INRIA, Evaluation of Theme Sym B

Project-team Algorithms

November 14-15, 2006

Project-team title: Algorithms (ALGO)

Scientific leader: Bruno Salvy

Research center: Rocquencourt

1 Personnel

Personnel (March 2002)

	Misc.	INRIA	CNRS	University	**Total**
DR (1) / Professors		3			**3**
CR (2) / Assistant Professors		1			**1**
Permanent Engineers (3)					
Temporary Engineers (4)					
PhD Students	5			1	**6**
Post-Doc.	1				**1**
Total	**6**	**4**		**1**	**11**
External Collaborators	2		2	2	**6**
Visitors ($>$ 1 month)	1				**1**

(1) "Senior Research Scientist (Directeur de Recherche)"
(2) "Junior Research Scientist (Chargé de Recherche)"
(3) "Civil servant (CNRS, INRIA, ...)"
(4) "Associated with a contract (Ingénieur Expert or Ingénieur Associé)"

Personnel (November 2006)

	Misc.	INRIA	CNRS	University	**Total**
DR / Professors		3			**3**
CR / Assistant Professor		2			**2**
Permanent Engineer					
Temporary Engineer	1				**1**
PhD Students	1			2	**3**
Post-Doc.	1				**1**
Total	**3**	**5**		**2**	**10**
External Collaborators	1		1		**2**
Visitors ($>$ 1 month)					

Changes in staff

DR / Professors CR / Assistant Professors	Misc.	INRIA	CNRS	University	total
Arrival		1			
Leaving					

Comments: Alin Bostan joined our team in September 2004.

Current composition of the project-team in November 2006:

- DR: Philippe Flajolet (DR 0), Mireille Régnier and Bruno Salvy (DR 2);
- CR: Frédéric Chyzak (CR 1), Alin Bostan (CR 2);
- PhD Students: Éric Fusy (deputation from Corps des Télécoms), Frédéric Giroire and Carine Pivoteau (with scholarships from the Ministry of Research and Higher Education);
- Temporary Engineer: André Derrick Balsa;
- Post-doc: Frédéric Meunier (deputation from Corps des Ponts);
- External collaborators: Philippe Dumas (Mathematics Teacher in a *classe préparatoire aux grandes Écoles*, Lycée Jean-Baptiste Say, Paris; Pierre Nicodème (CR CNRS, École polytechnique).

Current position of former project-team members (including PhD students during the March 2002 to November 2006 period):

Former PhD students:

- Alin Bostan, researcher, Algorithms project, Inria;
- Marianne Durand, analyst in computational financial mathematics, BNP-Paribas, Paris, France;
- Julien Fayolle, post-doc, LRI, University of Paris-Sud, France;
- Ludovic Meunier, software engineer, Murex (software development for trading, risk management and processing), Paris, France;
- Vincent Puyhaubert, mathematics teacher, *classe préparatoire aux grandes Écoles*, Lycée Fénelon, Paris, France;
- Mathias Vandenbogaert, researcher in biology, Institut Pasteur, Paris, France.

Former post-docs:

- Ha Le, half-time math developer, Waterloo Maple Inc., and half-time post-doc, University of Waterloo, Canada;
- José Luis Martins, administrator at the Commission of the European Communities, Luxembourg.
- Alexandre Sedoglavic, maître de conférences in computer science, University of Lille, France;
- Markus Vöge, researcher in computer science, department of media technology, EMPA research institute, Dübendorf, Switzerland.

Last INRIA enlistments

Alin Bostan, CR 2 since September 1st, 2004.

2 Work progress

2.1 Keywords

Analysis of algorithms, analytic combinatorics, random discrete structures, limit laws, computer algebra, special functions, Gröbner bases, combinatorics on words, biological sequences, genome.

2.2 Context and overall goal of the project

The primal objective of the project, inherited from the former century, is the field of analysis of algorithms. By this is meant a precise quantification of complexity issues associated to the most fundamental algorithms and data structures of computer science. Departing from traditional approaches that, somewhat artificially, place emphasis on worst-case scenarii, the project focuses on average-case and probabilistic analyses, aiming as often as possible at realistic data models. As such, our research is inspired by the pioneering works of Knuth.

The need to analyze, dimension, and finely optimize algorithms requires an in-depth study of random discrete structures, like words, trees, graphs, and permutations, to name a few. Indeed, a vast majority of the most important algorithms in practice either "make bets" on the likely shape of input data or even base themselves of random choices. In this area we are developing a novel approach based on recent theories of combinatorial analysis together with the view that discrete models connect nicely with complex-analytic and asymptotic methods. The resulting theory has been called "Analytic combinatorics". Applications of it have been or are currently being worked out in such diverse areas as communication protocols, multidimensional search, data structures for fast retrieval on external storage, data mining applications, the analysis of genomic sequences, and data compression, for instance.

The analytic-combinatorial approach to the basic processes of computer science is very systematic. It appeared early in the history of the project that its development would greatly benefit from the existence of symbolic manipulation systems and computer algebra. This connection has given rise to an original research programme that we are currently carrying out. Some of the directions pursued include automating the manipulation of combinatorial models (counting, generating function equations, random generation), the development of "automatic asymptotics", and the development of a unified view of the theory of special functions. In particular, the project has developed the Maple library Algolib, that addresses several of these issues.

2.3 Objectives for the evaluation period

Copy-pasted and translated from the previous evaluation document.

One of the first objectives is the completion of a vast synthesis on methods in *analytic combinatorics* (by P. Flajolet and R. Sedgewick, a Princeton success writer). Then the development of analysis of algorithms, along the natural evolution: worst-case (often obsolete) → average-case (current) → probabilistic analysis (i.e., in distribution, the future). We aim at: various hashing schemes; the dynamical analysis of algorithms (with

B. Vallée, Caen); random generation (thanks to a new idea with a physical origin: Boltzmann generators); the automation of the analysis of combinatorial structures by computer algebra; tree models and random graphs; problems exhibiting a *phase transition* in relation with analytic methods and coalescent saddle-points; elucidating the role of "holonomic" functions in the combinatorics of "hard" models, etc.

Our research projects in computer algebra are guided by two objectives. First, complete and make more systematic the automation of the analysis of combinatorial structures, in particular towards ordered structures and analyses in distribution. Second, take advantage of our specific strength in analysis to bring progress at the interface between computer algebra and numerical analysis. The idea is to bring together our work on special functions, recurrences and on asymptotic analysis so as to develop tools producing automatically efficient numerical code. This is the topic of L. Meunier's PhD thesis. It will be validated and focused by applications in physics. It will also be fueled by more research on the analysis of non-linear differential equations and on efficiency issues in commutative and non-commutative elimination.

Concerning the analysis of sequences, the recent trends, which will be pursued in the next years, concern the study of algorithms that search for sets of words or regular expression, and the extension to the (more realistic) Markovian model of analysis performed under the Bernoulli model. Biological applications reveal how crucial time and memory complexity of the algorithms are. Thus part of our research will aim at the efficient construction of minimal automata (with M. Raffinot) in the case of regular expressions and on the set of words neighboring a given word (with U. Bordeaux).

2.4 Evolution of the objectives

The long-term directions have not changed.

The unexpected discovery of the LogLog algorithm for approximate counting has focused some of our energy.

The hiring of Alin Bostan has allowed us to concentrate more on efficiency issues in computer algebra.

2.5 Objective 1 (Analysis of algorithms): Executive summary

The planned development of the book *Analytic Combinatorics* has been progressing at a fair pace, though the task of synthesizing a domain of such amplitude is quite labour-intensive. Random generation by means of Boltzmann models has been advancing and is gradually finding applications developed by partner teams in computational biology and software testing. Progress has been made in the analysis of phase transitions, in particular in connection with percolation on the random graph, which provides an entry point to the analysis of several hard optimization problems (e.g., constraint satisfaction). The perfection of the standard theory of analytic combinatorics has benefited from side studies related to urn processes, tree models, continued fractions, interconnection graphs, complex asymptotics, etc. New applications have been developed in the area of data compression (classical and new algorithms, e.g., by antidictionary), hashing, boolean function complexity, coding theory, and pattern occurrences in the context of intrusion detection. An unexpected development has been the revival of interest for probabilistic counting algorithms, which are currently the object of intense study by the network monitoring and data mining communities: the LogLog algorithm is currently the absolutely best algorithm available on the market (in terms of efficiency and robustness) for the fundamental task of estimating cardinalities of very large data sets of various sorts (router traces, genetic data, natural language corpuses, etc).

2.5.1 Personnel

Philippe Flajolet, Research Director, exceptional class, INRIA.
Bruno Salvy, Research Director, 2nd class, INRIA.
Frédéric Chyzak, Junior Research Scientist, 1st class, INRIA.
Marianne Durand, PhD student from École Normale Supérieure (Ulm).
Julien Fayolle, PhD student from University Paris VI.
Frédéric Giroire, PhD student from University Paris VI.
Vincent Puyhaubert, PhD student from École Normale Supérieure (Cachan).
Éric Fusy, PhD student from École Polytechnique and Corps des Télécommunications.

