

*Projet ALGO**Algorithmes**Rocquencourt*

THÈME 2B



*R*apport
d'Activité

1999

Table des matières

1	Composition de l'équipe	2
2	Présentation et objectifs généraux	3
3	Fondements scientifiques	5
3.1	Analyse d'algorithmes	5
3.2	Calcul formel	6
3.3	Algorithmiques des séquences	9
3.4	Algorithmique et modélisation des réseaux	10
4	Domaines d'applications	11
4.1	Panorama	11
5	Logiciels	12
6	Résultats nouveaux	13
6.1	Analyse d'algorithmes	13
6.2	Calcul formel	16
6.3	Algorithmiques des séquences	17
6.4	Algorithmique et modélisation des réseaux	18
7	Contrats industriels (nationaux, européens et internationaux)	20
7.1	Calcul formel	20
7.2	Algorithmique et modélisation des réseaux	21
8	Actions régionales, nationales et internationales	21
8.1	Actions nationales	21
8.2	Actions financées par la Commission Européenne	22
8.3	Relations bilatérales internationales	22
8.4	Accueils de chercheurs étrangers	22
9	Diffusion de résultats	23
9.1	Animation de la communauté scientifique	23
9.2	Enseignement universitaire	23
9.3	Participation à des colloques, séminaires, invitations	24
10	Bibliographie	25

1 Composition de l'équipe

Responsable scientifique

Philippe Flajolet [DR]

Responsable permanent

Mireille Régnier [DR]

Assistante de projet

Virginie Collette [TR]

Personnel Inria

Philippe Robert [DR]

Bruno Salvy [CR]

Frédéric Chyzak [CR]

Collaborateurs extérieurs

Philippe Dumas [professeur, Cl. Prépa. lycée Jean-Baptiste Say]

François Morain [Dret et École polytechnique]

Pierre Nicodème [DKFZ, Heidelberg]

Michèle Soria [professeur, Université Paris VI]

Brigitte Vallée [professeur, Université de Caen]

Chercheurs invités

Bernard Chazelle [professeur, Princeton University, à partir du 1er septembre 1998]

Doctorants

Cyril Banderier [Université de Paris VI]

Julien Clément [Université de Caen (temps partiel)]

Jean-François Dantzer [Prag à l'Université de Versailles St-Quentin]

Mostapha Haddani [Université de Versailles St-Quentin]

Elchanan Mossan [Université Hébraïque, Israël]

Stagiaires

Wenchang Chu [boursier CIES, Université de Rome, du 1er mars au 31 décembre 1998]

Juliette Dorléans [INSA Toulouse, mai 1999 à juillet 1999]

Cyril Germa [INSA Toulouse, mai 1999 à juillet 1999]

2 Présentation et objectifs généraux

L'objectif global du projet ALGO est l'analyse et l'optimisation fines de systèmes complexes discrets présentant une forte composante aléatoire. De nombreux problèmes de grande taille rentrent dans ce cadre tels l'étude quantitative d'algorithmes probabilistes sur des structures discrètes, ou l'optimisation de l'allocation de ressources dans les réseaux de communication. La réalisation de cet objectif passe par la compréhension en profondeur de l'aléa discret et des problèmes de nature mathématique que pose sa quantification. Cela suppose de dégager des méthodes générales pour obtenir des résultats présentés de manière exacte ou asymptotique. Ces résultats fournissent alors des informations très précises sur le comportement qualitatif ou quantitatif des systèmes étudiés.

Étant donné le caractère très systématique de l'approche poursuivie, des méthodes de décision réalisables en calcul formel font aussi partie des objectifs du projet. Cette approche est un moteur puissant de renouvellement qui conduit à la révision d'approches classiques dans le domaine des fonctions spéciales et des développements en séries. L'objectif est de disposer d'une algorithmique fiable et complète pour de grandes classes de problèmes précisément caractérisés ; voir déjà les bibliothèques GFUN et MGFUN assez largement utilisées dans la communauté combinatoire et présentes dans la dernière version de Maple. Les résultats sont susceptibles de nombreuses applications bien au delà du domaine de la modélisation combinatoire : ainsi une meilleure intégration des fonctions spéciales au calcul formel est visée, ce qui s'applique à de larges classes de problèmes dans les sciences de l'ingénieur.

Thématique de recherche

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. Ici, nous sommes dans le domaine des mathématiques discrètes. La combinatoire étant par définition l'étude des objets finis discrets, nous visons à développer une approche globale que l'on pourrait qualifier de "combinatoire statistique" par analogie avec la physique statistique. Le but est de décrire le comportement macroscopique visible de l'objet étudié, comme par exemple, le temps d'exécution sur un grand "ensemble" d'un certain algorithme, ou encore l'évolution d'un grand réseau vers ses régimes fondamentaux.

Il s'agit ainsi de caractériser les propriétés attendues (en moyenne, en probabilité ou 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, un tri de seulement 100

éléments met en jeu 10^{158} configurations possibles, qui obéissent, avec une écrasante probabilité, à des règles fort précises. Les problèmes d'aléa combinatoire interviennent de manière essentielle en algorithmique. 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".

Le rôle de l'analyse de l'aléa combinatoire est 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 un substitut souvent très efficace aux arbres de recherche, les signatures accélèrent considérablement la recherche textuelle, les "skip lists" remplacent graduellement les arbres équilibrés dans de nombreuses applications. Parmi d'autres applications célèbres de cette "aléatorisation", on peut citer la construction de cryptosystèmes à clés publiques qui utilise de manière très sûre des tests probabilistes ; ou encore la conception de protocoles de communication, faisant suite à Ethernet, mais avec une utilisation mieux contrôlée de l'aléa qui permettent de stabiliser et d'acheminer des trafics plus importants dans les réseaux. La problématique de l'algorithmique randomisée est poursuivie par exemple avec succès par le projet PRISME de Sophia-Antipolis dans le domaine de l'algorithmique géométrique et par le projet HIPERCOM de Rocquencourt dans le domaine des télécommunications.

Le développement récent du *calcul formel* a éclairé tout l'intérêt de cette approche généraliste. 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, cela a donné lieu aux thèses de B. Salvy et P. Zimmermann (en 1991). Le programme de recherche correspondant à ces aspects est loin d'être achevé. De nombreux problèmes de portée générale en calcul formel ont été mis en évidence. Citons principalement ici, comme directions nouvelles, l'asymptotique automatique (B. Salvy), les méthodes mixtes symboliques numériques (B. Salvy), et la preuve automatique d'identités combinatoires (F. Chyzak, B. Salvy).

C'est dans ce contexte qu'ont pu être résolues au fil des ans diverses conjectures correspondant à des analyses précises d'algorithmes tels que le dimensionnement en hachage dynamique (M. Régnier, Ph. Flajolet), 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), les protocoles en arbre (Ph. Flajolet et Ph. Jacquet), et le comportement probabiliste de grands réseaux (Ph. Robert).

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 HIPERCOM (analyse d'algorithmes et protocoles ; théorie analytique de l'information).

Elles nous valent de nombreuses coopérations internationales avec des universités comme celles de Barcelone, Newcastle, Princeton, Purdue, Stanford, Vienne, Washington, Waterloo, etc.

3 Fondements scientifiques

3.1 Analyse d'algorithmes

Participants : Cyril Banderier, Julien Clément, Philippe Flajolet, Michèle Soria, Brigitte Vallée.

Mots clés : analyse d'algorithmes, analyse combinatoire, combinatoire analytique, structures aléatoires discrètes, arbres d'index, algorithmes de hachage, algorithmes géométriques, lois limites.

Résumé : *L'analyse d'algorithmes vise à une quantification précise des principaux algorithmes et structures de données de l'informatique. Elle repose sur une étude approfondie de l'aléa discret. Une approche originale développée au sein du projet se fonde sur la "combinatoire analytique". Les applications en sont les principaux problèmes de recherche de l'information dans de grands volumes de données et l'algorithmique de la communication.*

L'analyse d'algorithmes ou de structures de données dépend très étroitement d'une évaluation et d'une quantification précise de l'aléa discret. Il s'agit en effet de caractériser le comportement des principaux paramètres de structures combinatoires d'arbres, de mots, de graphes ou d'allocations aléatoires. Dans ce domaine, les travaux du projet ALGO s'appuient sur une approche originale fondée sur

- les méthodes symboliques en analyse combinatoire qui permettent de disposer de méthodes systématiques de mise en équation par séries génératrices ;
- les méthodes asymptotiques complexes fondées notamment sur l'analyse de singularité.

L'application de ces idées aux séquences textuelles ou génétiques fait l'objet de la section 3.3. L'automatisation de ces méthodes visant à l'analyse complète de modèles complexes est traitée à la section 3.2.

