Rapport INRIA 1994 — Programme 2

Algorithmes

PROJET ALGO

3 mai 1995

PROJET ALGO

Algorithmes

Localisation : *Rocquencourt*

Mots-clés : Analyse d'algorithmes, Algorithmique, Analyse combinatoire et asymptotique, Évaluation de performances, Analyse automatique, Calcul formel, Arbres, Tri et recherche, Protocoles de communication et réseaux, Algorithmique des systèmes informatiques, Algorithmique géométrique, Algorithmique de la théorie des nombres

1 Composition de l'équipe

Responsable scientifique

Philippe Flajolet, directeur de recherche, Inria

Responsable permanent

Mireille Régnier, directeur de recherche, Inria

Secrétariat

Virginie Collette

Personnel Inria

Philippe Jacquet, détachement corps des Mines
Philippe Robert, directeur de recherche, à partir du 1er septembre 1994
Bruno Salvy, chargé de recherche

Chercheurs Extérieurs

Henry Crapo, université de Montréal
Philippe Dumas, Prof. Cl. Prépa. lycée Jean-Baptiste Say
Danièle Gardy, Prof., université de Versailles Saint-Quentin

Dominique Gouyou-Beauchamps, Prof., université Paris-Sud
Michèle Soria, Prof., université Paris 6
Jean-Marc Steyaert, M. de Conf., École polytechnique

Chercheurs doctorants

Frédéric Chyzak, boursier de l'École polytechnique
Vincent Dumas, boursier de l'École polytechnique, à partir du 1er septembre 1994
Xavier Gourdon, boursier de l'École polytechnique
Pierre Nicodème, société Copernique et Inria

Chercheur en mise à disposition

François Morain, Inria (40%) et École polytechnique (60%)

Stagiaires

Jacques Carette, Waterloo Maple Software et boursier du Gouvernement Canadien
Hsien-Kuei Hwang, École polytechnique et Inria
Eithne Murray, Inria, à partir du 1er mars 1994

2 Présentation du projet

L'objectif général du projet est la conception, l'évaluation et l'optimisation des algorithmes et structures de données fondamentaux de l'informatique. Dans ce contexte, l'activité principale se situe dans la construction de modèles analytiques pour l'évaluation de performances et l'optimisation fine ; c'est dans ce domaine, par nature fortement mathématisé, que se place l'essentiel des recherches du projet.

La ligne scientifique suivie part d'une théorie de larges classes de "processus combinatoires et algorithmiques" susceptibles de donner lieu au développement de modèles mathématiques qui permettent des prédictions très précises des fonctions de coût (en nombre d'opérations, en temps, en espace mémoire) d'algorithmes fondamentaux, soit en moyenne, soit en distribution. C'est là la version moderne du domaine de *l'analyse d'algorithmes* fondé par Knuth il y a une vingtaine d'années. Une approche unitaire fondée sur des développements récents en analyse combinatoire conjugués à des méthodes d'analyse asymptotique complexe a permis ainsi des caractérisations très précises (et large-

ment indépendantes de la technologie) de nombreux algorithmes dont l'évaluation résistait intrinsèquement à des techniques élémentaires.

L'approche suivie débouche alors naturellement sur une théorie globale de l'analyse en moyenne, voire en probabilités, de nombreux processus combinatoires. Cette direction de recherche est l'une des originalités du projet au plan international. Son caractère très systématique a conduit, il y a quelques années, à envisager notamment le développement d'un programme automatique d'aide à l'analyse d'algorithmes. Il s'agit du programme $\Lambda\Upsilon\Omega$ ("Lambda-Upsilon-Omega") qui repose de manière essentielle sur les fonctionnalités offertes par les systèmes de calcul formel contemporains. Le programme $\Lambda\Upsilon\Omega$ est constitué de grosses bibliothèques en calcul formel (20 000 lignes de programmes en MAPLE environ). Son développement est à l'origine d'un ensemble d'activités de portée générale en calcul formel.

Il se dégage en effet de ces travaux une très importante démarche applicable à plusieurs champs d'application dans l'*analyse de systèmes complexes* : une description formelle d'un système complexe est prise en compte par un système de calcul formel qui calcule automatiquement les équations mathématiques en décrivant le comportement du système. Ces équations du modèle peuvent alors être traitées automatiquement, avec l'aide de bibliothèques spécialisées, par le système de calcul formel qui en extrait les grandes caractéristiques interprétables par l'utilisateur.

Enfin, le développement des méthodes d'analyse, hors sa logique interne, suit de près l'évolution des techniques algorithmiques et des grands champs d'application. C'est ainsi que dans les années récentes ont été abordées les grandes classes de problèmes algorithmiques suivantes :

- Les arbres et les structures de données pour la recherche efficace d'informations ainsi que l'algorithmique générale des systèmes informatiques. Des caractérisations complètes permettant une utilisation quasi-optimale des ressources ont été obtenues pour les problèmes de recherche multidimensionnelle en mémoire centrale ou secondaire d'ordinateur (méthodes de hachage extensible, méthodes de grille, arbres d'index, arbres quadrants, arbres digitaux), certains de ces problèmes intéressant les bases de données.
- L'organisation de la communication sur un canal à accès multiples (protocoles pour réseaux locaux), et plus récemment réseaux mobiles radio. L'algorithmique de base est ici fondamentalement

probabiliste et les performances dépendent de manière cruciale d'optimisations fines dictées par des analyses de haute technicité.

- L'algorithmique des mots et séquences qui se relie naturellement aux activités récentes sur le génome humain.
- L'algorithmique en arithmétique et en théorie des nombres. C'est dans ce domaine que se situent des fonctionnalités nouvelles en cryptographie (authentification, systèmes à clefs publiques, par exemple), et le projet détient toujours le record du monde dans le domaine des tests de primalité (Atkin-Morain).

Contexte des recherches du projet Les recherches du projet sont menées tout d'abord en collaboration très étroite avec l'École polytechnique (J.-M. Steyaert, H-K. Hwang, J. van der Hoeven), l'université Paris-Sud à Orsay (D. Gouyou-Beauchamps, et l'équipe d'algorithmique), l'université de Versailles (D. Gardy nommée professeur en 1993) et l'université Paris 6 (M. Soria). Un séminaire commun se tient à l'Inria sur une base très régulière. En matière de chercheurs extérieurs, ce rapport présente de manière prépondérante les contributions de J.-M. Steyaert et M. Soria, lesquels sont présents à environ un quart de temps sur le site.

Diverses coopérations existent avec plusieurs projets au sein de l'Inria-Rocquencourt sur les thèmes de recherche intéressant le groupe ALGORITHMES. L'on peut citer le groupe de modélisation et d'évaluation de performances MEVAL (G. Fayolle, Ph. Robert : méthodes mathématiques), le groupe de systèmes temps réel REFLECS (G. Le Lann : protocoles de communication), le projet CODES (P. Camion et P. Charpin : cryptographie, calcul formel), les projet RODIN et VERSO (collaboration de M. Régnier avec É. Simon, et activités liées au GÉNOME) ainsi que le groupe META2 (J.-P. Quadrat, Cl. Gomez : calcul formel et automatique). Également, de nombreux échanges ont lieu avec plusieurs projets du Programme 2 (COQ, CAML, ICSLA, PARA) avec lesquels sont en particulier partagées des ressources informatiques.

D'autre part, après le départ de Paul Zimmermann à l'Inria-Nancy (Projet EURECA), la coopération se poursuit notamment en ce qui concerne l'analyse automatique d'algorithmes et la génération aléatoire.

Enfin, plusieurs travaux sont développés en collaboration avec des Universités et centres de recherche étrangers : AT&T Bell Laboratories à

Murray Hill, l'université technique de Vienne, Purdue University, Princeton University, etc. L'équipe, avec ses partenaires parisiens susnommés, fait partie d'un projet Esprit Basic Research intitulé ALCOM II et, avec ses partenaires Inria CODES et REFLECS, du projet ESPRIT III LAURA.

