBIO

Financement : 7Keuros / 65Keuros (INRIA)

Description :

Ce projet vise le développement d'une suite de logiciels de modélisation et de simulation en chimie du vivant et notamment des mécanismes enzymatiques. Les outils théoriques essentiels sont, la mécanique quantique et la mécanique statistique. Les codes, conçus pour une utilisation optimale (mais pas exclusive) sur architecture informatique parallèle, se basent sur l'utilisation d'algorithmes à "croissance linéaire" et sur des logiciels déjà écrits à Nancy et Bordeaux (QM/MM et QCPP). La partie appliquée de ce projet porte sur l'étude du mode de fonctionnement des enzymes Méthionine sulfoxyde réductases (Msr). Ces enzymes catalysent la réduction des Méthionine sulfoxydes (MetSO) au sein des protéines suite à un stress oxydant. Des études d'enzymologie très récentes ont démontré un nouveau mécanisme original de réduction de la fonction sulfoxyde par un thiol via une cystéine et passant par la formation d'un intermédiaire acide sulfénique.

Sécurité Java Card

Période : 2001-2003

Source : PROGSI

Financement : 99.8Keuros (LaBRI)

Description :

SERMA Technologies et le LaBRI ont collaboré à un projet intitulé “Sécurité Java Card”, centré sur l’analyse de la sécurité des cartes à puce Java Card. Ce projet labellisé “Société de l’Information” par le Secrétariat d’Etat à l’Industrie, avait pour but de mettre au point une expertise en défauts de sécurité logiciels sur ces produits. Il s’agissait de développer une bonne maîtrise de la structure des produits répondant aux spécifications Java Card, afin de pouvoir mettre en œuvre une gamme étendue d’attaques logicielles de haut niveau technique permettant de tester et de valider la qualité de l’implémentation des produits sécurisés soumis à une évaluation Critères Communs ou ITSEC.

5.6.4 Contrats publics avec instances locales ou régionales

Modélisation, conception et spécification d’algorithmes et de systèmes distribués

Période : 2005

Source : BQR

Financement : 80Keuros (LaBRI)

Description :

L’objectif de ce projet est de définir un cadre permettant de modéliser, de concevoir et de spécifier des algorithmes et des systèmes distribués en mettant en commun un savoir faire sur ces systèmes et ces algorithmes acquis dans différentes thématiques du LaBRI. Les thématiques concernées vont des aspects théoriques pour la modélisation du système distribué, de ses évolutions et des calculs qui s’y déroulent (équipe Combinatoire et Algorithmique et équipe LAAA) jusqu’au déploiement de plates-formes à base de cartes à puces Java ou à base de terminaux mobiles (thème Objets distribués) en passant par la maîtrise et l’optimisation des ressources grâce à la spécification d’une signalisation générique dans les réseaux (thème Comet).

Pôle de Recherche en Informatique

Période : 2005-2007

Source : Région Aquitaine

Financement : 556Keuros (LaBRI)

Description :

Ce dossier de financement concerne plusieurs projets de recherche du LaBRI et correspond au renforcement d’activités existantes ou à la mise en place de nouvelles activités. Il s’agit pour PARADIS des équipes Comet, Runtime et ScAlApplix.

Plate-forme pour le développement d’applications multimédia fiables sur terminaux communicants

Période : 2003-2004

Source : Région Aquitaine

Financement : 250Keuros (LaBRI)

Description :

Le monde des terminaux mobiles communicants (TMC) (téléphone, assistant personnel, ordinateurs) connaît une progression fulgurante depuis quelques années. Ces évolutions ont jusqu’à présent porté plus sur les capacités matérielles que sur les services offerts aux utilisateurs. Avec l’arrivée de nouveaux standards de communication (GPRS, UMTS, BlueTooth, WiFi), il devient encore plus important d’offrir des applications exploitant au mieux les ressources disponibles mais surtout de les fiabiliser.

Notre projet vise à développer une plate forme pour le développement de services multimédia fiables sur TMC. Il se base sur des compétences existant dans plusieurs équipes du

LaBRI et propose de développer un serveur graphique évolutif s'exécutant sur une grappe de micro-ordinateurs organisés autour d'un serveur spécialisé.

La contribution du projet Phoenix consiste à utiliser l'approche des langages dédiés pour faciliter la création de services dans cette plate-forme, sans sacrifier sa robustesse. La contribution du projet SOD concerne le support d'exécution distribué de la plate-forme.

Développement de services de téléphonie robustes

Période : 2004-2007

Source : Région Aquitaine

Financement : 69Keuros (LaBRI)

Description :

Ce projet concrétise notre intérêt pour le domaine des télécommunications. Il a pour objectif de développer des techniques permettant le développement de services de téléphonie robustes. Pour ce faire, différentes techniques sont à l'étude : des techniques langages (analyse de programmes, annotations, déclarations...) et des techniques systèmes (capacités, mécanismes de gestion de ressources...).

Composants systèmes à haute performance

Période : 2003-2005

Source : Région Aquitaine

Financement : 35Keuros (LaBRI)

Description :

Ce projet vise à combiner l'approche des langages dédiés, pour le développement de composants systèmes sûrs, et la spécialisation de programmes pour l'optimisation en taille et en temps. Ces travaux ont déjà fait l'objet de plusieurs publications.

GRINTA – Grappes et grilles de recherche pour l'interactivité et les applications scientifiques

Période : 2004-2005

Source : Région Aquitaine + ACI GRID + INRIA

Financement : 1150Keuros (Bordeaux 1 + INRIA)

Description :

Les objectifs principaux de GRINTA, qui est la composante bordelaise du projet national GRID5000, sont le déploiement, la gestion et l'utilisation algorithmique et applicative d'une grille à l'échelle du territoire français. Ce projet concerne Runtime, ScAIApplix, Sod et IPARLA qui fait partie de l'équipe Image et Son.

Mise en œuvre et sécurisation d'un environnement distribué pour la modélisation et la simulation de structures en biologie cellulaire

Période : 2003

Source : BQR

Financement : 90Keuros (LaBRI)

Description :

Ce projet a pour but la mise en œuvre d'un environnement permettant de manipuler des modèles et de réaliser des simulations dans le domaine de la biologie cellulaire. L'intervention de Paradis, et plus précisément de SOD concerne la sécurisation des calculs à l'aide de cartes à puce.

5.7 Animation de la recherche

5.7.1 Comités de rédaction

Parallel Computing – Numéro spécial pour la conférence PMAA 2004 – Parallel Matrix Algorithms and Applications

5.7.2 Comités de programmes de colloques

DNAC 2004 – 18ème Congrès DNAC : De Nouvelles Architectures pour les Communications
- Contrôle, Maîtrise et Autonomie des Réseaux - Qualité de Service, Sécurité, Mobilité

IWUC'05 – 6th International Conference on Enterprise Information Systems Ubiquitous Computing

NGNM 2005 – International Workshop on Next Generation Networking Middleware

MAN 2005 – IFIP Open Conference on Metropolitan Area Networks Architecture, protocols, control, and management

NETCON 2005 – IFIP International Conference on Network Control and Engineering for QoS, Security and Mobility

GRES 2005 – Colloque Gestion de Réseaux et de Services

CC 2005 – 14th International Conference on Compiler Construction

ECOOP 2004 – European Conference on Object Oriented Programming

PLDI 2002, 2004 – ACM Sigplan Conference on Programming Language Design and Implementation

JFDLPA 2004 – Première Journée Francophone sur le Développement de Logiciels Par Aspects

DECOR 2004 – Première Conférence Francophone sur le Déploiement et la Reconfiguration de Logiciels

CFSE 2003, 2005 – Conférence Française sur les Systèmes d'Exploitation

GPCE 2002 – ACM Sigplan Conference on Generative Programming and Component Engineering

ESOP 2002 – European Symposium on Programming

ASIA-PEPM 2002 – ACM Sigplan Asian Symposium on Partial Evaluation and Semantics-Based Program Manipulation

HiPC 2004 – International conference on High Performance Computing

HeteroPar 2004 – International conference on heterogeneous computing

IPDPS 2004 – IEEE International Parallel & Distributed Processing Symposium

VECPAR 2004 – 6th International Conference on High Performance Computing for Computational Science

RenPar 2002, 2003, 2004, 2005 – Rencontres francophones du parallélisme

PMAA 2002, 2004 – Parallel Matrix Algorithms and Applications Conference

IPDPS 2003 – IEEE International Parallel & Distributed Processing Symposium

ICGSN 2005 – International Conference on Grid Networking and Services

NOTERE 2004, 2005 – Nouvelles Technologies de la Répartition

5.7.3 Organisation de colloques, écoles jeunes chercheurs, ...

JDIR 2005 – Journées Doctorales en Informatique et Réseaux

Co-organisation en 2003 du séminaire Dagstuhl *Domain-Specific Program Generation*

Organisation en 2003 du premier workshop du *IFIP Working Group 2.11* à Saint-Émilion

Comité scientifique de l'école thématique (CNRS, INRIA, ACI GRID, GDR ARP) GRID 2002 – Calcul Distribuée, Méta-Computing, Globalisation des Ressources (Aussois)

Organisation en 2003 des journées algorithmique numérique sur la grille à Bordeaux

Co-organisation de la journée “les nouveaux enjeux de la carte à puce” en janvier 2002

Co-organisation de IWJPCD de 2001 à 2005 – International Workshop on Java for Parallel and Distributed Computing

Co-organisation de IPDPS en 2003 et 2004 – International Parallel and Distributed Processing Symposium

5.7.4 Administration de la Recherche

Création et responsabilité du thème Comet (*COntext-aware ManagEment and neTworking*)

Création et direction de l'équipe Phoenix (LaBRI - INRIA) – *Technologies langages pour les services de communications*

Membre du IFIP Working Group 2.11 sur la Génération de Programmes

Membre du groupe de réflexion Informatique Diffuse à l'Observatoire Français des Techniques Avancées (OFTA)

Expertise pour la sélection des sites du projet Grid 5000

Création et direction du projet commun LaBRI - MAB - INRIA Futurs ScAIApplix – *Schémas et Algorithmes Hautes Performances pour les Applications Scientifiques Complexes*

Vice-présidence du Comité des projets de l'INRIA futurs depuis janvier 2002 puis Présidence à compter d'octobre 2005

Membre de la Commission d'Évaluation nationale de l'INRIA

Participation au conseil scientifique d'EDF pour l'évaluation du programme de R & D “Simulation 2010”

Responsabilité du thème “Algorithmique et applications hautes performances” de l'ACI d'animation GRID2

Responsabilité du thème SOD Systèmes et Objets Distribués

Les membres de l'équipe ont participé à 22 jury extérieurs de thèses et 18 fois en tant que rapporteur. De même, nous avons participé à 3 jury extérieurs d'HDR et 2 fois en tant que rapporteur.

Projets transversaux

6.1 Projet VISIDIA : Visualisation de structures complexes évolutives et d'algorithmes distribués sur ces structures

6.1.1 Introduction

L'informatique a été marquée ces dernières années par le développement d'une information de plus en plus complexe et massivement distribuée sur les réseaux (comme Internet). Les progrès technologiques des postes de travail et des réseaux, notamment à haut débit, et les efforts de standardisation des services d'interopérabilité (comme CORBA) ont permis d'envisager de réelles applications distribuées. Ceci continue de stimuler une intense activité de recherche autour des systèmes et des applications distribuées, et de leur fondement que constitue l'algorithmique distribuée.

La compréhension, l'étude et l'analyse d'algorithmes distribués passent par la recherche de modèles qui permettent de les exprimer et de les coder de manière intuitive (ce qui aide à les développer et à les enseigner), formelle (ce qui permet de les prouver en utilisant des outils mathématiques) et implémentable (ce qui permet de les programmer par une simple conversion). Les systèmes de réécriture de graphes, et plus généralement les calculs locaux dans les graphes, constituent un modèle intuitif et formel puisqu'il permet de bénéficier de l'abstraction de haut niveau des systèmes de réécriture. Ceci permet une présentation simple ainsi que des preuves unifiées et plus faciles d'algorithmes distribués.

De plus, pour un système à échange de messages, nous avons dans le cadre du projet visidia élaboré une méthode générale qui permet de fournir, pour un algorithme décrit par un calcul local, un code source implémentable.

Le projet visidia poursuit dans le cadre de ce formalisme trois objectifs complémentaires. Le premier consiste en l'étude des transformations de graphes (d'étiquettes, de sommets, ou d'arêtes) pour les appliquer à la construction de classes de graphes et d'algorithmes, et tout particulièrement d'algorithmes distribués. Le deuxième objectif est de réaliser un environnement logiciel de simulation, de visualisation et de tests permettant d'aider à la mise au point et à la validation de ces algorithmes. Enfin, le troisième objectif consiste à implémenter réellement et à expérimenter ces algorithmes sur des plate-formes distribuées. Ces trois objectifs sont étroitement liés et couvrent un spectre large allant de l'étude formelle à la mise en oeuvre opérationnelle des applications réparties.