Traditionnellement, l'analyse d'algorithmes propose des prédictions quantitatives de comportement moyen, c'est-à-dire des analyses d'espérances de coûts. Une évolution importante qui commence au début des années 1990 avec la thèse de doctorat ès-sciences de Michèle Soria consiste à aborder des études en distribution. Ceci permet non seulement de connaître les comportements moyens attendus mais encore de prédire les profils de coûts sous la forme de distributions limites ou de bornes de grandes déviations. En d'autres termes, on parvient depuis peu à quantifier très précisément les compromis de type *risque-efficacité*. Du point de vue mathématique, les travaux établissent un pont intéressant entre approches probabilistes classiques et approches issues de la combinatoire analytique. Une monographie intitulée "Analytic Combinatorics" par P. Flajolet et R. Sedgewick est en préparation (500 pages sont déjà rédigées et publiées en rapport de recherche INRIA). Cette monographie fait partie d'un programme de recherche visant à établir un corpus cohérent de méthodes pour le domaine de la combinatoire analytique et pour la quantification de l'aléa discret. Les méthodes qui sous-tendent

ce programme ont de nombreuses applications en analyse d'algorithmes, parmi lesquelles sont traités spécifiquement dans le projet :

- les structures de données fondamentales comme les tables de hachage, les arbres de recherche uni- ou multi-dimensionnels (voir la section 6.1) ;
- l'algorithmique de la recherche textuelle et du traitement de séquences génomiques (voir la section 6.3) ;
- les arbres digitaux et les arbres liés aux protocoles en arbre (voir les sections 6.1 et 7.2).

Ces méthodes inspirent également une partie des recherches en calcul formel (section 6.2) par les problèmes que pose l'automatisation de la manipulation de grands modèles combinatoires.

3.2 Calcul formel

Participants : Frédéric Chyzak, François Morain, Bruno Salvy.

Mots clés : échelles asymptotiques, génération aléatoire, fonctions spéciales, élimination polynomiale, bases de Gröbner, courbes elliptiques.

Résumé : *Le projet ALGO développe des algorithmes de calcul formel qui permettent: le traitement de modèles combinatoires récurrents; l'analyse asymptotique automatique de nombreuses classes de problèmes; le traitement de fonctions spéciales et de suites combinatoires de manière unifiée, et notamment leur intégration ou leur sommation; une arithmétique efficace en très grande précision. Ces algorithmes sont dans la plupart des cas validés par des implantations. Leur développement est motivé par l'automatisation d'une approche combinatoire à l'analyse d'algorithmes.*

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 explique l'importance de l'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 été élaborés de nombreux algorithmes et programmes. Les domaines mathématiques permettant l'expression et la preuve de nos algorithmes sont la combinatoire, l'analyse asymptotique, l'algèbre différentielle (corps de Hardy et anneaux de polynômes non commutatifs) ou les propriétés des courbes elliptiques sur les corps finis. Notre travail en calcul formel est ainsi complémentaire de celui des projets SAGA et CAFE consacrés aux systèmes polynomiaux et aux équations différentielles.

Comme pour l'analyse d'algorithmes, l'approche du projet aux algorithmes du calcul formel est globalisante et unificatrice. La résolution de problèmes appliqués mettant en œuvre des cryptosystèmes, de la combinatoire ou de l'analyse d'algorithmes est abordée à un niveau

de généralité qui permet le développement d'une algorithmique de portée large. 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. Les outils développés dans le projet 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 asymptotique. La diversité de ce public est encore accrue par la présence de certains de nos programmes dans les bibliothèques du système de calcul formel MAPLE, l'un des deux principaux systèmes généralistes actuellement disponibles.

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 en arithmétique.

Structures combinatoires 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, il est possible de (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 ; (v) calculer des équations satisfaites par les séries génératrices d'énumération de ces objets — utiles pour la phase d'analyse asymptotique.

Le programme $\Lambda\Omega$ réalisé au début des années 90 par B. Salvy et P. Zimmermann fournissait une partie de ces fonctionnalités, mais sa portabilité et ses fonctionnalités étaient limitées par l'usage de CAML en conjonction avec MAPLE. L'objectif poursuivi depuis plusieurs années est de tirer parti de l'expérience acquise avec $\Lambda\Omega$ pour réaliser une version (COMBSTRUCT) entièrement intégrée en MAPLE, en mettant l'accent sur la modularité, la robustesse et la souplesse d'emploi.

Suites et fonctions spéciales 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 ne sont le plus souvent que décrites sous forme implicite 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, implantée en grande partie dans le *package* GFUN développé par B. Salvy et P. Zimmermann (projet POLKA). Le pendant multivarié de ces travaux est extrêmement prometteur, puisqu'il permet d'envisager une approche algorithmique de nombreux calculs avec des polynômes orthogonaux, des fonctions spéciales, ou plus généralement des fonctions, suites, séries ou distributions définies par un *système* d'équations linéaires à coefficients polynomiaux. Le travail mené par F. Chyzak considère des systèmes d'opérateurs linéaires qui peuvent être différentiels, aux différences, aux q -différences, ou de nombreux autres types. Il s'avère que les opérations d'addition, de produit, d'intégration ou de sommation peuvent s'effectuer algorithmiquement par le biais de techniques d'élimination sophistiquées dans des algèbres non commutatives d'opérateurs linéaires. Le *package* MGFUN développé par F. Chyzak valide et motive les recherches théoriques correspondantes.

Séries et échelles asymptotiques 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 (University of 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. Depuis lors, grâce à une série de travaux en commun de B. Salvy et J. Shackell, des classes de plus en plus vastes de fonctions ont reçu un algorithme permettant le calcul asymptotique de manière garantie. Ces progrès n'ont pas toujours été suivis d'implantation, mais une nouvelle structure de données pour les développements asymptotiques, les multiséries, a vu le jour récemment. L'algorithmisation de cette structure de données conduit à revisiter et prolonger les algorithmes anciens, et mène à la réalisation de prototypes permettant déjà de résoudre des problèmes du niveau de la recherche. Cette structure de données a fait l'objet d'une proposition aux développeurs du système MAPLE. Là encore, on s'attaque à une brique essentielle des systèmes de calcul formel, puisque les séries sont la structure de données efficace utilisée pour manipuler les polynômes, eux-mêmes à la base de nombreuses structures de données.

Arithmétique en très grande précision La cryptologie à clefs publiques repose principalement sur l'arithmétique des grands nombres. C'est un thème de base de l'algorithmique, et elle bénéficie de l'apport d'autres branches, comme par exemple la complexité intrinsèque et l'analyse d'algorithmes.

Un axe de recherche récent étudié par F. Morain concerne les aspects algorithmiques liés à l'utilisation des courbes elliptiques sur les corps finis. Ces objets prennent une importance grandissante dans les applications industrielles liées à la cryptologie et au commerce électronique. Une des retombées intéressantes des progrès faits dans les domaines évoqués précédemment, hors la cryptologie, concerne le calcul formel. Tous les systèmes de calcul symbolique disposent d'une arithmétique entière en précision arbitraire sur laquelle est ensuite développée tout un empilement de couches logicielles de plus en plus complexes. Un des exemples récents concerne la factorisation de polynômes dans les corps finis, qui a bénéficié de découvertes récentes de Shoup, faites en relation avec des algorithmes puissants traitant des courbes elliptiques.

F. Morain a pour rattachement principal le LIX (École Polytechnique) dans le cadre duquel sont développés et valorisés les aspects cryptographiques de son action (voir le rapport d'activité du LIX). Il participe à l'Action de Recherche Coordonnée (ARC) Courbes où les compétences sont mises en commun (avec le Projet CODES et l'Université de Limoges) pour traiter des problèmes liés à l'utilisation des courbes elliptiques et hyperelliptiques en cryptographie.

3.3 Algorithmiques des séquences

Participants : Frédéric Chyzak, Philippe Dumas, Philippe Flajolet, Pierre Nicodème, Mireille Régnier, Bruno Salvy.

Mots clés : combinatoire des mots, séquences, recherche de motifs, génome.

Résumé : *L'objet des recherches sur les séquences est la conception de nouveaux algorithmes, l'obtention de la complexité moyenne de ces algorithmes et l'application à l'algorithmique de certains résultats statistiques. Plus généralement, nous développons une théorie analytique de l'information qui s'appuie sur la combinatoire, les probabilités et l'analyse.*

L'algorithmique des séquences ou objets textuels couvre des domaines d'application variés (compression des données textuelles et séquences biologiques). Ce sujet comprend d'abord des recherches algorithmiques. Il s'agit de trouver efficacement un motif ou un ensemble de motifs dans un texte. Cet ensemble peut en particulier être l'ensemble des mots voisins d'un mot donné, à un nombre d'erreurs près, où le type et le nombre d'erreurs autorisées sont déterminés par l'application. Une nouvelle classe d'algorithmes dits d'extraction de motifs, émerge dans les recherches sur le génome. On extrait des motifs à signification biologique particulière, non connus à l'avance. Ce sont par exemple des mots à fréquence exceptionnelle : surreprésentés ou sousreprésentés pour un modèle probabiliste donné. On peut aussi rechercher une expression régulière caractéristique d'une famille donnée. Ces algorithmes testent sur une même séquence un grand nombre de candidats potentiels et la rapidité du calcul statistique, et sa précision, apparaissent essentielles.