3 Action de recherche

3.1 Analyse d'algorithmes

Participants : Philippe Dumas, Philippe Flajolet, Xavier Gourdon, Hsien-Kuei Hwang, Philippe Jacquet, Jean-Marc Steyaert, Michèle Soria

3.1.1 Méthodologie de l'analyse d'algorithmes

L'orientation principale est la quête de cadres toujours plus généraux et plus puissants pour l'analyse de *schémas* d'algorithmes opérant sur les structures combinatoires fondamentales (mots, arbres, graphes notamment). Les recherches concernent l'analyse combinatoire, outil de base pour les dénombrements, et l'analyse asymptotique ici fondée sur l'analyse du comportement dans le plan complexe des séries génératrices.

L'idée essentielle est qu'il est possible de définir un langage de spécification de structures combinatoires "constructibles" ou "décomposables" pour lequel existent des règles systématiques de traduction en séries génératrices. La puissance expressive de ce langage est grande : celui-ci couvre les langages réguliers et automates finis, les langages *context-free* et les arbres de termes ou de syntaxe, mais aussi les fonctions aléatoires de l'analyse cryptographique, les paramètres fondamentaux liés aux statistiques d'ordre comme par exemple l'algorithme *Quicksort* et diverses variétés d'arbre de recherche. Un ensemble cohérent de méthodes d'analyse complexe peut alors être mis en place pour extraire les informations de ces séries génératrices qui sont pertinentes à l'analyse en moyenne d'algorithmes.

Ces travaux fournissent de surcroît le cadre dans lequel peut se développer l'analyse automatique d'algorithmes. Ils se retrouvent encore en amont de recherches plus appliquées développées dans le projet et concernant les arbres, les chaînes, les bases de données, les protocoles de communication et le calcul formel. Ph. Flajolet a dégagé le cadre d'une telle démarche (qui constitue un programme de recherche qui s'étale sur plusieurs années) dans divers articles de portée générale. Une partie

de ce programme est menée en étroite collaboration avec R. Sedgewick (Princeton), et deux volumes sont en préparation sur ce thème.

3.1.2 Schémas et distributions

La théorie des schémas combinatoires et algorithmiques décrite ci-dessus s'applique aux propriétés énumératives des structures ainsi qu'aux évaluations en moyenne. Il découle en fait de travaux de D. Gardy, M. Soria, et Ph. Flajolet qu'il est possible de traiter à un niveau comparable de généralité les propriétés probabilistes de larges classes de paramètres. Cet axe de recherche établit ainsi un pont entre l'analyse d'algorithmes en moyenne et les analyses probabilistes qui est d'une grande généralité.Il apparaît notamment qu'un grand nombre de schémas combinatoires transverses conduisent à des phénomènes probabilistes communs : lois gaussiennes très souvent dotées de queues exponentielles, ce qui indique que les probabilités de forte déviation par rapport à la moyenne sont extrêmement faibles. Ainsi une même loi limite accompagnée d'une loi de grande déviation recouvre des situations aussi diverses que l'algorithme de Pollard en factorisation d'entiers, le comportement itératif des fonctions aléatoires en cryptographie, la factorisation des polynômes en calcul formel, et les statistiques d'ordre dans les arbres.

C'est dans ce cadre, que M. Soria a obtenu, en collaboration avec M. Drmota, une classification de schémas "produit" [11]. H.-K. Hwang a précisé les taux de convergence dans les théorèmes limites de Flajolet-Soria. Ses travaux portent sur la forme des développements asymptotiques des théorèmes centraux et locaux limites [20, 36] et sur les grandes déviations dans les cas des fonctions de répartition et des densités. Ces résultats sont intéressants car ils quantifient précisément les occurrences d'événements rares dans des processus combinatoires très généraux.

Dans le cadre de sa thèse, X. Gourdon travaille sur une théorie générale portant sur la taille des grandes composantes dans les structures combinatoires. Il s'agit d'étudier quantitativement des schémas recouvrant des situations pratiques aussi diverses que : propagation de retenues dans certains circuits d'addition asynchrone, apparition de régularités statistiques attendues dans des chaînes d'ADN, décryptage de système cryptographique par itération de la fonction de codage, etc. Un cas particulier a déjà donné lieu a une collaboration avec H. Prodinger (Technical University of Vienna) [43].

3.1.3 Algorithmes diviser-pour-régner

Un des schémas les plus classiques de résolution algorithmique efficace en informatique est le schéma appelé "diviser-pour-régner", lequel correspond à un partage récursif suivi d'une résolution indépendante en sous-problèmes, puis d'une recombinaison des résultats partiels. Ce schéma puissant trouve des applications constantes en algorithmique du tri et de la recherche ainsi qu'en géométrie algorithmique par exemple. Les fonctions de coûts de tels algorithmes s'expriment alors par des récurrences dites aussi "diviser-pour-régner" (*divide-and-conquer*, c'est-à-dire *divide ut imperes*). Celles-ci constituent un des chapitres de l'analyse d'algorithmes, classique, quoique traité universellement de manière superficielle. Les analyses, au delà d'estimations grossières, recouvrent en fait un domaine fort riche de structure et délicat. Des résultats récents dus aux membres du projet montrent pour la première fois la possibilité de prédictions quantitatives précises s'exprimant au moyen de fonctions fractales. Ces techniques ont été mises au point par Ph. Flajolet, M. Golin, Ph. Dumas et divers collaborateurs viennois [12, 13, 28]. H.-K. Hwang [44] a raffiné l'analyse en moyenne du tri fusion effectuée par Ph. Flajolet et M. Golin en donnant des formules exactes. Ces résultats sont les meilleurs possibles.

La technique de transformation de Mellin est d'un emploi délicat et encore peu maîtrisé dans le domaine. Ph. Flajolet, X. Gourdon et Ph. Dumas ont rassemblé les principes de cette technologie et de son application à l'analyse d'algorithmes dans [29].

3.1.4 Arbres et structures de données

Il a été dit souvent que la structure d'arbre est sans doute la plus importante de l'informatique. Elle permet une séparation des données selon divers critères permettant ensuite un accès très rapide.

Les travaux ont été principalement axés sur la recherche multidimensionnelle, les graphes fonctionnels [42], utiles dans de nombreux problèmes de théorie algorithmique des nombres et donc en cryptographie moderne, et les structures de données dynamiques.

La structure de tas (*heap*) est une structure fondamentale de l'algorithmique, permettant notamment la gestion efficace des files de priorités. Dans [35], H.-K. Hwang et J.-M. Steyaert ont résolu une classe d'équations de récurrences liées aux tas. En particulier, ils donnent une formule

asymptotique pour le nombre de tas de grande taille, pour le coût moyen et la variance de la construction d'un tas, et montrent que la loi limite du coût est asymptotiquement gaussienne.

Dans le domaine de la recherche multidimensionnelle, un effort tout particulier a porté sur l'analyse fine des arbres quadrants (*quadtrees*). Cette structure de données est présente sous des avatars diverses dans tous les domaines qui manipulent des structures "spaciales" et deux livres entiers de Samet y sont consacrés. Elles trouvent des applications en compactification d'images, gestion de bases de données géographiques, ou requêtes complexes portant sur des données multidimensionnelles. Le problème de base d'évaluation de ces structures, vieux de près de 20 ans et posé par leurs inventeurs (Bentley et Finkel 1974), a été finalement résolu par Ph. Flajolet en collaboration avec C. Puech (Grenoble), G. Gonnet (Zurich) et M. Robson (Canberra). Ces travaux ont ensuite été prolongés par Ph. Flajolet et T. Lafforgue [14] qui ont même pu fournir une analyse distributionnelle des coûts. Tout récemment, Ph. Flajolet et B. Salvy en collaboration avec une équipe de l'Uqam à Montréal (L. Laforest, G. Labelle) ont enfin établi l'existence de formules explicites de forme hypergéométrique pour ces questions [30]. Ceci permet notamment de quantifier très précisément les taux d'occupation mémoire dans un contexte paginé. Il résulte donc de tous ces travaux que le comportement des arbres quadrants peut être considéré comme pleinement résolu.

L'analyse dynamique de l'évolution de structures de données est un domaine classique de l'activité du projet. Les résultats obtenus par D. Gardy et G. Louchard (Bruxelles) permettent de caractériser l'évolution de la taille d'une relation au cours d'une suite d'insertions et de suppressions [33, 34, 16].