Depuis plus de dix ans, des équipes du LaBRI s'intéressent aux transformations de graphes décrites sous forme de réécritures, dont elles ont étudiés de façon plus particulière, deux catégories :

- les réécritures d’étiquetages de graphes qui ont permis de développer un modèle mathématique solide pour formaliser à un bon niveau d’abstraction divers algorithmes distribués
- les réécritures de constituants du graphe (sommet, arête, sous-graphes) conduisant à la transformation du graphe lui-même.

Dans ce cadre sont apparus divers problèmes dont on peut donner deux exemples :

- Si la modélisation d’algorithmes distribués à l’aide de transformations de l’étiquetage d’un graphe représente un modèle mathématique utile et performant, elle nécessite le développement de méthodes et d’outils pour mettre en oeuvre et implémenter ces algorithmes. La visualisation de l’exécution de ces algorithmes permet de mieux les comprendre, les appréhender et les prouver.
- Les transformations de graphes obtenues par réécriture d’arêtes ou de sommets utilisent des opérations comme le produit ou le produit fibré de graphes. Celles-ci nécessitent des calculs complexes dont les résultats exprimés sous forme abstraite ne sont pas toujours limpides et demandent impérativement à être représentés visuellement. Ces transformations permettent également de décrire des classes de graphes qui admettent des propriétés logiques ou algorithmiques intéressantes.

Le besoin d’outils permettant de visualiser des structures complexes de ce type ou de simuler l’exécution d’algorithmes distribués est encore plus évident lorsqu’on essaie d’enseigner ce type de formalisme et de techniques à des étudiants. Les éditeurs de graphes classiques ne sont pas adaptés à ce besoin car ils ne tiennent pas compte en général des transformations de graphes.

Le projet *visidia* vise donc le développement d’une méthodologie générale allant de la modélisation des graphes, de leurs transformations, et des algorithmes distribués jusqu’à une implémentation effective. Ce continuum entre la théorie et la pratique permet d’une part de valider les avancées théoriques en les confrontant à une vraie implémentation, et d’autre part suscite des nouvelles problématiques, voire des nouvelles conjectures, permettant ainsi d’affiner les modèles théoriques étudiés et de mieux comprendre leur puissance. Plusieurs thématiques sont abordés dans ce projet ; citons par exemple la construction de classes de graphes, la modélisation d’algorithmes distribués, l’utilisation d’outils combinatoires et probabilistes, la caractérisation de familles d’algorithmes par des langages logiques, la visualisation de graphes et des algorithmes. De fait, ce projet est transversal aux équipes L3A et Combinatoire dont la synergie permet de fédérer plusieurs compétences complémentaires et qui sont nécessaires à la réalisation des objectifs de *visidia*.

Mots-Clés Transformations de graphes, Algorithmes distribués, Systèmes de réécriture, Visualisation

6.1.2 Description des activités et principaux résultats

Systèmes de réécriture d’étiquettes et algorithmique distribuée

On considère un réseau de processeurs ayant une topologie quelconque. Il est modélisé par un graphe étiqueté, connexe, et non orienté, où les sommets correspondent aux processeurs et les arêtes aux canaux de communication. Les processeurs communiquent seulement par envoi de messages asynchrones. L’étiquette d’un sommet (resp. d’une arête) code l’état du processeur (resp. du lien de communication) correspondant. Un calcul sur un réseau, qui consiste donc à faire évoluer ces états, peut être décrit par des règles de réécriture de graphes. Chaque règle, définie par un graphe connexe et deux étiquetages de ce graphe, traduit un pas élémentaire du calcul effectué sur une portion locale du graphe.

Un système de réécriture est défini par un ensemble fini de règles et par une structure de calcul de contrôle local : ce contrôle autorise ou bien interdit l'application de certaines règles. Pour les règles avec priorité, une relation de priorité est une relation d'ordre partiel entre les règles. Lorsque deux règles se chevauchent sur des sommets et/ou des arêtes, c'est la règle de plus grande priorité qui est appliquée. Pour les règles avec contextes interdits, l'application d'une règle ne peut se faire que si dans son environnement local, aucun contexte interdit n'est présent.

Le système distribué sur lequel on implémente ces algorithmes est un système anonyme et asynchrone à échange de messages en mode asynchrone. Chaque processeur communique uniquement avec ses voisins. Les liens de communication acheminent les messages dans l'ordre. Afin d'avoir des solutions efficaces et pour éviter les problèmes de non déterminisme des algorithmes distribués, nous nous intéressons aux algorithmes probabilistes. Les solutions sont satisfaisantes d'un point de vue pratique, puisqu'elles permettent une implémentation automatique des règles de réécriture. L'idée de base est de synchroniser un processeur avec un de ses voisins (rendez-vous) ou avec tous ses voisins (synchronisation en étoile) selon le calcul local considéré. L'implémentation d'un algorithme distribué décrit par des calculs locaux repose sur des procédures probabilistes de synchronisation locale dont l'analyse est étudiée par l'équipe combinatoire.

Etant donné un algorithme distribué décrit par un système de réécriture. Pour l'implémentation, il faut d'abord choisir une procédure de synchronisation. Chaque processeur exécute les actions suivantes. Il tente de se synchroniser avec un ou tous ses voisins. Une fois synchronisé, il exécute une règle de réécriture ; ce qui correspond à une étape de l'algorithme. L'exécution d'une telle règle consiste, en général, à échanger des informations avec les voisins qui sont impliqués dans la synchronisation, et à faire une mise à jour de son état. A la fin de l'exécution de la règle, la synchronisation est interrompue et chaque processeur tente à nouveau d'entrer dans une nouvelle synchronisation. Il faut bien noter ici que, plusieurs synchronisations peuvent avoir lieu, sur des endroits disjoints du réseau, au même moment.

La méthode précédente a été utilisée pour mettre au point un catalogue de plusieurs exemples d'algorithmes distribués : calcul distribué d'arbres recouvrants, élection, nommage, etc. De plus, nous avons étudié dans le cadre de ce modèle, quelques paradigmes de l'algorithmique distribuée permettant de contribuer à l'enrichissement de ce dernier. Nous avons étendu le codage d'algorithmes distribués par des systèmes de réécriture aux graphes avec identités. En effet, nous avons introduit des règles génériques (ou méta-règles) qui tiennent compte de l'identité des noeuds. Un autre paradigme important des systèmes distribués que nous avons regardé concerne la détection locale de la terminaison globale. En effet, il n'est pas trivial pour un processeur de savoir et de décider si, globalement, tous les processeurs ont terminé. Nous avons conçu une méthode constructive et unifiée fondée sur la combinaison de systèmes de réécriture pour rajouter la terminaison à un algorithme distribué qui ne possède pas cette propriété.

Un des critères fondamentaux dans le développement d'algorithmes distribués est de savoir si le système est synchrone ou asynchrone. Un synchroniseur est un mécanisme qui permet, en produisant des séquences de pulsions locales, d'exécuter un algorithme conçu pour des systèmes synchrones dans des environnement asynchrones. Notre contribution est de spécifier et d'étudier les synchroniseurs à l'aide de systèmes de réécritures permettant ainsi d'élaborer

un modèle général pour la synchronisation. Nous avons mis au point plusieurs synchroniseurs pour différentes classes de graphes. En utilisant le codage de ces synchroniseurs par des systèmes de réécritures, nous avons pu fournir un codage automatique d'un algorithme synchrone dans un système distribué asynchrone.

Pour valider en pratique les approches décrites ci-dessus, nous avons développé un outil logiciel *visidia* en java qui offre une interface pour implémenter, tester et visualiser l'exécution d'un algorithme distribué défini par un système de réécriture. L'outil offre une interface graphique permettant de construire un graphe à l'aide de la souris, et de programmer un algorithme distribué en utilisant une bibliothèque de primitives de base. Il est également possible d'utiliser directement l'interface, sans programmer, pour construire à l'aide de la souris un système de réécriture qui sera traduit automatiquement en un programme. L'utilisateur pourra ainsi tester, expérimenter ou visualiser cet algorithme en utilisant l'environnement de *visidia*. Plusieurs algorithmes sont déjà implémentés et peuvent être visualisés directement par l'utilisateur. Cet outil est utile pour mettre au point les algorithmes distribués que nous sommes entrain de développer dans le cadre de nos travaux de recherches, et tout particulièrement pour aider à les prouver. A chaque noeud du graphe, le logiciel *visidia* associe automatiquement un processus léger (*thread*) et donne la possibilité à l'utilisateur de manipuler les communications entre les noeuds du graphe à l'aide des fonctions classiques d'envoi et de réception de messages, en masquant la complexité des communications de bas niveau. L'utilisateur peut ainsi prototyper un algorithme distribué à l'aide de ces fonctions et observer l'animation de l'exécution de cet algorithme : échange de messages, modification des attributs et des états des noeuds et des arêtes, progression des calculs.... Nous avons développé deux versions de *visidia* : l'une s'exécute sur une seule machine et l'autre utilise un réseau de machines physiquement réparties. La première version est destinée au prototypage et à la visualisation d'algorithmes, quant à la deuxième, elle donne la possibilité de manipuler des graphes de très grande tailles permettant ainsi le passage à l'échelle et l'expérimentation grandeur nature de l'exécution d'algorithmes distribués. Dans la version répartie de *visidia*, chaque machine héberge un sous ensemble de noeuds, ce qui permet d'augmenter considérablement les capacités de calcul. Signalons qu'en exploitant le polymorphisme en java, les deux versions de *visidia* utilisent le même code pour chaque algorithme. Nous utilisons l'environnement *visidia* comme support pédagogique pour l'enseignement en deuxième et troisième cycle. La visualisation de l'exécution d'un algorithme distribué permet aux étudiants de mieux comprendre les paradigmes fondamentaux des systèmes et des algorithmes répartis. Les travaux pratiques et les projets réalisés par les étudiants en utilisant cet environnement logiciel leur permet de prototyper des algorithmes distribués en manipulant des primitives de haut niveau. Il en résulte, grâce à la pratique, une meilleure acquisition des concepts de l'algorithmique distribuée qui sont généralement difficiles à enseigner.

Le développement de la version répartie de *visidia* nous a conduit au problème de décomposition de graphes pour obtenir des algorithmes distribués efficaces. Plus précisément, nous nous sommes intéressés aux algorithmes de décomposition d'un graphe en un ensemble de clusters ayant un petit rayon et faiblement connectés entre eux. Intuitivement, l'intérêt de telles décompositions est de minimiser les communications intra-cluster puisque les clusters sont de faibles rayons et inter-clusters puisque la connexité du graphe des clusters est faible. Nous avons étudié plusieurs algorithmes distribués qui permettent de construire cette décomposition en temps linéaire, et même sous linéaire. Pour certaines familles de graphes (les graphes circulants par exemple) la complexité est même logarithmique. Comme application de cette décomposition, on obtient un algorithme déterministe qui permet de construire de

manière distribuée un sous-graphe couvrant peu dense ayant un petit facteur d'étirement (*sparse graph spanner with low stretch*) et de complexité sous-linéaire en temps.

Afin de disposer d'un langage de description des algorithmes distribués codés par des systèmes de réécriture, nous avons développé un langage appelé Lidia. Ce langage, destiné à devenir un langage de programmation d'algorithmes distribués, est fondé sur une extension de la logique du premier ordre, appelé \mathcal{L}_∞^* , permettant d'exprimer les propriétés locales des graphes. Il est doté en particulier de quantificateurs de comptage et de fonctions arithmétiques. Un intérêt particulier de ce langage réside dans le fait qu'il permet de caractériser la puissance algorithmique des calculs locaux. Le modèle de calcul de Lidia utilise un système qui gère des transitions de la forme précondition-action, similaire à celui des automates IO. La particularité de Lidia par rapport aux langages des systèmes concurrents est que toutes les préconditions sont exclusivement décrites dans la logique \mathcal{L}_∞^* . Nous avons étudié la puissance de cette logique en montrant en particulier sa complétude. Nous avons également établi un catalogue d'algorithmes distribués implémenté à l'aide de Lidia. Cet aspect logique des calculs locaux est effectué dans L3A, et complète l'aspect combinatoire et algorithmique qui se situe plutôt dans l'équipe Combinatoire.

Décomposition de graphes symétriques par des transformations "pullback"

De nombreux travaux utilisent des propriétés structurelles des graphes, pour obtenir des algorithmes plus performants, testant certaines propriétés de graphes, ou testant l'appartenance des graphes à des classes connues. Dans ce domaine, la décomposition modulaire, les 2-structures, les systèmes d'équations récursives sont des outils largement répandus. Cependant, certaines classes de graphes complexes, comme les grilles, les graphes de De Bruijn, ne rentrent pas dans ces formalismes.

Nous nous sommes donc intéressés à des graphes présentant une structure fortement symétrique. Nous montrons que l'existence d'une symétrie dans un graphe permet de le définir comme le produit d'un facteur dépendant du graphe de départ et d'un facteur canonique dépendant uniquement de la nature de la symétrie. Par itération de ce procédé, un graphe présentant plusieurs symétries, même locales, peut se décomposer en un produit de facteurs, appelés "facteurs premiers". Le produit de graphes, appelé "produit fibré", utilisé dans cette décomposition, est une généralisation du produit de Kronecker. Dans chacun des facteurs, les sommets sont regroupés en classes d'équivalence, appelées "fibres". Le produit fibré est alors obtenu par des produits de Kronecker sur certains couples de fibres particuliers.

