Projet ALGO

Algorithmes

Localisation : *Rocquencourt*

Mots-clés : Algorithmes probabilistes, Algorithmique, Analyse combinatoire et asymptotique, Analyse d'algorithmes, Arbres, Calcul formel, Évaluation de performances, Mathématiques discrètes, Protocoles de communication, Réseaux sans fil, Séquences génétiques et textuelles, Structures de données, Systèmes complexes et analyse automatique, Théorie algorithmique des nombres

1 Composition de l'équipe

Responsable scientifique

Philippe Flajolet, directeur de recherche, Inria

Responsable permanent

Mireille Régnier, directeur de recherche, Inria

Secrétaire

Virginie Collette

Personnel Inria et détachements

Philippe Jacquet, détachement corps des Mines
Paul Mühletalher, détachement du corps de l'armement, projet REFLECS puis projet ALGO à partir d'octobre 1996
Philippe Robert, directeur de recherche, Inria
Bruno Salvy, chargé de recherche, Inria

Collaborateurs extérieurs

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

Chercheurs associés

Philippe Dumas, Prof. Cl. Prépa. lycée Jean-Baptiste Say
Xavier Gourdon, Dassault Systèmes, à partir d'octobre 1996
François Morain, DRET et École polytechnique

Chercheurs doctorants

Frédéric Chyzak, boursier de l'École polytechnique
Jean-François Dantzer, Prag., université de Versailles St-Quentin et boursier de l'École polytechnique
Vincent Dumas, boursier Inria, jusqu'à septembre 1996
Xavier Gourdon, boursier de l'École polytechnique, jusqu'à septembre 1996
Eithne Murray, boursière Inria, jusqu'à juillet 1996
Pierre Nicodème, Convention avec l'université Paris 7
Jean-Marc Wachter, boursier de l'ENS

Chercheur en mise à disposition

Jacques Carette, Ingénieur Waterloo Maple Software et boursier du gouvernement canadien, jusqu'à juillet 1996

Stagiaires

Cédric Adjih, Inria
Olivier Lecomte, Inria, d'avril à juin 1996
Amir Qayyum, boursier Inria CIES, de juin à octobre 1996

2 Présentation du projet

Le thème premier du projet ALGO est l'aléa combinatoire. Il s'agit de caractériser les propriétés attendues (en moyenne, en probabilité, en distribution) d'objets obéissant à des règles de combinaison finies, mais constituant de très grands ensembles. Ces situations se rencontrent sans cesse en informatique, par exemple les configurations d'ordre possibles lors d'un tri de seulement 100 éléments sont en nombre voisin de 10^{158}, tout en obéissant, avec une écrasante probabilité, à des règles fort précises.

Les problèmes d'aléa combinatoire interviennent de manière essentielle en algorithmique. D'abord, la conception de la plupart des algorithmes efficaces se fonde naturellement sur les cas attendus, en moyenne ou en probabilité, plutôt que sur une analyse pessimiste qui doit être réservée à des contraintes de type "temps réel". D'autre part, l'optimisation fine des algorithmes repose précisément sur une exploitation très serrée des lois du hasard. Ce rôle de l'analyse de l'aléa combinatoire se trouve encore renforcé par l'importance croissante des algorithmes dits "randomisés" (bien formalisés par Karp et Rabin depuis une vingtaine d'années) où il s'avère payant d'introduire volontairement le hasard dans le calcul. Ainsi, les tables de hachage constituent-elles une alternative souvent très efficace aux arbres de recherche, les signatures accélèrent-elles considérablement la recherche textuelle, les "skip lists" remplacent-elles graduellement les arbres équilibrés dans de nombreuses applications. Parmi d'autres applications célèbres de cette "aléatoirisation", la construction de cryptosystèmes à clefs publiques utilise de manière très sûre des tests probabilistes (domaine dans lequel F. Morain détient des records mondiaux) et il apparaît que des protocoles de communication, faisant suite à Ethernet, mais avec une utilisation mieux contrôlée de l'aléa, brisent infiniment mieux qu'Ethernet la symétrie dans les réseaux ; ceci permet d'acheminer des trafics plus importants de manière beaucoup plus robuste et les principaux phénomènes de stabilité, voire hyperstabilité, ont été mis en évidence dans le projet (Ph. Jacquet, Ph. Flajolet, *et al.*). La problématique d'algorithmique randomisée est poursuivie par exemple avec succès par le projet PRISME de Sophia-Antipolis dans le domaine de l'algorithmique géométrique.

Le projet ALGO se donne comme objectif l'analyse en profondeur de l'aléa combinatoire et la recherche de ses *lois* générales. Ce thème est voisin par ses objectifs, mais dual par ses méthodes, de la modélisation des systèmes informatiques, laquelle repose de manière prédominante sur des mathématiques *a priori* continues, calcul des probabilités en tête. Ici, nous sommes dans le domaine des mathématiques discrètes. La combinatoire étant par définition l'étude des objets finis discrets, nous visons à développer un domaine que l'on pourrait qualifier de "combinatoire statistique" par analogie avec la physique

statistique. Il s'agit en effet d'expliquer le comportement macroscopique visible (le temps d'exécution global et sur un grand "ensemble" d'un certain algorithme) en partant des règles d'agencement élémentaire, simples et connues, qui régissent les données, les méthodes de calcul, ainsi que les structures de données en jeu.

L'une des caractéristique du projet, au plan méthodologique, est la globalité de l'approche poursuivie, laquelle sera détaillée dans les sections suivantes. Disons seulement ici qu'au prix d'un saut dans l'abstraction (méthodes symboliques) et dans la géométrisation (singularités de séries génératrices), de nombreuses lois émergent qui sont transverses à des applications variées. Ainsi, les mêmes principes — manifestés par une loi gaussienne commune — régissent-ils aussi bien la pagination d'un arbre d'index, l'apparition attendue de motifs spécifiques dans un texte (par exemple une chaîne d'ADN), la factorisation de polynômes en calcul formel, ou encore, sous certaines conditions, la déconnexion partielle d'un réseau de communication.

Cette approche généraliste a trouvé une résonnance particulière lorsque confrontée au *calcul formel*. Il est apparu, en effet, vers le tournant des années 1990 qu'il était possible de décider mathématiquement de nombreuses propriétés de l'aléa combinatoire, ce par un calcul de nature essentiellement formelle. Au sein du projet, c'est ce qui a donné lieu aux thèses de B. Salvy et Paul Zimmermann (en 1991). Le programme de recherche correspondant à ces aspects est loin d'être achevé. Il est apparu que surgissaient de nombreux problèmes de portée générale en calcul formel. Citons principalement ici, comme directions nouvelles, l'asymptotique automatique (B. Salvy) les méthodes mixtes symboliques numériques (X. Gourdon, B. Salvy), et la preuve automatique d'identités combinatoires (F. Chyzak, B. Salvy).

Ces approches théoriques fournissent des repères puissants issus des modèles analytiques dont les résultats quantitatifs sont toujours très précis et qui s'appliquent même parfaitement, dans le contexte des algorithmes "randomisés". On en verra l'illustration précise dans le rôle joué par plusieurs membres du projet, Ph. Robert et Ph. Jacquet notamment, dans la conception et l'optimisation d'algorithmes et de protocoles de communication. Des solutions complètes et complexes ont été proposées pour les réseaux sans fils et pour le projet MulticablE qui vise à utiliser au mieux le câble de télévision comme support de connexion à l'Internet. Les algorithmes proposés sont complets et les membres du projet sont allés jusqu'au dépôt de brevet (Ph. Jacquet, P. Mühlethaler) et à l'acceptation de leurs solutions par les comités de normalisation concernés. En leur cœur, ces procédés efficaces reposent *in fine* sur une utilisation algorithmiquement très astucieuse et mathématiquement bien maîtrisée de quelques jets de dés. En passant interviennent quelques équations différentielles et aux différences non-linéaires, des transformation de Mellin, ou de délicates intégrales de contour.

Le cheminement qui sous-tend ces recherches permet d'aborder l'analyse de *modèles complexes* dans les domaines très variés sujets à modélisation combinatoires. C'est ainsi qu'ont pu être résolues au fil des ans diverses conjectures correspondant à des analyses précises de problèmes tels que le dimensionnement en hachage dynamique (M. Régnier, Ph. Flajolet), la redondance des algorithmes de compression de Lempel et Ziv (Ph. Jacquet), la performance des arbres quadrants pour la recherche multidimensionnelle (Ph. Flajolet, B. Salvy), ou le comportement probabiliste des meilleures méthodes de recherche de motifs (M. Régnier). D'autres applications typiques sont constituées par les algorithmes d'estimation probabiliste en bases de données, l'allocation mémoire partagée (Ph. Flajolet), et les protocoles en arbre pour réseaux locaux et réseaux sans fils (Ph. Jacquet, Ph. Flajolet). Les travaux du groupe sont assez largement repris dans la littérature scientifique spécialisée (livres ou articles). Plusieurs de ces recherches ont été menées en collaboration avec les projets VERSO (bases de données), REFLECS (protocoles et contraintes temps réel), MEVAL (modélisation) et Action AGCT (génomique). Elles nous valent de nombreuses coopérations internationales avec des universités comme celles de Barcelone, Princeton, Purdue, Stanford, Vienne, Washington, Waterloo, etc.

3 Actions de recherche

3.1 Analyse d'algorithmes

Participants : Philippe Dumas, Philippe Flajolet, Danièle Gardy, Xavier Gourdon, Philippe Jacquet, Jean-Marc Steyaert, Michèle Soria

La combinatoire statistique est le cadre naturel de développement de l'analyse d'algorithmes ; les méthodes sont alors de deux types :

— les méthodes probabilistes (dans le style d'Erdős et de l'École hongroise), fondées sur des approximations probabilistes *a priori* et bien adaptées aux problèmes de graphes aléatoires et d'optimisation combinatoire dans la classe complexité NP ;

— les méthodes analytiques (dont les livres de Knuth représentent la première génération) fondées sur des descriptions exactes relayées par une analyse asymptotique précise.

L'originalité du projet consiste à avoir poussé très loin les méthodes analytiques dans ce domaine et à fonder les bases d'une véritable *"combinatoire analytique"*.