En collaboration avec W. Szpankowski, Ph. Jacquet a étudié les arbres digitaux de recherche. L'arbre digital est l'outil fondamental pour l'algorithme de compression de Lempel et Ziv. Ph. Jacquet et W. Szpankowski ont mis en évidence des lois limites gaussienne sur le comportement asymptotique des arbres digitaux. Cette étude a permis de résoudre et d'élargir un problème ouvert par Aldous et Shields au sujet de l'évaluation de la loi limite de la compression sur des suites de tirages de bits indépendants.

L'étude s'appuie sur une évaluation asymptotique de la solution d'une équation différentielle non linéaire aux différences, à deux variables, du

type :

$$\frac{\partial}{\partial z} h(z,u) = h(p_1 z, u) \times \cdots \times h(p_V z, u) \ .$$

Il a ainsi été montré que la loi normale asymptotique s'appliquait à la fois pour la longueur de cheminement et pour le nombre des phrases compressées dans le texte. Seuls les ordres de croissance des moyennes et variances sont différents. Par ce résultat, on obtient une analyse très fine de la capacité de compression de l'algorithme Lempel et Ziv. La méthodologie développée s'applique aussi à des équations différentielles d'ordre supérieur. Elle permet aussi de résoudre le problème de la compression Lempel et Ziv lorsque le modèle probabiliste des données est de nature plus générales (lois de Markov). Ces nouveaux résultats donneront lieu à de nouvelles publications.

3.2 Calcul formel

Participants : Jacques Carette, Frédéric Chyzak, Philippe Flajolet, Xavier Gourdon, François Morain, Eithne Murray, Bruno Salvy

Le calcul scientifique, traditionnellement orienté vers le calcul numérique, accorde une place de plus en plus importante au calcul formel, c'est-à-dire à des logiciels permettant le calcul de *formules* mathématiques, et non plus simplement de nombres. Des systèmes comme MACSYMA, REDUCE ou MAPLE sont utilisés depuis longtemps par de nombreux mathématiciens, physiciens ou chimistes, et la publicité importante qui accompagne le système MATHEMATICA depuis cinq ans entraîne une augmentation rapide de la variété des applications du calcul formel en sciences de l'ingénieur. L'annonce récente par IBM et NAG de la commercialisation du système AXIOM montre que ce domaine a atteint maintenant une dimension industrielle.

Le rôle du calcul formel au sein du projet se situe à plusieurs niveaux. Tout d'abord au niveau du développement de programmes symboliques, ensuite au niveau de la recherche d'algorithmes du calcul formel et enfin au niveau de la diffusion.

3.2.1 Développement de programmes symboliques

Analyse automatique d'algorithmes. Les méthodes développées dans le projet pour l'analyse d'algorithmes et de structures de données, ainsi que les techniques d'analyse complexe utilisées pour obtenir des

estimations asymptotiques sont suffisamment générales pour se prêter à une implémentation dans un système de calcul formel. Une telle implémentation présente le double intérêt de faciliter l'expérimentation sur des cas concrets et de diriger l'attention vers des points de la théorie encore insuffisamment étudiés. L'activité du projet sur ce sujet se situe donc à la fois dans le domaine de la programmation et dans une recherche plus théorique sur les algorithmes du calcul formel nécessaires à l'automatisation de l'analyse d'algorithmes et de structures combinatoires.

L'analyse automatique d'algorithmes se décompose en deux étapes principales : l'analyse combinatoire et l'analyse asymptotique.

La première partie, d'analyse combinatoire, traduit des spécifications de types de données ou de manière équivalente de structures combinatoires en équations sur les fonctions génératrices de dénombrement. Les types de données analysés font partie d'un ensemble bien défini de types récursivement descriptibles en termes d'un jeu fixe de constructeurs combinatoires. Par ailleurs, tout algorithme de *descente* opérant sur ces structures combinatoires est traduit dans cette phase en équations sur les fonctions génératrices de complexité des différentes procédures. On peut montrer que tout algorithme écrit dans le langage de description d'algorithmes ADL admet une telle traduction.

La seconde étape de l'analyse est une étape d'analyse asymptotique. L'information asymptotique est à rechercher dans le comportement des fonctions génératrices au voisinage de leurs singularités. Il découle de travaux antérieurs que pour une grande classe de fonctions élémentaires, la recherche des singularités et de leur développement singulier peut s'obtenir de manière automatique.

Le système $\Lambda\Upsilon\Omega$ démarré il y a 6 ans implémente ce programme. Une première phase, écrite en CAML par P. Zimmermann effectue la traduction du langage ADL vers des équations de fonctions génératrices. Ceci représente à peu près 110 ko de sources. Ces équations sont ensuite confiées au système de calcul formel MAPLE que l'on aide si possible à résoudre. Une fois des formes explicites obtenues, la deuxième phase du système effectue la recherche de singularités et les développements singuliers nécessaires, avant de les traduire en une information asymptotique sur le problème de départ. Cette deuxième partie, écrite par B. Salvy, constitue le package GDEV de la share library de MAPLE qui consiste en environ 500 ko de sources (plus une partie qui est intégrée aux ver-

sions récentes de MAPLE). Une partie des fonctionnalités de la première a été traduite cette année en MAPLE par P. Zimmermann. L'objectif à moyen terme est de produire un $\Lambda\Upsilon\Omega$ entièrement en MAPLE. Ainsi libéré des difficultés de portage sur différentes architectures, et en particulier sur Macintosh ou PC, le système deviendrait plus facilement accessible, notamment en enseignement.

Génération aléatoire. Pour des programmes que leur structure rend impropre à l'analyse par $\Lambda\Upsilon\Omega$, l'analyse peut être éclairée par des *simulations*. De telles simulations nécessitent en premier lieu un bon générateur aléatoire d'entrées du programme. Le package GAIA, écrit en MAPLE par P. Zimmermann suite à des travaux menés avec Ph. Flajolet et B. Van Cutsem (Imag), permet de tirer aléatoirement des structures combinatoires de la classe de celles qui sont acceptées par $\Lambda\Upsilon\Omega$. Plus précisément, ce programme prend en entrée une description de structure combinatoire et donne en sortie une procédure de génération aléatoire uniforme d'objets de cette structure. Cette procédure produit un objet de taille n en $O(n \log n)$ opérations arithmétiques dans le cas le pire.

Cette année, un nouveau package COMBSTRUCT a été développé en MAPLE par E. Murray. Il intègre une version robuste de l'ancien GAIA dans un ensemble plus général qui vise à unifier le traitement des structures combinatoires décomposables en calcul formel. Dans le futur, les fonctionnalités de $\Lambda\Upsilon\Omega$ seront intégrées à COMBSTRUCT, qui fournira un environnement complet de manipulation et d'expérimentation pour la manipulation d'une classe générale et importante de modèles combinatoires.

Le package GFUN**.** Dans l'état actuel du système $\Lambda\Upsilon\Omega$, les équations calculées par analyse combinatoire doivent être résolues explicitement avant d'être fournies à l'analyse asymptotique. Ceci limite l'analyse automatique aux algorithmes conduisant à des séries génératrices ayant une forme explicite à l'aide des fonctions usuelles. Or bien souvent, l'accès à l'information est plus facile sur une équation que sur une de ses solutions. Ces considérations motivent l'intérêt pour un traitement direct d'équations de fonctions génératrices appartenant à une classe assez générale. Une classe de plus en plus importante en calcul formel combinatoire est la classe des *fonctions holonomes*, c'est-à-dire en gros solutions d'équations différentielles linéaires à coefficients polynomiaux. B. Salvy et P. Zimmermann ont écrit sur ce thème le système GFUN [22].