Inversement, la construction d'un graphe par applications successives de symétries (locales ou non) à un graphe de base, peut être formalisée en un système de réécriture par pullback dans la catégorie des graphes. Dans ce système, une règle de réécriture est la donnée d'un morphisme de graphes, p , dont l'image est appelée "alphabet" de la règle. Une occurrence de l'alphabet A dans un graphe G est la donnée d'un deuxième morphisme, x , de G dans A . La réécriture de G est alors obtenue en effectuant le produit pullback de p et de x dans la catégorie des graphes. Une règle permettra donc la substitution d'une occurrence d'un sommet v de l'alphabet par son image réciproque par p , et la substitution d'une occurrence de l'arête e de l'alphabet par son image réciproque par p .

Cette technique de réécriture unifie les techniques utilisant des remplacements de sommets et celles utilisant des remplacements d'arêtes. L'application de cette technique permet notamment d'obtenir une décomposition des grilles carrées en facteurs premiers, ou inverse-

ment un système de réécriture par pullback engendrant la famille des grilles carrées.

Opérations de flips sur les réalisateurs de graphes

Dans le cadre de recherche d'algorithmes efficaces pour dessiner des graphes, nous avons été amenés à nous intéresser aux réalisateurs. Un réalisateur (realizer) d'un graphe plan maximal est un ensemble de trois arbres recouvrant. Schnyder a montré en 1990 l'existence d'un réalisateur pour tout graphe plan maximal, qui peut être calculé en temps linéaire. Les réalisateurs sont utiles pour plusieurs algorithmes de graphes : dessin de graphes, codage de graphes, ordre canonique, 3-orientations, ou encore les arbres recouvrants ordonnés. Les réalisateurs peuvent également caractériser les graphes planaires en termes d'incidences ; i.e. un graphe est planaire si et seulement si la dimension de l'ordre d'incidence des sommets et des arêtes est au plus 3. Les réalisateurs d'un même graphe ont été déjà étudiés. Des opérations appropriées permettant de transformer un réalisateur en un autre d'un même graphe ont été introduites. Ceci a permis de caractériser une forme normale particulière très utile pour construire des algorithmes efficaces pour les graphes plans.

Nous avons étudié les réalisateurs de taille n , i.e. l'ensemble des réalisateurs de tous les graphes plans de taille n . Nous avons généralisé le théorème de Wagner aux réalisateurs. Wagner a montré en 1936 que deux graphes planaires maximaux sont équivalents par des transformations de diagonales, appelées *flips*. Un flip est une opération qui consiste à supprimer la diagonale (u_1, u_3) d'un quadrilatère (u_1, u_2, u_3, u_4) et la remplacer par la diagonale opposée (u_2, u_4) (voir Fig.). Ainsi, on peut obtenir tous les graphes plans maximaux de taille n par des flips diagonaux. Ce théorème a été étendu aux tores. Nous avons montré une extension de ce théorème aux réalisateurs de taille n . Pour ce faire, nous avons introduit deux nouvelles opérations : *deux flips diagonaux coloriés*. Nous avons montré que l'ensemble des réalisateurs de taille n est un poset.

Nous avons aussi caractérisé le nombre de noeuds internes d'un réalisateur. Plus précisément, nous avons prouvé que $(\xi_0 + \xi_1 + \xi_2 - \Delta = n - 1)$ où ξ_j est le nombre de noeuds internes de l'arbre T_i et Δ est le nombre de faces tricolores. Comme application de cet invariant, nous prouvons qu'un arbre recouvrant ordonné, d'un graphe plan maximal, avec au plus $\lfloor \frac{2n+1-\Delta}{3} \rfloor$ feuilles peut être calculé en temps linéaire. Nous avons également énuméré et généré aléatoirement les réalisateurs de taille n .

6.1.3 Projets et perspectives 2005-2008

Dans le cadre du projet visidia, nous continuons à enrichir et approfondir le modèle des systèmes de réécriture de graphes pour décrire, exprimer, étudier et implémenter les algorithmes distribués. Nous nous intéressons actuellement aux problèmes suivants :

- Nous avons supposé que les réseaux sur lesquels se déroulent les systèmes de réécriture (et donc les algorithmes) sont statiques. Nos travaux s'orientent actuellement vers la prise en compte de l'aspect dynamique du réseau. Cependant, La transcription directe des algorithmes distribués définis par des systèmes de réécriture dans un modèle où les sommets sont des entités autonomes de calcul semble être difficile. Un moyen d'introduire la dynamique consiste à utiliser les agents mobiles comme modèle d'exécution des systèmes de réécriture. En effet, on peut considérer les sommets comme des sites passifs et on se donne un nombre initial d'agents mobiles qui se déplacent de site en site et qui exécutent les règles de l'algorithme. Dans un tel modèle, la topologie du graphe peut changer puisque c'est un agent qui découvre localement le voisinage d'un sommet avant

- d'exécuter un pas de calcul.
- Une autre direction de nos travaux consiste à étudier les réductions de graphes dans un environnement distribué. Les algorithmes de réduction sont définis par un ensemble de règles de réduction et d'un ensemble fini de graphes. Chaque règle décrit une transformation topologique locale du graphe conduisant à une diminution de sa taille. L'application des règles continuent jusqu'à arriver à un graphe final qui appartient à l'ensemble fini de graphes. Certaines classes de graphes peuvent être caractérisées par des règles de réduction ; citons par exemple les graphes série-parallèles, ou les graphes de largeur arborescente bornée. Ces graphes admettent des solutions efficaces pour plusieurs problèmes difficiles dont certains sont NP-complets en général. En particulier, la résolution de problèmes de décision se fait de manière efficace à l'aide de règles de réduction. Notre objectif est d'étudier les algorithmes de réduction distribués ; c'est à dire que les règles peuvent être appliquées de manière distribuée. Il s'agit de fixer un modèle et de préciser les conditions qui permettent la résolution de problèmes de réduction de graphes de manière distribuée.
 - Un autre paradigme de l'algorithmique distribuée qui n'a pas encore été considéré par les systèmes de réécriture est celui de la *tolérance aux pannes*. Nous sommes entrain de rajouter à l'outil logiciel la possibilité de simuler et gérer une panne sur un sommet ou une arête, en le marquant par exemple. Parallèlement, la question encore ouverte concerne la prise en compte de la tolérance aux pannes par un système de réécriture. A l'aide d'étiquettes supplémentaires et sous certaines conditions, on peut transformer un système de réécriture en un autre qui gère les pannes. La tolérance aux pannes est aussi liée à l'*auto-stabilisation*. Un algorithme est auto-stabilisant s'il converge vers une configuration terminale, en partant de n'importe quelle configuration initiale. Les systèmes de réécriture peuvent être utilisés comme modèle de haut niveau de l'auto-stabilisation permettant l'utilisation des résultats des systèmes de réécriture pour l'étude des algorithmes auto-stabilisants.

6.1.4 Publications d'audience internationale

Revues avec Comité de lecture

- [905] BONICHON (N), LE SAEC (B) et MOSBAH (M). – Orthogonal drawings based on the stratification of planar graphs. *Discrete Mathematics*, vol. 276, n 1-3, 2004, pp. 43–57.
- [906] BONICHON (N) et MOSBAH (M). – Watermelon uniform random generation with applications. *Theor. Comput. Sci.*, vol. 307, n2, 2003, pp. 241–256.
- [907] METIVIER (Y), MOSBAH (M), WACRENIER (P) et GRUNER (S). – A distributed algorithm for computing a spanning tree in anonymous t-prime graphs. *Studia Informatica Universalis*, vol. Special Issue Vol.2 HS2, 2002, pp. 141–157.
- [908] MÉTIVIER (Y), SAHEB (N) et ZEMMARI (A). – Analysis of a randomized rendezvous algorithm. *Inform. and Comput.*, vol. 184, n1, 2003, pp. 109–128.

Colloques avec Comité de programme et Actes

- [909] BAUDERON (M) et CARRERE (F). – Decomposing graphs with symmetries. *In : International Conference on Graph Transformation (ICGT 2002)*. pp. 45–59. – Springer-Verlag, 2002.

- [910] BAUDERON (M) et MOSBAH (M). – A unified framework for designing, implementing and visualizing distributed algorithms. *In : Graph Transformation and Visual Modeling Techniques (GT-VMT 2002)*, éd. par Bottoni (P) et Minas (M). – Barcelona, Spain, 2003.
- [911] BONICHON (N), LE SAEC (B) et MOSBAH (M). – Optimal area algorithm for planar polyline drawings. *In : 28th International Workshop, Graph - Theoretic Concepts in Computer Science (WG)*. pp. 35–46. – Springer-Verlag, 2002.
- [912] BONICHON (N), LE SAEC (B) et MOSBAH (M). – Wagner’s theorem on realizers. *In : Proceedings of the 29th International Colloquium on Automata, Languages and Programming*. pp. 1043–1053. – Springer-Verlag, 2002.
- [913] DERBEL (B) et MOSBAH (M). – Distributing the execution of a distributed algorithm over a network. *In : 7th International Conference on Information Visualization (IV03)*. pp. 485–490. – London, England, 16-18 Juillet 2003.
- [914] DERBEL (M) et MOSBAH (M). – Distributed graph traversals by relabeling systems with applications. *In : Graph Transformation for Verification and Concurrency (GT-VC 2005)*, éd. par R. Heckel, B. König (A. R). pp. 25–40. – San Francisco, California, USA, 2005. A PARAÎTRE.
- [915] DUCHON (P), HANUSSE (H), SAHEB (N) et ZEMMARI (A). – Broadcast in the rendezvous model. *In : STACS : 21st International Symposium on Theoretical Aspects of Computer Science*. pp. 559–570. – Springer, 2004.
- [916] GODARD (E), MÉTIVIER (Y), MOSBAH (M) et SELLAMI (A). – Termination Detection of Distributed Algorithms by Graph Relabelling Systems. *In : Proc. of 1st International Conference on Graph Transformation*. pp. 106–119. – Springer-Verlag, 2002.
- [917] HAMID (B) et MOSBAH (M). – An implementation of a failure detector for local computations in graphs. *In : International Conference on Parallel and Distributed Computing Networks, (PDCN 2005)*. pp. 473–478. – Innsbruck, Austria, 15-17 Février 2005.
- [918] HAMID (B) et MOSBAH (M). – An automatic approach to self-stabilization. *In : the 6th ACIS International Conference on Software Engineering, Artificial Intelligence, Networking, and Parallel/Distributed Computing (SNPD2005)*. pp. 123–128. – Towson University, Maryland, U.S.A., 23 - 25 Mai 2005.
- [919] HAMID (B) et MOSBAH (M). – A formal model for fault-tolerance in distributed systems. *In : THE 24th International Conference on COMPUTER SAFETY, RELIABILITY AND SECURITY*. – Fredrikstad, Norvège, 28 - 30 Septembre 2005. A PARAÎTRE.
- [920] HAMID (B) et MOSBAH (M). – Visualization of self-stabilizing algorithms. *In : 9th International Conference on Information Visualization (IV05)*. pp. 550–555. – London, England, 6-8 Juillet 2005.
- [921] MÉTIVIER (Y), MOSBAH (M), OSSAMY (R) et SELLAMI (A). – Synchronizers for local computations. *In : International Conference on Graph Transformation*. pp. 271–286. – Springer-Verlag, 2004.
- [922] MÉTIVIER (Y), MOSBAH (M) et SELLAMI (A). – Proving distributed algorithms by graph relabeling systems : Examples of trees in networks with processor identities. *In : Applied Graph Transformations*, pp. 45–57. – Grenoble, 2002.

- [923] MOSBAH (M) et OSSAMY (R). – Checking global properties for local computations in graphs with applications to invariants testing. *In : Proceedings of the Fifth Mexican International Conference on Computer science, 20-24 September*, éd. par IEEE. pp. 35–42. – Computer Society, 2004.
- [924] MOSBAH (M) et OSSAMY (R). – A programming language for local computations in graphs : Computational completeness. *In : Proceedings of the 5th. Mexican International Conference in Computer Science Colima Mexico 20-24 September*, éd. par IEEE. pp. 12–19. – Computer Society, 2004.
- [925] MOSBAH (M) et SELLAMI (A). – Un environnement général pour étudier et implémenter les algorithmes distribués. *In : Eighth Maghrebian Conference on Software Engineering and Artificial Intelligence (MCSEAI'04)*. pp. 161–172. – Centre de Publication Universitaire, 2004.
- [926] OSSAMY (R). – A simple randomized k -local election algorithm for local computations. *In : WEA, 4th. International Workshop on Efficient and Experimental Algorithms*. pp. 290–301. – Springer-Verlag, 2005.

6.1.5 Publications d'audience nationale

Colloques avec Comité de programme et Actes

- [927] MOSBAH (M), SELLAMI (A) et ZEMMARI (A). – Résolution distribuée de conflits dans un réseau par les systèmes de réécritures. *In : Journées Scientifiques Franco-phones, JSF 2003*. – Tozeur, 2003.