3.1.1 Processus combinatoires

Le point de départ est constitué par les méthodes symboliques en analyse combinatoire qui permettent de traduire des modèles complexes. Il s'agit d'un courant "symbolique" en accord avec les tendances modernes en combinatoire, et dont le projet a montré la puissance dans le domaine de l'analyse d'algorithmes (calculs de complexité moyenne de Flajolet et Steyaert). Ceci est à la base de l'automatisation réalisée par la bibliothèque de calcul formel COMBSTRUCT décrite plus loin. Cette phase symbolique confronte alors l'analyste à une grande variété d'équations fonctionnelles, soit à une variable (dénombrements, moyennes, variance) soit à plusieurs variables (distributions probabilistes). Il s'agit ensuite d'extraire les informations utiles de ces séries génératrices, et se manifeste alors un phénomène de simplification asymptotique de portée considérable et qu'il s'agit de capturer. Ici encore, l'approche suivie vise à dégager quelques grands schémas généraux correspondant à une classification du domaine en "processus combinatoires" régis par des lois précises. Les méthodes s'appuient sur l'analyse de singularités de transformées, séries génératrices ou transformations intégrales. C'est là un courant "géométrique" original dont les bases élémentaires viennent d'être dégagées par Sedgewick et Flajolet dans un ouvrage introductif [278, 277] (tirages de 5000 et 2000 exemplaires, respectivement). Une monographie de synthèse par Flajolet et Sedgewick, au nom évocateur de *Analytic Combinatorics* est en préparation pour 1998.

Le schéma diviser-pour-régner est un grand classique de l'informatique qui échappait jusque récemment à une quantification précise, étant donné le caractère apparemment fort cahotique des temps de calcul observés. En fait, il est montré dans [318] que, sous des conditions très générales, ces phénomènes ont une structure fractale très exactement quantifiable (voir aussi le rapport d'activité du Projet FRACTALES), comme il apparaît grâce à des méthodes de transformations de Mellin voisines de la théorie analytique des nombres. Ces travaux sont poursuivis par Ph. Flajolet et Ph. Dumas [283]. Ils s'appliquent par exemple au problème de la recherche efficace de points visibles dans un nuage de points d'un espace de dimension quelconque.

Le schéma analytique des sommes harmoniques est également traité de manière complète dans le travail [318] au moyen de la transformation de Mellin. Ceci unifie et simplifie une bonne cinquantaine d'analyses portant sur des objets et algorithmes combinatoires très variés, tels : hachage extensible, compression de données à la Lempel et Ziv, protocoles de communication en arbre, factorisation de polynômes en calcul formel, algorithmes d'estimation probabilistes en bases de données, etc. Cet exemple est typique d'un processus combinatoire unique — le processus "d'arbre digital" — qui est conceptuellement très simple et fait surface dans de nombreuses applications, mais dont l'évaluation quantitative résiste (à cause de subtiles fluctuations inhérentes) à toute analyse élémentaire. Ces travaux trouvent

leur prolongement dans l'analyse des séquences (algorithmes de compression, par exemple) décrite à la section 3.3.

Les modèles d'urnes sont parmi les modèles les plus courants d'allocations aléatoires, et peuvent être utilisés pour représenter nombre de phénomènes informatiques (hachage, bases de données, etc). En particulier, l'apprentissage de fonctions symétriques, dans un cas où les données peuvent être erronées, a pu être modélisé et analysé dans le cas statique, conduisant à un coût d'apprentissage (l'erreur en généralisation) asymptotiquement gaussien [281].

Le schéma des "composantes maximales" est l'un des thèmes centraux de la thèse de X. Gourdon [280]. Ici encore, des lois très générales régissent l'apparition d'objets maximaux en combinatoire statistique. Ainsi, le plus grand facteur premier d'un entier en arithmétique, la plus grande composante d'une fonction aléatoire en cryptographie, ou le plus grand cycle d'une permutation obéissent-ils à une loi commune qui se caractérise par la classique fonction de Dickmann originaire de la théorie analytique des nombres. Un signe de la profondeur technique de la thèse [280] est la résolution d'une conjecture de Golomb-Knuth datant de la fin des années 1960 et portant sur la distribution de la longueur maximale de cycle dans une permutation aléatoire, ou encore la distance entre maxima successifs. Ces analyses s'appliquent notamment à la factorisation de polynômes en calcul formel, où X. Gourdon et Ph. Flajolet [302] ont obtenu pour la première fois une analyse complète de l'algorithme de factorisation tel est qu'il implanté dans les systèmes de calcul formel courants.

3.1.2 Applications algorithmiques

Les processus d'arbres liés aux ensembles ordonnés se retrouvent dans les arbres binaires de recherche, les arbres d'index, les arbres pour la recherche multidimensionnelle, dans les bases de données géographiques par exemple. Il a été montré par Ph. Flajolet et X. Gourdon [315] que les arbres d'index sont "sûrs" au sens que leur occupation mémoire (qui fluctue de manière gaussienne) est très étroitement centrée autour de ce que prédit le cas moyen. Les propriétés aléatoires (donc les performances des algorithmes) des arbres quadrants sont désormais bien comprises grâce à l'analyse de singularités de systèmes différentiels et à la théorie de la perturbation associée [303]. Au passage apparaissent certaines sommes dont la simplification pose de délicats problèmes de fonctions spéciales (polylogarithmes) et de calcul formel ; voir [317].

La dynamique d'un algorithme est jusqu'ici peu prise en compte par la théorie combinatoire. À terme, on pourrait espérer mieux comprendre les phénomènes où l'aléa n'est plus hérité par les sous-structures. Une première ouverture en direction de la théorie des systèmes dynamiques a ainsi été réalisée par Ph. Flajolet en collaboration avec Brigitte Vallée (université de Caen). Le rôle des opérateurs de transfert est apparu comme crucial à la compréhension des fonctions de coût de certains algorithmes arithmétiques (tels les très classiques développements en fraction continue et l'algorithme d'Euclide) ou géométriques (telles les méthodes numériquement stables pour déterminer l'orientation en algorithmique géométrique). Les premiers travaux dans ce sens sont en cours de publication [282, 319]. Par de nombreux égards, les opérateurs sont une généralisation naturelle des séries génératrices pour la prise en compte de tels phénomènes dynamiques en analyse d'algorithmes.

Enfin, M. Soria, Ph. Flajolet et H.-K. Hwang (ancien stagiaire du projet, désormais membre de l'*Academia Sinica* à Taïwan) ont commencé une synthèse d'envergure visant à expliquer en termes simples la fréquence extrême d'apparition de la loi de Gauss en combinatoire analytique.

La plupart des travaux mentionnés ont, implicitement ou explicitement, une portée unificatrice : rôle des lois de Poisson et de Gauss en combinatoire analytique, importance des phénomènes oscillatoires (fractals ou non) en algorithmique, lois transverses des maxima ou des modèles d'urnes, par exemple. Ils permettent de tirer un certain nombre de conclusions simples telles :

— la phase coûteuse de factorisation de polynômes est celle de la factorisation en degré distincts, sur laquelle doit alors porter l'optimisation principale (travaux en cours de X. Gourdon et D. Panario (Waterloo) sur les algorithmes de Shoup) ;

— les fonctions cryptographiques usuelles (de type DES, par exemple) doivent présenter un petit nombre de grandes composantes maximales et ainsi doivent être robustes à l'attaque par itération de la fonction de codage (voir le rapport du Projet CODES) ;

— la compression de type Lempel-Ziv est d'un ordre asymptotique moins redondante (donc, plus efficace) si l'on utilise la méthode à fenêtre.

L'analyse d'algorithmes a ainsi un rôle scientifique généraliste de clarification, prélude à l'optimisation fine. Elle sert de fondement aux études dans divers domaines d'application, dont la technicité mathématique va toujours croissante.

3.2 Calcul Formel

Participants : Frédéric Chyzak, Xavier Gourdon, François Morain, Eithne Murray, Bruno Salvy

Les trois étapes fondamentales de l'analyse d'algorithmes telle qu'elle est pratiquée au projet ALGO sont la modélisation combinatoire, la manipulation de séries génératrices et l'analyse asymptotique. Chacune de ces étapes requiert des capacités de calcul symbolique importantes, tant pour l'application des méthodes symboliques que pour l'expérimentation. Ce besoin a été la source d'une intense activité en calcul formel au sein du projet. L'objectif à long terme est de systématiser et d'automatiser ces trois étapes. Au cours des années ont ainsi été élaborés de nombreux algorithmes et programmes. Au fil de nos recherches, chacun des volets de ce triptyque s'est par ailleurs développé suffisamment pour fournir des outils de portée de plus en plus générale qui touchent maintenant un public assez large d'utilisateurs du calcul formel intéressés tant par la combinatoire que par les manipulations de séries ou par l'analyse. La diversité de ce public est encore accrue à la suite de l'introduction de certains de nos programmes [285] dans les bibliothèques du système MAPLE.

L'activité du projet en calcul formel se caractérise par une ambivalence fondamentale. En effet, les progrès algorithmiques sont souvent motivés par le désir de développer les capacités d'application de l'implantation, et inversement l'implantation est souvent utilisée pour valider et diffuser les développements théoriques. Notre domaine de recherche en calcul formel est complémentaire de celui étudié par le projet SAFIR, à savoir les systèmes polynomiaux. Il arrive que nos objectifs convergent, et c'est notamment le cas avec les travaux de X. Gourdon sur la résolution numérique de polynômes [280, 287]. On peut distinguer quatre grandes directions de travail dans notre activité en calcul formel : les structures combinatoires ; les suites et fonctions spéciales ; l'asymptotique automatique et l'algorithmique des nombres.

3.2.1 Calcul formel et combinatoire

La bibliothèque COMBSTRUCT a été conçue et développée par le projet ALGO (Ph. Flajolet, B. Salvy, E. Murray) en liaison avec le projet EURECA de Nancy (P. Zimmermann). Elle fait l'objet d'une collaboration régulière avec les groupes de Waterloo (Université et Compagnie WMS) et s'est vue intégrée totalement dans la dernière mouture du système MAPLE (1996, release 5.4). Elle permet actuellement la génération aléatoire ou exhaustive, le calcul automatique de dénombrements et de séries génératrices, et est à ce titre une aide de portée générale pour la simulation et le test systématique de modèles combinatoires. Disons qu'en l'état actuel, et sur son créneau, son expertise est de l'ordre de celle d'un étudiant en début de 3ème cycle.