Ce système a deux facettes. Un premier usage du système consiste à manipuler des équations différentielles linéaires et des récurrences linéaires. Ceci permet de prouver un grand nombre d'identités combinatoires ou de fonctions spéciales. Le second aspect du système est sa capacité à *deviner* des récurrences satisfaites par des suites ou des équations différentielles satisfaites par des séries. Les conjectures fournies par le système rendent un grand service pour guider l'intuition combinatoire et en font un outil puissant utilisé aujourd'hui par de nombreux combinatoriciens. Ce succès a motivé une récriture du package en vue d'en améliorer la robustesse et la maintenabilité. Par ailleurs, GFUN a été intégré à un serveur installé au BELL-LABORATORIES par N. Sloane. Par simple courrier électronique soumettant une suite de nombres, ce serveur répond par courrier électronique, indiquant si cette suite figure au *Handbook of Integer Sequences* et si cette suite ou des suites transformées satisfont une récurrence holonome.

Toujours dans cette orbite d'opérateurs linéaires différentiels ou aux différences, et à la suite de discussions avec le projet européen CATHODE, B. Salvy a également écrit un programme MAPLE permettant de manipuler les polynômes de Ore. À l'avenir, ce package devrait être utilisé pour récrire une partie de GFUN.

En outre, F. Chyzak [27] a commencé à étudier les opérateurs de Ore en plusieurs variables, ouvrant ainsi la voie au traitement automatique de suites de polynômes ou plus généralement de suites ou de fonctions à plusieurs variables satisfaisant des systèmes linéaires d'équations différentielles ou de récurrences. Le package réalisé permet la preuve automatique d'un grand nombre d'identités intervenant en combinatoire ou dans l'étude des fonctions spéciales.

Résolution de polynômes. À partir d'un algorithme dû à Schönhage, X. Gourdon a poursuivi le développement d'un programme MAPLE permettant d'approcher les racines d'un polynôme complexe. Son programme est d'une robustesse totale, ce qui n'est pas le cas des procédures réalisant la même tâche au sein des différents systèmes de calcul formel, au prix, bien sûr, d'un coût de calcul plus élevé. Malgré cela, la garantie des résultats numériques donnés permet de pallier le coût encore plus lourd de certaines opérations purement algébriques. Au cours de cette année, ce programme a été combiné à des procédures du noyau du système MAPLE, réalisant ainsi un programme à double niveau qui

permet d'avoir des temps de calcul très courts sur des polynômes "faciles" tout en garantissant les résultats. Cette fonctionnalité s'intégrera à terme dans le système $\Lambda\Upsilon\Omega$. X. Gourdon a également généralisé les techniques de Schönhage aux fonctions analytiques. Dans ce même cadre, une collaboration avec J.-Cl. Yakoubsohn (Laboratoire d'Analyse Numérique, Toulouse) a permis de développer une nouvelle technique pour approcher les modules extrêmes des racines d'un polynôme. Enfin, une collaboration est en cours entre X. Gourdon et J. Von zur Gathen (Univ. Toronto) sur l'analyse et l'optimisation d'algorithmes de factorisation de polynômes.

3.2.2 Algorithmique du calcul formel

Le système $\Lambda\Upsilon\Omega$ est un système généraliste. Il fournit un ensemble de fonctionnalités qui permettent de traiter routinièrement des cas courants. Cependant, une limitation du système réside dans la nécessité d'obtenir des formes closes pour les fonctions génératrices. Une direction de recherche consiste à déterminer directement à partir des équations satisfaites par les fonctions génératrices leur comportement asymptotique. Plusieurs types d'équations ont ainsi été étudiées cette année : équations différentielles [41], et certaines fonctions définies implicitement [23]. Ces recherches s'insèrent dans une problématique plus générale : il s'agit, dans le cas de l'*analyse de systèmes complexes*, de développer la capacité de traiter automatiquement des *équations de modèles implicites*.

3.3 Algorithmique

3.3.1 Protocoles de communication

Participant : Philippe Jacquet

Les activités de Philippe Jacquet dans le cadre du Projet concernent la communication informatique mobile et les réseaux radio. Une situation typique d'application est celle de robots mobiles dans un site de production. Ces communications sont plus exigeantes en terme de performances du réseau (débit, promptitude d'accès) que le téléphone cellulaire. La brique essentielle d'un tel système de communication est le réseau numérique local par radio. Au premier abord, un tel objet paraît simple à concevoir. Pourtant si cette brique est restée absente jusqu'à mainte-

nant du domaine des réalisations concrêtes, c'est parce qu'elle pose des problèmes ardus.

Le premier problème est d'ordre matériel. L'onde radio subit des dédoublements sur les obstacles qu'elle rencontre, ce qui en général altère fortement l'information qu'elle transporte quand le débit dépasse les millions de bits par seconde. Pour lutter contre ce phénomène, on introduit des redondances massives dans le code transporté ce qui nécessite le recours à des composants très puissants (aspects développés par le Projet Codes à l'INRIA).

Le deuxième problème est d'ordre algorithmique et concerne la communication des stations en réseau. Le problème est plus complexe que pour les réseaux locaux filaires. En effet le réseau est ouvert et mobile, le médium de communication est sujet à des perturbations constantes. Ces contraintes imposent la définition de protocoles d'accès d'un type particulier qui feront appel à des algorithmes nouveaux tout en respectant les interfaces et les logiciels de communications existant.

Le troisième problème concerne la normalisation. Le spectre électromagnétique est une ressource très demandée. L'attribution des fréquences est un processus arbitré aux niveaux européens et internationaux avec les implications politiques que l'on devine. Comme il est difficile d'imaginer plusieurs types de réseaux sans fil partageant le même milieu de communication, il est indispensable d'adopter l'approche la plus concertée possible.

Dans le cadre du programme ESPRIT III, la Communauté Européenne soutient et finance le projet de réseau local sans fil à haut débit LAURA (*cf*, Actions Industrielles). Le produit visé est un réseau sans fil comparable aux réseaux de la norme Ethernet (IEEE 802.3), avec un débit utile d'environ 10 millions de bits par seconde, en vue de raccorder les équipements compatibles avec Ethernet.

La partie recherche du projet est partagée entre deux compétences : le domaine de la radio et le domaine des réseaux. Les universités britanniques partenaires ont une solide expérience de la radio, l'INRIA a de nombreux résultats sur les réseaux et sur les protocoles à son actif. Dans l'institut, trois projets sont impliqués : le projet Algo, le projet Codes et le projet Reflecs.

Le projet Codes conçoit les codes détecteurs et correcteurs d'erreurs pour la transmission de données par radio. Ce travail est crucial car

l'application envisagée entraîne des taux d'erreurs bruts pratiquement incompressibles. Au contraire de nombreuses autres applications, augmenter la puissance de l'émetteur ne peut qu'aggraver la qualité de la transmission radio (l'onde est dédoublée sur un plus grand nombre d'obstacles à portée).

Le projet Reflecs définit les règles du nouveau protocole d'accès d'après les contraintes déjà citées. Le caractère mouvant de la propagation radio nécessite de faire appel à des solutions à la fois flexibles et robustes, comme par exemple les protocoles à compétition et à détection de collision, et les routeurs à apprentissage dynamique implantés dans chaque station.

Le projet Algo travaille à l'élaboration et l'optimisation des nouveaux algorithmes. On a vu que les contraintes des réseaux par radio induisent une déperdition d'efficacité supérieure à celle connue pour les réseaux équivalents cablés. En conséquence, il s'agit d'optimiser en terme de performances globales les algorithmes qui agiront dans les différentes couches du protocole de communication.

Tout cela doit permettre d'obtenir les performances visées à un coût permettant le raccordement direct de PC ou de portables. Aujourd'hui, nous participons à la construction d'un des tous premiers réseaux locaux numériques sans fil qui sera peut-être l'une des briques technologiques essentielle de demain.

3.3.2 Algorithmique en théorie des nombres

Participant : François Morain

Il s'agit là d'un axe fortement motivé au départ par la cryptographie et qui constitue principalement le domaine de recherche de F. Morain (mis à disposition du Laboratoire d'informatique de l'École polytechnique par l'Armement et associé au projet à 2/5 de son temps sur le site). Les problèmes de base du domaine sont ceux portant sur la décomposition des entiers en facteurs premiers. Ces problèmes sont au cœur de la réalisation de systèmes de chiffrement à clefs publiques, systèmes très robustes et dont l'essor ne fait que commencer. À ce jour, F. Morain dispose d'outils particulièrement performants pour mener à bien cette tâche : en particulier, factoriser des nombres de 100 chiffres ou bien prouver la primalité de nombres de 1000 chiffres est chose courante.