6.1.6 Autres publications

Langage Spidle, Développement d'applications de streaming, architecture client-serveur programmable, personnalisation de serveurs en fonction des utilisateurs.

Langage SPL, Conception et développement d'une machine virtuelle SIP, architecture logicielle pour le déploiement de services de téléphonie.

Tempo, Plateforme de spécialisation automatique pour divers systèmes Unix, développement d'un évaluateur partiel pour les programmes écrits en C.

Madeleine, Bibliothèque de communication multi-grappes qui implémente efficacement un concept de canal de communication pouvant être soit physique (configurations homogènes) soit virtuel (configurations hétérogènes).

Marcel, Bibliothèque de threads utilisateurs s'adaptant à la plateforme visée (architectures mono-processeur, SMP, NUMA) et intégrant un serveur d'événements améliorant le temps de réaction des applications aux entrées/sorties.

LinuxActivations, Extension du noyau Linux permettant de contrôler l'ordonnancement des threads utilisateurs lorsqu'ils effectuent des opérations d'entrées/sorties bloquantes dans le noyau.

MPICH/MADIII, Implémentation multithreadée et multi-protocole du standard de communication par passage de messages MPI.

PadicoTM, Déploiement d'intergiciels sur les réseaux des grilles de calcul.

PM2, Support exécutif générique de bas niveau qui intègre la gestion des threads, la gestion des communications et une gestion mémoire distribuée élémentaire.

EPSN, Plateforme couplée de pilotage interactif et de visualisation pour des simulations numériques parallèles et distribuées.

PaStiX, Bibliothèque de solveurs parallèles haute performance basés sur des factorisations complètes et incomplètes pour des très grands systèmes linéaires creux issus de problèmes 3D.

PHIDAL, Bibliothèque de solveurs parallèles itératifs basés sur des décompositions hiérarchiques et des factorisations multiniveaux.

Scotch, Bibliothèque pour le partitionnement et le placement de graphes et de maillages irréguliers de grande taille.

Jem, Plateforme de manipulation de ressources connectées à un réseau fixe.

JToe, Bibliothèque Java de transfert efficace d'objets.

Mandala, Plateforme pour le développement d'applications Java distribuées intégrant la notion de conteneur et d'appel de méthodes distantes asynchrones.

6.1.7 Formation par la recherche

Habilitations

[928] MOSBAH (M). – *Construction d'algorithmes et Transformation de graphes*. – Talence, France, Habilitation à diriger des recherches, Université Bordeaux I, 2002.

6.2 Algorithmes et Méthodes pour la Bioinformatique

6.2.1 Introduction

Le projet transversal *Algorithmes et Méthodes pour la Bioinformatique* regroupe des chercheurs des équipes CombAlgo et MVTsi, qui travaillent en collaboration quotidienne avec des chercheurs de l'Université Bordeaux 2 et de l'INRA Bordeaux sur l'application de techniques d'informatique théorique aux problèmes d'origine biologique. Ces collaborations ont pour cadre le Centre de Bioinformatique de Bordeaux, une structure régionale multi-établissement (CNRS, INRA, INSERM, Universités) dont le LaBRI est l'une des principales composantes.

Notre objectif dans le projet AMBI est de construire des vues globales et multi-échelles de systèmes vivants, grâce au développement de méthodes informatiques pour la comparaison de génomes, l'analyse de réseaux d'interactions, la modélisation du comportement cellulaire, l'intégration de données hétérogènes, et la modélisation de l'architecture des plantes.

Notre activité est organisée en 3 thèmes et ses recherches mettent en jeu des disciplines variées – biologie, botanique, mathématiques appliquées, informatique –, dont l’association constitue un enjeu scientifique et technique majeur pour la génomique moderne. Les méthodes et les outils que l’équipe élabore sont mis à la disposition des chercheurs et des étudiants accueillis au sein de ses équipes. Ils sont aussi diffusés largement auprès de la communauté scientifiques à travers à la fois des grandes revues de biologie et d’informatique.

Notre activité se résume à trois thèmes.

La *génomique comparée* utilise la comparaison de génomes complets (séquencés et annotés) pour comprendre à la fois les fonctions biomoléculaires de gènes individuels et les mécanismes d’évolution moléculaire qui ont pour résultat le gain et la perte de ces fonctions. C’est l’objectif du programme Génolevures. Nous associons à cette vue essentiellement statique du génome, une vue dynamique produite par l’étude des réseaux *d’interaction moléculaires*, extraits à partir de la compilation de résultats d’expériences en laboratoire et organisé dans la base de données européenne IntAct. Afin de développer des vues globales de systèmes vivants, il est nécessaire de définir des outils formels pour modéliser des processus biomoléculaires et pour raisonner sur leur dynamique. Nous développons une approche originale pour la modélisation discrète, stochastique et multi-échelles de systèmes vivants, inspirée d’approches qui ont fait leurs preuves en vérification formelle de gros systèmes industriels.

L’intégration de données hétérogènes cherche à développer, à l’échelle d’une cellule, une vue globale des multiples relations qui existent entre les composants qui la constituent. Ces composants sont pour la plupart des protéines, codées par les gènes. Leurs relations sont décrites par des multiples méthodes complémentaires, qui produisent des données expérimentales à haut débit. Le système BlastSets fournit une plateforme d’intégration de ces données, et un algorithme de recherche de groupes similaires munis d’une mesure de signification probabiliste.

L’analyse et la modélisation de structures biologiques arborescentes. L’essor des bases de données contenant des descriptions arborescentes de l’architecture de plantes ou de structures génomiques crée un important besoin de nouveaux outils de traitement automatique et d’analyse. Les travaux initiaux de comparaison d’architecture de plantes et de structures secondaires d’ARNs ont ainsi jeté les bases de nouvelles techniques algorithmiques exploratoires prenant en compte la nature structurelle des données. Les applications biologiques nous ont conduit à étendre cette démarche selon un axe théorique (extension des outils conceptuels et algorithmiques existants) et applicatif (application de ces techniques à différents domaines biologiques).

Mots-Clés Bioinformatique, Méthodes prédictives, Génomique comparée, Intégration de données, Modélisation des plantes, Modélisation multi-échelles stochastique

6.2.2 Description des activités et principaux résultats

Génomique comparée et interactions moléculaires Comprendre la structure et la fonction de génomes consiste en trois activités : répertorier l’ensemble de **composants** élémentaires codés par les génomes, identifier comment ces composants **travaillent ensemble** pour assurer les fonctions cellulaires, et comprendre comment les génomes **changent** et acquièrent des nouvelles fonctions.

Chaque génome, aussi apparenté qu’il soit à d’autres génomes, a sa propre histoire. Cette histoire est unique et non reproductible, le résultat d’innombrables hasards et adaptations opportunistes. Il est impossible de retracer l’histoire d’une espèce seule. Mais les espèces actuelles sont descendues d’ancêtres communs, et si l’on compare plusieurs espèces apparen-

tées, on peut mieux identifier les composants de chacun, mieux comprendre les mécanismes d'évolution, et mieux comprendre ces adaptations historiques. Le défi particulier de la génomique comparée est de trouver les relations généralisables entre les génomes étudiés tout en appréciant l'histoire individuelle de chacun.

Les relations qui peuvent exister entre gènes et entre génomes sont diverses : relations d'orthologie et de paralogie entre gènes, familles de gènes et familles de protéines, relations de synténie et de duplication entre segments chromosomiques, et le gain ou la perte de gènes et de fonctions. Les protéines engendrées par les gènes ont aussi des relations, dynamiques, qui décrivent leurs manière d'interagir et de coopérer pour réaliser les comportements cellulaires. Pour développer une *vue dynamique* de la fonction, il est nécessaire d'observer expérimentalement les interactions au sein de processus bio-chimiques, et de corrélérer l'expression et l'interaction avec les conditions expérimentales. Enfin, pour structurer les relations statiques et dynamiques ainsi identifiées, on fait appel à la *modélisation*, qui permet de décrire et à terme de simuler le comportement de cellules.

L'équipe AMBI s'est beaucoup investi dans cette démarche ambitieuse, et mène ses recherches au sein de trois projets :

- *Génolevures*, la génomique comparée à grande échelle de levures hémiascomycètes,
- *IntAct*, le développement de ressources européennes et outils pour les réseaux d'interaction de protéines,
- *S. siliceus*, la modélisation multi-échelles stochastique de processus biologiques. (Ce dernier est présenté dans le paragraphe "Modélisation du vivant" de la présentation de l'équipe MVTsi.)

Le projet national Génolevures Avec leurs génomes petits et compacts, les levures offrent une opportunité unique pour l'exploration de l'évolution des génomes eucaryotes par l'analyse comparative de plusieurs espèces. Le programme Génolevures est dédié à la comparaison à grande échelle de génomes de levures des différentes branches de la classe des Hémiascomycètes, dans le but d'adresser des questions fondamentales d'évolution moléculaire telles que le degré de conservation de gènes, l'identification de gènes spécifiques à une espèce, une clade ou une classe, la distribution de gènes parmi les familles fonctionnelles, la vitesse de divergence de séquences et de cartes, et des mécanismes de brassage chromosomique. Le LaBRI est responsable du recherche et développement bio-informatique et de la mise à disposition de données et d'outils pour le projet national Génolevures. Génolevures est une communauté active de chercheurs repartis dans neuf laboratoires français et associés en GDR CNRS sous la direction du professeur Jean-Luc Souciet de l'Université Strasbourg. Le site web Génolevures [458] (<http://cbi.labri.fr/Genolevures/>) fournit des outils pour la comparaison multi-espèces de génomes, un accès aux alignements et annotations, et un accès aux séquences complètes et aux autres analyses réalisées par le projet.

Génolevures est soutenu par l'ACI IMPBIO "Génolevures En Ligne" et par le contrat Région "Génotypage et Génomique Comparée".

La première phase du projet Génolevures, achevée en 2000, concernait une étude très large de 13 espèces de levures par une stratégie de séquençage aléatoire de fragments d'ADN génomique. Cette étude, qui a donné lieu à un numéro spécial de la revue *FEBS Letters*, a mise en évidence différents degrés de conservation de gènes et de synténie, et a permis de développer un modèle de l'évolution moléculaire des génomes des hémiascomycètes fondé sur la *duplication ségmentale*. Ces résultats comparatifs ont également permis de revisiter le génome de *Saccharomyces cerevisiae*. Ce travail a également permis des études spécifiques [476, 454, 492].

Sur la base de cette première étude, le Consortium Génolevures a séquencé et annoté

quatre génomes complets, choisis à la fois pour leur position phylogénétique et pour leur intérêt médical ou biotechnologique. Nous avons développé au LaBRI, en collaboration avec Lionel Frangeul de l'Institut Pasteur, les outils d'annotation utilisés par le réseau d'experts. Ce travail a permis d'identifier environ 29.000 nouveau gènes et d'y associer une annotation fonctionnelle prédite. Grâce à des techniques de clustering consensus développés au LaBRI, nous avons identifié 4.700 familles de protéines multi-espèces, qui ont permis l'estimation des gains et de perte de gènes et de fonctions ainsi que l'estimation des taux de substitution entre gènes orthologues et paralogues. La comparaison de redondances et de cartes chromosomiques a révélé que l'évolution des lignées différentes de levures est le résultat d'une combinaison riche et variée de mécanismes distinctes, y compris la formation de répétitions en tandem de gènes, la duplication ségmentale, la duplication massive et la perte de gènes à grande échelle. Un article significatif dans la revue *Nature* présente l'ensemble de ses résultats [449].

Le projet européen IntAct Le LaBRI est partenaire du projet européen IntAct, financé dans le FP 5 grâce au contrat QLRI-2001-00015 – TEMBLOR, coordonnée par l'Institut Européen de Bio-informatique (EBI). IntAct développe une base de données fédérée européenne d'interactions protéine-protéine. Les interactions entre protéines sont fondamentales à pratiquement tous les processus biologiques, et la description des réseaux d'interactions protéine-protéine correspondants à ces processus est une étape importante dans la compréhension de la physiologie cellulaire. Des données sur les interactions protéine-protéine sont produites à grande vitesse, grâce à une gamme complémentaire de nouvelles techniques artisanales et à haut débit. L'intégration et la comparaison de la masse de données résultante est un défi majeur de la bio-informatique. Le LaBRI est responsable pour la recherche et développement d'outils graphiques et d'une plate-forme logicielle extensible pour l'interrogation, l'analyse et l'annotation de données d'interaction protéine-protéine. Nous avons développé un logiciel de visualisation hautement interactif appelé **ProViz**, qui propose une interface orientée-biologiste qui intègre les données d'interaction, des ontologies existantes pour décrire les protéines, et des vocabulaires contrôlés pour les interactions et les méthodes expérimentales développées par le projet IntAct[455].

Le groupe IntAct a développé un modèle objet et une base de données pour les résultats expérimentaux, actuellement mise à disposition de la communauté scientifique sur le site de l'EBI mais conçu pour installation sous forme de nœuds locaux fédérés. Cette base répertorie au jour d'aujourd'hui environ 45.000 interactions multiples et 52.000 interactions binaires, portant sur environ 31.000 protéines et 1.600 expériences [453, 457]. Les formats d'échange et les vocabulaires contrôlés utilisés par IntAct ont été défini au sein du *Human Proteome Organisation Proteomics Standards Initiative* dans lequel nous sommes très actifs [452].