Un langage de description généralisant les grammaires *context-free* permet d'exprimer des objets aussi divers que permutations, arbres binaires, arbres généraux, partitions d'entiers ou d'ensembles, graphes fonctionnels ou molécules chimiques, par exemple carbures ou alcools. À partir d'une description de structure décomposable, les outils proposés dans le *package* COMBSTRUCT peuvent *(i)* compter efficacement le nombre d'objets d'une certaine taille répondant à la spécification ; *(ii)* produire des fonctions de génération aléatoire uniforme de faible complexité — utiles pour des tests statistiques ; *(iii)* produire

des fonctions de génération exhaustive de ces objets — utiles pour des tests de robustesse de procédures ; *(iv)* produire des itérateurs, c'est-à-dire des fonctions permettant d'accéder successivement à tous les objets d'une certaine taille, mais sans les stocker tous en mémoire (cette fonctionnalité est en cours d'implantation) ; *(v)* calculer des équations satisfaites par les séries génératrices d'énumération de ces objets — utiles pour la phase d'analyse asymptotique. L'objectif à terme est d'ajouter à ces fonctionnalités des capacités asymptotiques et des capacités d'analyse d'algorithmes opérant sur les structures combinatoires. Le programme $\tilde{\cdot}$ réalisé au début des années 90 par B. Salvy et P. Zimmermann fournissait déjà une partie de ces fonctionnalités, mais sa portabilité et ses fonctionnalités étaient limités par l'usage de CAML en conjonction avec MAPLE. Notre objectif est maintenant de tirer parti de l'expérience acquise avec $\tilde{\cdot}$ pour réaliser une version (COMBSTRUCT) entièrement intégrée en MAPLE, en mettant également l'accent sur la modularité et la souplesse d'emploi. La possibilité de produire des visualisations de grandes structures décomposables est par ailleurs étudiée au projet EURECA par P. Zimmermann.

3.2.2 Séries et échelles asymptotiques

Selon l'origine combinatoire du problème, les séries génératrices que l'on est amené à étudier peuvent être données sous des formes diverses. Elles peuvent être connues sous forme explicite, mais elles peuvent aussi être définies par une ou plusieurs équations, fonctionnelles, différentielles ou aux différences. De même, leurs coefficients peuvent vérifier des récurrences de natures diverses. Manipuler ces fonctions définies implicitement nécessite des innovations théoriques, ainsi qu'un important effort d'implantation. Ce thème de recherche touche aux fondements du calcul symbolique, où il apparaît qu'il est paradoxalement souvent plus facile de traiter une fonction lorsqu'elle est représentée comme solution d'équations que lorsqu'elle est représentée sous forme close. En particulier, les questions de simplification et de formes normales qui sont une des difficultés majeures rencontrées par l'utilisateur trouvent une bien meilleure réponse dans ce contexte.

Le cas des solutions d'équations différentielles ou de récurrences linéaires attire beaucoup l'attention de la communauté de combinatoire et de calcul formel. De nombreuses suites et fonctions spéciales sont définies par de telles équations, qui bénéficient d'une algorithmique très riche. Le *package* GFUN développé par B. Salvy et P. Zimmermann (projet EURECA) s'est enrichi cette année d'une nouvelle fonctionnalité produite par E. Murray qui prend en entrée la forme close d'une fonction et produit une équation différentielle linéaire dont cette fonction est solution (lorsqu'une telle équation existe). Cette fonctionnalité, qui effectue précisément le chemin inverse de celui vers lequel se précipitent nombre d'utilisateurs, permet ensuite de calculer des développements en série de manière plus rapide qu'avec la forme close ; elle permet également la localisation des singularités et le calcul des comportements au voisinage des singularités. Le *package* GFUN comporte actuellement 4400 lignes de code MAPLE. Il a fait l'objet d'une revue très positive dans *Computing Reviews* et est incorporé au *superseeker* de N. Sloane aux Bell Laboratories, accessible sur le Web et qui détermine de nombreuses suites d'après leurs premiers termes.

La thèse de F. Chyzak démarrée cette année s'inscrit également dans cette thématique, et s'attaque au problème multivarié, c'est-à-dire au cas des fonctions, suites, séries ou distributions définies par un *système* d'équations linéaires. Les opérateurs linéaires considérés ici peuvent être différentiels, aux différences, aux q-différences, ou de nombreux autres types. Le traitement porte en fait sur des polynômes de Ore, qui unifient les propriétés communes à tous ces opérateurs. Il s'appuie sur la théorie des fonctions holonomes de plusieurs variables qui garantit la clôture d'une très large classe de suites et de fonctions par des opérations simples comme la somme et le produit, mais aussi par les opérations plus compliquées de sommation et d'intégration. En pratique, l'étude de l'effectivité de ces propriétés de clôture dans le cas de la somme et du produit met en œuvre des algorithmes simples et proches d'algorithmes existant pour le traitement des nombres algébriques, alors que dans le cas de la sommation et de l'intégration, les algorithmes requièrent la mise en œuvre de moyens plus sophistiqués d'élimination dans des algèbres non commutatives d'opérateurs linéaires [313].

F. Chyzak a implanté ses algorithmes dans un package du nom de MGFUN pour le système MAPLE. À ce jour, ce package est constitué de 8400 lignes de code, de 4400 lignes de documentation et de 2600

lignes de tests utilisés pour la maintenance du programme. L'intégration d'une partie de MGFUN à la bibliothèque standard Maple est prévue pour 1997, le *Symbolic Computation Group* de l'université de Waterloo ayant invité F. Chyzak à cet effet.

Les besoins de la combinatoire analytique en matière de développements asymptotiques dépassent les capacités actuelles des systèmes de calcul formel. En effet, les calculs de coûts moyens et plus encore de variance donnent systématiquement lieu à des annulations non seulement dans les premiers termes des développements mais aussi dans l'ordre de grandeur exponentiel des croissances. La construction automatique des échelles asymptotiques nécessaires et le calcul avec ces échelles pose de nombreux problèmes sur lesquels le calcul formel est en progrès rapide. Les premiers travaux sur ce sujet datent des années 90. En 1988, G. Gonnet et K. Geddes (créateurs du système MAPLE) proposent un modèle permettant de traiter des formules de complexité proche de la formule de Stirling. Puis en 1990, John Shackell (université de Canterbury) publie un algorithme qui permet de déterminer *de manière garantie* la limite des fonctions exp-log (fonctions de base de l'asymptotique). L'année suivante, la thèse de B. Salvy propose une première implantation de développements asymptotiques dans des échelles asymptotiques générales.

Après les travaux fondateurs de J. Shackell, de nombreuses classes d'équations ont été étudiées. Les équations différentielles algébriques peuvent dans une certaine mesure se traiter algorithmiquement, mais la complexité des algorithmes est pour l'instant exponentielle [299]. Les équations implicites ont longtemps résisté à un traitement systématique, mais un algorithme vient d'être développé par B. Salvy et J. Shackell [322]. Cet algorithme généralise des travaux antérieurs de B. Salvy et J. Shackell sur le cas des inverses fonctionnels, mais les méthodes sont assez différentes. Ces progrès n'ont pas toujours été suivi d'implantation. Les idées s'étant clarifiées au fil des années, il est apparu utile de donner une présentation pédagogique du cas le plus simple — les fonctions données explicitement — afin d'y intéresser les concepteurs des systèmes. C'est l'objet de [311]. Parallèlement et dans le même esprit, une nouvelle structure de donnée pour les développements asymptotiques vient de faire l'objet d'une proposition aux développeurs du système MAPLE. Là encore, on s'attaque à une brique essentielle du système, puisque les séries en MAPLE sont la structure de donnée utilisée pour les polynômes, eux-mêmes base de nombreuses structures de données.

3.2.3 Arithmétique en très grande précision

Autre constituant nucléaire du calcul formel, l'arithmétique des grands nombres a énormément progressé ces dernières années, progrès motivé par l'application à la cryptologie à clefs publiques. La thèse de F. Morain (soutenue en 1990), portant sur l'utilisation des courbes elliptiques dans l'étude de la primalité des entiers, et qui lui vaut toujours d'être détenteur de records du monde dans le domaine, l'a naturellement conduit à approfondir les liens entre calcul formel, courbes algébriques, arithmétique et cryptologie. Il a poursuivi dans [293] et [294] certains des travaux de sa thèse, ce qui l'a entraîné dans des développements théoriques novateurs, concernant l'évaluation des sommes de caractères.

Une des tâches récentes de F. Morain a été de participer au mouvement qui a complètement "algorithmisé" un des principaux problèmes liés aux courbes elliptiques, celui du calcul de la cardinalité d'une courbe dans un corps fini de grande caractéristique. Grâce à ses travaux, il est désormais possible de construire des cryptosystèmes basés sur des courbes elliptiques, et qui sont plus robustes que leurs équivalents classiques, à taille de clefs comparables.

Poursuivant et amplifiant les travaux de Schoof, Atkin et Elkies, F. Morain [298] a simplifié et optimisé les algorithmes de calcul d'isogénies entre courbes elliptiques, qui sont au cœur des avancées récentes. Avec l'aide de J.-M. Couveignes (originellement à l'ENS, et maintenant à Bordeaux I), il a mis au point des raffinements de l'algorithme qui l'ont conduit à battre le record du monde une première fois. D'autres astuces, décrites dans [314], lui ont permis de garder son record, malgré une concurrence internationale féroce.

En petite caractéristique, Reynald Lercier (École polytechnique) et F. Morain ont décortiqué un algorithme inventé par Couveignes pour résoudre les problèmes spécifiques de la petite caractéristique [327, 328]. Outre l'étude approfondie des courbes elliptiques, F. Morain s'est intéressé également

aux problèmes liés aux nombres pseudopremiers (c'est-à-dire aux nombres composés résistant aux tests de primalité faibles), ce qui a fait l'objet d'un travail en collaboration [288].

Comme pour l'analyse d'algorithmes, l'approche du projet au calcul formel est globalisante et unificatrice. La résolution d'un problème appliqué de cryptosystèmes, de combinatoire ou d'analyse d'algorithmes est abordée à un niveau de généralité qui permet le développement d'une algorithmique à large portée. Ainsi, les travaux sur les cryptosystèmes mettent en jeu des courbes elliptiques et ont des retombées sur la factorisation d'entiers et les tests de primalité ; les travaux sur la combinatoire fournissent des générateurs aléatoires efficaces susceptibles de nombreuses applications ; les travaux sur l'analyse d'algorithmes ont abouti au développement d'une algorithmique d'échelles asymptotiques très générale, dont le besoin s'était fait sentir en intégration numérique et en physique mathématique.