2.5.2 Project-team positioning

We estimate that we have a leading position world-wide in the area of analytic methods for the quantification of properties of random discrete structures, where we have been conducting a long-haul joint research programme, in particular through a cooperation with Robert Sedgewick at Princeton. This area is directly relevant to the design, analysis, and optimization of fundamental algorithms and data structures of computer science, especially the ones of low (polynomial-time) complexity. The approach taken strongly benefits from competences present in the Algorithms Project, regarding computer algebra and special functions, which gives to the theory developed here a unique twist. This approach also benefits of being developed within a computer science research institute, where numerous problems arise (often unexpectedly) from discussions with groups that conduct research in networking, performance evaluation, and probabilistic methods.

The book manuscript *Analytic Combinatorics*, though currently available only as a free web edition (it should be published by Cambridge University Press in 2008, and comprise 750 pages) is clearly influential and has already been used in courses by us and others at Oxford, Princeton, Barcelona, Vienna, Montevideo, Ottawa, Berlin, Philadelphia, Paris, Caen, Berkeley, Graz, etc.

In the area of analysis of algorithms *per se*, a solid international community of about 100 researchers has been formed, thanks largely to our efforts in setting up a series of workshops, conferences, and research collections (special issues, books). Given this and the relative lack of permanent members in this area at INRIA, we can only regard ourselves nowadays as a major international player in this field.

In data mining applications, the best organized team is that of Motwani at Stanford, whose research treats in remarkable depth some of the types of algorithmic problems we are interested in. We cannot compete with them on all fronts in their range of activities, given that our global objectives are more "foundational" and broader, and given the small amount of personnel available. Still, we have achieved the world record regarding the design (and validation) of efficient cardinality estimators.

Our research seems to be well accepted by mathematicians as attested by an invited lecture at the International Congress of Mathematicians (ICM, 2002), and by an election at the French Academy of Sciences in 2003.

Within INRIA, we have a very strong cooperation on these topics with Project RAP (dir.: P. Robert) with which we are engaged in two joint activities (ACI and ANR project). We cooperate with Project MAESTRO of Sophia Antipolis (A. Jean-Marie). We have regular exchanges with Project HIPERCOM (P. Jacquet, L. Viennot). More generally in France, we are at the origin of the Aléa Group, which gathers about 80 participants nationwide (largely from universities and CNRS) and is recognized as a component of GDR IM ("Mathematical Informatics"): we consider that we continue to play a leading scientific role in this active, friendly, and highly cooperative federation that we founded.

Our attitude to research in this chapter is that of elaborating long-term projects with a strong methodological flavour. We feel that this approach is the one that has conditioned successes in the past (e.g., in communication protocols, data base algorithms, bioinformatics, symbolic algorithms). The angle we have adopted, being based on analysis (either combinatorial, complex, or asymptotic) rather than the more customary stochastic theories, has provided us with original points of view in the design and optimization of several fundamental algorithms of computer science. We regard theory, provided it keeps an alert eye on other areas of computer science, as an essential component of a research institute, which, like INRIA, aims at results of the best quality.

2.5.3 Scientific achievements

Progress in analytic combinatorics has been along three major axes: (*i*) combinatorial enumeration; (*ii*) complex asymptotics; (*iii*) random structures and probabilities. Point (*i*) addresses the issue of setting up equations that translate a given combinatorial-probabilistic model. It has been shown that lattice paths, generation trees, and an important class of urn models (of relevance in particular to data structures and algorithms) can be treated systematically within the framework. Point (*ii*) aims at developing general-purpose methods for extracting asymptotics on coefficients of generating functions. In our perspective, this problem is best placed within the framework of singularity analysis, a technology first developed by Flajolet and Odlyzko: tangible progress has been attained in enlarging the class of functions amenable to the method (e.g., closure through Hadamard products, with a quaint turn-of-the-1900's mathematical flavour), which has applications to a number of tree data structures. Point (*iii*) has been developed in the direction of classifying non-Gaussian phenomena in combinatorics: this is a difficult task left largely untouched by our predecessors, but we have been successful in "explaining" some of the basic phenomena in percolation theory over the random graph—the methods (based on Airy functions and coalescent saddle points) have also proved useful in elucidating some curious phenomena present in large polynomial systems (see the computer algebra section). The eventual goal is to attack the analysis of hard combinatorial problems and their "phase transitions".

A major success in the category of algorithmic design is the LogLog counting algorithm. With a few thousand bits of memory, it is possible to estimate the number of distinct elements in a stream of data of several tens of gigabytes, with an accuracy of few percent (typically 2% with 2048 bytes). Another algorithm, MinCount can even be adapted to maintain running counts in a "sliding window", without degradation of efficiency. In this context, no a priori probabilistic assumption is made on the nature of data—the algorithms are universal. These algorithms have applications in the area of data mining and especially network monitoring. Such developments arose *in a totally unexpected unplanned manner*, as outcomes of casual exchanges with colleagues at UCSD (themselves under contract from CISCO), and were only made possible thanks to previous heavy investments in analytic combinatorics (here: uses of Mellin transforms in discrete probabilistic models).

Random generation has made a sizeable leap forward: we now have a general framework that makes it possible to construct in a systematic fashion random structured objects of large sizes (till sizes in the range 10^5–10^8). It currently applies to maybe 200 different discrete types. Applications outside INRIA are developing in the area of genomics, for the fast simulation of secondary RNA structures and sequences that are constrained by patterns, as well as in the perspective of random testing in software engineering, where a former PhD student is currently developing algorithms along these lines in the course of

his postdoctoral period.

The precise probabilistic analysis of pattern occurrences in sequences is a valuable outcome of analytic combinatorics (see below). We note here a paper published in the highly selective *Journal of the ACM*, which predicts the statistically unavoidable subsequences to be observed in a random text, a problem originally motivated by intrusion detection in computer security.

2.5.4 Collaborations

The list in this section is limited to teams with which we have co-authored articles during the period.

Princeton University, Computer Science (R. Sedgewick): write up of *Analytic Combinatorics*; interactions between design and analysis of algorithms.

Johns Hopkins University, Mathematics (J. Fill, N. Kapur): asymptotic methods and tree models.

Polytechnic University of Catalonya, Computer Science (J. Gabarro) and Applied Mathematics (M. Noy, O. Gimenez): urn models, data structures, maps.

University of Caen, Informatics (B. Vallée): probabilistic models of random words and intrusion detection; dynamical systems.

University of Paris 6, Informatics (M. Soria, C. Pivoteau): random generation and Boltzmann models.

University of Versailles, Mathematics and Computer Science (B. Chauvin, D. Gardy, N. Pouyanne): asymptotic methods, boolean complexity, random tree models.

Purdue University, Computer Science (W. Szpankowski, M. Ward): digital structures and compression; coding theory; pattern analysis and networks.

École polytechnique (G. Schaeffer, P. Nicodème): maps and graphs, patterns and computational biology, Boltzmann models.

Inria project Mascotte, Sophia Antipolis (J.-C. Bermond): satellite networks and probabilistic design.

University of Bordeaux, Informatics (P. Duchon, M. Bousquet-Mélou): combinatorics of paths, random generation, Boltzmann model.

University of Orsay, Informatics (A. Denise, D. Gouyou-Beauchamps): combinatorics of paths and random generation.

Free University of Brussels, Informatics (G. Louchard): random generation and Boltzmann models.

Patras University, Computer Science (P. Spirakis et al.): interconnection networks and probabilistic analysis.

University of Villetaneuse, Informatics (C. Banderier): lattice paths and enumerative combinatorics.

Technical University of Vienna (M. Drmota, B. Gittenberger): random trees, boolean complexity, asymptotic methods.

Humbold Universität zu Berlin, Informatik (M. Bodirsky, M. Kang): graphical enumeration.

2.5.5 External support

ALCOM-FT, Long term research Esprit project of the European community.

ACI ACPA within the programme "Mathematics and interfaces" (ACPA: Trees and Paths, Probabilty and Algorithms).

ACI FLUX within the programme "Massive Data" (FLUX: "Data Flows"), jointly with Project RAP, INRIA Rocquencourt, involving two other teams.

ANR Project NT SADA (Random Discrete Structures and Algorithms), also jointly with Project RAP, INRIA Rocquencourt, involving three other teams.

2.6 Objective 2 (Computer algebra): Executive summary

We now view the automatic analysis of algorithms as being best handled through attribute grammars, that admit of a direct translation into multivariate generating function equations. This has been implemented in our combstruct package (part of Algolib, see software section).

As part of our work on special functions, we have produced an automatically generated encyclopedia of special functions, the ESF, that gets approximately 27,000 http requests per month. The same tools that are used to generate the ESF have also been extended to generate symbolic Maple routines for the computation of series and asymptotic expansions. These routines are part of our MultiSeries package, whose aim is to replace all the asymptotic routines in Maple by more powerful ones.

Our research program on efficiency issues has produced the fastest known algorithms for evaluation and interpolation of polynomials, for the evaluation of linear recurrences, with applications to the resolution of linear differential and difference equations and to Zeilberger's algorithm. Also, several extensions of Newton's method have led us to new and fast algorithms for power series solutions of differential systems, for the computation of isogenies between elliptic curves, and a numerical algorithm with quadratic convergence to multiple roots and clusters of roots.