D'un point de vue méthodologique, des 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. Les formules obtenues sont valables dans le domaine fini. Parallèlement, nous développons des outils de calcul dans le modèle probabiliste markovien. 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. Ceci équivaut à une agrégation des états de l'automate associé. Les résultats s'appliquent à la compression et à la recherche de motifs avec erreurs ou de motifs exceptionnels dans les textes (DosDNA, reconnaissance de gènes, etc.). Ils permettent aussi d'établir les domaines d'efficacité des différents algorithmes de recherche de similarités dans des bases de données protéiques.

3.4 Algorithmique et modélisation des réseaux

Participants : Jean-François Dantzer, Mostafa Haddani, Philippe Robert.

Mots clés : grands réseaux, méthodes de renormalisation, vitesse de convergence.

Résumé : *L'étude quantitative des modèles probabilistes de réseaux, complexe dès lors que le nombre de nœuds est grand, est envisagée du point de vue des techniques de renormalisation de la physique statistique. L'état du réseau est renormalisé avec un paramètre (la taille, la capacité du réseau, l'intensité du trafic) qui tend vers une valeur critique. Asymptotiquement, l'étude du comportement du réseau se ramène à la résolution d'équations différentielles quasi-déterministes.*

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 distribution à l'équilibre 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. Le processus limite obtenu de cette façon 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 "limites" peuvent présenter des solutions pathologiques qu'il convient d'éliminer (les physiciens le font avec des arguments d'entropie). La renormalisation présente l'intérêt de mettre en évidence les modes de fonctionnement fondamentaux. Ce programme d'étude a déjà largement été entamé dans la thèse de V. Dumas et développé par J.-F. Dantzer dans le cadre des réseaux multi-classe et par M. Haddani dans une thèse en cours sur les questions d'allocation de bande passante.

Les problèmes de vitesse de convergence constituent un aspect important de notre étude. Ils se situent en amont des questions de renormalisation. Schématiquement, l'étude d'un système donné peut se décrire de la façon suivante : étude de la renormalisation, propriétés gaussiennes

autour des trajectoires renormalisées, existence d'un principe de grandes déviations, et enfin étude de la convergence à l'équilibre des modèles considérés. Jusqu'à présent, les résultats obtenus dans ce domaine sont principalement des estimations du taux de convergence exponentielle par rapport au temps, pour obtenir des bornes du type $Ke^{-\lambda_2 t}$. La constante K , qui est généralement négligée, joue cependant assez souvent un rôle majeur dans la convergence à l'équilibre. Par exemple, l'étude de la seconde valeur propre λ_2 suggère souvent un temps de relaxation (le temps pour que la distance à l'équilibre soit inférieur à $1/e$) de l'ordre de $1/\lambda_2$. Il est cependant très fréquent que l'ordre de grandeur du véritable temps de relaxation soit complètement différent. L'accumulation de valeurs propres au voisinage de λ_2 semble être à l'origine de ces phénomènes encore mal compris. Notre objectif principal dans ce cadre est de donner des *bornes* aussi simples et précises que possible sur la distance entre l'état d'un système à un instant donné et ce même système à l'équilibre. L'intérêt de tels résultats est de pouvoir donner une description du comportement transitoire d'un système qui n'est pas à l'équilibre. Si l'état d'équilibre a, en général, une expression difficile à utiliser, le cas transitoire est en revanche pratiquement toujours inconnu. Ce type d'étude a été mené avec succès, avec des techniques d'analyse de Fourier. Les marches aléatoires sur des graphes ayant une structure assez régulière s'étudient avec ces méthodes, voir par exemple le livre de Diaconis "Group representations in probability and statistics" sur ce sujet. Les techniques envisagées pour cette étude sont essentiellement probabilistes. Notre objectif est d'étudier dans un premier temps des réseaux simples et de dégager des méthodes d'investigation. Dans ce cadre aussi, nous avons recours à une variable de renormalisation qui permet de dégager le mode majeur de convergence. Outre les bornes explicites sur la distance à l'équilibre, ce type d'étude permet de définir, dans certaines situations, le temps d'atteinte de l'équilibre ; avant ce temps, il est possible de trouver un état initial pour lequel la distance est maximale, et après la distance est nulle.

4 Domaines d'applications

4.1 Panorama

Résumé : *Les applications de nos travaux sur les structures combinatoires sont la modélisation et l'étude des systèmes discrets complexes et des réseaux de télécommunication.*

Les applications visées par l'analyse d'algorithmes sont la conception de méthodes d'accès rapide à des informations structurées, une algorithmique rapide du calcul formel, le traitement statistique des séquences biologiques, et les protocoles de communication.

Nos domaines de recherche en calcul formel sont : les structures combinatoires, les suites et fonctions spéciales, l'analyse asymptotique et la théorie des nombres. Les applications de nos travaux sur les structures combinatoires sont la modélisation et l'étude de systèmes discrets complexes. Nos résultats sur les suites et fonctions spéciales débouchent sur la manipulation par le calcul formel de fonctions spéciales intervenant de manière classique en physique mathématique et en mécanique. Nos travaux sur l'asymptotique devraient permettre à long terme de faire le pont entre le calcul numérique et le calcul formel : le calcul numérique, robuste

en l'absence de singularités pourrait être complété par une étude formelle fine au voisinage des singularités débouchant sur de la production de code numérique robuste dans ces zones sensibles. Par ailleurs les travaux de F. Morain sur les courbes elliptiques ont des applications en cryptologie.

Les travaux sur les réseaux sont naturellement liés aux problèmes de dimensionnement, de routages des messages, ou encore d'optimisation des paramètres de qualité de service (délai, taux de rejet, etc.). Les grandes tailles des réseaux, les hauts débits utilisés et l'hétérogénéité des trafics (vidéo, voix, données) qui y circulent rendent de plus en plus nécessaire une utilisation accrue de modèles mathématiques.

5 Logiciels

La bibliothèque COMBSTRUCT a été conçue et développée par le projet ALGO (Ph. Flajolet, B. Salvy, E. Murray) en liaison avec le projet POLKA de Nancy (P. Zimmermann). Elle fait l'objet d'une collaboration régulière avec les groupes de Waterloo (Université et Compagnie WMI) et une version assez récente est intégrée au système Maple. 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. La dernière version de COMBSTRUCT comporte un système de marquage combinatoire, développé par Marni Mishna, qui permet de calculer automatiquement la moyenne et la variance de certains paramètres combinatoires. La prochaine version incorporera une routine permettant l'asymptotique des suites rationnelles, implantée par Frédéric Cazals (projet PRISME), qui est garantie indépendamment du degré, ainsi qu'une routine due à Cyril Chabaud calculant l'asymptotique des suites algébriques positives, également de manière garantie indépendamment du degré ou de l'existence d'une forme close. La version actuelle représente environ 6 000 lignes de code (500 ko).

Le *package* GFUN développé par B. Salvy et P. Zimmermann (projet POLKA) fournit de nombreux outils de manipulations de suites et de fonctions, à commencer par une fonction de traduction qui part de 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 environ 4 400 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. Une extension de GFUN pour produire automatiquement des procédures efficaces d'évaluation numérique en précision arbitraire a fait l'objet du travail de Juliette Dorleans cette année.

La bibliothèque GDEV pour l'asymptotique des fonctions génératrices développée par Bruno Salvy a été entièrement remise à jour. Ceci a permis la mise au point par Stéphanie Petit d'une mini-encyclopédie des structures combinatoires reposant sur la combinaison de COMBSTRUCT,

GFUN et GDEV et dotée d'un mode d'interrogation via le Web à l'adresse <http://algo.inria.fr/encyclopedia>.

Les algorithmes développés par Frédéric Chyzak sont implantés dans une bibliothèque Maple du nom de MGFUN partiellement intégrée dans la distribution grand public de Maple. Depuis la dernière version, les utilisateurs de Maple disposent ainsi d'une bibliothèque pour la manipulation d'opérateurs linéaires, ainsi que d'une nouvelle bibliothèque pour calculer des bases de Gröbner, capable de traiter aussi bien des polynômes que des opérateurs non commutatifs. C'est désormais cette dernière qui est employée pour réaliser l'élimination polynomiale dans tout le logiciel Maple. Par ailleurs, la simple présence dans Maple de routines de calcul sur les opérateurs linéaires a ouvert la voie à l'implantation, jusqu'alors impossible, de toute une nouvelle génération d'algorithmes pour la sommation et l'intégration symboliques, et plus généralement d'algorithmes récents pour la manipulation de représentations implicites de suites et fonctions spéciales. Suite à l'intégration de MGFUN, la société WMI qui développe Maple a accentué son effort dans cette direction. L'ensemble de la réalisation logicielle correspondant à la bibliothèque MGFUN est constitué de 12 000 lignes de code (430 ko), accompagné d'une quantité équivalente de documentation et de jeux de tests ; elle apporte plus d'une cinquantaine de nouvelles fonctions à l'utilisateur. Afin de rendre l'utilisation de MGFUN plus transparente aux utilisateurs dont les problèmes ne requièrent pas toute la puissance du *package*, Cyril Germa a réalisé cette année une extension de MGFUN. Cette extension détermine automatiquement l'algèbre et les opérateurs avec lesquels effectuer les calculs de somme ou d'intégrale demandés par l'utilisateur. Le *package* MGFUN a été choisi pour ses primitives de calcul de bases de Gröebner pour le *package* DESING développé au RISC (université de Linz, Autriche) et disponible sur le Web (<http://www.risc.uni-linz.ac.at/projects/basic/adjoints/blowup/>).