3.3 Recherche dans les séquences

Participants : Philippe Jacquet, Pierre Nicodème, Mireille Régnier

L'algorithmique des séquences ou objets textuels, couvre des domaines d'application variés (compression, images, biologie, ...). Ce sujet comprend d'abord des recherches algorithmiques : recherche d'un motif dans un texte, recherche multidimensionnelle, compression, ... On y analyse aussi l'aléa combinatoire qui est au cœur de ces problèmes. Tant du point de vue de la conception que de l'analyse, la structure des mots — leurs périodes — apparaît essentielle. Par exemple, tout algorithme raisonnable de recherche de motif est conçu pour exploiter certaines régularités [309] et on relie son domaine d'efficacité à des structures de langages particulières [321, 296]. Notre approche probabiliste a deux aspects. D'une part, des grands théorèmes probabilistes trouvent des applications naturelles dans l'étude des séquences. Plus précisément, nous avons mis en évidence différents types de processus de renouvellement, la loi limite étant généralement gaussienne ; le calcul effectif des paramètres de coût peut être très délicat et les outils combinatoires et analytiques permettent pour cette classe de problèmes les calculs effectifs des distributions. D'autre part, certaines applications (biologie, compression) se ramènent à des recherches de similarité, où des résultats statistiques sur les mots ou sous-séquences d'un texte sont nécessaires pour évaluer la pertinence des similarités mises à jour par les algorithmes.

D'un point de vue méthodologique, nous cherchons à développer une *théorie analytique de l'information*. La "dépoissonisation analytique" est un procédé de portée générale (inventé par Ph. Jacquet et M. Régnier) qui est sans analogue probabiliste direct, car il revient à considérer un processus de Poisson de taux complexe. L'année 1996 a vu une synthèse des théorèmes de dépoissonisation menée par P. Jacquet et Wojciech Szpankowski (université de Purdue). La dépoissonisation trouve de nombreuses applications en combinatoire statistique (arbres digitaux, par exemple). Fondée sur les intégrales de col, elle apparaît en aval des analyses de Mellin et permet d'apercevoir les bases d'une théorie analytique de l'information ; elle permet en effet des développements asymptotiques complets sur des paramètres comme l'entropie et la compressibilité d'un texte que les théoriciens de l'information avaient du mal à estimer auparavant. Nous développons aussi des outils de calcul pour le cas markovien [330, 292]. Les résultats s'appliquent à la compression et à la recherche de motifs exceptionnels dans les textes (DosDNA, prédiction de structures secondaires, ...). Ils permettent aussi d'établir les domaines d'efficacité des différents algorithmes de recherche de similarités dans des bases de données protéiques.

3.3.1 Motifs

Dans l'étude des variantes de l'algorithme de Knuth-Morris-Pratt, nous avons prouvé l'existence de points de renouvellement fréquents : au moins un dans chaque fenêtre de la taille du motif. On en déduit la linéarité de la complexité de ces algorithmes et des propriétés de convergence presque sure ; les constantes de linéarité elles-mêmes ont été calculées dans de précédents travaux. Pour les algorithmes classiques du type Boyer-Moore, nous avons montré que le renouvellement est presque sûr. Il s'ensuit que le coût, sur l'ensemble des textes possibles, a une distribution gaussienne. Des calculs

combinatoires, utilisant fortement les périodes du mot recherché permettent de caractériser la moyenne, la variance et la vitesse de convergence vers l'état stationnaire.

La complexité des évaluations de performances dans le cas markovien provient du nombre de cas différents à considérer. Nous définissons pour chaque problème des langages caractéristiques dont la contribution au coût total de l'algorithme est calculable. Ainsi, la constante de linéarité des variantes de Knuth-Morris-Pratt a été obtenue en agrégeant des états de l'automate associé. L'évaluation du nombre d'occurrences d'un motif donné dans un texte se ramène à des inversions de systèmes d'équations algébriques satisfaits par les séries génératrices et la généralisation au cas markovien est faite dans [330] sans augmenter la taille du système. Ce travail est en cours d'extension à la recherche approchée d'un motif.

Dans le domaine de la recherche multidimensionnelle, L. Rostami et M. Régnier ont appliqué leurs résultats théoriques sur les périodes en 2D pour implanter des algorithmes de recherche 2D. En effet, un motif périodique est engendré par un ensemble (minimal) de mots. Ceci permet d'utiliser en dimension 2 les procédures de recherche efficaces en 1D utilisant les périodes (duel, témoin, ...). Elles ont ainsi proposé un algorithme de recherche de témoins [309].

Les travaux de Ph. Jacquet ont permis de traiter en profondeur les lois limite de l'algorithme de compression de Lempel et Ziv lorsque la taille du texte à compresser croît. De ces résultats a découlé la caractérisation du facteur de redondance de l'algorithme de Lempel et Ziv par rapport à la compression entropique optimale, fermant par là même un problème ouvert depuis 1978 dans la communauté de la théorie de l'information. Ces résultats sont relatifs au modèle statistique dit de Bernoulli où chaque caractère possède une distribution indépendante des caractères qui le précèdent ou lui succèdent dans le texte. Notre objectif est de résoudre le problème de l'extension markovienne d'ici un ou deux ans avec les moyens dont nous disposons dans la boîte à outils de la *théorie analytique de l'information*.

Dans un domaine plus pratique, les algorithmes de compression de type Lempel et Ziv peuvent être étendus à des média autres que le texte, par exemple l'image et le son. Des expériences concluantes ont été menées sur des images fixes à l'université de Purdue, en liaison avec le projet ALGO. L'intérêt de l'approche à la Lempel et Ziv réside dans le fait que la décompression est très rapide (simple lecture en ligne d'un fichier) et ne nécessite que peu de ressources de calcul. Une application prometteuse est dans les techniques de communication Internet où on ne peut pas exiger du client trop de ressources instantanées pour la décompression en ligne. On reviendra sur cette problématique dans la section consacrée aux réseaux.

3.3.2 Séquences génétiques

Les calculs statistiques évoqués ci-dessus fournissent des formules exactes pour la probabilité d'occurrence d'un mot donné. Appliquées au DosDNA, petits motifs répétés qui sont la trace de l'instabilité génétique dans les séquences d'ADN, elles permettent de tester la signification statistique des répétitions. En effet, pour les tailles de séquences considérées, ces formules sont calculables grâce au package MAPLE GFUN dû à B. Salvy et P. Zimmermann.

Une autre application est la prédiction de structures secondaires, déterminées par l'appariement de mots avec leurs images inverses. On définit un seuil de pertinence pour de tels appariements qui représente la longueur au dessus de laquelle un appariement n'est plus un simple effet du hasard statistique. Choisi comme point d'ancrage, un tel appariement permet une approche "diviser pour régner" qui a donné de bons résultats sur l'ARN 16S et 23S. La complexité de la prédiction dépend du nombre de structures qu'il est possible d'associer à une séquence donnée. Fariza Tahi, du projet VERSO et M. Régnier ont généralisé des travaux d'énumération dus à M. de Chaumont, X. Viennot et M. Waterman afin de prendre en compte des contraintes biologiques [310].

P. Nicodème poursuit son travail sur la recherche de similarités entre protéines. Il s'agit ici d'organiser les bases de données de séquences protéiques connues en regroupant les protéines partageant des fonctions biologiques proches. Il participe à l'INRA-Toulouse au développement de la base ProDom29 où les familles de protéines sont regroupées suivant leurs domaines fonctionnels. Une similarité entre

protéines s'exprime via une fonction de score. Pour définir des seuils de pertinence, P. Nicodème utilise les formules de Karlin-Iglehart qui associent une pertinence probabiliste aux scores obtenus par les algorithmes de recherche de similarités tels Blast. Il a entrepris de regrouper les différents développements mathématiques permettant d'obtenir ces formules ; ce travail précisera en particulier les hypothèses nécessaires pour leur utilisation. Les constantes des formules appliquées à chaque domaine ont été calculées à partir de séquences aléatoires dans une collaboration avec Jean-Jacques Codani de l'action GÉNOME et le logiciel de recherche de similarité BlastMultAl est maintenant opérationnel sur ProDom29. Il ressort du travail de comparaison avec d'autres méthodes que BlastMultAl est d'une sensibilité équivalente à celle des Profiles, et supérieure à celle obtenue avec les consensus. BlastMultAl ayant permis de trouver des similarités distantes nouvelles, cette méthode, qui est utilisable avec des familles contenant très peu de séquences, peut être considérée comme complémentaire des Profiles. On étudie une importante extension aux familles contenant des trous (insertions et suppressions autorisées). Les formules de Karlin-Iglehart precédemment citées ne s'appliquent pas ; on utilisera la méthode d'approximation de Poisson récemment proposée par Waterman et Vingron.

L'objet des recherches sur les séquences est ici non seulement l'obtention de la complexité moyenne des algorithmes, mais aussi l'application en retour à l'algorithmique de résultats probabilistes et énumératifs difficiles. Ceci s'insère dans un cadre plus général visant à développer une théorie analytique de l'information qui s'appuie sur la combinatoire, les probabilités et l'analyse.

3.4 Algorithmique et modélisation des réseaux

Participants : Vincent Dumas, Philippe Jacquet, Olivier Lecomte, Paul Mühlethaler, Amir Qayyum, Philippe Robert, Jean-Marc Wachter

Cette section concerne la conception et l'analyse des algorithmes qui sont essentiellement utilisés dans les réseaux de télécommunication. Nous nous intéressons par exemple aux réseaux qui seront créés à partir des nouveaux media comme la radio hautes fréquences ou encore le câble TV. Nous sommes particulièrement attachés à garder un lien étroit entre des activités théoriques comme la modélisation, et des activités plus appliquées comme l'expérimentation de plateformes. Cette double compétence s'avère nécessaire pour aborder le domaine de la standardisation en télécommunication.

Nous poursuivons nos activités dans les trois domaines suivants :

a) La conception des algorithmes sur le réseau étudié, la modélisation et l'analyse des modèles probabilistes associés ;

b) L'implantation des algorithmes en vraie grandeur au travers d'une collaboration (via un contrat européen par exemple) ;

c) La participation aux activités des groupes internationaux de normalisation concernés (ETSI, IEEE).

3.4.1 Étude de modèles probabilistes de réseaux

Participants : Vincent Dumas, Philippe Robert, Jean-Marc Wachter

Le cadre général de cette recherche concerne les propriétés de renormalisation des réseaux de communication. Le processus de Markov décrivant l'état d'un réseau est en général complexe, même si la loi stationnaire de celui-ci est connue. La combinatoire des expressions ne permet pas une évaluation qualitative de ces réseaux, pour les problèmes de dimensionnement notamment, lorsque leur nombre de nœuds est significatif. Une méthode intéressante, issue de la physique des particules, consiste à renormaliser le processus à la fois en temps et en espace par un petit paramètre ε et faire tendre celui-ci vers 0. Un processus limite ainsi obtenu conserve les caractéristiques essentielles du réseau, schématiquement la partie "bruit" autour des trajectoires est éliminée. Un processus limite est quasi-déterministe, les points de discontinuité de la dynamique conservant une part d'aléatoire. Toute la difficulté de l'étude

consiste à identifier et caractériser les limites possibles. Il peut y avoir plusieurs limites et les équations "limite" peuvent présenter des solutions pathologiques qu'il convient d'éliminer (les physiciens le font avec des arguments d'entropie dans leur cadre). Ce programme a déjà largement été entamé dans la thèse de V. Dumas [279] dans le cadre des réseaux multi-classe. J.-M. Wachter a démarré une thèse sur ce sujet dans le cas des réseaux avec perte, en septembre 1996 (directeur Ph. Robert).