The techniques of analytic combinatorics have been applied with great success to the "generic" complexity of Gröbner basis computations. In particular, we derived the asymptotic behavior of the highest degree of polynomials entering a Gröbner basis for so-called semi-regular systems and showed that those involved such unexpected quantities as the largest roots of Hermite polynomials and the largest zero of the Airy Ai function.

2.6.1 Personnel

Bruno Salvy, Research Director, 2nd class, INRIA.
Frédéric Chyzak, Junior Research Scientist, 1st class, INRIA.
Alin Bostan, Junior Research Scientist, 2nd class, INRIA.
Philippe Flajolet, Research Director, exceptional class, INRIA.
Ludovic Meunier, PhD Student from École polytechnique.

2.6.2 Project-team positioning

The originality of our team in computer algebra comes largely from our long term research program on the applications of computer algebra to the analysis of combinatorial structures and special functions. This provides us with concrete examples that point to directions where new or faster algorithms are needed. Along the way, we design algorithms and implementations of wide applicability. This work is well recognized in the computer algebra community as is attested by our contributions to the program committees and editorial boards of the field.

We have a leading position on the use of Ore algebras in summation and integration. Our work has stimulated a lot of research both on non-commutative elimination and on stating algorithms that operated previously either on differential or on recurrence equations at the level of Ore polynomials. Interesting work is thus being done in Russia (Tsarev), in China (Ziming Li) and in several other teams (Waterloo, Moscow, Linz,

...) bringing new ideas from various backgrounds. We are also in touch with an active community in effective differential algebra, D-module theory and algebraic analysis. Their algorithms often provide the basis for nontrivial generalizations to Ore algebras.

Our recent steps towards revisiting the algorithms on linear recurrences and differential equations from the point of view of complexity are original to the team. We build upon all the recent progress in fundamental algorithms for algebraic computation (notably the works of Pan and Kaltofen), and bring our own improvements to some of the most classical algorithmic questions.

2.6.3 Scientific achievements

Our activity in computer algebra is largely centered around the effective manipulation of complex functions and functional systems. An original angle is the importance given to D-finiteness and its applications. This consists in exploiting linear operators (differential, difference, q-difference and more generally Ore polynomials) as a data structure to represent their solutions. In the multivariate case, the operation of *creative telescoping*, due to D. Zeilberger and extended by us to ideals of Ore algebras, is a non-commutative elimination that computes definite integrals and sums. There are several ways to perform this elimination, either based on non-commutative Gröbner bases or on an algorithm due to Chyzak that relies on rational solutions of systems of linear operators. In practice both these operations are very demanding in terms of computational power.

We have made progress along three main directions: extending the applications of D-finiteness; improving the efficiency of fundamental operations of computer algebra; improving the efficiency of commutative elimination.

Applications of D-finiteness In the univariate setting, we have gathered many of the classical algorithms on D-finite series and sequences in one showcase application: the automatically generated Encyclopedia of Special Functions (ESF). Moreover, the generator of this encyclopedia has been tailored to produce symbolic routines so that in our MultiSeries package devoted to symbolic asymptotics, all the code dealing with special D-finite functions (Bessel, Struve, Anger, polylogarithms, Airy, Weber,...) has been generated automatically. While many classical functions and sequences belong to the D-finite universe, there are strong constraints of analytic origin that prevent simple sequences from being D-finite. We showed how some of these constraints lead to direct proofs that sequences like the sequence of prime numbers or the logarithms of integers cannot satisfy linear recurrences with polynomials coefficients. In the multivariate context, we showed how the computation of Gröbner bases in Ore modules applies to the identification of structural properties of linear control systems. Finally, in the context of D-finiteness in "an infinite number of variables", we made operations on Gessel's D-finite symmetric functions effective, with applications to k-regular graphs and k-uniform Young tableaux. For a variety of generating sequences of objects of combinatorial origin, this provides a way to compute systems of linear differential equations that their generating function satisfy. From there, asymptotic estimates can then be derived by analytic combinatorics.

Efficiency of fundamental algorithms In the period, we obtained the best algorithms known for many fundamental operations of computer algebra: polynomial evaluation and interpolation; polynomial matrix multiplication; evaluation of linear recurrent sequences (improving an algorithm of the Chudnovsky's) and, as a by-product, deterministic factorization of integers (improving a 1970 algorithm due to Strassen); bivariate polynomial factorization (the first algorithm whose complexity is less than quadratic); addition and

multiplication of algebraic numbers in quasi-optimal complexity; computation of power series solutions of differential linear and non-linear systems, also in quasi-optimal complexity. All these algorithms lie at the heart of many of the operations that are needed throughout computer algebra and in particular in the manipulation of D-finite sequences and series. We thus applied these improvements to design fast algorithms for polynomial and rational solutions of linear difference and differential equations, and to improve on the complexity of Zeilberger's algorithm for hypergeometric summation. We also considered several related problems in positive characteristic, improving the efficiency of computations of the number of points on elliptic curves.

Commutative elimination As a preliminary step towards achieving a good complexity for the non-commutative elimination needed in creative telescoping, we studied several of its commutative counterparts. In two variables, we exhibited quasi-optimal algorithms for special types of bivariate resultants that play an important role in the manipulation of algebraic numbers. In an arbitrary number of variables and in the zero-dimensional case, we showed how a baby step–giant step technique applies to reduce the complexity of linear algebra in the quotient rings of zero-dimensional ideals. In the analytic counterpart of this setting, we considered how multiplicity and near-multiplicity (clusters of roots) can be attacked numerically, by designing a Newton iterator that converges quadratically to such roots, with a complexity that does not blow up with the multiplicity.

Finally, we studied the complexity of commutative Gröbner bases. It is well known that the worst case is doubly exponential, but the generic case is much better, being only singly exponential (or even polynomial if one measures the complexity in terms of the size of the output). It is then an important question to characterize the exponent of this generic complexity. First, in the regular case, we give a precise upper bound for the complexity of Faugère's F5 algorithm, which is the basis of the fastest current implementation of Gröbner bases. Then in the overdetermined case (which is important in many applications and in particular in cryptanalysis), the techniques of analytic combinatorics enabled us to show that this complexity exponent was related to the largest root of a Hermite polynomial (few more equations than unknowns) or to the largest zero of the Airy function (even more equations).

2.6.4 Collaborations

The list in this section is limited to teams with which we have co-authored articles during the period.

École polytechnique, Computer Science (M. Giusti, É. Schost, F. Ollivier): Polynomial systems, evaluation and interpolation, bivariate factorization, numerical deflation, fast power series for differential systems, Newton sums in small characteristic.

Risc, Linz (P. Paule): chapter on Computer Algebra for the Digital Library of Mathematical Functions (NIST).

RWTH, Aachen (D. Robertz): Ore modules in control theory.

Simon Fraser University, Vancouver, Mathematics (M. Mishna): D-finite symmetric functions.

University of Angers, Mathematics (M. Loday-Richaud): linear differential operators.

University of Cantabria, Santander, Mathematics, Statistics and Computation (L. Gonzalez-Vega): Newton sums in small characteristic.

University of Lille, Computer Science (A. Sedoglavic): multi-series, fast power series for differential systems.

University of Limoges, Mathematics (M. Barkatou): linear differential operators.

University of Toulouse, Mathematics (J.-C. Yakoubsohn): numerical deflation.

University of Versailles, Mathematics and Computer Science (G. Lecerf): bivariate factorization, numerical deflation.

Inria development team Scilab, Rocquencourt (C. Gomez): chapter on computer algebra in a French journal for engineers.

Inria project Café, Sophia-Antipolis (A. Quadrat, T. Cluzeau): Ore modules in control theory, solutions of linear differential and difference equations.

Inria project Salsa, Rocquencourt (J.-C. Faugère, M. Bardet): complexity of Gröbner bases.

Inria project Tanc, Futurs (F. Morain, P. Gaudry): algorithms for elliptic curves.

2.6.5 External support

ANR Project Gecko (A Geometric Approach to Complexity and its Applications), coordinator: B. Salvy, jointly with École polytechnique, U. Nice, U. Toulouse and G. Villard, G. Lecerf, J. van der Hoeven, O. Ruatta.

Waterloo Maple Inc. The context is described in the section on External Funding below. In this period, our activity on this contract has been devoted to maintaining and integrating the Groebner package; to extensions of the gfun package; and to the new MultiSeries package for asymptotic expansions.

2.7 Objective 3 (Analysis of sequences) : Executive summary

Our goal is to combine analytic combinatorics to establish general mathematical results and algorithms that provide efficient optimized implementations for specific data constraints.

We have derived expressions for the probability of rare words in random large texts and the waiting time of sets of words in small texts. To achieve this, we developed a formalism that applies to the Bernoulli model as well as to the more realistic Markov model, with a minimum increase of computational complexity. A formal study of statistical criteria commonly used by bioinformaticians has been performed.

The knowledge of secondary structures folding is a crucial step to predict 3D folding and to study phylogeny. A new interest arose recently, with the discovery of miRNA properties. Our heuristics combines phylogeny and probability in a Divide-and-Conquer approach. The more recent version, P-DCfold, deals efficiently with pseudo-knots, the most expensive constraint to deal with.

2.7.1 Personnel

Mireille Régnier, Research Director, 2nd class, INRIA.
Mathias Vandenbogaert, PhD Student from University of Bordeaux.
Julien Fayolle, PhD student from University Paris VI.