6 Résultats nouveaux

6.1 Analyse d'algorithmes

Participants : Cyril Banderier, Julien Clément, Philippe Flajolet, Michèle Soria, Brigitte Vallée.

L'analyse d'algorithmes et de structures de données dépend très étroitement d'une évaluation et d'une quantification précise de l'aléa discret. Il s'agit en effet de caractériser le comportement des principaux paramètres de structures combinatoires d'arbres, de mots, de graphes ou d'allocations aléatoires. Dans ce domaine, les travaux du projet Algorithmes s'appuient sur une approche originale et systématique qui a été baptisée "Combinatoire Analytique". La combinatoire analytique repose sur deux ingrédients mathématiques essentiels, dont le développement a largement pris place dans le projet :

- les méthodes symboliques en analyse combinatoire et analyse d'algorithmes qui permettent de disposer de principes systématiques de mise en équation par séries génératrices ;
- les méthodes asymptotiques complexes fondées notamment sur l'analyse de singularité et les transformations intégrales.

Les recherches du projet Algorithmes se placent dans ce cadre général, et la partie décrite ici constitue le “tronc commun” à de nombreux autres travaux conduits dans l'équipe. Ces études soutiennent donc les applications aux séquences textuelles ou génétiques (section 6.3) ainsi qu'à certains des problèmes de modélisation de réseaux (voir la section 6.4). Enfin, l'automatisation de méthodes visant à l'analyse complète de modèles complexes est intimement liée au calcul formel et traitée à la section 6.2.

Combinatoire analytique. Un premier lot de travaux conçus ou publiés en 1999 a trait à l'approfondissement des méthodes fondamentales de la combinatoire analytique. Deux aspects sont couverts : l'aspect combinatoire qui repose sur une élaboration de méthodes de dénombrement, et l'aspect analytique, c'est-à-dire asymptotique, qui permet d'extraire des prédictions quantitatives à la fois précises et simples dans leurs formes.

Deux articles [6] et [8] (en liaison avec l'Université de Barcelone) montrent l'applicabilité du cadre général de la combinatoire analytique au dénombrement de configurations de base de l'algorithmique géométrique telle que développée au Projet PRISME de l'INRIA Sophia-Antipolis. Il s'agit d'abord des triangulations pour lesquelles la distribution du degré maximal est caractérisée dans [6]. Mais surtout, l'article [8] montre que de nombreuses distributions de probabilités des configurations planaires aléatoires succombent simplement aux méthodes de la combinatoire analytique. Les résultats obtenus concernent les triangulations, les dissections, les arbres, les graphes, et les graphes connexes qui sont “sans croisement”. En un certain sens, ces travaux établissent les bases d'une théorie de la percolation portant sur ces graphes sans croisement. A ce titre, ils se rattachent à l'action de recherche incitative ALCOPHYS de la direction scientifique de l'INRIA, et il est de fait que les premiers travaux dans ce domaine sont issus d'un besoin d'élucider certains développements perturbatifs liés à des changements de phase en physique statistique.

Des préoccupations déjà anciennes du projet quant à la génération aléatoire uniforme de configurations combinatoires complexes ont été renouvelées par la communication [21] où interviennent Cyril Banderier et Philippe Flajolet. Dans une certaine mesure l'objectif se relie à la génération de jeux de tests pour la validation de programmes ou de conjectures, certains de ces aspects (algorithmique rapide) étant poursuivis par le projet POLKA de l'UR de Nancy. Ici, ce qui est visé est la génération aléatoire d'objets qui n'obéissent pas à des règles de décomposition simple se prêtant à la “méthode récursive” élaborée par Flajolet, Zimmermann et Vancustem et reprise entièrement dans le système MAPLE. Le travail [21] fournit des réponses conséquentes quant à la nature des objets combinatoires qui sont générables par arbres de possibilités et chemins aléatoires.

L'article [10] est initialement motivé par le cadre de la “théorie analytique de l'information” développé par W. Szpankowski et P. Jacquet (Projet HIPERCOM) dans une longue série de travaux. L'objectif en est d'obtenir des évaluations très précises en théorie de l'information par des méthodes inspirées de la combinatoire analytique. Les problèmes d'analyse d'entropie, de taux de compression, de variances sont dans ce domaine souvent très délicats, à cause par exemple de phénomènes d'annulation difficiles présents dans les sommes combinatoires et difficiles à contourner. Il est fourni dans [10] une méthode générale qui élargit sensiblement le domaine d'application des méthodes d'analyse de singularités introduites par Flajolet et Odlyzko (ATT Research), il y a une dizaine d'années. L'article [26] montre encore la richesse

d'applicabilité des méthodes de combinatoire analytique à la prédiction de plausibilité de motifs dans des séquences génétiques. Par son implantation, il se rattache au calcul formel (section 6.2) tandis que ses objectifs et méthodes sont inspirés de l'analyse des séquences (section 6.3).

La note [9] propose un cadre original unificateur pour les nombres de Stirling. Ces “nombres spéciaux” apparaissent dans de nombreux modèles d'allocation aléatoire (hachage, gestion de cache) ou de recherche (arbres de recherche, arbres quadrants, par exemple) et les identités combinatoires correspondantes sous-tendent un grand nombre d'analyses d'algorithmes. L'article [7] issu d'une collaboration entre le projet et F. Guillemin (CNET Lannion) fournit un cadre nouveau à l'étude de certaines files d'attente à fort volume de trafic. Ce travail qui relève *in fine* de l'analyse de réseaux est décrit à la section 6.4. On notera ici qu'il établit un pont entre théorie combinatoire des chemins et de nombreux modèles probabilistes classiques, tout en conduisant à des analyses nouvelles.

Ainsi, à travers une certaine diversité technique, ces travaux participent à l'entreprise globale qui consiste à délimiter précisément le champ des méthodes de la combinatoire analytique. Les objets traités sont des objets de base de l'informatique: arbres, chemins, allocations aléatoires, mots et quantité d'information, hachage, gestion de caches, etc.

Analyse d'algorithmes. Les méthodes d'accès rapide à l'information se classent grosso modo en deux catégories. D'une part, les méthodes fondées sur des comparaisons des objets entre eux (le tri rapide Quicksort, les arbres de recherche, par exemple), d'autre part les méthodes opérant par calculs d'adresses (le hachage, les méthodes digitales). Les comportements dans le cas le pire, très étudiés jusqu'à une période récente, ne constituent pas un critère de choix valable, car ils ne détectent que des pathologies dont la probabilité est souvent infime.

Un article [13] dont les auteurs sont H. Mahmoud (U. de Washington), P. Jacquet (Projet HIPERCOM), ainsi que M. Régnier et P. Flajolet met en place tout un ensemble de méthodes pour l'analyse en moyenne et en distribution des principaux algorithmes de tri et de sélection fondées sur les méthodes de calcul d'adresse. Il est ainsi établi qu'existe une algorithmique rapide (par exemple, pour l'extraction de “quantiles” en traitement de données statistiques) valable pour des classes de données naturelles.

Dans le cadre des méthodes digitales, qui sont aussi liées aux méthodes de calcul d'adresses, on retiendra cette année une étude de 61 pages [4] qui unifie considérablement le cadre d'analyse des structures de *trie*. La structure de *trie* ou arbre digital est l'une des plus fondamentales pour le traitement de données textuelles. Le modèle aléatoire correspondant peut être considéré comme un modèle probabiliste majeur issu de l'informatique, puisqu'il se retrouve dans la conception de protocoles, les algorithmes d'élection de leader en calcul distribué, les méthodes d'accès à de très grandes bases de données ou encore le calcul formel (factorisation de polynômes). Le point de vue qui est adopté se rattache à la théorie des opérateurs de transfert issue des systèmes dynamiques et du “cadre thermodynamique” (Ruelle). Ce point de vue original introduit en analyse d'algorithmes par B. Vallée (U. de Caen, membre associé au projet) est développé conjointement avec J. Clément (doctorant co-encadré, U. de Caen) qui doit prochainement soutenir une thèse sur ces questions. L'idée consiste très simplement à relier les paramètres d'arbres digitaux à la géométrie d'intervalles fondamentaux (concept issu de la théorie métrique des fractions continues), puis à quantifier de la sorte les propriétés stochastiques des paramètres fondamentaux déterminant la complexité des algorithmes grâce à

l'analyse d'objets associés (séries génératrices, séries de Dirichlet, transformées de Mellin) par opérateurs de transfert. Les paramètres de hauteur, taille, longueur de cheminement sont couverts par ces méthodes ; les modèles de Bernoulli, Markov, ou encore les sources "dynamiques" (par ex., les fractions continues) y deviennent ainsi susceptibles d'un traitement transparent. La méthode est très "versatile" et les nombreuses ramifications en sont poursuivies par B. Vallée et son équipe à l'Université de Caen.