Au cours de sa thèse, F. Morain s'est intéressé au problème de la primalité exacte des entiers. Au contraire des tests probabilistes, les algorithmes de primalité exacts fournissent des résultats exacts qui s'accompagnent de preuves de primalité au moyen de "certificats" faciles à vérifier. Grâce à ses travaux sur l'algorithme ECPP, qui utilisent la théorie des courbes elliptiques sur les corps finis, F. Morain détient toujours le record du monde de primalité sur des nombres quelconques avec un nombre de 1505 chiffres décimaux. Pour des applications à la cryptographie, le programme a été optimisé en vue de pouvoir tester des nombres premiers de 100 chiffres le plus rapidement possible. À ce jour, il ne faut que 2 minutes de temps CPU sur une Dec Station 3100 pour avoir une preuve de primalité pour un entier de cette taille.

Pour arriver à ces résultats, F. Morain a développé un ensemble de logiciels très complets dont la construction utilise les ressources de langages divers tels que Le-Lisp, C et MAPLE. L'ensemble en l'état actuel comporte environ 10 000 lignes de programmes et représente un effort de près de 3 homme-ans de programmation. Ces programmes ont subi une refonte récente et forment désormais le système CESAR (Courbes ElliptiqueS et ARithmétique). Ces programmes ont été mis en accès public par FTP anonyme et ont été portés sur une dizaine d'architectures.

En vue d'accélérer les temps de réponse, F. Morain a modifié son programme afin de pouvoir tester des nombres de manière distribuée. À ce jour, il utilise simultanément un ensemble d'une dizaine de stations pour mener à bien ses expériences. Cette idée a permis de gagner un facteur 4 dans le délai nécessaire pour tester des nombres de 500 chiffres décimaux. Le processus de distribution des calculs est réalisé à l'aide de scripts Unix qui peuvent être réutilisés à d'autres fins de calcul réparti. Le record de primalité de 1505 chiffres décimaux a ainsi été obtenu en moins d'un mois de temps écoulé au prix d'un effort cumulé de calcul de 4 ans CPU sur machines de puissance moyenne 5 MIPS.

Plus récemment, F. Morain a travaillé sur l'attaque des systèmes de chiffrement RSA, en étudiant les algorithmes de factorisation d'entiers. Par une utitlisation adaptée de calculs distribués, il a ainsi pu factoriser des entiers de plus de 100 chiffres décimaux (l'un de 101 chiffres, tiré du projet Cunningham, l'autre de 100 chiffres du *challenge* des nombres de partition de RSA, Inc.). Il a fallu pour cela un mois de CPU sur plusieurs sites (Lix, Inria, Ensta), puis une algorithmique sophistiquée afin de mener à bien une étape cruciale de calcul matriciel difficile. En

effet, la phase finale des algorithmes de factorisation moderne (crible quadratique) nécessite la recherche de relations de dépendance dans des matrices booléennes creuses de taille 50 000 et plus. Pour pouvoir "rentrer" ces matrices en mémoire centrale sur un ordinateur, il faut utiliser des algorithmes récents dus notamment à Odlyzko, Pomerance, Lenstra, etc., ("élimination Gaussienne intelligente"), algorithmes qui réduisent la taille de la matrice par un pré-traitement utilisant peu de mémoire. À cette occasion, F. Morain a pu développer des algorithmes tirant partie du processeur vectoriel du Vax du Centre Informatique de l'École polytechnique, pour gagner du temps : un gain d'un facteur 3 a ainsi été possible dans cette phase du calcul. Une implantation sur machine parallèle (ncube) est en cours à l'École Polytechnique.

F. Morain a également continué à travailler sur un des thèmes centraux de la théorie algorithmique des nombres : la théorie des courbes elliptiques sur les corps finis. À travers une convention entre l'École polytechnique et le Centre Électronique de l'Armement (Celar), il s'est ainsi intéressé au calcul du nombre de points sur une courbe elliptique dans un corps fini [39] et [24]. Ces travaux sont un prélude nécessaire à la mise en œuvre de systèmes de chiffrement arithmétique à clefs publiques plus résistants que le système RSA.

Plus récemment, F. Morain a apporté deux contributions à l'étude des sommes de caractères liées aux courbes elliptiques à multiplication complexe. Ces contributions ont fait l'objet de deux articles [37, 38]

3.3.3 Algorithmique géométrique

Participant : Henry Crapo

Les méthodes de la théorie des invariants sont essentielles à la création des langages de haut niveau en géométrie. Des telles méthodes servent pour la modélisation (graphique informatique) et la démonstration automatique, et s'avèrent plus efficaces que des algorithmes construits dans le contexte de l'algèbre informatique (bases standards, méthode de Wu), qui sont plutôt super-exponentiels. H. Crapo et Jürgen Richter-Gebert (TU Berlin) ont exposé des algorithmes de démonstration automatique en géométrie projective et en géométrie euclidienne, et se servant seulement de l'algèbre linéaire.

Le développement d'un outil informatique expérimental pour le graphique, incorporant les méthodes de la théorie des invariants et de la

géométrie projective orientée, est le but d'un effort commun de H. Crapo et J. Richter-Gebert (TU, Berlin). Ce logiciel est construit sur NextStep, l'outil d'interface inventé par J.-M. Hullot et ses collègues à l'Inria (1985-1987).

Une première version est opérationnelle : des configurations projectives (images euclidiennes, sphériques ou descriptions ascii) sont définies, étudiées, manipulées. Des théorèmes projectifs sont détectés par la persistance d'incidences pendant le mouvement relatif des composants de la configuration. Des preuves algébriques sont trouvées. Dans cette première version du logiciel, une configuration consiste en la donnée de points, de droites, de cercles et autres courbes coniques dans le plan projectif.

La deuxième étape de leur projet a démarré : une architecture ouverte de logiciel permettra à l'utilisateur de définir des "macros" géométriques et de spécifier un contexte géométrique, ainsi que de traiter des configurations de dimension supérieure à 2.

L'idée d'une telle architecture est de fournir pour la géométrie les outils que TEX fournit pour la typographie, c'est-à-dire, un vocabulaire librement extensible pour le calcul et pour la modélisation. Des contextes spéciaux (géométrie euclidienne, hyperbolique, conforme, ...) sont alors fournis.

3.3.4 Chaînes, compressions, ADN

Participants : Pierre Nicodème, Mireille Régnier

(Voir le rapport d'activité de l'Action Génome).

P. Nicodème a poursuivi son travail concernant le multiplexage de la réaction de polymérisation en chaîne PCR. L'utilisation de cette réaction, qui fait partie des outils usuels des biologistes moléculaires permet d'amplifier un grand nombre de fois un segment de gène. Un certain nombre de maladies comme le cancer peuvent souvent être caractérisées par la perte d'une ou de plusieurs portions d'ADN, empêchant les gènes correspondants de jouer le rôle qu'ils jouent dans le métabolisme de cellules saines. Le multiplexage de PCR permet de faire à la fois des réductions de coûts et des gains de temps appréciables. Il pose des problèmes de type combinatoire entre les amorces utilisées pour faire les amplifications. Un modèle théorique applicable à des locis non polymorphes (identiques chez tous les individus) prédit la possibilité d'amplifier simultanément

plus de deux cents loci. Les applications pratiques à des locis fortement polymorphes sont beaucoup plus modestes, limitant le nombre d'amplifications simultanées à trois ou quatre. Néanmoins la technique récente de PCR utilisant des amorces fluorescentes de couleurs différentes doit permettre d'amplifier simultanément une quinzaine de locis, ce qui confirme l'intérêt du travail informatique réalisé.