2.7.2 Project-team positioning

The design of statistical software to extract exceptional words from texts was very hot during the last few years. An international competition of statistical regulation-oriented software was organized in 2004, and our group was the only French group to participate, presenting its library, QuickScore, of C procedures. While most bioinformatics teams design their own software, somehow neglecting the precision of the statistical aspect, the

originality of our group is to formally establish precise criteria and their validity domains, reusable in other software. A second originality is our approach that provides mathematical results under the most general assumptions and specifies afterwards optimized algorithms for specific applications.

U. Keich, at Cornell University, is heading the closest team. Their goal is (also) a rigorous analysis from a mathematical point of view of heuristic models used in bioinformatics. Their research is more focused on alignment methods than sites prediction.

Word counting and pattern matching is a very active area in France, and members of the group are active in national networks in this area, notably "Word combinatorics, algorithms for texts and genomes" of the CNRS GDR "MathInfo", headed by G. Kucherov. For pattern matching, Marne-la-Vallée is a leading group at the international level. M. Régnier and P. Nicodème collaborate with them on automata and word counting. The Inria project Helix with Marie-France Sagot is a leading group for motif algorithms on the genome. The group ABISS at INRA, with S. Schbath and S. Robin, and "Statistics and genome" at Evry made important contributions to word counting, with an approach that is more statistical than algorithmic.

The Research and Training Center on Bioinformatics, headed by M. Gelfand, is a leading group at international level in comparative genomics and methods and algorithms for identification of regulatory signals.

2.7.3 Scientific achievements

Our group studies combinatorial properties of words in order to improve searching algorithms and extract unexpectedly frequent or rare motifs from random sequences. Our main motivation is to relate such motifs to biological functions on the genome. Ultimately, statistical or probabilistic results are to be integrated into string searching algorithms or software. Starting from exact counting formulae derived by the tools of analytic combinatorics, we focus on efficient implementations and tight approximation derivations. On the one hand, efficient implementations have been realized for sets of words, that rely on automata theory. Our main application is the study of words, also called sites on the genome, that may be recognized by different proteins, in a competing or synergistic manner. On the other hand, approximations have been derived, with a tight error bound. We have shown that the probability of rare events, commonly called *p-value* by bioinformaticians, is computable with a linear algorithm, a drastic improvement on the existing exponential algorithm. Tightness of these *p-value* expressions allows to compute conditional probabilities, and extract weak signals hidden by stronger signals. As a whole, systematic comparisons of commonly used statistical criteria for exceptional words, and precise statements on their validity domains, have been realized. A special attention has been paid to Tandem Repeats. The originality of the software TandemSWan, developed jointly with NIIGenetika, is the implementation of a procedure to assess the statistical significance of repeats. Such results can be implemented in other software, including Mreps developed at Inria Futurs (Lille) by G. Kucherov.

P-value computation was implemented in software, including RSA-tools developed at ULB in Bruxelles or ScanSeq developed at NIIGenetika, with whom we collaborate. D. Papatsenko (UC Berkeley) confirmed experimentally some predictions of sites on *Drosophila* embryos. TandemSwan results led to establish a relationship between tandem repeats and regulatory mechanisms.

RNA secondary structures prediction is a hard problem, as a given sequence has an exponential number of potential foldings. It is even worse when pseudo-knots are taken into account. Finding a structure, e.g., a set of helices, that is common to sequences

from phylogenetically closed organisms turned out to be the most efficient way. This is meaningful as sequencing is often realized now for closely related genomes. Our heuristic relies on a few simple probability criteria to chose common helices, and led recently to a variant, P-DCfold, that deals with pseudo-knots.

2.7.4 Collaborations

The list in this section is limited to teams with which we have co-authored articles during the period.

Évry University (F. Tahi): secondary structures folding.

Orsay University (A. Denise): random strings and genome.

Marne-la-Vallée University (J. Clément): string searching and automata.

École polytechnique (P. Nicodème): word counting and tandem repeats.

Institut Pasteur (M. Vandenbogaert): regulation and prediction In Silico.

NYU and now UC Berkeley (D. Papatsenko): experimental validation of prediction on biological data.

NIIGenetika (V. Makeev): algorithms for biopolymers.

Purdue University (W. Szpankowski): suffix trees.

A collaboration with NIIGenetika led to the joined development of TandemSwan and the implementation in ScanSeq of our combinatorial results.

2.7.5 External support

GenomAl, joint project between Algo, the Computer Science laboratories at École polytechnique (Lix) and the University of Bordeaux (Labri). Applications are provided by biology labs, including ULB, NYU and NIIGenetika, and mainly related to motifs search or structures folding.

3 Knowledge dissemination

3.1 Publications

	2002	2003	2004	2005	2006
PhD Thesis		1	2	1	2
H.D.R (*)					
Journal	9	5	8	10	10
Conference proceedings (**)	5	8	11	9	11
Book chapter				1	
Book (written)					1
Book (edited)	2	1	1	1	
Patent					
Technical report				1	5
Deliverable					

(*) HDR Habilitation à diriger des Recherches

(**) Conference with a program committee

Indicate the major journals in the field and, for each, indicate the number of papers coauthored by members of the project-team that have been accepted during the evaluation period.

By alphabetical order:

1. Algorithmica (0);
2. Discrete Mathematics & Theoretical Computer Science (2);
3. Journal of Symbolic Computation (3);
4. Journal of the ACM (1);
5. Random Structures & Algorithms (0);
6. The Electronic Journal of Combinatorics (3);
7. Theoretical Computer Science (7).

Indicate the major conferences in the field and, for each, indicate the number of papers coauthored by members of the project-team that have been accepted during the evaluation period.

By alphabetical order:

1. AofA: Analysis of Algorithms (5);
2. FPSAC: Formal Power Series and Algebraic Combinatorics (2);
3. ICALP: International Colloquium on Automata, Languages and Programming (2);
4. ISSAC: International Symposium on Symbolic and Algebraic Computation (8);
5. SODA: ACM-SIAM Symposium on Discrete Algorithms (4).

3.2 Software

The Algolib library is a set of Maple routines that have been developed in the project for more than 10 years. Several parts of it have been incorporated into the standard library of Maple, but the most up-to-date version is always available for free from our web pages. This library provides: tools for combinatorial structures (the combstruct package), this includes enumeration, random or exhaustive generation, generating functions for a large class of attribute grammars; tools for linear difference and differential equations (the gfun package), which have received a very positive review in Computing Reviews and have been incorporated in N. Sloane's superseeker at Bell Labs; tools for systems of multivariate linear operators (the Mgfun package), including Gröbner bases in Ore algebras, that also treat commutative polynomials and have been the standard way to solve polynomial systems in Maple for a long period (although the user would not notice it); Mgfun has also been chosen at Risc (Linz) as the basis for their package Desing.

We also provide access to our work to scientists who are not using Maple or any other computer algebra system in the form of automatically generated encyclopedias available on the web. The Encyclopedia of Combinatorial Structures thus contains more than 1000 combinatorial structures for which generating series, enumeration sequences, recurrences and asymptotic behavior have been computed automatically. It gets more than 16,000 hits per month. The Encyclopedia of Special Functions gathers around 40 special functions for which identities, power series, asymptotic expansions, graphs,... have been generated automatically, starting from a linear differential equation and its initial conditions. The underlying algorithms and implementations are those of gfun and Mgfun. All the production process being automated, the difficult and expensive step of checking each formula

individually is suppressed. Available on the web (http://algo.inria.fr/esf/), this encyclopedia also plays the role of a showcase for part of the packages developed in our project. It gets 27,000 hits per month.

A new package, MultiSeries has been developed during the period. It implements so-called multi-series, that are series in general asymptotic scales, each of whose coefficient is itself potentially a new series. This makes it possible to handle in a transparent and dynamic way the problems of finding the proper asymptotic scale for an expansion and of dealing with the indefinite cancellation problem. This package is designed in such a way that it can take the place of the existing series, asympt and limit Maple functions, in a totally transparent manner.

Tandem repeats are short repetitions that are hotspots for genome recombinations and are also related to some genetic diseases. TandemSWAN searches for degenerate tandem repeats without insertions and deletions, but with a high substitution rate. It is based on calculation of the repeat statistical significance and identifies the length of the repeated unit and the number of repetitions. It allows for the identification of weak clustered sites, that are needed for the analysis of several important regulatory systems such as nitrate/nitrite switch. It is written in C language, using some C++ features.

P-DCFold implements in Java a heuristic algorithm for the prediction of RNA secondary structures including all kinds of pseudoknots. It is based on the comparative approach, and its input is a small set of RNA sequences. It has been applied to *tmRNA* and *RnaseP* sequences.

3.2.1 Valorization and technology transfert

Part of our code ends up in the Maple library, see sections 2.6.5 and 4.

3.3 Teaching

Alin Bostan, Frédéric Chyzak and Bruno Salvy together with Marc Giusti and Éric Schost (the latter 2 at École polytechnique) have set up and teach a course on computer algebra in the Parisian Master of Research in Computer Science. They each give approximately 12h/year in this course.