L'intégration de données hétérogènes La connaissance des génomes complets d'un nombre croissant d'organismes et l'accumulation de résultats expérimentaux produits par des approches 'post-séquençage' à l'échelle cellulaire et à haut débit doivent permettre de mieux comprendre l'articulation entre les mécanismes moléculaires et les fonctions cellulaires. C'est l'intégration de données à la fois volumineuses et hétérogènes qui permettra de progresser vers une vision plus intégrée du fonctionnement de la cellule. Cette intégration est l'enjeu scientifique et technologique qui est au coeur de notre travail.

Ce travail s'inscrit dans le cadre du projet BlastSets (<http://cbi.labri.fr/BlastSets>), projet porté par le CBiB - Centre de Bioinformatique de Bordeaux - et coordonné par Antoine de Daruvar, professeur de l'Université Bordeaux 2. Le projet est soutenu pour trois ans dans le cadre de l'ACI IMPBIO 2003.

Nous proposons le développement d'une nouvelle méthode qui permet de rapprocher dynamiquement les informations hétérogènes disponibles à l'échelle des génomes (ou protéomes) entiers. Il s'agit d'un mode d'intégration des données qui vise à faire apparaître et à exploiter des correspondances nouvelles entre des informations variées. Ce projet repose sur un parti pris dans la représentation des connaissances biologiques qui consiste à ramener systématiquement les connaissances au niveau de groupes de séquences (gènes ou protéines). Notre projet consiste à spécifier, concevoir, produire et valider un système informatique. Ce travail fait appel à des compétences dans différentes disciplines : informatique, mathématique et biologie.

Notre travail se situe à l'intersection de deux approches d'intégration de données biologiques. Ces approches sont

- la mise en oeuvre de références croisées entre les banques de données biologiques existantes, sous forme de liens statiques à travers lesquels on peut naviguer,
- la modélisation des processus biologiques mis en jeu au sein de la cellule, discipline en plein essor, mais qui souffre encore souvent de la complexité du formalisme choisi et de la difficulté de représenter les données à l'aide de ce formalisme.

Nous proposons une nouvelle approche qui exploite le concept de voisinage dont le principe est de s'intéresser aux relations entre objets biologiques plutôt qu'aux objets pris indépendamment. Les principes de base de la stratégie BlastSets sont les suivants :

- utiliser une représentation simple homogène et efficace pour les différentes relations de voisinages : le groupe de séquences (gènes ou protéines). Un groupe de gènes correspond donc à un ensemble de gènes voisins pour un certain critère ou une certaine information biologique,
- utiliser un modèle probabiliste pour comparer des groupes de gènes d'origines diverses et décider s'ils sont similaires ou non. Ce modèle se base uniquement sur la composition des groupes, c'est-à-dire la probabilité d'observer un certain nombre d'éléments en commun entre deux groupes.

La conception et le développement d'un premier prototype du système BlastSets ainsi que sa validation biologique ont été réalisés en étroite relation avec des biologistes et des bioinformaticiens du CBiB.

En particulier, une étude approfondie des sources diverses de données et des critères de regroupement considérés a permis de formaliser la notion de voisinage et d'aboutir à une représentation uniforme sous forme de DAG pour les ensembles de groupes issus de chaque voisinage. Cette représentation joue un rôle central aussi bien pour le stockage des données que pour les résultats fournis par le système. Elle est également au coeur des perspectives d'évolution et d'optimisation du système discutées ci-après.

En termes de mise en oeuvre, le système informatique BlastSets a nécessité en particulier :

- la constitution effective d'une base de données de groupes de séquences,
- la définition d'un format d'importation standard qui permette de rajouter facilement de nouveaux groupes,
- le développement d'un système qui permet de formuler des requêtes consistant à soumettre un ou plusieurs groupes (groupes requêtes), et qui identifie les groupes similaires présents dans la base de données (groupes cibles).

L'application et la base de données sont opérationnelles et disponibles en ligne sous forme de pages et de services WEB sur le site du CBiB (<http://cbi.labri.fr/BlastSets>).

Les principes du système BlastSets ont fait l'objet d'un article publié en 2004 dans *Nucleic Acids Research* [447].

Modélisation et analyse de structures biologiques arborescentes Les méthodes usuelles de l’agronomie, de la foresterie et de l’écologie, considèrent généralement la production globale des plantes (nombre de fruits, volume de bois, masse de carbone, etc.). Des avancées récentes dans différents champs des mathématiques appliquées permettent de refonder ces méthodes en intégrant à la fois la dynamique de développement et la structure de la plante. Notre but est ainsi d’analyser le développement de la structure de la plante aux échelles macroscopiques. Cet objectif nécessite de tenir compte de la structure arborescente des données dans les différentes méthodes d’analyse. On retrouve là un parallèle avec l’analyse de données génomiques où les spécificités structurelles de ces données conditionnent très fortement les méthodes d’analyse proposées.

Dans ce contexte, le développement de méthodes d’investigation de l’architecture des plantes ou de structures génomiques revêt une importance majeure. En modélisation des plantes d’une part, la comparaison structurelle d’individus permet d’aborder sous un angle nouveau le problème de l’évaluation des modèles de simulation de la croissance. D’autre part, la comparaison d’ARNs trouvent de nombreuses applications en biologie. Il est, en effet, couramment admis que si deux ARNs ont une structure spatiale proche, ils auront des fonctions biologiques proches.

L’organisation structurelle des entités d’une plante, i.e. son architecture, ou la structure secondaire d’un ARN sont souvent représentés par un graphe arborescent. Ainsi comparer deux structures de plantes ou deux structures secondaires se ramène formellement à la comparaison de deux arborescences. La famille d’algorithmes que nous étudions consiste à reconstruire un graphe cible en appliquant à un graphe initial différentes opérations structurelles élémentaires, appelées opérations d’édition. On peut montrer que la définition d’une telle distance se ramène à la recherche d’une solution optimale dans un espace dont la taille croît exponentiellement avec la taille des objets à comparer. Ainsi, pour définir des algorithmes efficaces à la fois en temps et en espace, la solution optimale est construite de façon incrémentale en utilisant des techniques de programmation dynamique.

Les applications biologiques envisagées nous ont amené à étudier trois grands types d’extensions algorithmiques afin de mieux prendre en compte les spécificités structurelles des objets biologiques analysés.

- *Description multi-échelles.* Pour mettre en évidence les différents niveaux d’organisation dans la description de l’architecture des plantes un formalisme, basé sur la notion de graphe quotienté, a récemment été introduits : les structures arborescentes multi-échelles. Nous avons repris cette décomposition en graphe multi-échelles pour la représentation de structures secondaires d’ARNs [247]. La connaissance de cette information multi-échelles dans les algorithmes d’édition permet de contraindre biologiquement la recherche d’une solution optimale. Un premier travail nous a ainsi permis de définir une méthode de comparaison entre deux plantes utilisant deux échelles dans leurs descriptions [65].
- *Edition locale.* L’édition de deux arborescences permet de mettre en correspondance les sous-parties homologues des deux arborescences. Cependant dans de nombreux cas d’applications réelles, on assiste à un phénomène “d’éclatement” des structures lors de la recherche de l’appariement optimal. Il est alors difficile d’interpréter biologiquement cette mise en correspondance. Le problème algorithmique sous-jacent consiste à définir les méthodes permettant de minimiser le nombre de composantes dans le graphe induit tout en conservant une solution optimale [65]. Ce même éclatement peut se traduire par une dissimilarité marquée entre les deux individus comparés. Pourtant, ces deux structures peuvent faire apparaître localement des parties très similaires. Nous avons développés des méthodes d’*édition locale* [295, 244], sur le principe de celles appliquées aux

séquences, consistant à rechercher le meilleur appariement entre deux sous-structures extraites des arborescences comparées. Ces méthodes permettent d'identifier des motifs arborescents à l'intérieur de plantes ou de structures secondaires.

- *Evaluation de l'autosimilarité.* La botanique introduit des notions comme les complexes réitérés (duplication de parties de la plante) et l'âge physiologique des méristèmes qui traduisent une forme d'autosimilarité des arborescences. Les premiers résultats [243, 931] ont ouvert de nombreuses pistes algorithmiques pour quantifier ces notions botaniques considérées comme fondamentales dans la compréhension du développement de la plante.

Ces outils sont mis en œuvre au sein de la plateforme logicielle ALEA et sont évalués sur des bases de données architecturales en collaboration avec l'UMR AMAP (Cirad - INRA) ou sur des bases de données de structures secondaires avec l'aide du Centre de BioInformatique de Bordeaux (CBiB).

L'ensemble de ce travail s'inscrit dans le projet ARBORESCENCES, soutenu pour trois ans par l'ACI Nouvelles Interfaces des Mathématiques et dans l'Action Spécifique CNRS AReNa.

6.2.3 Projets et perspectives 2005-2008

Génomique comparée et interactions moléculaires Le GDR CNRS du Consortium Génolevures, en plus d'être renouvelé pour la deuxième fois, vient d'obtenir du Génoscope l'accord sur le séquençage de cinq nouvelles espèces. Ces espèces occupent des positions phylogénétiques clés et leur analyse permettra de résoudre un certain nombre de questions ouvertes sur les mécanismes d'évolution moléculaire, notamment les rôles respectifs de la duplication massive et la duplication segmentale. Grâce au retour sur l'expérience de l'annotation des quatre espèces précédents et l'identification de règles spécifiques aux levures, l'équipe bordelaise réalise des nouveaux outils algorithmiques pour rendre plus efficace l'annotation de nouvelles espèces par les membres du Consortium.

La question fondamentale qui reste est : comment apparaissent et comment disparaissent les gènes ? Toutes les espèces que nous étudions comportent un nombre significatif de gènes spécifiques, qui n'ont aucun semblable dans les génomes actuels. Ils viennent d'où ? En étudiant les différents moyens de créer des gènes (fusions, recombinaisons, transferts) mais aussi les reliques des gènes en voie de disparition, nous espérons mieux comprendre les mécanismes impliqués.

Pour les réseaux d'interaction moléculaires, nos travaux actuels portent sur les différentes méthodes d'extraction de réseaux à partir des masses de données expérimentales. Nous développons des méthodes d'extraction dites *dictées par une politique* qui permettent de définir à la fois les critères d'identification et de combinaison d'observations, et les structures différentes des graphes qui pourront en être construites. L'objectif à long terme est de travailler en amont des activités de modélisation et de simulation stochastique.

L'intégration de données hétérogènes Sur la base des résultats obtenus, il s'agit d'une part d'utiliser le système BlastSets afin d'étudier des hypothèses biologiques et d'autre part d'améliorer le système en intégrant certaines optimisations.

Un premier travail en cours porte sur l'évaluation d'un certain nombre de méthodes de clustering appliquées aux données d'expression des gènes issues d'expériences de puces à ADN. Ces données sont capitales pour la compréhension de la régulation de l'expression des gènes et leur analyse (clustering suivi de l'annotation des clusters obtenus) est rendue difficile du fait des très nombreuses méthodes de clustering disponibles et des choix qu'elles impliquent. Dans cette étude, nous allons utiliser le système BlastSets afin d'évaluer et

de comparer différentes méthodes de clustering en confrontant les clusters produits à des complexes protéiques publiés dans la littérature.

Nous travaillons également sur l'application du système BlastSets dans le cadre de l'étude du voisinage des génomes. En effet, le nombre de génomes disponibles en constante augmentation permet à présent de les regrouper selon certains critères. Le but est d'identifier des corrélations entre certaines propriétés des génomes comme par exemple la présence d'une voie métabolique par rapport à l'utilisation préférentielle d'acides aminés dans les protéines.

Des travaux théoriques récents ont été réalisés sur l'identification d'ensembles pertinents pour leur comparaison aux ensembles requêtes. Plutôt que de stocker un très grand nombre d'ensembles dans une base de données, il est plus intelligent de générer les ensembles à la volée sur la base d'une représentation moins explicite. L'aboutissement de ses travaux devrait permettre de proposer des optimisations significatives pour le traitement des requêtes, ainsi que d'améliorer la qualité globale des résultats renvoyés.

Les réflexions sur l'identification des ensembles pertinents pour les comparaisons ainsi que l'utilisation du système nous ont conduit à considérer un type de requêtes particulièrement 'difficiles' : il s'agit de la comparaison de voisinages correspondant à des graphes tels que les réseaux d'interactions des protéines et les enchaînements de réactions dans les voies métaboliques. La comparaison de tels voisinages (ensemble d'ensembles de séquences) correspond à l'identification de paires de sous-graphes maximisant la mesure de similarité entre ensembles. Ce problème est NP-difficile et l'approfondissement des recherches sur l'identification des ensembles pertinents pour les calculs de comparaison permet d'envisager des améliorations algorithmiques ayant un fort potentiel en terme d'application biologique.