Pierre Nicodème collabore à l'avancement du projet de base de données de domaines protéiques PRODOM. Une protéine est en général constituée de plusieurs domaines ; chacun de ceux-ci présente des caractéristiques propres au niveau fonctionnel et au niveau des structures secondaires et tertiaires (repliements respectifs en deux et trois dimensions). Dire que deux protéines ont deux domaines communs, du point de vue des séquences, signifie que ces deux protéines possédent deux sous-séquences très voisines, ou homologues (au sens biologique). Une théorie mathématique s'appuyant sur des propriétés de files d'attentes et de marches aléatoires permet de bien comprendre les résultats obtenus par des méthodes de scores quand on cherche les homologies entre une séquence et une base de séquences. Les domaines de PRODOM ont été constitués en comparant deux à deux toutes les séquences de la base de séquences protéiques SWISSPROT, et en effectuant l'aggrégation des sous-séquences homologues par associativité. De ce fait, les domaines de PRODOM peuvent contenir jusqu'à soixante sous-séquences. P. Nicodème cherche, dans ce contexte, à mettre au point des méthodes qui permettent de rechercher l'homologie entre une séquence et un domaine, en tirant partie de l'information riche contenue dans un domaine. Dans un premier temps, des méthodes de recherche d'homologies sans insertions et suppressions dans les séquences sont étudiées. L'extension au cas des insertions et suppressions sera faite ensuite.

4 Actions industrielles

4.1 Calcul formel et actions MAPLE

Par le biais du développement intensif nécessité par le système $\Lambda\Upsilon^{\Omega}$ et ses cousins, une expertise importante de la pratique et de la programmation en MAPLE a été acquise par divers membres du projet qui contribuent à la diffusion du domaine. Ainsi P. Zimmermann et B. Salvy participent régulièrement à des enseignements du calcul formel, et sont également sollicités avec régularité par d'autres chercheurs de l'Inria ou

d'autres institutions pour résoudre divers problèmes liés au calcul formel. Cette année, en collaboration avec le projet META2, une école Inria sur l'utilisation du calcul formel et en particulier du système MAPLE a été organisée. Cette école a été un succès : vingt personnes (limite du nombre de places) d'origine industrielle et universitaire sont venus y participer.

Cette relation privilégiée avec le système MAPLE, qui possède des fonctionnalités voisines et des performances souvent nettement supérieures aux autres systèmes (excepté AXIOM, dont la philosophie et l'objet sont un peu à part), est par ailleurs prolongée par un accord conclu cette année avec *Waterloo Maple Software*, qui commercialise le produit. L'Inria est ainsi considéré comme un utilisateur privilégié du système, ce qui permet un accès gratuit au système à tous les projets Inria, dont une douzaine utilise à des degrés divers de tels outils de calcul formel. Cette collaboration s'est accrue grâce à la participation de J. Carette aux activités du projet. En effet, plusieurs faiblesses du logiciel MAPLE qui gênaient particulièrement son utilisation efficace ont pû être identifiées et corrigées. Cette participation a permis d'influer sur les priorités de certains projets de développement de *Waterloo Maple Software* de façon à favoriser son utilisation à l'Inria. Dans cette perspective, B. Salvy est entré cette année au *Maple Design Team*, groupe de huit personnes qui se réunit semestriellement pour décider à haut niveau de l'évolution des fonctionnalités et de l'implémentation du système.

4.2 Réseaux radio et Projet Européen LAURA

(Voir aussi 3.3.1)

L'action dans le domaine des protocoles de communication s'est essentiellement déroulée dans le cadre du projet européen LAURA. Le projet ESPRIT III LAURA a pour but de réaliser et commercialiser un réseau local sans fil de performances similaires à celles du réseau Ethernet. Il a débuté en 1992 et terminera en 1995. Le projet LAURA regroupe deux industriels majeurs, Dassault Automatique (F) et Elettronica (I), une société d'ingénierie, Symbionics (GB), et trois centres de recherche institutionnels, l'Inria et les universités de Bradford et Bristol (GB). Au sein de l'Inria le projet LAURA concerne trois équipes : ALGO, CODES et REFLECS.

L'avancement des travaux se déroule conformément au calendrier prévu. Notre projet est essentiellement impliqué au niveau de la définition des algorithmes d'accès radio et de routage dans le réseau. Le responsable de cette action au sein du projet est Ph. Jacquet.

L'équipe a surtout travaillé sur la détection et la résolution de collisions en radio communication et sur les algorithmes de routage adaptatifs. Les principes de base reposent sur des brevets préalables pris au nom de l'Inria par Ph. Jacquet et P. Mühlethaler (projet REFLECS).

Le projet LAURA a permis à l'Inria de participer d'une manière soutenue à des comités de normalisation européens où se réunissent la plupart des industriels qui comptent dans le domaine (AT&T, IBM, Apple). En effet le réseau LAURA devra se conformer aux normes strictes qui sont en cours d'élaboration à l'Institut Européen des Normes de Télécommunication (ETSI) sous le nom d'HIPERLAN (comité RES10). La norme devrait voir le jour d'ici 1995.

4.3 Convention Digital Equipment

Une convention passée avec Digital Equipment, de sorte à assurer les échanges de technologie liés aux processeurs arithmétiques rapides, et incluant P. Bertin (détachement du corps des Mines) se termine fin 1994, suite à la dissolution du Laboratoire DEC PRL.

5 Actions nationales et internationales

5.1 Actions nationales

Les actions nationales mettent en jeu :

- la direction du PRC Mathématiques et Informatique (MATHINFO) ;
- la participation à l'action GÉNOME du Ministère de l'Enseignement Supérieur et de la Recherche.

Le Programme de Recherches Coordonnées *Mathématiques et Informatique* met en jeu une quarantaine d'équipes en milieu universitaire et CNRS principalement, pour un budget global d'environ 4MF sur deux ans (sources MESR). Ph. Flajolet assisté de V. Collette assume depuis fin 1991 la direction de ce groupement qui a, dans le même temps, un statut de formation CNRS (GDR 967). Ce programme a une forte allure

interdisciplinaire et couvre les aspects fondamentaux de l'informatique ainsi que les méthodes mathématiques qui s'y rattachent. Les thèmes de recherche concernent l'algorithmique, les automates, le calcul formel, la combinatoire et la logique.

Le développement actuel des recherches sur le génome humain entraîne une collaboration entre informaticiens et biologistes. M. Régnier participe au premier des trois axes définis pour cette collaboration dans le cadre du MESR : "analyse des séquences", et notamment au groupement "Méthodes Formelles pour l'analyse du Génome" du GDR "Informatique et Génomes". Une collaboration régulière est établie avec l'université Claude Bernard à Lyon (M. Gouy), dans les domaines de la détermination de structures secondaires des ARN et de la recherche de phylogénie. On souhaite automatiser et systématiser des recherches actuellement faites empiriquement par les biologistes. P. Nicodème a travaillé sur un outil d'aide au multiplexage des réactions de polymérisation en chaîne et développé un logiciel, *multipcr*, correspondant à cette problématique. Ce travail a été effectué en collaboration avec G. Thomas (Institut Curie) et J.-M. Steyaert (Lix). D'autres collaborations s'établissent avec l'Inra-Toulouse.

5.2 Actions internationales

Les collaborations scientifiques internationales du projet sont nombreuses et concrétisées par deux projets ESPRIT, ALCOM II (Basic Research) et LAURA (Project). En 1994, les collaborations scientifiques ont donné lieu à plusieurs publications et travaux communs avec divers centres de recherche, dont : Technische Universität Wien, Massachussets Institute of Technology, University of Chicago, University of Hong-Kong, Waterloo University, Eidgenössische Technische Hochschule (ETH, Zürich), Princeton University, Purdue University, Université du Québec à Montréal, etc. (Voir la liste de publications du projet.)

Basic Research, ALCOM II. Le projet ALGO est l'un des partenaires du projet ESPRIT Basic Research ALCOM (*ALgorithms and COMplexity*). L'action regroupe 12 partenaires de 8 pays de la CEE ; en France y participent également le groupe de P. Rosenstiehl à l'EHESS et le groupe de J.-D. Boissonnat à Sophia Antipolis, à l'étranger les universités de Saarbrücken, Warwick, Dublin, Barcelone, Utrecht, Patras et Berlin. Les thèmes de recherche en sont l'algorithmique et les structures de don-

nées, l'algorithmique géométrique, parallèle et distribuée ainsi que les problèmes correspondants d'évaluation de performance. Le démarrage effectif de ce projet qui fait suite à un projet homologue, ALCOM I, a eu lieu en juillet 1992.