Frédéric Chyzak has been spending 96h/year in teaching part time at École polytechnique, till mid-2006. There he taught several computer science courses, including one on computer algebra.

Philippe Flajolet gives on average 24h/year at the Parisian Master of Research in Computer Science, on the analysis of algorithms.

Mireille Régnier gives 35h/year of courses in the Bioinformatics Master of Évry and Orsay and in École Centrale, Paris, on combinatorics and algorithms in genomics.

3.4 Visibility

Algorithms Seminar The ALGO project runs a biweekly seminar. Several partner teams in the grand Paris area attend on a regular basis. Summaries of the talks are written by volunteers in the audience and the proceedings are edited and published [2, 3, 5].

Alin Bostan has been member of the poster committee for the International Symposium on Symbolic and Algebraic Computation ISSAC'05 (Beijing, China) and of the program committee for ISSAC'06 (Genoa, Italy).

He was an invited speaker at the workshop "Algebraic Complexity Theory meets Algorithmic Differentiation" organized at Humboldt Universität zu Berlin, Germany, in 2004.

Frédéric Chyzak has been a member of the program committee for ISSAC'05 (Beijing, China). He is one of the three co-organizers of the French computer algebra meeting in Luminy, 2007.

Together with Peter Paule (Risc, Linz), he has been invited to write the methodological chapter on computer algebra for the Digital Library of Mathematical Functions (the new version of the Abramowitz and Stegun) edited by the National Institute for Standard and Technology.

He is a member of the recruiting committee of Univ. Limoges, in mathematics.

Philippe Flajolet serves as Chair of the Steering Committee of the international series of Conferences and Workshops called "Analysis of Algorithms". The yearly edition (Strobl, Austria, 2002; San Miniato, Italy, 2003; MSRI, Berkeley, 2004; Barcelona, 2005; Alden Biesen, Belgium, 2006; Juan-les-Pins, France, 2007) attracts some 80 specialists of the area. He serves in a similar capacity as founder and chair of the French Working Group Aléa supported by CNRS: the yearly meetings are held at Luminy near Marseilles, and the participation nears 80 every year.

He has been program committee chair of the Colloquium on Mathematics and Computer Science, for the 2002 edition, Versailles, France; and regular member for the 2004 edition, Vienna, Austria. He has also been in the program committee of ICALP 2003, Malaga, Spain.

He is an editor of the journal Random Structures and Algorithms, an honorary editor of Theoretical Computer Science, and an honour member of the French association SPECIF. He also serves as one of the three editors of Cambridge University Press' prestigious series "Encyclopedia of Mathematics and its Applications".

Philippe Flajolet has been invited Plenary Speaker at the "KnuthFest", Stanford, 2002. In 2002, he has also been invited speaker at the International Congress of Mathematicians, Beijing, China. He has been invited Plenary Speaker at the First Workshop on Analytic Algorithmics and Combinatorics (ANALCO04), New Orleans, January 2004. He has been one of the four invited speakers at ICALP'04 (Turku, Finland, July 2004), which is the major European conference in theoretical computer science in Europe. He has been (together with Don Knuth and Persi Diaconis) one of the keynote speakers at the Tenth Seminar on Analysis of Algorithms AofA'04 (Berkeley, USA, June 2004). He has also given the Opening Keynote Address at the Ninth Asian Computing Science Conference ASIAN'04 (Chiang-Mai, Thailand, December 2004). He has been one of the invited speakers (in 2005) at Séminaire Lotharingien de Combinatoire, which despite its name is an international event. In 2006, he has invited speaker at the International Symposium on Theoretical Aspects of Computer Science (STACS 2006) Marseille, France; at the Colloquium on Mathematics and Computer Science : Algorithms, Trees, Combinatorics and Probabilities, Nancy, France; at the conference Gascom'06 on random generation, Dijon, France; at the ACM-SIAM Symposium on Discrete algorithms (SODA'07), New Orleans.

Philippe Flajolet has been invited to teach three postgraduate courses of 10 to 12 lectures each in Barcelona (April 2004, Polytechnic University of Catalonia, doctoral programme), Berkeley (June 2004, under the auspices of the Mathematical

Sciences Research Institute), and Chiang-Mai, Thailand (the ASIAN'04 postconference school).

In 2003, he has served on the Committee on the Teaching of Mathematics commissioned by the Ministry of Education.

In 2004, he has been awarded the Silver Medal of CNRS for his contributions to research in computer science. (Such a distinction is awarded only every second year to a computer scientist in France.)

In 2005, for its 30th Anniversary, the journal Theoretical Computer Science listed its 100 most cited articles. Among those, two were authored by P. Flajolet: Analytic Models and Ambiguity of Context-Free Languages (1987); Mellin transforms and asymptotics: Harmonic sums (1995), with Ph. Dumas and X. Gourdon.

In 2006, for its 35th Special Anniversary Issue, the journal Discrete Mathematics reprinted a selection of 23 papers published in its 37 years, (Volume 306, Issue 10–11). Among them is Philippe Flajolet's 1980 "Combinatorial Aspects of Continued Fractions".

Philippe Flajolet is also an external member of the Recruiting Committee for computer science at the École polytechnique. In 2005, he served as a member of the College of Reviewers for the Canada Research Chairs Program (mathematics and computer science), as well as an external reviewer for chairs (full professorships) at the Universities of Vienna and Turku. He has also served as member of the evaluation committee of ACI-NIM, a concerted action of the French Ministry of Education dedicated to the new interfaces of mathematical sciences.

Since 2005, Philippe Flajolet has assumed the somewhat heavy responsibility of chairing the Scientific Committee for Mathematics of the newly formed National Research Agency (ANR), which implied the heavy responsibility of launching a programme of some 5 million Euros.

He is a member of the French Academy of Sciences and of the Academia Europaea.

Mireille Régnier has co-organized the new Moscow Conference on Computational Molecular Biology (MCCMB) in 2003 and 2005, that was supported by ERCIM and INRIA. This new conference is held on odd years, while the older BGRS in Novosibirsk is held on even years. Mireille Régnier has been on the program committee of MCCMB in 2003. She has also been on the program committee of a satellite meeting on Regulation of RECOMB'04, 05 and 06.

She was an invited speaker at the Drug and Discovery Bioinformatics conference in Mainz, Germany, 2003.

Mireille Régnier has been invited at the colloquium "Entangling Mathematics, Life Sciences and Information Technology : Biomatics, a Science of multi-scale complexity", Novossibirsk, Russia.

She organized in Erevan (Armenia) a 10 days school "Combinatorics and Genome" in 2004. She has been invited by INTAS to BGRS'04 for a prospective workshop on the collaboration between EC and NEI in Life Science.

M. Régnier participated to the organization of the new computer science option at the French *agrégation* of mathematics. She served in the committee in 2006.

Bruno Salvy has been a member of the steering committee of ISSAC in the period 2000–2003. He was on its program committee in 2002 (Lille, France) and in 2004 (San-

tander, Spain). He was one of the three invited speakers in the 2005 edition, Beijing, China (the other two were Bruno Buchberger and Wen-Tsun Wu). He has also been on the program committee of the French-Canadian Congress on Mathematical Sciences Toulouse, 2004, and of the conference "Computational Geometry and Applications", Nice, 2006 and he is on the program committee for the conference Formal Power Series and Algebraic Combinatorics, Talca, Chile, 2008. He is organizing the working group Computer Algebra of the CNRS GDR IM ("Mathematical Informatics").

Before being defeated by spam, he has been maintaining and animating the French mailing list on Maple for many years.

He is member of the editorial board of the Journal of Symbolic Computation and of the Journal of Algebra (section Computational Algebra).

He is or has been in the period a member of the recruiting committees of the University of Lille (in computer science), of the University of La Rochelle (in mathematics), as well as a member of the scientific council of the University of Versailles.

He has also been engaged in a number of recruiting committees at Inria Rocquencourt: junior researchers in 2005; promotion committee for research technicians (a category of administrative staff) in 2005; he has organized the committees for postdocs and for the recruitment of researchers from other institutes and universities (détachements and délégations) in 2005 (with A. Sulem) and 2006 (with S. Gaubert).

4 External Funding

(k euros)	2002	2003	2004	2005	2006
National initiatives					
PAI Amadeus	3	3			
GénomAl		23			
ACI ACPA		4.5	4.5	4.5	1.5
ECO-NET			19.5	19	16
ACI FLUX			16	16	16
ANR GECKO					23
ANR SADA					12
European projects					
ALCOM-FT	46	23			43
IST INTAS					10
Industrial contracts					
Waterloo Maple Inc.	28	28	28	28	28
Total External Contracts	**77**	**81.5**	**68**	**67.5**	**149.5**
Scholarships					
AI†	17.2	43			
PhD AMX, École polytechnique	60	60	52.5		
PhD AMN, École normale supérieure	60	60	40	7.5	
PhD MESR	37.5	90	67.5	60	30
PhD, Corps des Télécoms			24.7	49.3	49.3
Post Doc INRIA		12.4	9.3		
Post Doc ERCIM			9.3	27.9	
Post Doc Corps des Ponts					22
Total Scholarships	**174.7**	**265.4**	**203.3**	**144.7**	**101.3**
Total	**251.7**	**346.9**	**271.3**	**212.2**	**250.8**

† junior engineer supported by INRIA

Apart from the amounts described here, we are supported by Inria in the form of salary for the permanent members and the following amounts for our expenses in the period:

(k euros)	2002	2003	2004	2005	2006
	58	29.5	51	2.5	−49

The last amount being negative means that the income from our contracts was higher than our expenses.