Le cadre est alors celui d'un réseau de télécommunications, les appels réservent plusieurs liens pendant un temps aléatoire, si un des liens nécessaire à l'appel est déjà complet (i.e. à capacité maximale), l'appel est perdu. La renormalisation utilisée consiste à augmenter de façon conjointe la capacité des liens du réseau ainsi que le trafic. Cette normalisation diffère de celle utilisée dans [279] où l'espace d'état était indépendant de la normalisation, ce n'est pas le cas dans ce cadre. Hunt et Kurtz ont montré la relative compacité des trajectoires renormalisées ainsi que certaines propriétés de leurs limites. Une étude générale détaillée nous semble, comme dans le cas des réseaux multi-classe, hors d'atteinte pour l'instant. Le programme sera donc dans un premier temps d'étudier les questions de convergence sur des exemples assez simples et certaines topologies symétriques.

Avec la normalisation mentionnée plus haut, Ph. Robert et Danièle Tibi (université de Paris VII), se sont intéressés aux problèmes d'estimation de la vitesse d'atteinte de l'équilibre de files d'attente à un serveur. Le but général est de donner, s'il existe, l'estimation asymptotique de l'instant de *cut-off* : avant cet instant, le processus est très près de l'état initial et après celui-ci l'état stationnaire est atteint. La distance utilisée ici entre les distributions est celle de la convergence en variation totale. Une activité importante se dégage actuellement autour de ces questions (Diaconis, Salff-Coste, Stroock, ...). Les outils utilisés sont principalement géométriques (méthodes de chemins dans des graphes, inégalités de Poincaré, Cheeger, ...). La seule méthode connue dans le domaine des files d'attente consiste à calculer explicitement les transitoires des processus incriminés ; ce n'est effectivement possible que pour de très rares cas.

Les estimations données par ces méthodes dans notre cadre ne sont pas satisfaisantes. Nous nous sommes intéressés tout d'abord au problème de l'estimation de la seconde valeur propre. En utilisant une formule variationnelle et un théorème de convergence en loi (du type central limite, loi des grands nombres ou une convergence vers une loi de Poisson dépendant du modèle étudié), il est possible de donner l'asymptotique de la seconde valeur propre. Il suffit pour cela de résoudre un problème classique de minimisation d'un opérateur sur un sous-espace de fonctions de type L_2 qui se résout en choisissant la base orthonormée idoine (Fourier, les polynômes d'Hermite, ...). Cette méthode simple donne l'ordre de grandeur de façon assez immédiate, le problème de minimisation ne concerne que la valeur exacte de la constante numérique devant l'ordre de grandeur. Nous nous sommes ensuite intéressés au problème plus délicat des temps de *cut-off*. Nous avons montré que ce phénomène existe pour les files simples avec pertes (un cas frontière reste cependant non résolu pour l'instant). La méthode employée utilise un couplage pour chaque cas. L'étape actuelle consiste à généraliser ces méthodes au cas des réseaux.

3.4.2 Conception et étude de protocoles sur le réseau câblé

Participants : Philippe Jacquet, Paul Mühlethaler, Philippe Robert

Le projet ALGO participe au projet MulticablE (contact Inria : Olivier Muron) pour la définition d'un protocole d'accès sur le réseau câblé. Le projet MulticablE réunit en consortium la Lyonnaise des Eaux, France Telecom, Cap Gemini Sogeti, la SAT, et l'Inria autour d'une expérimentation de réseaux à accès distribué sur le réseau câblé de Paris. Ce projet fait partie des projets agréés pour les autoroutes de l'information et est partiellement financé par le Ministère de l'Industrie. Les principales caractéristiques du réseau câblé sont :

— l'existence d'un canal montant permettant l'interactivité ;

— un débit, 1-6 Mbps en montée et 10-20Mbps en descente, assurant la transmission haut débit.

Dans le cadre de cette activité l'Inria a effectué une campagne de mesures sur le réseau de Paris pour fournir à la Lyonnaise des indications sur l'utilisation du réseau. C. Adjih prépare une étude de trafic similaire sur le réseau câblé du Mans.

Le projet ALGO a aussi démarré une réflexion sur les problèmes touchant aux couches basses du protocole (la couche MAC) qui avaient été identifiés dès le début du projet MulticablE, par exemple le problème de la symétrie entre flux montant et flux descendant. À ce titre l'Inria participe aux réunions du comité de normalisation internationale IEEE 802.14 (voir actions internationales). Des contributions ont ainsi été proposées concernant les protocoles d'accès, et les mécanismes de réservation et de priorité pour gérer les qualités de service. Par exemple un mécanisme particulier d'accès qui entrelace d'une manière naturelle les *slots* de requêtes avec les *slots* de données, permet d'obtenir une latence réduite à faible charge et une capacité maximale optimale. Un protocole en arbre légèrement modifié pour tenir compte du *feedback* différé et de l'entrelacement des réservations a été également présenté au comité.

Une action européenne de R&D dans ce domaine est de plus ressentie comme une nécessité. Il est en effet très possible que le comité IEEE (essentiellement américain) ne converge pas vers un standard adapté au réseau câblé européen. Par conséquent il faudra utiliser les outils développés dans le comité IEEE 802.14 et par d'autres comités en Europe (DAVIC, DVB, ETSI) pour construire à partir de ces briques de base, un protocole MAC performant compatible avec les standards européens. Notons par exemple, que la technologie de voie descendante DVB est déjà disponible sous forme de composants. L'ampleur du mouvement industriel sur ce sujet, comme la très vive concurrence dans le domaine laisse la place à un important travail de synthèse, travail qui profitera également de l'expérience acquise à l'Inria depuis plus de dix ans sur les protocoles d'accès. Il y a là un appel à une collaboration européenne au sujet du développement d'une technologie du modem câble.

L'exploitation des applications qui pourraient déboucher de la généralisation des expériences de type MulticablE offre une autre piste intéressante. Il s'agirait par exemple d'étudier, sur le réseau câblé, les éléments d'un serveur multimédia performant résidant chez l'abonné. Les variations de capacité disponible sur le réseau devront être gérées sur le serveur (mécanismes adaptatifs, méthodes de décompression rapide). Une proposition Esprit/Multimédia a été rédigée en ce sens avec Dassault Électronique et un opérateur du câble concernant la conception et la mise au point d'un serveur multimédia sur le réseau câblé actuel (c'est-à-dire avec les modems câble du commerce). Une autre proposition plus axée sur la partie réseau proprement dite (c'est-à-dire concernant la partie modem câble, tête de station) devrait être déposée d'ici à la fin de l'année.

3.4.3 L'activité réseaux sans fils

Participants : Philippe Jacquet, Olivier Lecomte, Paul Mühlethaler, Amir Qayyum

Le projet ESPRIT LAURA, financé par la Commission européenne, s'est officiellement achevé en 1995. Le projet consistait à définir et à construire un réseau local sans fil compatible avec l'architecture Ethernet et aux performances similaires (10 Mbps). Le but a été atteint avec succès. La normalisation européenne du réseau sans fil HiPeRLAN qui faisait pendant au projet LAURA et auquel l'Inria a activement contribué s'est prolongée jusqu'en 1996, année qui a connu la conclusion de l'enquête publique à l'ETSI et le vote final de la norme qui a eu lieu en août.

La norme a été acceptée sous le numéro ETS 300-652 avec plus de 92 % des voix pondérées des pays représentés à l'ETSI (en fait seule la Suède a voté contre). La norme sera promulguée comme standard technique européen, et reproduite par tous les états membres de l'ETSI (ensemble de l'Europe plus Turquie et Russie). La bande de fréquence 5.15 - 5.25 GHz a été allouée exclusivement aux réseaux HiPeRLAN suite à la confirmation de la décision de l'ERO (TR 22-06). Une extension de cette bande de fréquence jusqu'à 5.3 GHz est d'ores et déjà disponible dans la plupart des pays de la Communauté.

Il est aussi question d'ouverture de bandes similaires aux États-Unis et au Japon pour des réseaux compatibles avec HiPeRLAN. Par exemple, suite à l'initiative WINForum et NII, la FCC (régulation US des fréquences) a décidé d'allouer 300 MHz à 5.2 GHz pour des réseaux compatibles avec HiPeRLAN. Tous les éléments sont réunis pour faire de cette norme la base d'un standard mondial.

HiPeRLAN est une version améliorée du réseau LAURA, ce dernier représentant une sorte de validation technique préalable du premier. HiPeRLAN aura un débit de 24 Mbps sur chacun des cinq canaux dont il dispose dans la bande 5.2 GHz (soit une capacité brute agrégé de 120 Mbps). En plus des éléments déjà développés pour le réseau LAURA, HiPeRLAN dispose aussi d'un mécanisme de gestion des qualités de service pour des trafics multimédia, rendant le produit interfaçable avec ATM par exemple.

L'Inria a contribué avec succès à la définition de la partie protocole de la norme HiPeRLAN. Notre équipe a été représentée à toutes les réunions de normalisation HiPeRLAN organisées pour l'ETSI ; elle était constituée principalement de Ph. Jacquet, P. Mühlethaler puis de Pascale Minet et Nicolas Rivierre du projet REFLECS. L'activité de cette équipe a été déterminante dans trois domaines clefs du protocole.

Premièrement elle a été à l'origine de la définition du protocole d'accès à signalement actif qui marque l'originalité du standard HiPeRLAN.

Deuxièmement, elle a aussi défini le protocole de routage interne qui permet au réseau HiPeRLAN de fonctionner indépendamment de toute infrastructure câblée, sans restriction de topologie et avec une fiabilité accrue. D'ailleurs, à titre anecdotique, lors de l'enquête publique, l'équipe Inria a défendu avec succès le système de routage basé sur l'expérience scientifique acquise depuis Arpanet et Internet.