6.2 Calcul formel

Participants :

Frédéric Chyzak, Bruno Salvy.

Juliette Dorléans a réalisé une extension de GFUN permettant de produire automatiquement une procédure optimisée d'évaluation numérique en précision arbitraire de solutions d'équations différentielles linéaires. Pour l'instant la procédure produite s'applique à l'intérieur du disque de convergence de la série associée. Il s'agit là d'un premier pas important vers l'interface entre le calcul symbolique et le calcul numérique. À titre d'exemple, le calcul de π à quelques milliers de décimales est plus rapide par les nouvelles procédures produites automatiquement par GFUN et interprétées que par les routines spécialisées de Maple, pourtant en C compilé.

Bruno Salvy et John Shackell (University of Canterbury, UK) mènent depuis de nombreuses années un travail sur l'algorithmisation du calcul asymptotique. Cette année est ainsi paru [18] un algorithme permettant le calcul du comportement asymptotique des inverses fonctionnels, avec des applications combinatoires comme le développement asymptotique de la variance du nombre de partitions d'un ensemble de taille n , dont les termes au-delà du premier sont inaccessibles à la main. Par ailleurs, B. Salvy a écrit cette année une synthèse sur l'asymptotique automatique [19], à paraître comme chapitre d'un livre sur le calcul formel chez Hermès.

Depuis quelques années, Bruno Salvy collabore avec l'équipe GAGE (École polytechnique) pour comprendre comment des développements récents sur l'algorithmique des *straight-line programs* peuvent être interprétés de manière à produire des implantations efficaces pour des problèmes de nature géométrique dans des systèmes de calcul formel du type de Maple. Les premiers résultats sont prometteurs, puisqu'une implantation prototype en Maple permet de résoudre numériquement des systèmes polynomiaux de grande taille. Ces travaux ont été présentés à la conférence *Foundations of Computational Mathematics* à Oxford et font l'objet d'une soumission au *Journal of Complexity*.

Une extension de COMBSTRUCT vers les automates et les expressions régulières a été réalisée par Pierre Nicodème. Les premières applications à la recherche de motifs dans l'ADN [26] montrent qu'une combinaison de modélisation combinatoire et d'asymptotique automatique (fournie par GFUN) permet le traitement de problèmes de très grande taille en des temps modestes.

Frédéric Chyzak a poursuivi sa recherche sur la manipulation des suites et des fonctions spéciales par le calcul formel lors de son séjour post-doctoral à l'université de Linz (Autriche). Ce travail a fourni des algorithmes pour le calcul d'équations fonctionnelles linéaires caractérisant la somme ou l'intégrale d'une expression faisant intervenir des suites et fonctions spéciales [3]. L'objectif est maintenant d'améliorer ces algorithmes de façon à assurer la minimalité de

l'ordre des équations obtenues.

F. Chyzak en collaboration avec Peter Paule du RISC (Linz, Autriche) et Ivan Gutman de l'université de Kragujevac (Yougoslavie) ont obtenu de premières indications sur un problème ouvert difficile d'énumération combinatoire en chimie mathématique [2]. Ce travail repose de manière fondamentale sur une expérimentation à l'aide du *package* GFUN développé au projet ALGO. Cette recherche s'oriente maintenant vers l'utilisation de méthodes de physique statistique.

Une autre collaboration de Frédéric Chyzak avec le RISC (P. Paule et Burkhard Zimmermann) et une autre équipe de l'université de Linz (Otmar Scherzer et Armin Schoisswohl de l'Industrial Mathematics Institute) a donné une nouvelle construction pour des ondelettes d'un certain type. Une étape cruciale est la détermination sous forme explicite par le calcul formel d'une famille de coefficients de filtrage.

Le stage de Cyril Germa a permis de doter le *package* MGFUN de F. Chyzak d'une interface grand public. Celle-ci sera intégrée dans la prochaine version diffusée sur le Web (<http://algo.inria.fr/chyzak/mgfun.html>).

6.3 Algorithmiques des séquences

Participants : Frédéric Chyzak, Philippe Dumas, Philippe Flajolet, Pierre Nicodème, Mireille Régnier, Bruno Salvy.

Le projet a poursuivi des recherches sur la statistique des mots. Il s'agit de compter le nombre d'occurrences d'un motif donné ou d'un ensemble de motifs dans un texte sous des contraintes variées (distance minimale entre les mots, types de chevauchement autorisés, etc.). Un tel comptage se ramène à l'évaluation des inversions de systèmes d'équations algébriques satisfaits par les séries génératrices.

Pierre Nicodème, chercheur post-doctorant dans le département de bioinformatique du DKFZ à Heidelberg, recruté au CNRS, Evry, en Octobre 99, Bruno Salvy et Philippe Flajolet ont considéré les expressions régulières. Un algorithme permet de calculer de manière exacte ou asymptotique espérance et variance du nombre d'occurrences de tout motif dans des textes de grande taille dont les lettres obéissent à une distribution de Markov. À part quelques cas de dégénérescence, la distribution limite est gaussienne. L'implantation correspondante a été réalisée par Pierre Nicodème en Maple, et appliquée aux statistiques d'occurrences de motifs dans les protéines.

Une approche plus formelle, reliant les calculs de moyenne ou de variance à l'étude de langages particuliers est présentée dans [14]. Elle conduit à des formules explicites. Il s'ensuit [31] que la complexité du calcul dépend de la taille minimale d'un automate de "recouvrement" plus petit que l'automate de recherche. La recherche d'un algorithme de calcul de l'automate de recouvrement est l'objet d'un travail en cours. On s'intéresse en particulier aux expressions régulières et aux motifs approchés.

Les résultats de [14] s'appliquent à la recherche de certains sites fonctionnels de l'ADN. Ils ont été utilisés par l'UVSQ pour la recherche de DosDNA, petits motifs répétés qui sont la trace de l'instabilité génétique dans les séquences d'ADN. Ils font l'objet de différentes collaborations. Avec l'Université d'Orsay, M. Régnier étudie les phases ouvertes de lecture

(Open Reading Frames). Il s'agit d'adapter et implémenter les formules [14] pour distinguer sur le chromosome d'un organisme (la levure) les gènes réels des faux gènes. Le modèle probabiliste est markovien.

Avec l'Institut Pasteur, M. Régnier recherche un algorithme de reconnaissance de signaux multiples dans les promoteurs. L'évaluation de la pertinence statistique des motifs extraits est une des phases potentiellement coûteuses de l'algorithme. L'obtention de formules simplement calculables pour des ensembles de motifs similaires selon une distance donnée est en cours.

De nombreux problèmes combinatoires se posent en biologie moléculaire. Dans [30], on étudie une distance particulière, la distance de Levenshtein. Les recherches de similitudes entre les séquences sont algorithmiquement coûteuses. Différentes méthodes de filtrage ont été développées. Un travail conjoint de R. Smythe (Oregon State University) et M. Régnier étudie, en s'appuyant sur des résultats combinatoires, la distribution de la "distance statistique" sur des séquences aléatoires.

Frédéric Chyzak et Philippe Dumas ont poursuivi leur collaboration avec Mikhail Atallah (Université de Purdue, USA) sur la recherche de motifs avec erreurs [1].

6.4 Algorithmique et modélisation des réseaux

Participants : Jean-François Dantzer, Mostapha Haddani, Philippe Robert.

Allocation de bande passante dans un réseau On s'intéresse ici à un modèle d'allocation de ressources : divers types de trafic arrivent avec des demandes de débit très différentes ; il est toujours possible par un mécanisme de réservation et en segmentant ces trafics de leur allouer la totalité de la bande passante. L'inconvénient de ce type de méthode réside dans le traitement ultérieur qui est requis pour recomposer les trafics initiaux. Ces problèmes peuvent se formuler en terme de *bin packing* dynamique : une boîte de taille 1 (la capacité maximale du réseau) et des arrivées de pièces dont la taille est inférieure à 1, chaque pièce demandant à rester dans la boîte le temps de son service (autrement dit dans le cas des réseaux, la requête occupe une fraction de la bande passante). Les questions naturelles qui se posent concernent le débit maximal d'un tel système et la politique d'allocation qui maximise celui-ci. Ce cadre a été étudié, pour des raisons très différentes, dans un article de Kipnis et Robert (1990) ; le débit maximal avait été calculé dans le cas où les temps de service sont exponentiels, les tailles des pièces indépendantes équidistribuées et la discipline de service premier arrivé-premier servi. Les problèmes d'allocation de bande passante ont mis en valeur tout l'intérêt de ce type de modèle, voir les travaux récents de Coffman et Stolyar (Bell laboratories, Lucent Technology).