National initiatives

PAI Amadeus: This is a bilateral action between France and Austria, funded by the French Ministry of Foreign Affairs. It involves the Algo project, the Vienna University of Technology, and RISC (University of Linz). Several joint work have been completed thanks to this, notably with PhD student Stefan Gerhold (RISC, Austria) on holonomic sequences, with Michael Drmota (Vienna) on the distribution of the number of occurrences of patterns in random trees and on the extraction of coefficients in bivariate generating series by double saddle-point contours.

GénomAl: joint project between Algo, the Lix at École polytechnique and the Labri at the Bordeaux University. The scientific objective is to study randomness in the genomes and implement efficient algorithms to compute mathematical results and predict relevant motifs. Applications are provided by biology labs, including ULB, NYU and NIIGenetika.

ACI ACPA: For the period 2003–2006, the Algorithm project has participated in a national research programme exploring New Interfaces of Mathematics. In this context, we have taken part in the ACPA project dedicated to paths and trees, probabilities and algorithms, this jointly with the Universities of Versailles, Bordeaux, and Nancy.

ECO-NET: This is a joint project with an Armenian, a Georgian and a Russian team funded by the French Ministry of Foreign Affairs, entitled "Regulatory signals: in Silico prediction and experimentation".

ACI FLUX: this project involves the Rap project at Inria as well as the University of Montpellier and has been funded for a three year period by the national action ACI-MD relative to massive data: our objective is to develop high performance algorithms for the quantitative analysis of massive data flows, an important problem in the monitoring of high speed computer networks.

ANR GECKO: For the period 2006–2009, the Algorithm project participates in a programme funded by the National Research Agency (ANR) entitled GECKO for "A Geometric Approach to Complexity and its Applications". Four teams are involved: Algo (coordinator) and teams at the École polytechnique, the Universities of Toulouse and Nice. The project concentrates on three classes of objects: (i) univariate and multivariate polynomials (Newton process, factorization, elimination); (ii) structured matrices (whose coefficients can be polynomials); (iii) linear differential operators (noncommutative elimination, integration). The aim is to improve significantly the resolution of systems of algebraic or linear differential equations that appear in models, by taking geometry into account.

ANR SADA: The goal of this project is to investigate fundamental properties of random discrete structures and algorithms. The project duration is 3 years (Dec. 2005–Dec 2008). It involves five teams: Algo and Rap from Inria Rocquencourt, the Universities of Caen, Versailles, and Bordeaux (coordinator), as well as the Laboratory for Computer Science of the École polytechnique (LIX).

European projects

ALCOM-FT: For a period of three years (2000–2003), we have been one of the components of the project ESPRIT "Long Term Research" ALCOM-FT. This project gathered ten leading groups in the field of algorithmic research in Europe. The objective was to find new algorithmic concepts and identify transverse key algorithms of many applications. Four directions of work have been identified: (i) Massive data sets; (ii) Communication systems; (iii) Optimisation in production and planning; (iv) Methodological and experimental algorithmic research. Work of our project was mainly in the axes (ii) and (iv).

IST INTAS: This is a grant to work on comparative genomics of bacteria: functional annotation, metabolic reconstruction, evolution of metabolic pathways andregulatory systems. This project is joint with TUM (Munchen), CNR-ITB (Milan), NII-Genetika (Moscow) and Moscow University. Its tasks are: (i) Algorithm and soft-

ware development (ii) gene recognition (iii) comparative evolution of orthologous sites (iv) identification of clustered regulatory sites (v) a server for prediction of RNA secondary structures. Our work is mainly in axes (i), (iii) and (v).

Industrial contracts

Waterloo Maple Inc. (WMI). The Algorithms Project and WMI have developed a collaboration based on reciprocal interests. It is obviously interesting for the company to integrate functionalities at the forefront of the current research in computer algebra. Reciprocally, this integration makes our programs and our research visible to a very wide audience.

Numerous exchanges have thus taken place between the project and the company over the years. After more than 3 years within the project, J. Carette has been for several years Product Development Director at WMI, before going back to the academic world. Similarly, E. Murray, who worked for two years in the project developing the combstruct package is now working at WMI.

Thanks to all this activity, the company WMI considers Inria as a special partner and grants it a free license for all of its research units. Moreover, a cooperation agreement has been signed between WMI and Algo in 2001. In particular, one of the objectives is to replace all the routines dealing with asymptotic and series expansions in Maple by implementation of new algorithms dealing with very general classes of asymptotic scales.

5 Objectives for the next four years

We plan to complete the synthesis work "Analytic Combinatorics" in 2008, a task that still involves a fair amount of surrounding research of a fundamental nature regarding limit distributions in combinatorics. Another objective is to investigate by analytic methods several model of random graphs that appear in the internet, the web, and in peer-to-peer networks, and at the same time develop data mining algorithms on such graphs that build upon our previous expertise in low-cost fast probabilistic estimation algorithms. On a methodological front, we propose to start investigating some of the probabilistic phenomena, known as threshold phenomena or phase transitions, attached to hard (NP-complete) optimization problems (e.g, boolean satisfiability and constraint satisfaction), a possible starting point being the representation of quantities of interest by means of large multidimensional structured integrals, somewhat in the spirit of statistical physics.

Our aim in computer algebra is increase the speed of definite summation and integration so as to attack the questions of large size being sent to us by physicists and chemists. The automatic generation of numerical approximation routines is a long term objective. This will be achieved through new algorithms and a systematic study of the related complexity issues. On the sizes we have in mind, the theoretical complexities are well reflected by computational timings and we will spend time to provide accordingly good implementations. Our language of choice might switch from Maple to Magma. The showcase provided by the ESF is a direction we believe in; it will be extended to integral transforms and if implementation power permits, we will add interactivity to it.

Our research on pattern matching and word counting will be pursued. We will develop a mathematical study of heuristic models in biology, and notably the Probability Weight Matrix Model, very much in favour. Combinatorial results on sets of words should be extended to multisets. Results on joint probabilities for multisets are relevant, as a clustering

of regulatory sites on genomes has been experimentally observed. This is also relevant to the study of metabolic networks. A second objective is the study of RNA, including newly discovered miRNA. One possible (combinatorial) direction is a classification of secondary structures according to their probability of appearance. A second direction is the search of repeats, such as mini-satellites or Tandem Repeats, in RNA, in order to establish a relationship with functional properties. These objectives are two main axes of a new Inria project, joint with J.-M. Steyaert (LIX, École polytechnique) that M. Régnier intends to propose within the next few months.

6 Bibliography of the project-team

6.1 Books and Monographs

[1] Chauvin, Brigitte, Flajolet, Philippe, Gardy, Danièle, and Mokkadem, A., Eds. *Mathematics and Computer Science II: Algorithms, Trees, Combinatorics and Probabilities*. Trends in Mathematics. Birkhäuser Verlag, Basel, 2002. 560 pages. Proceedings of a Colloquium held at Versailles, September 2002.

[2] Chyzak, Frédéric, Ed. *Algorithms Seminar, 2000–2001* (March 2002), vol. 4406 of *Research Report*. 199 pages. Institut National de Recherche en Informatique et en Automatique.

[3] Chyzak, Frédéric, Ed. *Algorithms Seminar, 2001–2002* (November 2003), vol. 5003 of *Research Report*. 191 pages. Institut National de Recherche en Informatique et en Automatique.

[4] Drmota, M., Flajolet, P., Gardy, D., and Gittenberger, B., Eds. *Mathematics and Computer Science III: Algorithms, Trees, Combinatorics and Probabilities*. Trends in Mathematics (Mathematics, Computer Science). Birkhäuser Verlag, 2004. 554 pages.

[5] Chyzak, Frédéric, Ed. *Algorithms Seminar, 2002–2004* (April 2005), vol. 5542 of *Research Report*. 120 pages. Institut National de Recherche en Informatique et en Automatique.

[6] Flajolet, Philippe, and Sedgewick, Robert. *Analytic Combinatorics*. Cambridge University Press, April 2006. Chapters I–IX of a book to be published by Cambridge University Press, 717p.+x, available electronically from P. Flajolet's home page.

6.2 Doctoral dissertations and "Habilitation" theses

[7] Bostan, Alin. *Algorithmique efficace pour des opérations de base en Calcul formel.* PhD thesis, École polytechnique, 2003.

[8] Durand, Marianne. *Combinatoire analytique et algorithmique des ensembles de données*. PhD thesis, École polytechnique, 2004.

[9] Vandenbogaert, Mathias. *Algorithmes et mesures statistiques pour la recherche de signaux fonctionnels dans les zones de régulation*. PhD thesis, University of Bordeaux, 2004.

[10] Puyhaubert, Vincent. *Modèles d'urnes et phénomènes de seuils en combinatoire analytique*. PhD thesis, École polytechnique, 2005.

[11] Fayolle, Julien. *Compression de données sans perte et combinatoire analytique*. PhD thesis, University of Paris VI, 2006.