Troisièmement, l'équipe Inria a tenu un rôle moteur dans la conception du système de priorités et d'ordonnancement qui se trouve à la base de la gestion des qualités de service d'HiPeRLAN. Ce dispositif s'avère indispensable dans la perspective de l'adaptation d'HiPeRLAN à des trafics multimédia autres que des transferts de données. L'ensemble des activités de cette équipe s'est concrétisé par la rédaction de plus d'une cinquantaine de documents techniques sur deux ans au sein du comité de normalisation. Trois articles d'introduction à ses trois principaux domaines d'investigation ont été acceptés dans la revue *Wireless Personal Communications* [291, 290, 289].

Suite aux résultats positifs du projet LAURA, il a été décidé de mener une action de valorisation de ces résultats en portant la norme HiPeRLAN sur des réseaux radio à 2.4 GHz. Ces réseaux auront des performances moindres de celle d'HiPeRLAN (de 1 à 2 Mbps de débit au lieu de 24 Mbps) mais reposent sur une technologie mûre et meilleur marché. La nouvelle norme ETS 300-328 sur la libération de la bande 2.4 GHz fournit le cadre juridique à cette action, car elle permet la réalisation de réseaux sans fil sans exclusivité de fréquences et de protocoles dans la bande 2.4 GHz.

Les cartes réseaux radio à 2.4 GHz existent dans le commerce. Nous avons sélectionné la carte Wavelan de NCR. Dans le cadre de la collaboration avec l'action PRAXITÈLE, ces cartes servent aussi à la liaison sans fil entre les voitures électriques et une base au sol. D'ailleurs PRAXITÈLE sera l'un des premiers bénéficiaires du réseau lorsque celui-ci sera étendu à l'ensemble des fonctions HiPeRLAN.

La valorisation se déroule en deux phases : une phase *software* et une phase *hardware*. La phase *software* consiste à réaliser des *drivers* HiPeRLAN au dessus des cartes 2.4 GHz. Le *driver* contrôle la gestion du routage interne et les fonctions de passerelles pour des stations mixtes air-câble. Ainsi le *driver* sans fil est capable de suivre en temps réel les changements topologiques qui interviennent dans le réseau et le cas échéant de prendre un raccourci par le câble si les conditions de connectivité l'exigent. Deux stagiaires, O. Lecomte de Supélec et A. Qayyum du DEA de parallélisme d'Orsay, ont mené à bien cette entreprise (en détectant d'ailleurs une faute dans la pré-norme HiPeRLAN). Le *driver* est actuellement installé sous LINUX et un portage sous Windows est à l'étude.

La phase *hardware* consiste à programmer une carte réseau intégrant un modem radio à 2.4 GHz. La carte, dont la réalisation est financée par l'intermédiaire de PRAXITÈLE, permettra d'implanter le protocole d'accès par signalement actif et de mettre en œuvre le système de priorité d'HiPeRLAN. Ces fonctions ne pouvaient pas être contrôlées au niveau du *driver* en *software* car elles nécessitent une redéfinition du protocole d'accès déjà intégré sur la carte du commerce. Des fonctions de routage seront aussi transportées sur la carte afin d'alléger le *driver* et de libérer le CPU pour des applications temps réel comme celles de PRAXITÈLE.

4 Actions industrielles

4.1 HiPeRLAN

Le travail de portage du standard HiPeRLAN sur 2.4 GHz est effectué en collaboration avec l'action de développement PRAXITÈLE. L'objectif est de réaliser un réseau sans fil multimédia entre les voiturettes électriques et des structures fixes (voir la section 3.4.3).

4.2 MulticablE

Lyonnaise Communication est notre principal interlocuteur dans l'action que nous menons sur ce sujet. Notre collaboration avec cet industriel est décrite dans la section 3.4.2.

4.3 Calcul formel et actions MAPLE

Le projet ALGO et la compagnie *Waterloo Maple Software* ont développé une collaboration très étroite fondée sur des intérêts réciproques. D'une part il est intéressant pour la compagnie d'intégrer des fonctionnalités à la pointe de la recherche en calcul formel (voir la section 3.2). D'autre part cette intégration fournit aux programmes réalisés par les membres du projet un grand nombre d'utilisateurs d'origines très diverses. Cette relation étroite nous permet également de participer aux choix effectués par les développeurs du système.

De nombreux échanges ont ainsi lieu entre le projet et la compagnie. En juillet 1996, J. Carette est retourné à la compagnie WMS, après une participation pendant plus de trois ans au projet ALGO. De même, E. Murray, après avoir passé plus de deux ans au projet ALGO à programmer le *package* COMBSTRUCT de MAPLE poursuit maintenant ce développement au *Symbolic Computation Group* de l'Université de Waterloo, où le système MAPLE a été créé. Le co-directeur de ce groupe, George Labahn, a passé deux mois et demi au printemps 96 au projet ALGO. À cette occasion, il a invité F. Chyzak au SCG pour y incorporer certains de ses programmes dans la bibliothèque MAPLE.

L'arrivée de MAPLE dans l'enseignement en classes préparatoires nécessite un important travail de formation des enseignants. Les membres du projet participent activement à cet effort. Claude Gomez (projet META2), B. Salvy et P. Zimmermann (projet EURECA) ont écrit un livre l'an dernier sur l'utilisation de MAPLE auquel une mise à jour pour la nouvelle version a été ajoutée cette année. Ph. Dumas, X. Gourdon et F. Chyzak ont formé environ 400 professeurs de classes préparatoires au cours de stages à l'École des mines Nantes et à l'ESTP. En outre, une formation au calcul formel et à MAPLE pour industriels a été coorganisée par P. Zimmermann, Cl. Gomez et B. Salvy à l'Inria-Lorraine, où des membres du projet SAFIR sont également intervenus.

Grâce à ces nombreuses activités autour de MAPLE, la compagnie WMS considère l'Inria comme un partenaire privilégié et lui accorde une licence site gratuite couvrant l'ensemble des centres. Une quinzaine de projets utilisent ce système à des degrés divers.

5 Actions nationales et internationales

5.1 Actions nationales

Ph. Flajolet, qui a dirigé le GDR/PRC "Mathématiques et Informatique" jusqu'en 1994 (50 équipes, 300 membres), participe activement au GDR/PRC "AMI" qui lui fait suite. À ce titre, il est membre du comité de direction et responsable d'un groupe de travail baptisé "Aléa" qui regroupe une quinzaine d'équipes en France. Une réunion générale a été coorganisée avec B. Ycart (Grenoble), et a rassemblé une quarantaine de participants. Ceci procède d'un effort de structuration de la communauté algorithmique en France dans le cadre d'AMI. Ph. Flajolet participe par ailleurs à plusieurs conseils

scientifiques, tels, en 1996, celui du LIX (École polytechnique), du CIMPA (Nice, Éducation Nationale), et celui de l'UFR des sciences de l'Université de Versailles/Saint-Quentin (Prism). Ph. Flajolet reste membre correspondant de l'Académie des Sciences et, à ce titre, participe à ses travaux.

Le séminaire du projet attire régulièrement la communauté d'analyse d'algorithmes de la région parisienne (les universités de Paris 6, Paris-Sud, VersaillesSaint-Quentin, l'École polytechnique). Y ont été présentés cette année, des exposés de synthèse tout comme des travaux récents de nombreux chercheurs tant français (Inria Rocquencourt et Sophia Antipolis, École polytechnique, universités de Paris VI, VII et X, Versailles St-Quentin, Bordeaux et Lyon) qu'étrangers (universités d'Ann Arbor, Canterbury, Oslo, Purdue, Santiago, Vienne et Waterloo, Academia Sinica de Taïwan, Max Planck Institut). Les actes de ces séminaires sont regroupés dans [323].

5.2 Actions internationales

Les collaborations internationales sont nombreuses et le rapport de cette année liste des travaux communs effectués avec les universités de Barcelone, Canterbury, Francfort, Montréal, Princeton, Purdue, Toronto, Vienne, Waterloo. Le projet ALGO est par ailleurs, pour la période de trois ans 1996–1998, l'une des composante du nouveau projet ESPRIT "Long Term Research" ALCOM-IT (*ALgorithms and COMplexity in Information Technology*), sélectionné à l'automne 1995 lors d'une compétition difficile. Les thèmes de recherche en sont l'algorithmique et les structures de données, l'algorithmique parallèle et distribuée, l'optimisation combinatoire, le calcul formel, ainsi que les problèmes correspondants d'évaluation de performance. Les recherches du projet y ont une place centrale, notamment au titre des bibliothèques de combinatoire, analyse d'algorithmes, et calcul formel dont nous assurons la maîtrise d'œuvre.

L'année 1996 voit la fin d'une action franco-slovène entre le projet et l'université de Ljubljana (Slovénie) dans le cadre du programme PROTEUS du ministère des affaires étrangères. Dans le même temps, dans le cadre du programme ALLIANCE du ministère des affaires étrangères, démarre une action franco-britannique avec l'université de Canterbury.

Ph. Jacquet et P. Mühlethaler ont participé aux réunions de l'ETSI concernant l'enquête publique de la norme HiPeRLAN. Ils participent aux réunions du comité en charge de la rédaction des tests de conformance attachés à la norme HiPeRLAN. Ph. Jacquet, P. Mühlethaler et Ph. Robert ont participé aux réunions du comité de normalisation IEEE 802.14 à Montréal (novembre 1995), La Jolla (mars 1996), Los Angeles (mai 1996) et Twente (juillet 1996), pendant lesquelles ils ont présenté quatre documents techniques sur les protocoles de résolution de collisions et de réservation dans les réseaux CATV.

Ph. Flajolet est l'un des organisateurs principaux ainsi qu'initiateur d'un séminaire de haut niveau "*Analysis of Algorithms"*, et a préparé la tenue de la prochaine réunion à Dagstuhl (Allemagne) en juillet 1997. Il s'agit là de contribuer au développement de la communauté internationale en analyse d'algorithmes et aléa combinatoire, thèmes sur lesquels le projet ALGO occupe une place de premier plan. Ph. Flajolet est également membre de l'*Academia Europaea*. (L'*Academia* comporte 1500 membres élus, dans les disciplines littéraires, scientifiques, médicales et juridiques, dont une trentaine au total dans la section d'informatique). Il a été membre du comité de programme d'ESA'96 qui est la conférence européenne d'algorithmique tenue à Barcelone en 1996. Ph. Flajolet est membre des comités d'édition de *Random Structures and Algorithms* (WILEY), *Theoretical Computer Science* (ELSEVIER) et *Maple Technical Newsletter*.

F. Morain a fait partie du Comité de Programmes d'ANTS II (Algorithmic Number Theory Symposium) qui a eu lieu à Bordeaux en mai 1996. Les proceedings de ce congrès sont parus dans la série *Lecture Notes in Computer Science*.