J.-F. Dantzer, M. Haddani et Ph. Robert étudient actuellement le cas de la discipline *First Fit* : sous les hypothèses probabilistes mentionnées ci-dessus, un message est alloué si la place résiduelle dans la boîte le permet et si aucun des messages arrivés avant lui dans la file d'attente ne peut être alloué. La difficulté de ce modèle vient du fait que la suite des tailles des pièces dans la file d'attente n'est pas une suite indépendante, en raison des prélèvements successifs à l'intérieur de celle-ci. L'espace d'états décrivant le modèle est de dimension dénombrable, donc délicat à manipuler. Les problèmes techniques soulevés sont similaires à ceux, non résolus pour l'instant, des réseaux de files d'attente multi-classe FIFO.

L'étude d'un modèle simple pour lequel il n'y a que deux tailles de messages possibles, 1 et a , a été complétée [5]. M. Haddani a ensuite généralisé les résultats d'ergodicité au cas où il y a n tailles différentes $1, a, \dots, a^{n-1}$. Ce travail se place dans le cadre d'une collaboration avec Fabrice Guillemin (CNET-Lannion) pour étudier le cas où la taille de la bande passante vaut 512Mo et les tailles des messages sont 32ko, 128ko (petites transactions WEB) et 1Mo, 4Mo (flux multi-media). Une proposition d'étude de l'allocation de bande passante dans les réseaux a été déposée dans le cadre des consultations thématiques du CNET.

Algorithmes de contrôle d'admission dans les réseaux Cédric Adjih, Philippe Jacquet (projet HIPERCOM) et Philippe Robert ont étudié le contrôle d'admission dans un réseau à intégration de services. Le modèle considéré est le suivant : deux flots de messages notés respectivement P (Premium) et M (Medium) se partagent la bande passante (BP) d'un lien dans un réseau. Les messages de type P ont un degré de priorité supérieur à ceux de type M. Un message qui ne peut accéder immédiatement au lien par manque de BP disponible est rejeté. On fait l'hypothèse, vérifiée en pratique, que la BP moyenne demandée par ces deux flots est plus grande que la BP offerte. Nous avons étudié la part b_i de la BP inutilisée quand il n'y a pas de contrôle d'admission. Avec l'hypothèse de flots poissonniens, en utilisant un modèle renormalisé et des techniques de grandes déviations nous avons montré que la variable b_i converge en distribution vers une loi exponentielle dont le paramètre est déterminé par une équation de point fixe. Ce résultat est sur le plan théorique l'analogue continu d'un résultat obtenu par Kelly pour un modèle discret. À la différence de Kelly qui utilise une méthode de point de selle pour le résultat de convergence, notre preuve est entièrement probabiliste. Elle présente en outre l'avantage de bien comprendre le rôle des événements rares dans cette convergence.

L'étape suivante a été de concevoir un algorithme permettant de privilégier les messages de type P sans complètement ignorer ceux de type M, i.e.

- a) le taux de rejet des requêtes P est borné par une constante indépendante de l'intensité de trafic des requêtes M.
- b) L'impact de l'algorithme sur le taux de rejet des requêtes de type M doit être minimal.

L'algorithme proposé est adaptatif, il est caractérisé par deux constantes $\alpha > 1$ et $\beta < 1$ et une fonction ($p(t)$). Une requête de type P est toujours acceptée si la bande passante disponible est suffisante. Si une requête de type M arrive à l'instant t , elle est déclarée non-admissible avec probabilité $p(t)$, et par conséquent rejetée indépendamment de l'occupation de la bande passante. Si la requête est admissible et si la bande passante disponible n'est pas suffisante elle est rejetée et $p(t)$ est multiplié par α , autrement elle est acceptée et $p(t)$ est multiplié par β . Les simulations montrent que lorsque α et β sont proches de 1, l'algorithme remplit les conditions a) et b). En utilisant une approximation et le résultat exact du cas sans contrôle d'admission, on conjecture que la quantité $p(t)$ converge en distribution (i.e. l'algorithme est stable) et que la part de la bande passante inutilisée converge vers une distribution exponentielle.

Phénomènes d'attente dans les réseaux à fort trafic Philippe Flajolet et Fabrice Guillemin (CNET, Lannion) ont développé un cadre unifié d'analyse des très classiques processus

de vie et de mort. Ces processus multiformes incluent les modèles de files d'attente $M/M/1$ ou $M/M/\infty$ et permettent ainsi de rendre compte de phénomènes élémentaires de perte de trafic et de surcharge dans des noeuds ATM. L'intérêt méthodologique de l'approche introduite dans le rapport [7] consiste notamment en ce qu'elle jette un pont entre le domaine des modèles continus (les processus stochastiques) et les modèles discrets de la combinatoire analytique (chemins et histoires). La connexion s'établit via les fractions continues, les polynômes orthogonaux, et la transformation de Laplace. Il en découle un calcul systématique voire formel de nombreuses caractéristiques exprimables par des conditions géométriques incluant les passages et traversées de bandes, les volumes de trafic excédant un certain niveau, etc.

Étude d'un modèle généralisé d'urne de Pólya Samuel Kotz, Hosam Mahmoud (George Washington University) et Philippe Robert ont étudié une généralisation d'un modèle classique de la littérature combinatoire probabiliste. Une urne contient des boules de plusieurs couleurs, l'une d'entre elles est tirée au hasard et suivant sa couleur un certain nombre de boules de diverses couleurs sont rajoutées. L'opération est répétée indéfiniment. On s'intéresse à la convergence des proportions respectives de chacune des couleurs dans l'urne. Les résultats obtenus dans ce domaine supposent que le nombre de boules rajoutées à chaque opération est constant. Noter que même dans ce cas certains des phénomènes sont non triviaux. Nous nous sommes intéressés au cas où il n'y a que deux couleurs (rouge et blanc par exemple) mais le nombre de boules rajoutées dépend de la couleur. La distribution du nombre de boules rouges au bout de N opérations a été obtenue en utilisant une relation de récurrence et une propriété d'arborescence du modèle. La forme explicite de cette distribution ne permet apparemment pas une étude asymptotique quand N tend vers l'infini. Dans un cadre continu, en utilisant une méthode de poissonnisation (les boules sont tirées aux instants d'un processus de Poisson) et en renormalisant le modèle, nous avons montré la convergence en distribution de la proportion de chacune des couleurs. En particulier, nous montrons que les résultats asymptotiques conjecturés dans le cas discret sont vrais dans ce cadre.

Divers Jean-François Dantzer a rédigé sa proposition de thèse « Stabilité des réseaux de files d'attente et limites fluides stochastiques ». La soutenance de la thèse est prévue en janvier 2000 devant le jury constitué de F. Charlot (Université de Rouen), B. Chauvin (Université de Versailles), Ph. Flajolet, R. Iasnogorodski (Université d'Orléans), A. Mokkadem (Université de Versailles) et Ph. Robert

Ph. Robert a rédigé trois chapitres [15, 16, 17] pour le livre « Réseaux et files d'attente : méthodes probabilistes » à venir chez Springer-Verlag dans la collection Mathématiques et applications de la SMAI.

7 Contrats industriels (nationaux, européens et internationaux)

7.1 Calcul formel

Le projet ALGO et la compagnie *Waterloo Maple Inc.* 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. J. Carette est retourné à la compagnie WMI, après une participation de plus de trois ans au projet ALGO. Il y est maintenant *Product Development Director* et s'occupe de tous les aspects techniques de Maple. De même, E. Murray, après avoir passé plus de deux ans au projet ALGO à programmer le *package* COMBSTRUCT de Maple, travaille maintenant à la compagnie WMI.

L'arrivée de Maple dans l'enseignement en classes préparatoires a nécessité un important travail de formation des enseignants. Les membres du projet ont participé activement à cet effort. Claude Gomez (projet META2), B. Salvy et P. Zimmermann (projet POLKA) ont écrit un livre il y a quatre ans sur l'utilisation de Maple. Depuis deux ans, Ph. Dumas consacre un effort important à la rédaction des solutions des exercices de ce livre et à leur mise à la disposition de tous sur le Web. Ph. Dumas, X. Gourdon et F. Chyzak ont formé environ 400 professeurs de classes préparatoires au cours de stages à l'École des Mines de Nantes et à l'École Spéciale des Travaux Publics.

Grâce à ces nombreuses activités autour de Maple, la compagnie WMI 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.

Un club des utilisateurs francophones de Maple a été créé il y a deux ans. La liste de diffusion et le site Web de ce club sont hébergés par le projet ALGO, et gérés en commun avec Gilles Villard et Claire Di Crescenzo de l'IMAG.