Modélisation et analyse des structures biologiques arborescentes Les travaux initiaux ont jeté les bases de nouvelles techniques algorithmiques exploratoires prenant en compte la nature structurelle des données. Les applications biologiques nous conduisent à étendre cette démarche selon deux axes :

Analyse de la variabilité d'un ensemble d'individus

- *Modification des algorithmes globaux.* Du fait de la possibilité de croissance des structures que nous étudions, il nous paraît, par exemple, important d'étudier les algorithmes d'alignement à extrémités libres (au niveau des feuilles des arborescences). L'édition à extrémités libres revient à aligner une arborescence entière avec le préfixe (ou le suffixe) d'une autre arborescence, les extrémités libres étant donc soit les feuilles, soit la racine.
- *Ajout d'informations biologiques structurelles aux méthodes de comparaison.* Par souci de généralité Les travaux initiaux traitant de la comparaison de plantes se sont focalisés sur les méthodes de comparaison entre arborescences non-ordonnées. Cependant, il est possible de définir un ordre partiel sur les nœuds fils en distinguant l'entité suivante, construite par le même méristème, des entités portées dérivant par ramification. La prise en compte d'un ordre partiel sur les nœuds fils d'un nœud donné de l'arborescence pose de nouveaux problèmes algorithmiques. Dans le cas de l'analyse de structures secondaires d'ARNs, la définition de nouvelles opérations d'édition issue de l'expérience biologique est également un champ majeur d'investigation.
- *Contraintes multi-échelles.* La prise en compte de l'information multi-échelles est un moyen important de contraindre biologiquement la recherche des solutions optimales. Dans le cas le plus général, il reste nous donc à développer les méthodes permettant de d'utiliser l'information topologique à plus de deux échelles.

Analyse structurelle d'un individu

- *Étiquetage automatique.* Ce problème consiste à retrouver dans un individu toutes les occurrences d'un motif structurel particulier afin de constituer par exemple un échantillon homogène. Actuellement, seul un repérage manuel de ces entités, occurrence par occurrence est possible.
- *Identification de motifs et autosimilarité.* L'identification de motifs est un problème voisin, mais techniquement très différent, qui consiste à identifier les sous-parties homologues au sein d'une ou plusieurs arborescences. L'étiquetage résulte de l'identification algorithmique de structures homologues. Une telle technique sera utilisée pour établir des schémas « d'auto ou d'inter-corrélation » sur la structure des plantes : à quels endroits une plante répète-t-elle les mêmes motifs arborescents dans sa structure globale? La réponse à une telle question doit nous permettre d'envisager de nouvelles techniques de modélisation de plantes basées sur la quantification d'un degré d'autosimilarité des structures botaniques.

La combinaison des différents types d'alignements (global/à extrémités libres/local) avec les différentes structures d'ordre possibles sur les arborescences (non-ordonnée/partiellement ordonnée/ordonnée), la minimisation du nombre de composantes connexes et enfin la prise en compte d'une, deux ou plusieurs échelles dans les arborescences induit une combinatoire d'algorithmes dont beaucoup nécessitent une étude spécifique (e.g. étude de la complexité algorithmique associée). Dans ce contexte, nous chercherons à mettre en place un formalisme permettant de décrire tous ces algorithmes de manière unifiée, en vue de faciliter la mise en évidence de leurs propriétés.

6.2.4 Publications d'audience internationale

Reuves avec Comité de lecture

- [929] BARRIOT (R), POIX (J), GROUPI (A), BARRE (A), GOFFARD (N), SHERMAN (D), DUTOUR (I) et DE DARUVAR (A). – New strategy for the representation and the integration of biomolecular knowledge at a cellular scale. *Nucl. Acids Res.*, vol. 32, n12, 2004, pp. 3581–3589.
- [930] DULUCQ (S) et TICHIT (L). – RNA secondary structure comparison : exact analysis of the Zhang-Shasha tree edit algorithm. *Theoret. Comput. Sci.*, vol. 306, n1-3, 2003, pp. 471–484.
- [931] FERRARO (P), GODIN (C) et PRUSINKIEWICZ (P). – Toward a quantification of self-similarity in plants. *Fractals*, vol. 13, n2, 2005, pp. 91–109.
- [932] SHERMAN (D), DURRENS (P), BEYNE (E), NIKOLSKI (M) et SOUCIET (J. G. C). – Génolevures : comparative genomics and molecular evolution of hemiascomycetous yeasts. *Nucleic Acids Research*, vol. 32, January 2004, pp. D315–D3128. – Database issue.

6.2.5 Publications d'audience nationale

6.2.6 Autres publications

Langage Spidle, Développement d'applications de streaming, architecture client-serveur programmable, personnalisation de serveurs en fonction des utilisateurs.

Langage SPL, Conception et développement d'une machine virtuelle SIP, architecture logicielle pour le déploiement de services de téléphonie.

Tempo, Plateforme de spécialisation automatique pour divers systèmes Unix, développement d'un évaluateur partiel pour les programmes écrits en C.

Madeleine, Bibliothèque de communication multi-grappes qui implémente efficacement un concept de canal de communication pouvant être soit physique (configurations homogènes) soit virtuel (configurations hétérogènes).

Marcel, Bibliothèque de threads utilisateurs s'adaptant à la plateforme visée (architectures mono-processeur, SMP, NUMA) et intégrant un serveur d'évènements améliorant le temps de réaction des applications aux entrées/sorties.

LinuxActivations, Extension du noyau Linux permettant de contrôler l'ordonnancement des threads utilisateurs lorsqu'ils effectuent des opérations d'entrées/sorties bloquantes dans le noyau.

MPICH/MADIII, Implémentation multithreadée et multi-protocole du standard de communication par passage de messages MPI.

PadicoTM, Déploiement d'intergiciels sur les réseaux des grilles de calcul.

PM2, Support exécutif générique de bas niveau qui intègre la gestion des threads, la gestion des communications et une gestion mémoire distribuée élémentaire.

EPSN, Plateforme couplée de pilotage interactif et de visualisation pour des simulations numériques parallèles et distribuées.

PaStiX, Bibliothèque de solveurs parallèles haute performance basés sur des factorisations complètes et incomplètes pour des très grands systèmes linéaires creux issus de problèmes 3D.

PHIDAL, Bibliothèque de solveurs parallèles itératifs basés sur des décompositions hiérarchiques et des factorisations multiniveaux.

Scotch, Bibliothèque pour le partitionnement et le placement de graphes et de maillages irréguliers de grande taille.

Jem, Plateforme de manipulation de ressources connectées à un réseau fixe.

JToe, Bibliothèque Java de transfert efficace d'objets.

Mandala, Plateforme pour le développement d'applications Java distribuées intégrant la notion de conteneur et d'appel de méthodes distantes asynchrones.

6.2.7 Formation par la recherche

6.3 Visualisation de Structures Combinatoires

6.3.1 Introduction

Les outils de visualisation ont eu un rôle déterminant ces dernières années pour rendre l'informatique accessible au plus grand nombre. En effet, la possibilité de visualiser plus aisément les informations permet de transmettre d'énormes quantités de données sous des formes compréhensibles et facilement manipulables. Bien que d'autres techniques puissent contribuer à cet effort, comme par exemple l'extraction de données ou les techniques d'intelligence artificielle, la visualisation jouera toujours un rôle central. La visualisation de graphes est un chapitre important de la recherche menée en visualisation d'informations. De nombreuses structures de données sont modélisables par des graphes et des outils mathématiques efficaces permettent à des chercheurs et ingénieurs de divers secteurs d'employer ces structures avec succès. Les propriétés et comportements de divers graphes peuvent souvent être déduits automatiquement et manipulés par des outils de calcul appropriés. La recherche traditionnelle porte sur divers algorithmes mathématiques, notamment de dessin, qui ne sont souvent testés que sur des graphes relativement petits alors qu'aujourd'hui, les problèmes pratiques se posent sur de grands graphes de plusieurs dizaines de milliers de nœuds. L'interaction avec la représentation visuelle des graphes, dépendante du champ d'application, est peu traitée. Ceci est un des éléments majeurs de notre champs de recherche. Par exemple, même si la représentation graphique d'une structure biologique comme les ARNs est celle d'un graphe, l'habitude de dessin des biologistes doit être respectée et l'interaction entre le graphe, la base de données et l'utilisateur est essentielle. La visualisation d'informations est très développée en Amérique du Nord et peu en Europe. Lorsqu'on parle de visualisation d'informations en liaison avec les graphes, la problématique peut se décomposer en plusieurs sous thèmes :

- construction et classification de paramètres relevants,
- dessins de grands graphes,
- fragmentation de graphes,
- structure de données et algorithmique efficace,
- adaptation aux champs d'application.

Mots-Clés Visualisation, Masse de données, Interaction, Algorithmes et Dessins de Graphes, Logiciel de manipulation de graphes

6.3.2 Description des activités et principaux résultats

L'activité du projet "Visualisation de Structures Combinatoires" porte sur l'ensemble du domaine de "la visualisation d'informations", de l'étude des paramètres structuraux à la navigation dans des grandes structures. De plus, les recherches vont de travaux fondamentaux jusqu'au développement d'applications s'articulant autour de la plateforme logicielle "Tulip". Les domaines applicatifs de ces travaux sont nombreux, et l'on peut citer le multimédia, la gestion de fichiers, la bio-informatique et les bases de données. La plateforme "Tulip" [953] tend à devenir un des 5 logiciels de référence pour la manipulation de graphes. Elle a été spécialisée à travers les logiciels "EVAT" (comparaison d'arbres) et "ARNA" (aide visuelle à l'alignement d'ARNs). Les travaux de ce projet ont été classés par deux fois au concours Invoviz [947, 942].

Le projet est également impliqué dans l'ACI masse de données "Navgraph" ainsi que dans l'ACI jeune chercheur "visualisation de datacubes".

Construction et classification de paramètres relevant La structure intrinsèque du graphe contient une grande partie de la sémantique des données. Par exemple, pour un site web le nombre de références dans une page qui se traduit par l'arité d'un nœud peut contribuer à la caractérisation de cette page (bibliographie, pages de liens, etc). Autre exemple, pour les systèmes de fichiers, la recherche de leurs modifications passe par le calcul des parties qui ont une structure quasi-similaire.

L'équipe a étudié les propriétés structurelles des graphes pour les exploiter dans la chaîne de visualisation. Par exemple, un des paramètres structurels importants sur les arbres binaires est le nombre de Strahler. A l'origine, ce nombre, défini par des hydrologues, permet de caractériser les bassins fluviaux. Nous avons étendu la notion de nombre de Strahler aux arbres quelconques ainsi qu'aux cartes ou graphes valués. Dans le cas des arbres généraux nous avons étudié les propriétés asymptotiques de ce paramètre, et dans le cas des graphes nous avons établi un parallèle entre les nombres de Strahler et les programmes récursifs.

Ce paramètre a été utilisé pour optimiser la navigation interactive. En effet, les nombres de Strahler sont une fonction croissante des feuilles vers la racine, ils représentent l'importance du flux d'information. Lorsqu'un utilisateur navigue à l'intérieur d'une structure, il s'intéresse au flux d'information. Partant de cette remarque, nous avons optimisé la navigation dans le flux d'information en commençant à dessiner en premier les sommets présentant le plus fort Strahler. Ceci permet d'afficher en temps réel un squelette du graphe correspondant à son allure générale [933].

Le paramètre Strahler est également un des paramètres structurels des plus intéressants lorsqu'il s'agit de comparer deux structures quasi-équivalentes.

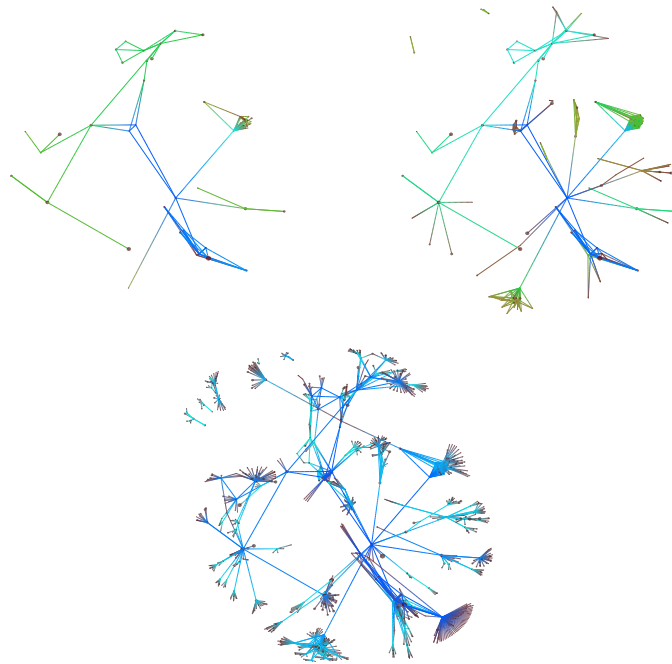


FIG. 6.1 Affichage incrémental de graphe.