B. Salvy est membre des comités éditoriaux du *Journal of Symbolic Computation* et de la *Maple Technical Newsletter* et membre du comité de programme de FPSAC'97 (conférence internationale de combinatoire algébrique, Vienne, Autriche).

6 Diffusion des résultats

Frédéric Chyzak participe au tronc commun d'informatique à l'École polytechnique. Avec Ph. Dumas, il anime des stages d'initiation et de perfectionnement à MAPLE organisés par l'École des Mines de Nantes et l'ESTP. Ces stages sont destinés à des professeurs de classes préparatoires, dans le cadre de la réforme de l'enseignement. Invité pendant un mois fin 1995 au *Symbolic Computation Group* de l'université de Waterloo (Canada), F. Chyzak a développé des liens avec l'équipe de développement du logiciel MAPLE ainsi qu'avec l'entreprise qui distribue ce système. Une nouvelle visite à l'université de Waterloo doit avoir lieu au cours de l'hiver 1996, avec pour but l'intégration du logiciel MGFUN dans la distribution de MAPLE. Lors d'une visite au *Research Institute for Symbolic Computation*, à Linz (Autriche), F. Chyzak a rapporté ses travaux dans une série d'exposés. Il a également présenté ses recherches à l'université de Marne-la-Vallée.

Philippe Dumas. Depuis l'an dernier les programmes d'enseignement en classes préparatoires aux grandes écoles spécifient l'utilisation de logiciels de calcul formel. Le projet ALGO, soucieux de la qualité de cet enseignement et de l'impact du calcul formel dans le monde scientifique, a décidé de participer activement à la formation des professeurs de ces classes. Trois de ses membres, Ph. Dumas, X. Gourdon et depuis peu F. Chyzak, ont dispensé cette année cinq actions de formation sur l'emploi du système de calcul formel Maple. Ces cours ont touché environ deux cents professeurs. Un livre s'adressant à ce public est en préparation.

Vincent Dumas a été invité au congrès de la SMAI sur les grandes déviations à Toulouse. Il a soutenu sa thèse en décembre 1995 à l'École polytechnique (directeur Ph. Robert). V. Dumas est depuis le 1er octobre 1996 en année post-doctorale au CWI dans l'équipe d'O. Boxma qui étudie les phénomènes de corrélation dans le trafic des réseaux de communication à haut débit.

Philippe Flajolet a été conférencier invité principal à la "Asian Computing Science Conference" (Bangkok, décembre 1995) : *"The digital tree process"*, au "Discrete Mathematics Day" à Carleton University (Ottawa), ainsi qu'à la 7ème conférence "Fibonacci" consacrée aux mathématiques discrètes (Graz, juillet 1996, voir [319]). En 1996, il a participé à plusieurs jurys de thèse ou habilitations, souvent comme rapporteur : M. Bousquet, combinatoire Bordeaux, P. Paule, calcul formel, Linz, D. Pointcheval, cryptographie, Caen, C. Kenyon, complexité, Lyon, I. Dutour, combinatoire, Bordeaux, X. Gourdon, analyse d'algorithmes, Palaiseau, V. Dumas, probabilités, Palaiseau, etc. Il a donné diverses conférences à Grenoble, Versailles, Caen, Barcelone, Paris, Villetaneuse, Bordeaux en nombre réduit cependant cette année par suite de la préparation des ouvrages de synthèse [278, 277], et de leur successeur en cours d'élaboration avec R. Sedgewick (Princeton) ; voir [318]. Il a enseigné avec B. Salvy l'une des options du DEA d'algorithmique , commun aux Grandes Écoles et Universités parisiennes.

Xavier Gourdon a été invité en hiver 95 à Sydney (Australie) par l'équipe des développeurs du système MAGMA pour implanter son algorithme de recherche de racines de polynômes dans leur système et pour décrire ses travaux sur la factorisation de polynômes [302]. Il a également présenté ses travaux au projet SAFIR à Sophia-Antipolis. Il participe avec Ph. Dumas à des formations au calcul formel pour professeurs de classes préparatoires. Il est aussi intervenu dans le tronc commun d'informatique de l'École polytechnique. X. Gourdon a soutenu en juin sa thèse [280] portant sur "Combinatoire, Algorithmique et Géométrie des Polynômes". À la suite de cette thèse, il a été contacté par Dassault Systèmes qui l'a embauché pour travailler sur le logiciel Catia.

Philippe Jacquet a présenté l'activité MulticablE au séminaire Aristote de l'École polytechnique. Il a été invité au séminaire de travail organisé par la commission européenne au sujet d'HiPeRLAN en juillet à Bruxelles. Ph. Jacquet a présenté une conférence au séminaire sur la mécanique statistique des grands réseaux qui s'est tenue à l'Inria en octobre.

Olivier Lecomte a présenté son rapport de stage en juin [326].

Paul Mühlethaler a participé aux réunions du comité AFNOR en charge de l'enquête publique et du vote national sur la norme HiPeRLAN ETS 300-652.

François Morain est Chef de Travaux Pratiques à l'École polytechnique depuis 1992, et chargé à ce titre de faire des TP de programmation (Pascal et C) aux élèves de première année. F. Morain donne

depuis 1991 un cours de Théorie algorithmique des nombres (20 h), d'abord au sein du DEA Informatique Mathématique et Applications, puis maintenant dans le DEA Algorithmique. Depuis septembre 1994, F. Morain encadre le travail de thèse de R. Lercier, qui doit soutenir dans le courant du printemps 1996. F. Morain est un habitué des congrès annuels EUROCRYPT, ainsi que des workshops en théorie algorithmique des nombres qui ont lieu tous les ans aux États-Unis (MSRI, DIMACS), ou bien à Luminy, Dagstuhl, Oberwolfach. En septembre 1995, F. Morain a été conférencier invité au congrès "Computational Perspectives on Number Theory : A conference in honor of A.O.L. Atkin" à Chicago. Son intervention fait l'objet de [327]. Pendant trois semaines en mars 1996, F. Morain a été invité au Newton Institute (Cambridge, Royaume Uni) au sein du Programme "Computer Security, Cryptology and Coding Theory" organisé par Ross Anderson. F. Morain a été conférencier invité à Edinburgh au congrès international "Curves and Computations" fin mars 1996.

Pierre Nicodème a été invité par Martin Vingron, à la suite du Symposium International "Theoretical and Computational Genome Research" qui s'est tenu à Heidelberg du 25 au 27 mars 1996, il a présenté ses travaux sur la recherche de similarités entre séquences protéiques au séminaire de bio-informatique théorique du DKFZ-Heidelberg (Centre de Recherche Allemand sur le Cancer) au mois de juin 96.

Amir Qayyum a présenté la norme HiPeRLAN à la quatrième École d'été distribuée (*International distributed Summer School on Advanced Broadband Communication : ABC'96*), à Bruxelles. Il a présenté son rapport de stage en septembre [329].

Mireille Régnier a participé au Symposium "Theoretical and Computational Genome Research" à Heidelberg et présenté une communication au "Workshop on Mathematical Analysis of Biological Sequences" (Trondheim, août 96). Elle s'est rendue à l'université de Purdue à laquelle nous relie un contrat OTAN (Ph. Jacquet, W. Szpankowski). Elle a fait un exposé aux journées Aléa de Biviers (mai 96). M. Régnier a participé à deux jurys de thèse en bio-informatique et enseigne un cours de DEA à Marne-la-Vallée sur le thème "Génome et Probabilités".

Philippe Robert a été invité au congrès de la SMAI sur les grandes déviations à Toulouse. Il donne un cours de DEA (30h.) intitulé "Files d'attente et modèles probabilistes discrets" au laboratoire de probabilités de l'université de Paris VI. Ph. Robert a été rapporteur de la thèse de M. Brilman, université de Grenoble, soutenue le 30 septembre 1996 et de M. Nafidi, université de Rouen (soutenance prévue en janvier 1997).

Bruno Salvy a été invité aux journées "Exponential Asymptotics" à Luminy, où il a donné un exposé sur "Asymptotics of Implicit Functions and Computer Algebra". Il a participé à ISSAC'96 (conférence internationale de calcul formel), où il a présenté "Asymptotic Expansions of Exp-log Functions" [311]. Il a coorganisé une école Inria de Calcul Formel pour l'industrie à Nancy avec P. Zimmermann (projet EURECA) et Cl. Gomez (projet META2). B. Salvy a fait un exposé sur des algorithmes rapides de manipulation de séries aux universités de La Rochelle, de Rennes, et de Marne-la-Vallée, ainsi qu'à Linz (Autriche) et à Ljubljana (Slovénie). Ses activités d'enseignement comportent une participation au tronc commun d'informatique à l'École polytechnique, un cours d'analyse d'algorithmes en commun avec Ph. Flajolet au DEA Algorithmique (École polytechnique, ENS, Paris VI, VII et X) et un cours sur les fonctions spéciales en calcul formel au DEA Informatique Fondamentale et Applications de l'université de Marne-la-Vallée.

7 Publications

Livres et monographies

[277] R. SEDGEWICK, P. FLAJOLET, *Introduction à l'analyse des algorithmes*, International Thomson Publishing, France, 1996, 492 pages. A translation of the original English version (ISBN 2-84180-957-9).

[278] R. SEDGEWICK, P. FLAJOLET, *An Introduction to the Analysis of Algorithms*, Addison-Wesley Publishing Company, 1996, 512 pages. (ISBN 0-201-4009-X).

Thèses

[279] V. DUMAS, *Approches fluides pour la stabilité et l'instabilité de réseaux de files d'attente stochastiques à plusieurs classes de clients*, thèse de doctorat, École polytechnique, décembre 1995.

[280] X. GOURDON, *Combinatoire, Algorithmique et Géométrie des Polynômes*, thèse de doctorat, École polytechnique, juin 1996.

Articles et chapitres de livre

[281] S. BOUCHERON, D. GARDY, «An urn model from learning theory», *Random Structures & Algorithms*, 1997, Special issue on Analysis of Algorithms. To appear.

[282] H. DAUDÉ, P. FLAJOLET, B. VALLÉE, «An Average-case Analysis of the Gaussian Algorithm for Lattice Reduction», *Combinatorics, Probability and Computing*, 1996, 30 pages. To appear.

[283] P. DUMAS, P. FLAJOLET, «Asymptotique des récurrences mahleriennes : le cas cyclotomique», *Journal de Théorie des Nombres de Bordeaux 8*, 1, juin 1996, p. 1–30.

[284] P. DUMAS, B. SALVY, «Maple and the Putnam Competition», *Maple Technical Newsletter 2*, 2, 1995, p. 63–68.