6 Diffusion des résultats

Frédéric Chyzak, dernier doctorant du projet a développé un ensemble de bibliothèques destinées à l'analyse de séries génératrices multivariées. Il a obtenu un Prix d'option de l'École polytechnique en 1994, prix qui couronne ses travaux dans le projet.

Henry Crapo a donné une conférence invitée sur son travail en homologie géométrique au congrès de la Société Mathématique de Belgique à Anvers. Il a donné deux conférences invitées (une, sur la théorie des invariants avec Gian-Carlo Rota, l'autre sur la démonstration automatique avec Jürgen Richter-Gebert) au congrès "Invariant Methods in Discrete and Computational Geometry". Il a donné un cours intitulé *Synthetic Projective Geometry* dans le cadre du projet ERASMUS (Mathematical Methods of Robotics) à l'Universidad d'Almería. Il a participé au colloque sur les Séries Formelles et Combinatoire Algébrique, à Firenze. Il a donné des conférences sur la démonstration automatique au colloque sur Polytopes and Computational Geometry et au Seminaire Lotharingien de Combinatoire, Bergakademie Freiberg.

Philippe Dumas a présenté au Séminaire Eureca Protheo (InriaNancy) et au Laboratoire de l'informatique du parallélisme de l'École Normale Supérieure de Lyon ses résultats sur les automates cellulaires.

Philippe Flajolet a été élu membre correspondant de l'Académie des Sciences en Juin 1994, a reçu le Prix Michel Monpetit décerné par l'Académie en novembre 1994, et s'est vu décerner un Doctorat *Honoris Causa* de l'Université Libre de Bruxelles en décembre 1994. Il a participé à divers Comités et Conseils Scientifiques dont l'Observatoire de la Recherche en Informatique, et a été Directeur du Programme de Recherches Coordonnées Mathématiques et Informatique pour la période 1992-1994. Ph. Flajolet est éditeur de *Theoretical Computer Science* (Elsevier, Amsterdam), de la *Maple Newsletter* et de *Random Structures and Algorithms* (Wiley).

Xavier Gourdon a présenté au CIRM Luminy à Marseille lors des journées MEDICIS en décembre 1993 l'algorithme de Schönhage sur la recherche des zéros d'un polynôme. Il a exposé à l'université de Rennes le 18 mars 1994 des travaux communs avec B. Salvy sur l'asymptotique des récurrences linéaires. Il a présenté ses travaux portant sur l'analyse d'un algorithme de factorisation de polynômes sur les corps finis à l'université de Toronto le 14 avril 1994. Le 20 mai 1994, il a été invité à Limoges pour présenter ses recherches sur l'approximation numérique des modules extrêmes d'un polynôme. Le 5 Septembre 1994 avec Ph. Dumas, il a donné un cours sur le logiciel de calcul formel MAPLE dans le cadre d'un stage à l'École des Mines de Nantes, destiné à un public composé de 35 professeurs de classes préparatoires.

Philippe Jacquet a participé activement aux réunions du projet LAURA, ainsi qu'aux réunions du comité RES10 de l'ETSI. La fréquence de ces réunions a été cette année portée à un rythme mensuel afin de parvenir à la norme HIPERLAN définitive pour la mi 1995.

François Morain a donné un cours de Théorie algorithmique des nombres (20 heures, octobre 1993 – janvier 1994) au sein du DEA Informatique Mathématique et Applications, organisé par l'École polytechnique et l'École normale supérieure. Il a dirigé des travaux pratiques de Pascal pour les élèves de première année de l'École polytechnique (septembre à décembre 1993). Il encadre le travail de thèse de R. Lercier. F. Morain a été examinateur d'informatique au concours d'entrée à l'École normale supérieure, du 1er au 13 juillet 1994. F. Morain a fait partie du jury de thèse de F. Arnault en décembre 1993, de celui de J.-M. Couveignes en juillet 1994. Du 28 avril au 9 mai, F. Morain a fait un séjour aux États-Unis, comprenant une conférence sur invitation au Fields Institute (Waterloo, Ontario), où il a parlé de "Using small prime powers in Schoof's algorithm" ; puis la conférence ANTS 1 à Cornell, où il présentait un article sur le même thème avec J.-M. Couveignes. Du 9 au 14 octobre, il a participé à la conférence sur invitation "Computational Number Theory" qui s'est tenue à Dagstuhl (Allemagne). Il y a parlé de "character sums".

Bruno Salvy a rédigé en commun avec Cl. Gomez (projet Meta2) et P. Zimmermann (projet Eureca) un livre sur l'utilisation du calcul formel, basé sur le système Maple. Il a été co-organisateur d'une école Inria intitulée "Calcul formel pour les applications. Utilisation du système Maple". Il a aussi été invité à donner un exposé à la "Maple retreat",

réunion d'une quarantaine de développeurs du système. Depuis cette année, il fait partie du "Maple Design Team", groupe de huit personnes qui se réunit semestriellement pour décider de l'évolution des fonctionnalités et de l'implémentation du système. Il a fait un exposé sur l'analyse des quadtrees à l'université du Québec à Montréal, un exposé sur la combinatoire et le calcul formel lors d'une réunion franco-russe à l'institut Lyapunov (Moscou), et deux exposés d'introduction aux fonctions holonomes à l'université du Chili, à Santiago du Chili. Cette année, B. Salvy a été confirmé dans sa position de chef de travaux pratiques en informatique à l'École Polytechnique, où il enseigne la programmation en Pascal et en C. Par ailleurs, dans le cadre du DEA "Informatique Mathématique et Applications", il donne en commun avec Ph. Flajolet le cours d'analyse d'algorithmes.

7 Publications

Livres et monographies

[1] X. GOURDON, *Les maths en tête, mathématiques pour M', Algèbre*, Ellipse, 1994, 288 pages.

[2] X. GOURDON, *Les maths en tête, mathématiques pour M', Analyse*, Ellipse, 1994, 416 pages, À paraître.

Thèses

[3] H.-K. HWANG, *Théorèmes Limites pour les Structures Combinatoires et les Fonctions Arithmétiques*, thèse de doctorat, École Polytechnique, Palaiseau, France, décembre 1994.

Articles et chapitres de livre

[4] H. CRAPO, R. PENNE, «Chirality and the Isotopy Classification of Skew Lines in Projective 3-Space», *Advances in Mathematics 103*, 1994, p. 1–106.

[5] H. CRAPO, J. RICHTER-GEBERT, *Invariant Methods in Discrete and Computational Geometry*, Kluwer Academic Publishers, 1995, ch. Automatic Proving of Geometric Theorems, To appear.

[6] H. CRAPO, G.-C. ROTA, *Invariant Methods in Discrete and Computational Geometry*, Kluwer Academic Publishers, 1995, ch. The Resolving Bracket, To appear.

[7] H. Crapo, W. Whiteley, «Spaces of Stresses, Projections and Parallel Drawings for Spherical Polyhedra», *Contributions to Algebra and Geometry*, 1995, To appear.

[8] H. Crapo, *Combinatorial Theory, Collected Works of Gian-Carlo Rota, 1*, Birkhaüser Verlag, 1995, ch. Rota's "Combinatorial Theory", To appear.

[9] H. Crapo, «On the Generic Rigidity of Plane Frameworks Structures in the Plane», *Advances in Mathematics*, 1995, To appear.

[10] H. Crapo, «Projective Configurations: their Invariants and Homology», *Proceedings of the Royal Irish Academy*, 1995, Special issue "The Mathematical Heritage of Sir William Rowan Hamilton". To appear.

[11] M. Drmota, M. Soria, «Marking in Combinatorial Constructions and Limit Distributions», *Theoretical Computer Science*, 1994, To appear.

[12] P. Flajolet, M. Golin, «Mellin Transforms and Asymptotics: The Mergesort Recurrence», *Acta Informatica 31*, 1994, In press.

[13] P. Flajolet, P. Grabner, P. Kirschenhofer, H. Prodinger, R. Tichy, «Mellin Transforms and Asymptotics: Digital Sums», *Theoretical Computer Science 123*, 2, 1994, p. 291–314.