Dessins de structures Les algorithmes de dessin de graphes sont un élément primordial en visualisation d'information. En effet, le dessin est l'élément qui permet d'appréhender

la structure et donc de pouvoir l'explorer et de mettre en évidence certaines propriétés. Par exemple, si on considère un même système de fichiers, l'utilisation d'un algorithme hiérarchique comme Reingold et Tilford souligne la structure hiérarchique du système de fichiers, alors qu'un algorithme comme le TreeMap souligne plus aisément l'importance des répertoires en considérant la taille de la sous arborescence. Un autre exemple où l'algorithme de visualisation est un élément prépondérant concerne le dessin des structures secondaires d'ARN. Dans ce domaine, les biologistes ont des habitudes particulières pour dessiner ces structures. Il faut donc que le dessin proposé soit fidèle aux habitudes des utilisateurs.

Une autre contrainte des algorithmes de dessin est leur rapidité. Pour des raisons d'interactivité, il faut être capable de dessiner de manière quasi temps-réel des structures de données dont le nombre d'éléments avoisine le million. Cette contrainte impose que les algorithmes de dessin aient une complexité maximale en $n \log(n)$ en taille et en mémoire pour être exploitable.

Concernant l'amélioration de la complexité des algorithmes de dessin, nous avons diminué la complexité en mémoire de l'algorithme de Sugiyama, la faisant passer de quadratique mn en linéaire $m + n$ en fonction du nombre de sommets n et du nombre d'arêtes m .

Dans le cadre de ce projet, nous avons développé deux nouveaux algorithmes de dessin, l'algorithme ballon pour la visualisation d'arbres (avec Sébastien Grivet) [949] et un algorithme spécifique au dessin de structures secondaires d'ARN (avec Serge Dulucq)[954].

L'algorithme ballon est un algorithme de dessin d'arbre qui préserve la planarité de la structure et dont la complexité est en $n \log(n)$. L'idée principale de cet algorithme est de dessiner de manière radiale la structure de données mais en utilisant le moins de place possible. Pour cela nous nous sommes inspirés de l'algorithme de Walker mais en calculant pour chaque nœud de l'arbre le plus petit cercle englobant qui permet le dessin sans croisement des arêtes. Le même procédé est ensuite appliqué récursivement. Cet algorithme a été comparé avec l'algorithme de Walker et l'algorithme de dessin radial en considérant deux critères esthétiques que sont : la distribution de la longueur des arêtes et la résolution. Les résultats ont montré que cet algorithme était le meilleur compromis maximisant ces deux critères.

Dans le cadre de la bio-informatique, un algorithme de dessin de structures secondaires d'ARN a été développé. Cet algorithme intègre les contraintes des dessins réalisés par les biologistes, par exemple, boucles dessinées sur des cercles, échelles dessinées en ligne droite, etc. Les premiers résultats obtenus suscitent l'intérêt des biologistes.

Interaction Les algorithmes de calcul exacts sur les graphes sont très fréquemment membres de la classe des problèmes NP. Il est donc souvent nécessaire d'utiliser dans les applications concrètes des heuristiques ou bien des approximations. Le problème majeur de ces méthodes est que les résultats ne sont pas exacts. D'autre part, la taille des données que nous cherchons à visualiser est gigantesque. De ce fait, même quand nous disposons d'algorithmes de calcul exacts, l'analyse des résultats produits n'est pas directe. Les deux points présentés ci-dessus montrent clairement que dans un système de visualisation d'informations, il est nécessaire de permettre à l'utilisateur une interaction directe avec les algorithmes et les métaphores visuelles mais aussi de lui permettre une modification manuelle des résultats produits. Cette dimension de la visualisation d'informations est une des préoccupations essentielles de nos recherches. Dans ce cadre nous avons proposé [939] une méthode permettant la navigation interactive dans les grands graphes. Cette méthode utilise les nombres de Strahler. Ces nombres nous permettent de construire une vue schématique de la structure de données que nous utilisons pour préserver un temps de réaction de l'ordre de 50ms pendant les interactions de l'utilisateur. En 2004 [941], nous avons proposé un algorithme de fragmentation

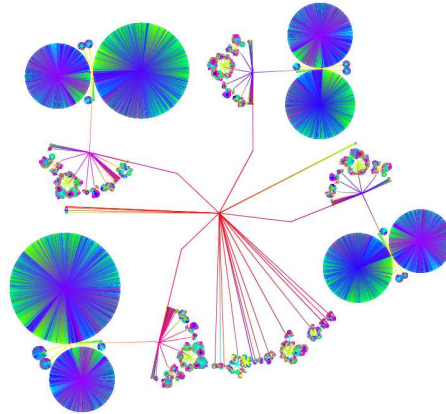


FIG. 6.2 Algorithme ballon pour un système de fichiers

de graphe. Cet algorithme calcule automatiquement une partition du graphe puis permet à l'utilisateur d'affiner interactivement son résultat.

Graphes Petits mondes Récemment, il a été montré que de nombreux graphes modélisant des données réelles issues de domaines aussi variés que la topologie de réseaux, les réseaux d'interaction protéines-protéines ou les réseaux sociaux partagent de nombreuses propriétés communes : densité globale faible, densité locale forte, petit diamètre, ... Les graphes possédant ces propriétés sont dits *petits-mondes*. Nous nous intéressons donc à des méthodes de construction ou de génération de telles familles de graphes ainsi qu'aux implications algorithmiques pour des applications sur le routage ou le partitionnement de graphes. En effet, il n'existe quasiment pas jusqu'à ce jour d'algorithmes dédiés aux graphes petits-mondes. Kleinberg a montré en 2000 qu'il est facile et rapide de router dans une famille de graphes petit-monde. Si un tel graphe possède ces propriétés, il est appelé graphe petit-monde *navigable*. La famille originale est constituée d'un graphe de base (ici une grille) et d'un petit ensemble d'arêtes construites de manière aléatoire (mais pas uniforme). En supposant que chaque sommet ne connaisse que ses voisins dans le graphe ainsi que la distance aux autres sommets dans le graphe de base, un routage glouton trouve sa cible en un nombre polylogarithmique de sauts.

Avec Philippe Duchon, nous avons étudié le routage dédié aux graphes petits-monde ainsi que des méthodes pour transformer un graphe quelconque en graphe petit-monde [936]. Nous avons montré que le résultat de Kleinberg est généralisable pour de nombreux graphes de base. En effet, au lieu de prendre la grille, nous pouvons prendre n'importe quel graphe pour lequel la croissance des boules est modérée (croissance polynomiale ou intermédiaire). Nous avons décrit la loi de distribution des arêtes aléatoires à rajouter en considérant que nous augmentons le graphe de base à l'aide d'un arc aléatoire par sommet.

Bases de données L'activité bases de données au sein du projet visualisation s'effectue autour de deux axes

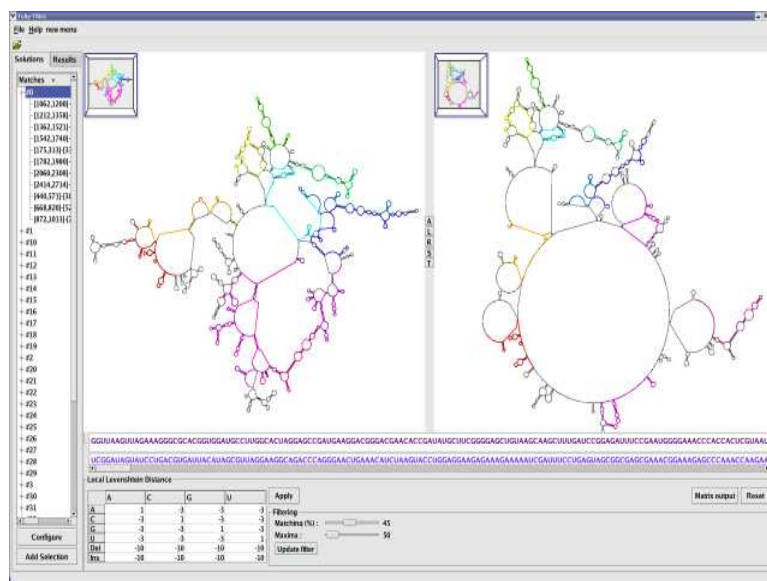


FIG. 6.3 Visualisation avec ARNA de structures secondaires d'ARN

- XML et contraintes d'intégrité,
- XML et norme de structuration de documents et de données.

Sous l'influence de chercheurs en bases de données, des questions ont été posées et des solutions ont été proposées quant à l'organisation "correcte" de ces documents. En effet, des formes normales, inspirées de ce qui a été fait dans les base de données relationnelles et objet, ont été proposées afin de réduire la redondance et pallier aux problèmes de mise à jour. Dans ce cadre, une thèse a commencé en novembre 2004 dont le sujet principal porte sur l'élaboration d'algorithmes efficaces pour l'extraction de dépendances fonctionnelles à partir de documents XML. Le résultat de ces algorithmes permettra de réorganiser les documents en entrée de façon plus optimale. Les cubes de données, ou datacubes, sont un cas particulier de bases de données structurés sous une forme multidimensionnelle (contrairement aux BD relationnelles où les données sont vues sous forme tabulaire, ie bi-dimensionnelle). Un datacube peut être représenté par une structure de treillis. Chaque noeud représente un tuple. L'ordre partiel définissant ce treillis est la relation "calculé à partir de". Ainsi $t1 < t2$ ssi $t1$ est calculé en utilisant $t2$. Nous nous sommes d'abord intéressé au calcul efficace de ce treillis en proposant un algorithme de faible complexité permettant de réaliser le produit de partition en un temps et un espace inférieur à ce qui est proposé dans l'état de l'art [952].

Logiciel Les algorithmes de visualisation d'informations que nous mettons en place requièrent une expérimentation pour être validés. Dans ce cadre, nous avons mis en place une plate-forme logicielle réutilisable appelée Tulip [953, 950]. Cette plate-forme intègre la plupart de nos résultats de recherche ainsi que de nombreux algorithmes d'autres équipes. Elle nous permet de faire des comparaisons entre les méthodes mais aussi de diffuser nos travaux à un large public. Sous licence GPL, ce logiciel fonctionne sous Linux et Windows XP. Les statistiques de fréquentation du site Web montrent clairement l'intérêt de la communauté scientifique pour ce genre d'outils. Depuis la création du site nous avons comptabilisé 43.000 connections et depuis le 27 avril 2005, date de mise en ligne de la version 2.0.2, 1.059 connections de 58 pays différents. Les proportions des différents pays sont représentées sur

le diagramme de la figure 6.4.



FIG. 6.4 Téléchargement du logiciel Tulip suivant les pays.

Fragmentation Dès que les graphes deviennent très gros (plusieurs centaines de milliers d'éléments), il est impossible de proposer à l'utilisateur d'explorer ses données. Ceci est dû d'une part à la taille de l'écran mais aussi à la capacité visuelle de l'utilisateur. Ainsi la fragmentation est essentielle au processus de visualisation. De nombreuses approches existent, notre groupe s'est focalisé sur l'utilisation de paramètres combinatoires. L'étude combinatoire sur le nombre de feuilles [934] d'un sous arbre de taille donnée en fonction de l'arité et de la longueur des segments a permis de développer un algorithme de fragmentation. En effet, un arbre aléatoire est un arbre très filiforme pour lequel la tendance asymptotique de ce paramètre est bien connue (loi normale). Les arbres issus d'applications réelles sont très différents et notamment d'arité faible et peu filiforme (arbre de syntaxe d'un compilateur, système de gestion de fichiers). Nous avons proposé un algorithme permettant de fragmenter efficacement un arbre. Ce résultat a été utilisé par Ivan Herman (CWI Amsterdam) pour rendre plus performante une interface de visualisation d'aide à la mise au point de compilateurs auto-parallélisants.

Avec Y. Chiricota (Université de Chicoutimi), nous avons utilisé des techniques de traitement du signal (convolution) pour fragmenter des cartes planaires pointées à partir de la valuation des sommets par des nombres de Strahler [941]. Une application est la possibilité de fragmenter un graphe d'inclusion de programmes (par exemple graphe des classes en C++) en obtenant comme fragments la structure du logiciel. Une autre application a été faite pour retrouver des séquences homogènes en couleur dans une vidéo [?].

La recherche de motifs identiques dans les arbres est un problème bien connu et de nombreux algorithmes existent. Dans la réalité, ce que l'on recherche dans un amas de données ce sont plutôt des structures qui se ressemblent sans être forcément totalement identiques. C'est le cas notamment dans les systèmes de gestion de fichiers. Nous avons développé une heuristique basée sur plusieurs paramètres (dont les nombres de Strahler) qui permet à l'utilisateur de découvrir visuellement des fragments de données que nous disons ϵ -isomorphes, l'utilisateur pouvant régler le seuil de tolérance. Cette heuristique a été utilisée avec succès sur le concours de la conférence Infovis'03 (2ième prix) [942] pour découvrir l'évolution d'un

système de gestion de fichiers. Plus récemment, nous l'avons utilisé en Bioinformatique pour le dessin de RNA afin de positionner les portions similaires à la même place et la même orientation sur le dessin [954]. Avec Y. Chiricota (Université de Chicoutimi), G. Mélançon et F. Jourdan (LIRMM), nous avons exploré la fragmentation de graphes *petit monde*[940]. Nous avons construit une métrique qui permet de détecter les arêtes les plus faibles. Ceci permet d'obtenir un algorithme efficace permettant à l'utilisateur de naviguer au sein de fragments significatifs. Nous avons montré son efficacité sur le graphe *d'Hollywood* (graphe des acteurs).