[12] Giroire, Frédéric. *Réseaux, algorithmique et analyse combinatoire de grands ensembles*. PhD thesis, University of Paris VI, 2006.

6.3 Articles in referred journals and book chapters

[13] Banderier, Cyril, Bousquet-Mélou, Mireille, Denise, Alain, Flajolet, Philippe, Gardy, Danièle, and Gouyou-Beauchamps, Dominique. Generating functions of generating trees. *Discrete Mathematics 246*, 1-3 (March 2002), 29–55.

[14] Banderier, Cyril, and Flajolet, Philippe. Basic analytic combinatorics of directed lattice paths. *Theoretical Computer Science 281*, 1-2 (2002), 37–80.

[15] Cori, Robert, Rossin, Dominique, and Salvy, Bruno. Polynomial ideals for sandpiles and their Gröbner bases. *Theoretical Computer Science 276*, 1 (2002), 1–15.

[16] Flajolet, Philippe, Hatzis, Kostas, Nikoletseas, Sotiris, and Spirakis, Paul. On the robustness of interconnections in random graphs: A symbolic approach. *Theoretical Computer Science 287*, 2 (2002), 513–534.

[17] Flajolet, Philippe, and Szpankowski, Wojtek. Analytic variations on redundancy rates of renewal processes. *IEEE Transactions on Information Theory 48*, 11 (2002), 2911–2921.

[18] Nicodème, Pierre, Salvy, Bruno, and Flajolet, Philippe. Motif statistics. *Theoretical Computer Science 287*, 2 (2002), 593–618. Extended version of an article published in the proceedings of 7th Annual European Symposium on Algorithms ESA'99, Prague, July 1999.

[19] Papatzenko, D., Makeev, V., Lifanov, A., Régnier, M., Nazina, A., and Desplan, C. Extraction of functional binding sites from Unique Regulatory Regions: The *drosophila* Early Developmental Enhancers. *Genome Research 12* (2002), 470–481. Preliminary version in Drosophila Workshop, Washington 2001.

[20] Sedoglavic, Alexandre. A probabilistic algorithm to test local algebraic observability in polynomial time. *Journal of Symbolic Computation 33*, 5 (May 2002), 735–755.

[21] Tahi, F., Gouy, M., and Régnier, M. Automatic rna secondary structure prediction with a comparative approach. *Computers and Chemistry 26*, 5 (2002), 521–530.

[22] Bostan, Alin, Salvy, Bruno, and Schost, Éric. Fast algorithms for zero-dimensional polynomial systems using duality. *Applicable Algebra in Engineering, Communication and Computing 14*, 4 (2003), 239–272.

[23] Chassaing, Philippe, and Flajolet, Philippe. Hachage, arbres, chemins, et graphes. *Gazette des Mathématiciens* (2003), 29–49.

[24] Durand, Marianne. Asymptotic analysis of an optimized quicksort algorithm. *Information Processing Letters 85* (2003), 73–77.

[25] Durand, Marianne, and Taylor, Stephen. Emerging behavior as binary search trees are symetrically updated. *Theoretical Computer Science 297* (2003), 425–445.

[26] Vandenbogaert, M., and Makeev, V. Analysis of bacterial rm-systems through genome-scale analysis and related taxonomic issues. *In Silico Biology 12*, 3 (2003), 127–143. Special Issues of Volume 3: Bioinformatics on Genome Regulation and Structure (BGRS 2002, Novossibirsk), Published as online journal by Bioinformation Systems e. V.; regularly issued as printed version by IOS Press.

[27] Abramov, S. A., Carette, J. J., Geddes, K. O., and Le, H. Q. Telescoping in the context of symbolic summation in maple. *Journal of Symbolic Computation 38*, 4 (October 2004), 1303–1326.

[28] Bostan, A., and Schost, É. On the complexities of multipoint evaluation and interpolation. *Theoretical Computer Science 329*, 1–3 (December 2004), 223–235.

[29] Chauvin, Brigitte, Flajolet, Philippe, Gardy, Danièle, and Gittenberger, Bernhard. And/Or Trees Revisited. *Combinatorics, Probability and Computing 13*, 4–5 (2004), 501–513. Special issue on Analysis of Algorithms.

[30] Duchon, Philippe, Flajolet, Philippe, Louchard, Guy, and Schaeffer, Gilles. Boltzmann samplers for the random generation of combinatorial structures. *Combinatorics, Probability and Computing 13*, 4–5 (2004), 577–625. Special issue on Analysis of Algorithms.

[31] Flajolet, Philippe, Salvy, Bruno, and Schaeffer, Gilles. Airy phenomena and analytic combinatorics of connected graphs. *The Electronic Journal of Combinatorics 11*, 1 (May 2004), R34. 30 pages.

[32] Lescot, M., and Régnier, M. Motif statistics on plants datasets. *Biophysics 48*, 1 (2004), 1–6. Special issue on the Proceedings of the Moscow Conference on Computational Molecular Biology, MCCMB'03.

[33] Puyhaubert, Vincent. Generating functions and the satisfiability threshold. *Discrete Mathematics & Theoretical Computer Science 6*, 2 (2004), 425–436.

[34] Régnier, Mireille, and Denise, Alain. Rare events and conditional events on random strings. *Discrete Mathematics & Theoretical Computer Science 6*, 2 (2004), 191–214.

[35] Abramov, S. A., and Le, H. Q. On the order of the recurrence produced by the method of creative telescoping. *Discrete Mathematics 298*, 1–3 (August 2005), 2–17.

[36] Bostan, A., and Schost, É. Polynomial evaluation and interpolation on special sets of points. *Journal of Complexity 21*, 4 (August 2005), 420–446. Festschrift for the 70th Birthday of Arnold Schönhage.

[37] Chyzak, Frédéric, Mishna, Marni, and Salvy, Bruno. Effective scalar products of D-finite symmetric functions. *Journal of Combinatorial Theory, Series A 112*, 1 (October 2005), 1–43. Extended version of an article published in the proceedings of the 14th Conference of Formal Power Series and Algebraic Combinatorics FPSAC'02, Melbourne, July 2002.

[38] Chyzak, Frédéric, Quadrat, Alban, and Robertz, Daniel. Effective algorithms for parametrizing linear control systems over Ore algebras. *Applicable Algebra in Engineering, Communication and Computing 16*, 5 (November 2005), 319–376.

[39] Fill, James A., Flajolet, Philippe, and Kapur, Nevin. Singularity analysis, Hadamard products, and tree recurrences. *Journal of Computational and Applied Mathematics 174* (February 2005), 271–313.

[40] Flajolet, Philippe, Gabarró, Joaquim, and Pekari, Helmut. Analytic urns. *Annals of Probability 33*, 3 (2005), 1200–1233.

[41] Flajolet, Philippe, Gerhold, Stefan, and Salvy, Bruno. On the non-holonomic character of logarithms, powers, and the nth prime function. *The Electronic Journal of Combinatorics 11*, 2 (April 2005). A2, 16 pages.

[42] Giusti, Marc, Lecerf, Grégoire, Salvy, Bruno, and Yakoubsohn, Jean-Claude. On location and approximation of clusters of zeroes of analytic functions. *Foundations of Computational Mathematics 5*, 3 (July 2005), 257–311.

[43] Huet, Gérard, and Flajolet, Philippe. ch. Mathématiques et Informatique, pp. 215–237, In *Les mathématiques dans le monde scientifique contemporain*, vol. 20 of *Rapport sur la science et la technologie*. Académie des sciences, Paris, 2005. Edited by Jean-Cristophe Yoccoz.

[44] Tahi, Fariza, Engelen, Stefan, and Régnier, Mireille. P-dcfold or how to predict all kinds of pseudoknots in RNA secondary structures. *International Journal on Artificial Intelligence Tools 5*, 14 (2005), 703–716.

[45] Tompa, M., Li, N., Bailey, T.L., Church, G.M., De Moor, B., Eskin, E., Favorov, A.V., Frith, M.C., Fu, Y., Kent, J.W., Makeev, V.J., Mironov, A.A., Noble, W.S., Pavesi, G., Pesole, G., Régnier, M., Simonis, N., Sinha, S., Thijs, G., van Helden, J., Vandenbogaert, M., Weng, Z., Workman, C., Ye, C., and Zhu, Z. An assessment of computational tools for the discovery of transcription factor binding sites. *Nature Biotechnology 23*, 1 (January 2005), 137–144.

[46] Boeva, V., Makeev, V., Papatsenko, D., and Régnier, M. Short fuzzy tandem repeats in genomic sequences, identification, and possible role in regulation of gene expression. *Bioinformatics 22*, 6 (2006), 676–684.

[47] Bostan, Alin, Flajolet, Philippe, Salvy, Bruno, and Schost, Éric. Fast computation of special resultants. *Journal of Symbolic Computation 41*, 1 (January 2006), 1–29.

[48] Conrad, Eric van Fossen, and Flajolet, Philippe. The Fermat cubic, elliptic functions, continued fractions, and a combinatorial excursion. *Séminaire Lotharingien de Combinatoire 54*, B54g (2006), 1–44.

[49] Flajolet, Philippe, Nebel, Markus, and Prodinger, Helmut. The scientific works of Rainer Kemp (1949–2004). *Theoretical Computer Science 355*, 3 (April 2006), 371–381.