[14] P. Flajolet, T. Lafforgue, «Search Costs in Quadtrees and Singularity Perturbation Asymptotics», *Discrete and Computational Geometry 12*, 4, 1994, p. 151–175.

[15] P. Flajolet, P. Zimmerman, B. Van Cutsem, «A Calculus for the Random Generation of Labelled Combinatorial Structures», *Theoretical Computer Science 132*, 1-2, 1994, p. 1–35.

[16] D. Gardy, G. Louchard, «Dynamic analysis of some relational data bases parameters», *Theoretical Computer Science, Series A*, 1995, To appear.

[17] D. Gardy, «Join sizes, urn models and normal limiting distributions.», *Theoretical Computer Science, Series A 131*, août 1994, p. 375–414.

[18] D. Gardy, «Some results on the asymptotic behaviour of coefficients of large powers of functions», *Discrete Mathematics*, 1995, To appear.

[19] X. Gourdon, B. Salvy, «Computing one million digits of $\sqrt{2}$», *Maple Technical Newsletter*, 10, 1993, p. 60–65.

[20] H.-K. Hwang, «Asymptotic expansions for Stirling's number of the first kind», *Journal of Combinatorial Theory,* Series A, 1995, To appear.

[21] S. P. Lipshitz, T. C. Scott, B. Salvy, «On the Acoustic Impedance of Baffled Strip Radiators», *Journal of the Audio Engineering Society 43*, 1995, To appear.

[22] B. Salvy, P. Zimmermann, «Gfun: a Maple package for the manipulation of generating and holonomic functions in one variable», *ACM Transactions on Mathematical Software 20*, 2, 1994, p. 163–177.

[23] B. Salvy, «Fast Computation of Some Asymptotic Functional Inverses», *Journal of Symbolic Computation 17*, 1994, p. 227–236.

Communications à des congrès, colloques, etc.

[24] J.-M. Couveignes, F. Morain, «Schoof's algorithm and isogeny cycles», *in : Proceedings ANTS'94*, janvier 1994. To appear.

[25] H. Daudé, P. Flajolet, B. Vallée, «An Analysis of the Gaussian Algorithm for Lattice Reduction», *in : Algorithmic Number Theory Symposium*, *Lecture Notes in Computer Science*, 1994. Proceedings of *ANTS'94*. In press, 15 pages.

[26] D. Gardy, G. Louchard, «Dynamic analysis of the sizes of relations», *in : STACS'95*, mars 1995.

Rapports de recherche et publications internes

[27] F. Chyzak, «Holonomic Systems and Automatic Proofs of Identities», *Research Report n° 2371*, Institut National de Recherche en Informatique et en Automatique, 1994.

[28] P. Dumas, P. Flajolet, «Asymptotique des récurrences mahleriennes: le cas cyclotomique», *Research Report n° 2163*, Institut National de Recherche en Informatique et en Automatique, janvier 1994, 23 pages. Accepted for publication in *Journal de théorie des nombres de Bordeaux*.

[29] P. Flajolet, X. Gourdon, P. Dumas, «Mellin Transforms and Asymptotics : Harmonic Sums», *Research Report n° 2369*, Institut National de Recherche en Informatique et en Automatique, octobre 1994, 55 pages. To appear in *Theoretical Computer Science*.

[30] P. Flajolet, G. Labelle, L. Laforest, B. Salvy, «Hypergeometrics and the cost structure of quadtrees», *rapport de recherche n° 2249*, Institut National de Recherche en Informatique et en Automatique, avril 1994, 26 pages. Accepted for publication in *Random Structures and Algorithms*.

[31] P. Flajolet, R. Sedgewick, «The Average Case Analysis of Algorithms: Saddle Point Asymptotics», *Research Report n° 2376*, Institut National de Recherche en Informatique et en Automatique, 1994, 55 pages.

[32] P. Flajolet, R. Sedgewick, «Mellin transforms and asymptotics: finite differences and Rice's integrals», *rapport de recherche n° 2231*, Institut National de Recherche en Informatique et en Automatique, mars 1994, 21 pages. To appear in *Theoretical Computer Science*.

[33] D. Gardy, G. Louchard, «Dynamic analysis of some relational data base parameters I: projections.», *rapport de recherche n°94-6*, Laboratoire Prism, University of Versailles, février 1994.

[34] D. Gardy, G. Louchard, «Dynamic analysis of some relational data base parameters II: equijoins and semijoins.», *rapport de recherche n°94-7*, Laboratoire Prism, University of Versailles, février 1994.

[35] H.-K. Hwang, J.-M. Steyaert, «On the number of heaps and the cost of heap construction», *Research Report n°LIX/RR/93/07*, Laboratoire d'Informatique de l'École Polytechnique, 1993, Submitted to *Theoretical Computer Science*.

[36] H.-K. Hwang, «On the convergence rates and asymptotic expansions in the central and local limit theorems for combinatorial structures», *rapport de recherche n°LIX/RR/93/08*, Laboratoire d'Informatique de l'École Polytechnique, 1993.

[37] A. Joux, F. Morain, «Évaluation des sommes de caractères liées aux courbes elliptiques à multiplication complexe par l'anneau des entiers d'un corps quadratique imaginaire de nombre de classes 1», *Rapport de Recherche n°LIX/RR/93/11*, Laboratoire d'Informatique de l'École Polytechnique, 1993, To appear in *J. Number Theory*.

[38] F. Leprévost, F. Morain, «Revêtements de courbes elliptiques à multiplication complexe par des courbes hyperelliptiques et sommes de caractères», *Rapport de Recherche n°LIX/RR/94/07*, Laboratoire d'Informatique de l'École Polytechnique, août 1994.

[39] F. Morain, «Calcul du nombre de points sur une courbe elliptique dans un corps fini : aspects algorithmiques», Submitted for publication of the Actes des Journées Arithmétiques, mars 1994.

[40] B. Salvy, (Editor), «Algorithms Seminar, 1993–1994», *Research Report n°2381*, Institut National de Recherche en Informatique et en Automatique, octobre 1994, 142 pages.

[41] J. Shackell, B. Salvy, «Asymptotic Forms and Algebraic Differential Equations», *Research report n°2319*, Institut National de Recherche en Informatique et en Automatique, août 1994, To appear in the *Journal of Symbolic Computation*.

Divers

[42] M. Drmota, M. Soria, «Marking in Combinatorial Constructions: Images and Preimages in Random Mappings», submitted to SIAM Journal on Discrete Mathematics, février 1994.

[43] X. Gourdon, H. Prodinger, «Random Subgraphs of the n-cycle», Submitted to *Discrete Mathematics*, 1994.

[44] H.-K. HWANG, «Asymptotic expansions of mergesort recurrences», Submitted to *Information Processing Letters*.

8 Abstract

The general objective of the Algorithms Project is the design, analysis, and optimization of major algorithms and data structures of computer science. The main activity revolves around the construction of analytic models permitting a precise performance evaluation and fine optimization of algorithms.

A unified theory for a large class of combinatorial and algorithmic processes has been built over the past few years. It is based in part on combinatorial analysis and discrete mathematics structures and in part on asymptotic analysis through complex function theory. In this way, very precise complexity characterizations can be obtained concerning the average–case or probabilistic behaviour of fundamental algorithms.

This systematic approach connects itself nicely with computer algebra. It has become in particular possible to develop a large computer algebra program that can assist the analysis of structurally complex algorithms. On this occasion, interest has also developed for computer algebra and symbolic manipulation systems.

The approach taken there is prototypical of the possibility of developing computer algebra libraries dedicated to the analysis of complex systems.

At the same time, all members of the project are engaged in research dealing with specific fields of application in computer science listed below.

- Trees and data structures for fast retrieval of information in the context of multidimensional data or secondary storage systems.
- Communication protocols for either local area networks or mobile radio networks.
- Strings and pattern matching algorithms with implications for DNA sequence processing.
- Computational number theory in relation to cryptography.

The project is a component of the LAURA ESPRIT Project and the ALCOM Basic Research Action.

Table des matières