7.2 Algorithmique et modélisation des réseaux

Catserver L'avant-projet HIPERCOM (Ph. Jacquet, P. Mühlethaler) et le projet ALGO (Ph. Robert) participent au projet Esprit/Multimedia intitulé CATSERVER. Il concerne la conception et la mise au point d'un serveur multimédia sur le réseau câblé actuel (i.e., avec les modems câble du commerce); voir la section correspondante du rapport d'activité d'HIPERCOM. Le projet ALGO contribue principalement à la partie compression et à l'algorithmique de la gestion des qualités de service. Les partenaires de cette proposition sont : Thomson-CSF Detexis — Leader (France), Brutélé (Belgique), Intrasoft (Greece), Université de Thessalonique (Grèce), Barco (Luxembourg), Lyonnaise Communications (France) et TOPNET (Luxembourg). Ce projet est programmé pour une durée de deux ans à partir du 1er décembre 1997.

8 Actions régionales, nationales et internationales

8.1 Actions nationales

Philippe Flajolet anime le groupe ALEA du GDR ALP (Algorithmes, Langages, Programmation) créée en 1998. Le groupe fédère une douzaine d'équipes nationales et il tient une réunion annuelle avec une quarantaine de participants (à Bordeaux en 1999, où la réunion était jumelée avec les Journées Combinatoire organisées par le LABRI de Bordeaux). Son but est de regrouper l'ensemble des chercheurs travaillant dans le domaine des structures combinatoires finies,

étudiées pour leurs propriétés aléatoires (probabilistes, en moyenne, en distribution), ce en liaison avec l'algorithmique séquentielle ou distribuée ainsi qu'avec les structures de données de base de l'informatique. Y participent une quinzaine d'équipes de probabilités appliquées, théorie des nombres, combinatoire et algorithmique.

En 1999 a démarré une Action de Recherche Coopérative de la direction scientifique de l'INRIA intitulée ALCOPHYS. Les maîtres d'œuvre en sont le LABRI de Bordeaux (R. Cori) et le Projet Algorithmes, et y participent une équipe du LRI (C. Kenyon) et une équipe de physiciens de l'Université Paris-Sud. Le thème en est le rapprochement et la confrontation entre méthodes probabilistes, combinatoire analytique et physique statistique dans le but d'analyser les systèmes complexes de grande taille.

Mireille Régnier participe à l' Action de Recherche Coopérative de l'INRIA REMAG, dont le thème est la Recherche et l'Extraction de Motifs pour l' Analyse Génomique.

8.2 Actions financées par la Commission Européenne

Le projet ALGO est pour la période de trois ans 1996–1999, l'une des composantes du projet ESPRIT "Long Term Research" ALCOM-IT (*ALgorithms and COMplexity in Information Technology*). Les partenaires sont issus de douze universités et centres de recherche publics et proviennent de huit pays européens. 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. Le projet ALCOM-IT regroupe l'essentiel des forces dans le domaine de l'algorithmique et de la complexité en Europe de l'Ouest. 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.

8.3 Relations bilatérales internationales

Le projet entretient un grand nombre de collaborations bilatérales, certaines formalisées, d'autres non, avec des universités et centre de recherche étrangers. En 1999, les travaux conjoints mettent en jeu par exemple les universités de Purdue, Montréal, Barcelone, Johannesburg, Cantorbery, Princeton, Bruxelles, Washington, et les centres Bell Labs de Lucent ou le Centre du cancer (DKFZ à Heidelberg). Une action AMADEUS avec l'Autriche se termine en 1999 et a donné lieu à des échanges croisés. Un accord ALLIANCE avec le Royaume Uni (Oxford, Cardiff) sous pilotage de l'Université de Caen a démarré en 1999 et a déjà donné lieu à un workshop conjoint.

8.4 Accueils de chercheurs étrangers

Ilan Vardi, IHES, Bures-sur-Yvette, a séjourné dans le projet une semaine mi-octobre et un mois du 24 novembre au 24 décembre ; Yoshiaki Itoh, Institute of Statistical Mathematics, Tokyo, Japon, de mi-avril à fin mai ; Mamoru Hoshi, University of Electro-Communications, Japan, 1 mois en octobre ;

Edward A. Bender, University of California at San Diego, USA, Ed Coffman, BELL LABS Murray Hill, USA, Daniel Panario, University of Toronto, Canada ; John Shackell, University

of Kent at Canterbury ; Vsevolod Makeev, Engelhardt Institute of Molecular Biology, Moscow, Russia, ont séjourné une semaine dans le projet.

Kevin Compton, EECS, University of Michigan, USA, Steve Lalley, Purdue University, West Lafayette, Indiana, USA, Lou van den Dries, University of Illinois at Chicago, USA, Marko Petkovšek, University of Ljubljana, Slovenia, et Cristopher Moore, The Santa Fe Institute, New Mexico, USA, ont fait des exposés au séminaire du projet.

9 Diffusion de résultats

9.1 Animation de la communauté scientifique

Philippe Flajolet est membre des comités d'édition des revues internationales suivantes : *Maple Technical Newsletter* (Birkhäuser), *Theoretical Computer Science* (Elsevier), *Random Structures & Algorithms* (Wiley). Il est aussi membre du comité éditorial des collections de monographies LACIM (Québec), de l'*Encyclopedia of Mathematics* publiée par Cambridge University Press, et des *Comptes Rendus de l'Académie des Sciences*. Il est également l'un des fondateurs et membre permanent du comité de pilotage des réunions internationales "*Analysis of Algorithms*" dont la cinquième édition s'est tenue à Barcelone en juin 1999. Il est enfin responsable du groupe de travail national ALEA au sein du GDR A.L.P (*Algorithmique, Langage et Programmation*) dont la réunion tenue à Bordeaux en mars 1999 a attiré plus de 70 scientifiques majoritairement français.

Mireille Régnier a été présidente du Comité d'organisation de RECOMB'99 (REsearch in Computational Biology) qui a réuni à Lyon trois cents chercheurs du domaine.

Bruno Salvy est membre des comités éditoriaux du *Journal of Symbolic Computation* de la *Maple Technical Newsletter*. Cette année il a aussi été président du comité de programme de l'*International Symposium on Symbolic and Algebraic Computation* (ISSAC), la principale conférence du domaine, qui s'est déroulée à Vancouver.

Cyril Banderier anime le Colloquium Junior de l'Inria Rocquencourt et a fait une intervention le 18 octobre 1999 au lycée de la plaine de Neauphle, à Trappes, dans le cadre de la Semaine de la Science.

9.2 Enseignement universitaire

Cyril Banderier enseigne, en tant que moniteur, à l'université de Paris-Nord (P13) en premier cycle universitaire (première et deuxième année, option maths-info) différentes matières informatiques.

Jean-Francois Dantzer enseigne en licence de mathématiques à l'université Versailles Saint-Quentin.

Philippe Flajolet est responsable avec B. Vallée d'un cours de filière "Modèles combinatoires" au sein du DEA Algorithmique de la région parisienne (20h).

Mostafa Haddani enseigne en DEUG STPI et en DEUG Economie à l'université Versailles Saint-Quentin.

François Morain est Maître de Conférences à l'École polytechnique depuis 1997. Il est responsable de l'Enseignement d'Approfondissement *Cryptologie et Arithmétique* au sein de la Majeure 1 Mathématiques et Informatique (50 h). Il donne depuis 1991 un cours de Théorie algorithmique des nombres (20 h) dans le DEA Algorithmique (École polytechnique, ENS, Paris VI, VII et X).

Mireille Régnier enseigne un cours de DEA sur le thème "génomique et probabilités" à Marne-la-Vallée.

Philippe Robert donne un cours de DEA intitulé "Théorèmes limites pour l'étude des files d'attente" au Laboratoire de Probabilités de l'Université de Paris VI.

Bruno Salvy participe au tronc commun d'informatique à l'École polytechnique en tant que chargé de travaux pratiques.

Michèle Soria est Professeur à l'Université Paris VI. Elle enseigne dans différentes formations des trois cycles universitaires.

9.3 Participation à des colloques, séminaires, invitations

Frédéric Chyzak a à plusieurs reprises présenté les résultats de sa thèse : Leoben (Autriche), Rouen, Nice, Université du Québec à Montréal (Canada), et à la conférence ISSAC'99 (Vancouver, Canada). Lors de son séjour post-doctoral à l'université de Linz (Autriche), il a donné une série d'exposés sur le sujet (15 h) au sein du séminaire de combinatoire du RISC [28].

Philippe Flajolet a été en 1999 l'objet d'invitations pour visites et exposés de travaux de recherche aux centres scientifiques suivants : à l'Institut des Hautes Études Scientifiques (Bures; Journée informatique, janvier 1999), aux Universités de Marne-la-Vallée, Purdue, Princeton, et Freiburg (Allemagne), au Centre de Recherche Mathématique de Barcelone, à l'Université d'Oxford, au Centre International de Rencontre Mathématiques (Systèmes dynamiques; Luminy, novembre 1999), au Colloque "*Random Structures and Algorithms*" (Poznań), à l'atelier Européen *RAND2*. Il a été professeur à l'International Centre for Theoretical Physics (Trieste, août 1999) et "keynote speaker" au Colloque I.E.E.E "*Information Theory and Networking*" (Metsovo, Grèce, juin 1999). En 1999, Philippe Flajolet a été membre de jurys de thèse ou d'habilitation en probabilités (Nancy), combinatoire (Bordeaux, Nancy), cryptographie (ENS Ulm), algorithmique (Caen, Marne-la-Vallée), et réseaux (Paris).