[285] P. FLAJOLET, B. SALVY, «Computer Algebra Libraries for Combinatorial Structures», *Journal of Symbolic Computation 20*, 1995, p. 653–671.

[286] P. FLAJOLET, *Encyclopedia of Mathematics*, Kluwer Academic Publishers, Dordrecht, 1996, ch. Adaptive Sampling, To appear.

[287] X. GOURDON, B. SALVY, «Effective asymptotics of linear recurrences with rational coefficients», *Discrete Mathematics 153*, 1–3, 1996, p. 145–163.

[288] D. GUILLAUME, F. MORAIN, «Building pseudoprimes with a large number of prime factors», *Applicable Algebra in Engineering, Communication and Computing 7*, 4, 1996, p. 263–277.

[289] P. JACQUET, P. MINET, P. MÜHLETHALER, N. RIVIERRE, «Data transfer in HiPeRLAN», *Wireless Personal Communications*, 1996, To appear.

[290] P. JACQUET, P. MINET, P. MÜHLETHALER, N. RIVIERRE, «Increasing reliability in cable-free Radio LANs: Low level forwarding in HiPeRLAN», *Wireless Personal Communications*, 1996, To appear.

[291] P. JACQUET, P. MINET, P. MÜHLETHALER, N. RIVIERRE, «Priority and collision detection with active signaling: the channel access mechanism of HiPeRLAN», *Wireless Personal Communications*, To appear, 1996.

[292] P. JACQUET, W. SZPANKOWSKI, «Analytical depoissonization and its applications», *Theoretical Computer Science*, 1996, 68 pp.

[293] A. JOUX, F. MORAIN, «Sur les sommes de caractères liées aux courbes elliptiques à multiplication complexe», *Journal of Number Theory 55*, 1, novembre 1995, p. 108–128.

[294] F. LEPRÉVOST, F. MORAIN, «Revêtements de courbes elliptiques à multiplication complexe par des courbes hyperelliptiques et sommes de caractères», *Journal of Number Theory*, 1996, To appear.

[295] S. P. LIPSHITZ, T. C. SCOTT, B. SALVY, «On the Acoustic Impedance of Baffled Strip Radiators», *Journal of the Audio Engineering Society 43*, 7/8, 1995, p. 573–580.

[296] H. MAHMOUD, M. RÉGNIER, R. SMYTHE, «Analysis of Boyer-Moore-Horspool String-Matching Heuristic», *Random Structures & Algorithms*, 1996, To appear.

[297] F. MORAIN, J. SHALLIT, H. C. WILLIAMS, «La machine à congruences», *La revue du Musée des Arts et Métiers 14*, mars 1996, p. 14–19.

[298] F. MORAIN, «Calcul du nombre de points sur une courbe elliptique dans un corps fini : aspects algorithmiques», *Journal de Théorie des Nombres de Bordeaux 7*, 1995, p. 255–282.

[299] J. SHACKELL, B. SALVY, «Asymptotic Forms and Algebraic Differential Equations», *Journal of Symbolic Computation 20*, 1995, p. 169–177.

Communications à des congrès, colloques, etc.

[300] J.-P. DEDIEU, X. GOURDON, J.-C. YAKOUBSOHN, «Computing the Distance from a Point to an Algebraic Hypersurface», *in : Proceedings of the American Mathematical Society, seminar of Park City on Mathematics of Numerical Analysis: Real Number Algorithms*, J. Renegar, M. Shub, S. Smale (réd.), juillet 1996. 8 pages. In press.

[301] V. DUMAS, «Analysis of the stability of the Cambridge ring», *in : Journées SMAI modélisation aléatoire et statistique*, SMAI, Toulouse, septembre 1996.

[302] P. FLAJOLET, X. GOURDON, D. PANARIO, «Random Polynomials and Polynomial Factorization», *in : Automata, Languages, and Programming*, F. Meyer auf der Heide, B. Monien (réd.), *Lecture Notes in Computer Science*, 1099, p. 232–243, 1996. Proceedings of the 23rd ICALP Conference, Paderborn, July 1996.

[303] P. FLAJOLET, «Analytic Variations on Quadtrees», *in : Notes of the Seminar on Probabilistic Methods in Algorithmics*, *Quaderns, Centre de Recerca Matemàtica*, 5, p. 44–53, Barcelona, 1996. (Summary written by Nicola Galesi).

[304] P. JACQUET, P. MÜHLETHALER, P. ROBERT, «Slotted multiple access MAC with collision detection proposal with priority management on Cable TV», *in : IEEE plenary 802 meeting*, IEEE 802.14, Montréal, novembre 1995.

[305] P. JACQUET, P. MÜHLETHALER, P. ROBERT, «Performant implementations of tree collision resolution on CATV network», *in : IEEE plenary 802 meeting*, IEEE 802.14, Los Angeles, mai 1996.

[306] P. JACQUET, P. MÜHLETHALER, P. ROBERT, «Simulation of a stack algorithm with priorities and reservation», *in : IEEE plenary 802 meeting*, IEEE 802.14, Twente, juillet 1996.

[307] P. JACQUET, P. MÜHLETHALER, P. ROBERT, «A unified approach for CATV networks capable of mixing ATM and non ATM traffics with CBR and non CBR traffics», *in : IEEE plenary 802 meeting*, IEEE 802.14, Twente, juillet 1996.

[308] JACQUET, PH. AND MÜHLETHALER, P. AND ROBERT, PH., «Performant implementations of tree collision resolution with large feedback time», *in : International workshop on Mobile Communications*, Thessalonique, 1996.

[309] M. RÉGNIER, L. ROSTAMI, «A Simple (but Optimal) 2D-Witness Algorithm», *in : WSP'96*, *International Informatics Series*, 4, 1996. Third South American Workshop on String Processing.

[310] M. RÉGNIER, F. TAHI, «Enumeration and Asymptotics in Computational Biology», *in : Mathematical Analysis for Biological Sequences*, 1996. Proceedings of a workshop held in Trondheim, Norway.

[311] D. RICHARDSON, B. SALVY, J. SHACKELL, J. VAN DER HOEVEN, «Asymptotic Expansions of exp-log Functions», *in : ISSAC'96*, Y. N. Lakshman (réd.), ACM Press, p. 309–313, 1996. Proceedings of the 1996 International Symposium on Symbolic and Algebraic Computation. July 24–26, 1996. Zurich, Switzerland.

[312] P. ROBERT, «Synchronisation d'un système de ressources et problèmes reliés», *in : Journées SMAI modélisation aléatoire et statistique*, SMAI, Toulouse, septembre 1996.

Rapports de recherche et publications internes

[313] F. CHYZAK, B. SALVY, «Non-commutative Elimination in Ore Algebras Proves Multivariate Holonomic Identities», *Research Report n°2799*, Institut National de Recherche en Informatique et en Automatique, février 1996, Submitted.

[314] J.-M. COUVEIGNES, L. DEWAGHE, F. MORAIN, «Isogeny cycles and the Schoof-Elkies-Atkin algorithm», *Research Report n°LIX/RR/96/03*, LIX, avril 1996.

[315] P. FLAJOLET, X. GOURDON, C. MARTÍNEZ, «Patterns in random binary search trees», *Research Report n°2997*, Institut National de Recherche en Informatique et en Automatique, octobre 1996, 23 pages. Submitted to *Random Structures & Algorithms*.

[316] P. FLAJOLET, R. KEMP, H. PRODINGER, R. SEDGEWICK (EDITORS), «Average Case Analysis of Algorithms», *Dagstuhl Seminar Reports n°119*, IBFI GmbH Schloß Dagstuhl, 1996, Summary of talks presented at a seminar, Wadern, Germany, July 3–7, 1995.

[317] P. FLAJOLET, B. SALVY, «Euler Sums and Contour Integral Representations», *Research Report n°2917*, Institut National de Recherche en Informatique et en Automatique, juin 1996, 23 pages. Submitted to the *Journal of Experimental Mathematics*.

[318] P. FLAJOLET, R. SEDGEWICK, «The Average Case Analysis of Algorithms: Mellin Transform Asymptotics», *Research Report n°2956*, Institut National de Recherche en Informatique et en Automatique, 1996, 93 pages.

[319] P. FLAJOLET, B. VALLÉE, «Continued Fraction Algorithms, Functional Operators, and Structure Constants», *Research Report n°2931*, Institut National de Recherche en Informatique et en Automatique, juillet 1996, 33 pages. (Invited lecture at the 7th Fibonacci Conference, Graz, July 1996.).

[320] P. NICODÈME, «BlastMultAl, a Blast Extension for Similarity Searching with Alignment Graphs», *Research Report n°2911*, Institut National de Recherche en Informatique et en Automatique, juin 1996.

[321] M. RÉGNIER, W. SZPANKOWSKI, «Exact Complexity of Sequential Pattern Matching Algorithms», *Research Report n°2549*, Institut National de Recherche en Informatique et en Automatique, 1996.

[322] B. SALVY, J. SHACKELL, «Symbolic Asymptotics: Functions of Two Variables, Implicit Functions», *Research Report n°2883*, Institut National de Recherche en Informatique et en Automatique, mai 1996.

[323] B. SALVY (EDITOR), «Algorithms Seminar, 1995–1996», *Research Report n°2992*, Institut National de Recherche en Informatique et en Automatique, septembre 1996.

Divers

[324] L. DEVROYE, P. FLAJOLET, F. HURTADO, M. NOY, W. STEIGER, «Random Triangulations», avril 1996, Submitted. 10 pages.

[325] P. JACQUET, P. MÜHLETHALER, «Data transmission device for random access network, with improved collision resolution, and corresponding method», 1996, US patent, No 5 517 501.

[326] O. LECOMTE, «Réseaux sans fil: Adaptation d'HiPeRLAN type 1 sur Wavelan», 1996, Mémoire de DEA, Supelec.

[327] R. LERCIER, F. MORAIN, «Algorithms for computing isogenies between elliptic curves», To appear in the Proceedings of the Atkin conference, avril 1996.

[328] F. MORAIN, «Classes d'isomorphismes des courbes elliptiques supersingulières en caractéristique ≥ 3», Submitted, mars 1996.

[329] A. QAYYUM, «Wireless Networks: HiPeRLAN», 1996, Mémoire de DEA, Université de Paris-Sud.

[330] M. RÉGNIER, W. SZPANKOWSKI, «A Last Word on Pattern Frequency Occurences in a Markovian Sequence», 1996, Submitted.

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.
- Computer algebra, especially new functionalities related to combinatorial libraries, revues and asymptotics.

The project is a component of the MULTICABLE Project and the ALCOM-IT Basic Research Action of the European Union.