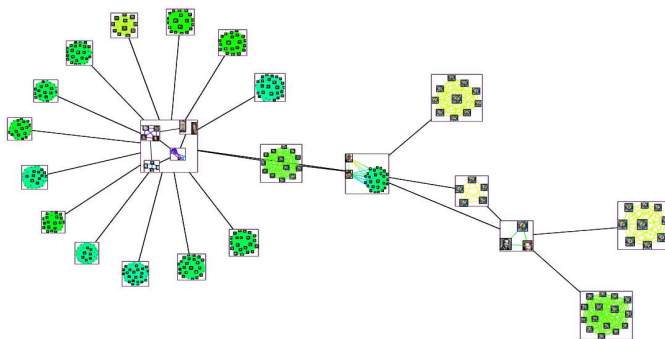


FIG. 6.5 Exemple de fragmentation sur le graphe Hollywood.

6.3.3 Projets et perspectives 2005-2008

Logiciel Nous prévoyons d'enrichir la plate-forme Tulip non seulement à travers les nouvelles méthodes que nous développons mais également en intégrant dans la distribution les contributions des utilisateurs externes au LaBRI. Ce travail n'a pu être réalisé jusque là faute d'ingénieur support.

Par ailleurs, nous allons dans les mois à venir constituer une documentation propre du logiciel ainsi qu'un tutorial. Ce tutorial nous servira lors des formations qui nous sont demandées par plusieurs groupes en France.

Dessins de structures Cet objectif rejoint une autre perspective de recherche actuellement développée en collaboration avec l'Université de British Columbia. Il s'agit de partitionner un graphe initial en sous-graphes possédant des caractéristiques structurelles particulières. Ces caractéristiques structurelles sont associées à un algorithme permettant de les dessiner au mieux. Chaque sous graphe est dessiné avec un algorithme adapté à sa structure et un processus centralisé contrôle le recollement des différents dessins.

Fragmentation de graphes L'observation d'un graphe dynamique passe par la reconnaissance de motifs quasi-similaires. C'est par exemple le cas lorsqu'on examine un ensemble

d'images à travers les graphes d'adjacence obtenu par segmentation. Nous souhaitons approfondir notre heuristique basée sur les Strahler :

- étudier l'interaction entre les paramètres intervenant dans le calcul des distances,
- faire une étude de complexité plus précise,
- examiner sa parallélisation.

Par ailleurs, nous l'exploiterons dans des graphes dynamiques issus de la chimie et/ou de la vidéo.

Structure de données et algorithmique efficace Nous espérons généraliser nos constructions de graphes petit-monde et répondre en partie à la question : peut-on transformer tout graphe en petit-monde avec une opération d'augmentation qui ne modifie que de très peu le graphe de base. Nous pensons ainsi utiliser les graphes petit-monde comme structure d'indexation et de navigation visuelle. Nous envisageons d'utiliser cette structure comme base pour naviguer dans les bases de données d'images ou vidéo. Le principe consiste à transposer le mécanisme de routage local en proposant à l'utilisateur un échantillonnage d'images qui permet de se rapprocher d'une image cible de manière quasi-certaine (l'utilisateur choisit une image dans l'échantillon qu'il juge comme intermédiaire pour sa cible) et d'accélérer ainsi le temps de navigation avec l'introduction de structures petit-monde.

Une partie de notre recherche portera sur la visualisation des cubes de données. Actuellement nous nous intéressons à l'extraction de connaissances à partir des datacubes. Cette représentation permet, en principe, d'extraire des connaissances en s'intéressant aux tendances qui y sont présentes. Ceci peut se faire en utilisant des algorithmes dits de datamining. L'approche que nous avons adoptée pour l'extraction de l'information consiste à utiliser les techniques de visualisation. Ce travail se fait dans le cadre de l'ACI jeunes chercheurs "visualisation de datacubes".

Interaction Nous travaillons actuellement sur de nouvelles méthodes d'interactions avec les métaphores visuelles. Ces travaux se font en collaboration avec le laboratoire CLIPS de Grenoble avec lequel nous avons déjà encadré deux DEA. L'objectif est de produire un système collaboratif pour la visualisation de grands graphes. En utilisant uniquement un vidéo projecteur et une caméra, ce système doit permettre à terme de manipuler, éditer et analyser des données à plusieurs utilisateurs. Une des applications envisagées est l'utilisation de cet outil de visualisation dans le cadre de développement logiciel. Une autre partie de nos recherches porte sur la mise au point de méthodes permettant l'édition manuelle des résultats obtenus par des algorithmes de fragmentations automatiques, cette édition manuelle permettra à l'utilisateur d'éditer efficacement les résultats produits par les heuristiques de calcul.

Adaptation aux champs d'application Un premier travail portera sur le dessin de RNA. Les biologistes associent des dessins de référence à certaines structures secondaires clairement identifiées. Une perspective d'amélioration consiste à intégrer dans l'algorithme les dessins de référence lorsqu'une structure secondaire répertoriée est reconnue. Les autres séquences étant dessinées automatiquement.

En ce qui concerne les travaux liés à la vidéo qui sont maintenant assez développés, une perspective consistera à permettre, à partir de la plate-forme Tulip, d'utiliser directement des graphes, dynamiques ou non, issus des bases de données vidéo.

Le groupe graphes et applications a mené des recherches sur les réseaux sémantiques, cartographiant les éléments d'un système d'information et leurs relations. Plus particulièrement,

ont été examinés des réseaux sémantiques codés à l'aide de la norme Topic Map, en coopération avec la société Mondeca (Paris). D'un point de vue mathématique, une topic map peut être vue comme un hypergraphe, les nœuds représentant les sujets d'information et les hyperarêtes leurs relations. La visualisation en liaison avec la norme Topic Map fournira un nouveau cadre applicatif.

6.3.4 Publications d'audience internationale

Revues avec Comité de lecture

- [933] AUBER (D). – Using strahler numbers for real time visual exploration of huge graphs. *International Journal of Applied Mathematics and Computer Science*, 2002, p. to appear.
- [934] AUBER (D) et DELEST (M). – A clustering algorithm for huge trees. *Advances in Applied Mathematics*, vol. 31, n1, 2003, pp. 46–60.
- [935] DELEST (M), DON (A) et BENOIS-PINEAU (J). – Dag based visual interfaces for navigation in index vidéo content. *Multimedia Tools and Applications*, 2005. – à paraître.
- [936] DUCHON (P), HANUSSE (N), LEBHAR (E) et SCHABANEL (N). – Could any graph be turned into a small world? *Theoretical Computer Science*, To appear. – Special Issue on Complex Networks.
- [937] IRAGNE (F), NIKOLSKI (M), MATHIEU (B), AUBER (D) et SHERMAN (D). – Proviz : protein interaction visualization and exploration. *Bioinformatics*, vol. 21, n2, 2005, pp. 272–274.

Colloques avec Comité de programme et Actes

- [938] AUBER (D). – Tulip. In : *9th Symp. Graph Drawing*, éd. par Mutzel (P), Jünger (M) et Leipert (S). pp. 335–337. – Springer-Verlag, 2001.
- [939] AUBER (D). – Using strahler numbers for real time visual exploration of huge graphs. In : *International Conference on Computer Vision and Graphics*, pp. 56–69. – 2002.
- [940] AUBER (D.), CHIRICOTA (Y), JOURDAN (F) et MELANCON (G). – Multiscale visualization of small world networks. In : *IEEE Symposium on Information Visualisation*, pp. 75–81. – 2003.
- [941] AUBER (D), DELEST (M) et CHIRICOTA (Y). – Strahler based graph clustering using convolution. In : *8th International Conference on Information Visualisation, IV*. pp. 44–51. – IEEE Computer Society, 2004.
- [942] AUBER (D), DELEST (M), DOMENGER (J), FERRARO (P) et STRANDH (R). – EVAT : Environment for visualization and analysis of trees. In : *IEEE Symposium on Information Visualisation Contest*, pp. 124–126. – 2003. Concours InfoVis - Deuxième place.
- [943] AUBER (D) et JOURDAN (F). – Interactive refinement of multi-scale network clusterings. *9th International Conference on Information Visualisation, IV*, 2005, p. to appear.
- [944] DELEST (M), DON (A) et BENOIS-PINEAU (J). – Graph-based visual interfaces for navigation in indexed video content. In : *CBMI*, éd. par Gabbouj (M). pp. 49–55. – Rennes, France, Septembre 2003.

- [945] DELEST (M), DON (A) et BENOIS-PINEAU (J). – Intuitive color-based visualization of multimedia content as large graphs. *In : EI'04*. SPIE, pp. 65–74. – 2004.
- [946] DELEST (M), LEBLANC (B), MARSHALL (M) et MELANCON (G). – Graph visualization and navigation as an interface to data exploration. *In : VSST'2001*, éd. par ESCORSA (P), pp. 279–284. – Barcelone, 2001.
- [947] DELEST (M), MUNZNER (T), AUBER (D) et DOMENGER (J). – Exploring infovis publication history with tulip. *In : InfoVis 2004*. IEEE. – 2004. Concours InfoVis - Deuxième place.
- [948] GAINANT (G) et AUBER (D). – Arna : Interactive comparison and alignment of rna secondary structure. *In : IEEE Symposium on Information Visualisation*. p. to appear. – IEEE Computer Society, 2004.
- [949] GRIVET (S), AUBER (D), DOMENGER (J) et MELANCON (G). – Bubble tree drawing algorithm. *In : International Conference on Computer Vision and Graphics*, p. to appear. – 2004.

Chapitres d'ouvrages

- [950] AUBER (D). – Tulip : A huge graph visualisation framework. *In : Graph Drawing Softwares*, éd. par Mutzel (P) et Jünger (M), pp. 105–126. – Springer-Verlag, 2003.

6.3.5 Publications d'audience nationale

Colloques avec Comité de programme et Actes

- [951] BBLANC (B. L), DION (D), AUBER (D) et MELANCON (G). – Constitution et visualisation de deux réseaux d'associations verbales. *In : Colloque Agents Logiciels, Coopération, Apprentissage et Activité Humaine (ALCAA)*, pp. 37–43. – 2001.
- [952] NOVELLI (N) et MAABOUT (S). – Algorithme efficace de calcul du produit de partitions et ses applications. *In : 19ème conférence Bases de Données Avancées (BDA '03)*, pp. pages 343–362. – 2003.

6.3.6 Autres publications

Notices d'emplois de logiciels

- [953] AUBER (D) et AL. – *Plate-forme Tulip*. – Rapport technique, LaBRI. www.tulip-software.org.

Rapports internes et autres publications

- [954] AUBER (D), DELEST (M), DOMENGER (J) et DULUCQ (S). – *Efficient drawing and comparison of RNA secondary structure*. – Rapport technique nRR-1325-04, LaBRI, 2004.

Langage Spidle, Développement d'applications de streaming, architecture client-serveur programmable, personnalisation de serveurs en fonction des utilisateurs.

Langage SPL, Conception et développement d'une machine virtuelle SIP, architecture logicielle pour le déploiement de services de téléphonie.

Tempo, Plateforme de spécialisation automatique pour divers systèmes Unix, développement d'un évaluateur partiel pour les programmes écrits en C.

Madeleine, Bibliothèque de communication multi-grappes qui implémente efficacement un concept de canal de communication pouvant être soit physique (configurations homogènes) soit virtuel (configurations hétérogènes).

Marcel, Bibliothèque de threads utilisateurs s'adaptant à la plateforme visée (architectures mono-processeur, SMP, NUMA) et intégrant un serveur d'évènements améliorant le temps de réaction des applications aux entrées/sorties.

LinuxActivations, Extension du noyau Linux permettant de contrôler l'ordonnancement des threads utilisateurs lorsqu'ils effectuent des opérations d'entrées/sorties bloquantes dans le noyau.

MPICH/MADIII, Implémentation multithreadée et multi-protocole du standard de communication par passage de messages MPI.

PadicoTM, Déploiement d'intergiciels sur les réseaux des grilles de calcul.

PM2, Support exécutif générique de bas niveau qui intègre la gestion des threads, la gestion des communications et une gestion mémoire distribuée élémentaire.

EPSN, Plateforme couplée de pilotage interactif et de visualisation pour des simulations numériques parallèles et distribuées.

PaStiX, Bibliothèque de solveurs parallèles haute performance basés sur des factorisations complètes et incomplètes pour des très grands systèmes linéaires creux issus de problèmes 3D.

PHIDAL, Bibliothèque de solveurs parallèles itératifs basés sur des décompositions hiérarchiques et des factorisations multiniveaux.

Scotch, Bibliothèque pour le partitionnement et le placement de graphes et de maillages irréguliers de grande taille.

Jem, Plateforme de manipulation de ressources connectées à un réseau fixe.

JToe, Bibliothèque Java de transfert efficace d'objets.

Mandala, Plateforme pour le développement d'applications Java distribuées intégrant la notion de conteneur et d'appel de méthodes distantes asynchrones.

6.3.7 Formation par la recherche

Thèses

- [955] AUBER (D). – *Outils de visualisation de larges structures de données*. – Thèse de PhD, University Bordeaux I, December 2002.