Mireille Régnier a participé aux Journées Aléa & Gascom (Bordeaux), à CPM'99 (Warwick) et RECOMB'99 (Lyon). Elle a fait des exposés à Padoue, Kazan, Barcelone et Rouen (WORDS'99).

Philippe Robert a participé à la conférence INFORMS à Ulm (Allemagne) du 26 au 28 juillet pour présenter les travaux avec Christine Fricker et Danielle Tibi sur la vitesse de convergence [25]. Il a en outre donné un exposé sur le processus d'Ornstein-Ühlenbeck discret aux séminaires des départements de mathématiques de l'Université de Cergy et de l'Université de Rouen.

Bruno Salvy a donné un exposé sur un nouvel algorithme de calcul d'intégrales définies à la conférence *Foundations of Computational Mathematics* (Oxford), un exposé sur le calcul avec des fonctions holonomes à Limoges, un autre sur les statistiques de motifs biologiques à Caen et à Barcelone.

Cyril Banderier a assisté aux Journées Aléa & Gascom (Bordeaux) puis présenté son travail de DEA à l'École Jeunes Chercheurs de Bordeaux. Il a présenté un article à la conférence internationale sur les Séries Formelles et la Combinatoire Algébrique (SFCA/FPSAC) à Barcelone puis assisté à la réunion "Analysis of Algorithms". Il a fait des exposés à Rouen et au Colloquium Junior de l'INRIA Rocquencourt.

10 Bibliographie

Articles et chapitres de livre

- [1] M. J. ATALLAH, F. CHYZAK, P. DUMAS, « A Randomized Algorithm for Approximate String Matching », *Algorithmica*, 1999, à paraître.
- [2] F. CHYZAK, I. GUTMAN, P. PAULE, « Predicting the number of hexagonal systems with 24 and 25 hexagons », *Communications in mathematical and in computer chemistry* 40, 1999, p. 139–151.
- [3] F. CHYZAK, « An Extension of Zeilberger's Fast Algorithm to General Holonomic Functions », *Discrete Mathematics*, 1999, à paraître.
- [4] J. CLÉMENT, P. FLAJOLET, B. VALLÉE, « Dynamical Sources in Information Theory: A General Analysis of Trie Structures », *Algorithmica*, 1999, à paraître. Version préliminaire dans le rapport de recherche INRIA n^0 3645, <http://www.inria.fr/RRRT/RR-3645.html>.
- [5] J.-F. DANTZER, M. HADDANI, P. ROBERT, « On the stability of a bandwidth packing algorithm », *Probability in the Engineering and Informational Sciences*, 2000, à paraître. Version préliminaire dans le rapport de recherche INRIA n^0 3605, <http://www.inria.fr/RRRT/RR-3605.html>.
- [6] L. DEVROYE, P. FLAJOLET, F. HURTADO, M. NOY, W. STEIGER, « Random Triangulations », *Discrete and Computational Geometry*, 1999, p. 105–117.
- [7] P. FLAJOLET, F. GUILLEMIN, « The Formal Theory of Birth-and-Death Processes, Lattice Path Combinatorics, and Continued Fractions », *Journal of Applied Probability*, 1999, 23 pages. A paraître. Version préliminaire dans le rapport de recherche INRIA n^0 3667, <http://www.inria.fr/RRRT/RR-3667.html>.
- [8] P. FLAJOLET, M. NOY, « Analytic Combinatorics of Non-crossing Configurations », *Discrete Mathematics* 204, 1-3, 1999, p. 203–229, (Selected papers in honor of Henry W. Gould).

- [9] P. FLAJOLET, H. PRODINGER, « On Stirling Numbers for Complex Argument and Hankel Contours », *SIAM Journal on Discrete Mathematics* 12, 2, 1999, p. 155–159.
- [10] P. FLAJOLET, « Singularity analysis and asymptotics of Bernoulli sums », *Theoretical Computer Science* 215, 1-2, 1999, p. 371–381.
- [11] C. FRICKER, P. ROBERT, D. TIBI, « On the rates of convergence of Erlang's model », *Journal of Applied Probability* 36, 4, 1999, p. 1–18.
- [12] S. KOTZ, H. MAHMOUD, P. ROBERT, « On Generalized Pólya Urn Models », *Statistics and Probability Letters*, 1999, à paraître.
- [13] H. MAHMOUD, P. FLAJOLET, P. JACQUET, M. RÉGNIER, « Analytic Variations on Bucket Selection and Sorting », *Acta Informatica*, 1999, à paraître.
- [14] M. RÉGNIER, « A Unified Approach to Word Occurrences Probabilities », *Discrete Applied Mathematics*, 1999, à paraître; version préliminaire à RECOMB'98.
- [15] P. ROBERT, *Files d'attente et réseaux : méthodes probabilistes : Loi de Poisson et événements rares*, Springer-Verlag, 1999, 17 pages. Chapitre d'un livre en préparation dans la collection Mathématiques et applications.
- [16] P. ROBERT, *Files d'attente et réseaux : méthodes probabilistes : Méthodes de renormalisation*, Springer-Verlag, 1999, 39 pages. Chapitre d'un livre en préparation dans la collection Mathématiques et applications.
- [17] P. ROBERT, *Files d'attente et réseaux : méthodes probabilistes : Vitesse de convergence à l'équilibre*, Springer-Verlag, 1999, 28 pages. Chapitre d'un livre en préparation dans la collection Mathématiques et applications.
- [18] B. SALVY, J. SHACKELL, « Symbolic Asymptotics: Multiseries of Inverse Functions », *Journal of Symbolic Computation* 27, 6, juin 1999, p. 543–563.
- [19] B. SALVY, *Asymptotique automatique*, Hermes, Paris, 1999, à paraître comme chapitre d'un livre sur le calcul formel. Version préliminaire dans le rapport de recherche INRIA n° 3707, <http://www.inria.fr/RRRT/RR-3707.html>.

Communications à des congrès, colloques, etc.

- [20] C. ADJIH, P. JACQUET, P. ROBERT, « Differentiated admission control in large networks », in : *INFOCOM 2000*, Juin 2000. A paraître.
- [21] C. BANDERIER, M. BOUSQUET-MÉLOU, A. DENISE, P. FLAJOLET, D. GARDY, D. GOUYOU-BEAUCHAMPS, « On Generating Functions of Generating Trees », in : *Formal Power Series and Algebraic Combinatorics*, O. S. C. Martínez, M. Noy (éditeur), p. 40–52, 1999. Proceedings of FPSAC'99, Universitat Politècnica de Catalunya. Version préliminaire dans le rapport de recherche INRIA n° 3661.
- [22] E. G. COFFMAN, P. ROBERT, A. L. STOLYAR, « The Interval Packing Process of Linear Networks », in : *Workshop of 1999 ACM Sigmetrics conference*, April 1999. A paraître dans Performance Evaluation Review.

-
- [23] J.-F. DANTZER, I. MITRANI, A. PUHALSKII, P. ROBERT, « Large Scale and Heavy Traffic Asymptotics for Systems with Unreliable Servers », *in* : *INFORMS AP99*, V. Schmidt (éditeur), Ulm, Juillet 1999.
- [24] P. FLAJOLET, B. VALLÉE, « Continued Fractions, Comparison Algorithms, and Fine Structure Constants », *in* : *Foundations of Computational Mathematics*, Oxford, 1999. Semi-plenary lecture; à paraître dans *Proc. Canadian Math. Society*.
- [25] C. FRICKER, P. ROBERT, D. TIBI, « Convergence to equilibrium of finite Markov chains », *in* : *INFORMS AP99*, V. Schmidt (éditeur), Ulm, Juillet 1999.
- [26] P. NICODÈME, B. SALVY, P. FLAJOLET, « Motif Statistics », *in* : *Algorithms, ESA '99*, J. Nešetřil (éditeur), *Lecture Notes in Computer Science, 1643*, p. 194–211, 1999. Proceedings of 7th Annual European Symposium on Algorithms ESA'99, Prague, Juillet 1999.
- [27] B. SALVY, « Convolution Integrals, Integral Transforms and Symbolic Computation », *in* : *Foundations of Computational Mathematics*, Oxford, 1999. Semi-plenary lecture.

Divers

- [28] F. CHYZAK, « Holonomic functions in computer algebra, Slides for a series of talks », Research Report no. 99–6 of the Austrian Project SFB F013, Numerical and Symbolic Scientific Computing, 1999.
- [29] M. GIUSTI, G. LECERF, B. SALVY, « A Gröbner Free Alternative for Polynomial System Solving », Soumis à *Journal of Complexity*, juillet 1999, 51 pages.
- [30] D. HIRSCHBERG, M. RÉGNIER, « Tight Bounds on the Number of String Subsequences », 1999, soumis; version préliminaire à CPM'99.
- [31] M. RÉGNIER, « Efficient Computation of Unusual Words Expectation », 1999, présenté à WORDS'99.