[50] Flajolet, Philippe, Szpankowski, Wojciech, and Vallée, Brigitte. Hidden word statistics. *Journal of the ACM 53*, 1 (January 2006), 147–183.

[51] Fusy, Éric. Counting d-polytopes with (d+3) vertices. *The Electronic Journal of Combinatorics 13*, 1 (2006).

[52] Giusti, Marc, Lecerf, Grégoire, Salvy, Bruno, and Yakoubsohn, Jean-Claude. On location and approximation of clusters of zeroes: Case of embedding dimension one. *Foundations of Computational Mathematics 6*, 3 (July 2006), 1–57.

[53] Régnier, M., and Vandenbogaert, M. Comparison of statistical significance criteria. *Journal of Bioinformatics and Computational Biology 4*, 2 (2006), 537–551.

[54] Bostan, A., Gaudry, P., and Schost, É. Linear recurrences with polynomial coefficients and application to integer factorization and Cartier-Manin operator. *SIAM Journal on Computing* (To appear).

[55] Fusy, Éric, Poullahon, Dominique, and Schaeffer, Gilles. Dissections, orientations and trees, with applications to optimal mesh encoding and to random sampling. *Transactions on algorithms* (To appear).

6.4 Publications in Conferences and Workshops

[56] Chyzak, Frédéric, Mishna, Marni, and Salvy, Bruno. Effective d-finite symmetric functions. In *14th Conference of Formal Power Series and Algebraic Combinatorics* (Melbourne, Australia, July 2002), University of Melbourne, pp. 19.1–19.12. Extended abstract.

[57] Duchon, Philippe, Flajolet, Philippe, Louchard, Guy, and Schaeffer, Gilles. Random sampling from Boltzmann principles. In *Automata, Languages, and Programming* (2002), P. Widmayer, Ed., no. 2380 in Lecture Notes in Computer Science, Springer Verlag, pp. 501–513. Proceedings of the 29th ICALP Conference, Malaga, July 2002.

[58] Flajolet, Philippe. Singular combinatorics. In *Proceedings of the International Congress of Mathematicians* (2002), L. Tatsien, Ed., vol. III, World Scientific, pp. 561–571. Invited lecture, ICM02, Beijing, China, 20–28 August 2002.

[59] Matera, Guillermo, and Sedoglavic, Alexandre. The differential Hilbert function of a differential rational mapping can be computed in polynomial time. In *Proceedings of the 2002 International Symposium on Symbolic and Algebraic Computation* (Lille, France, July 7–10 2002), T. Mora, Ed., Association for Computing Machinery, ACM press, pp. 184–191.

[60] Ollivier, François, and Sedoglavic, Alexandre. Algorithmes efficaces pour tester l'identifiabilité locale. In *Actes de la Conférence Internationale Francophone d'Automatique 2002* (Nantes, France, July 8–10 2002), IEEE, pp. 811–816.

[61] Barkatou, Moulay, Chyzak, Frédéric, and Loday-Richaud, Michèle. Remarques algorithmiques liées au rang d'un opérateur différentiel linéaire. In *From Combinatorics to Dynamical Systems* (2003), F. Fauvet and C. Mitschi, Eds., vol. 3 of *IRMA Lectures in Mathematics and Theoretical Physics*, de Gruyter, pp. 87–129. In French. ISBN 3-11-017875-3. Proceedings of Journées de Calcul Formel, Strasbourg, March 22-23, 2002.

[62] Bostan, A., Lecerf, G., and Schost, É. Tellegen's principle into practice. In *Symbolic and Algebraic Computation* (2003), J. R. Sendra, Ed., ACM Press, pp. 37–44. Proceedings of ISSAC'03, Philadelphia, August 2003.

[63] Brönnimann, Hervé, Cazals, Frédéric, and Durand, Marianne. Randomized jumplists: A jump-and-walk dictionary datastructure. In *Proceedings of the STACS'03 Conference, Berlin, February 2003* (2003), vol. 2607 of *Lecture Notes in Computer Science*, Springer Verlag, pp. 283–294.

[64] Chyzak, Frédéric, Quadrat, Alban, and Robertz, Daniel. Linear control systems over Ore algebras: effective algorithms for the computation of parametrizations. In *Time-Delay Systems* (2003). Electronic proceedings of an IFAC Workshop, INRIA-Roquencourt (France).

[65] Durand, Marianne, and Flajolet, Philippe. Loglog counting of large cardinalities. In *Annual European Symposium on Algorithms (ESA03)* (September 2003), G. Di Battista and U. Zwick, Eds., vol. 2832 of *Lecture Notes in Computer Science*, pp. 605–617.

[66] Meunier, Ludovic, and Salvy, Bruno. ESF: An automatically generated encyclopedia of special functions. In *Symbolic and Algebraic Computation* (2003), J. R. Sendra, Ed., ACM Press, pp. 199–205. Proceedings of ISSAC'03, Philadelphia, August 2003.

[67] Régnier, M., and Denise, A. Statistiques extrêmes sur les mots. In *XXXV-ièmes Journées de Statistique* (2003), vol. 2, pp. 799–802. Proceedings XXXV-ièmes Journées de Statistique, Lyon.

[68] Tahi, F., Engelen, S., and Régnier, M. A fast algorithm for RNA secondary structure prediction including pseudoknots approach. In *BIBE'03* (2003), IEEE Computer Society, pp. 11–17. Proceedings BIBE'03, Washington DC.

[69] Abramov, S. A., and Le, H. Q. Utilizing relationships among linear systems generated by Zeilberger's algorithm. In *Formal Power Series and Algebraic Combinatorics* (2004), pp. 29–38. Proceedings of FPSAC'04, Vancouver, June 2004.

[70] Bardet, Magali, Faugère, Jean-Charles, and Salvy, Bruno. On the complexity of Gröbner basis computation of semi-regular overdetermined algebraic equations. In *International Conference on Polynomial System Solving* (November 2004), pp. 71–74. Proceedings of a conference held in Paris, France in honor of Daniel Lazard.

[71] Bostan, A., Gaudry, P., and Schost, É. Linear recurrences with polynomial coefficients and computation of the Cartier-Manin operator on hyperelliptic curves. In *Fq7, International Conference on Finite Fields and Applications (Toulouse, France, May 5-9, 2003)* (2004), H. S. Gary L. Mullen, Alain Poli, Ed., vol. 2948 of *Lecture Notes in Computer Science*, Springer–Verlag, pp. 40–58.

[72] Bostan, Alin, Lecerf, Grégoire, Salvy, Bruno, Schost, Éric, and Wiebelt, Bernd. Complexity issues in bivariate polynomial factorization. In *Symbolic and Algebraic Computation* (2004), J. Gutierrez, Ed., ACM Press, pp. 42–49. Proceedings of ISSAC'04, Santander, July 2004.

[73] Chyzak, Frédéric, Quadrat, Alban, and Robertz, Daniel. Oremodules, a symbolic package for the study of multidimensional linear systems. In *Sixteenth International Symposium on Mathematical Theory of Networks and Systems* (Leuven, Belgium, July 2004), Katholieke Universiteit Leuven. Proceedings MTNS2004, Katholieke Universiteit Leuven, Belgium, July 5–9, 2004.

[74] Fayolle, Julien. An average-case analysis of basic parameters of the suffix tree. In *Mathematics and Computer Science* (2004), M. Drmota, P. Flajolet, D. Gardy, and B. Gittenberger, Eds., Birkhäuser, pp. 217–227. Proceedings of a colloquium organized by TU Wien, Vienna, Austria, September 2004.

[75] Flajolet, Philippe. Counting by coin tossings. In *Proceedings of ASIAN'04 (Ninth Asian Computing Science Conference)* (2004), M. Maher, Ed., vol. 3321 of *Lecture Notes in Computer Science*, pp. 1–12. Text of Opening Keynote Address.

[76] Geddes, Keith, Le, Ha, and Li, Ziming. Differential rational normal forms and a reduction algorithm for hyperexponential functions. In *Symbolic and Algebraic Computation* (2004), J. Gutierrez, Ed., ACM Press, pp. 183–190. Proceedings of ISSAC'04, Santander, July 2004.

[77] Le, Ha, and Li, Ziming. Differential rational normal forms and representations of hyperexponential functions. In *Rhine Workshop on Computer Algebra* (2004), pp. 3–12. Proceedings of RWCA'04, Nijmegen, March 2004.

[78] Lipmaa, Helger, Wallén, Johan, and Dumas, Philippe. Differential probability of exclusive-or. In *Fast Software Encryption 2004* (2004), B. Roy and W. Meier, Eds., vol. 3017 of *Lecture Notes in Computer Science*, Springer-Verlag, pp. 317–331. Delhi, India, February 5–7, 2004.

[79] Régnier, Mireille. Mathematical tools for regulatory signals extraction. In *Bioinformatics of Genome Regulation and Structure* (2004), N. Kolchanov and R. Hofestaedt, Eds., Kluwer Academic Publisher, pp. 61–70. Preliminary version at BGRS'02.

[80] Bardet, M., Faugère, J.-C., Salvy, B., and Yang, B.-Y. Asymptotic behaviour of the degree of regularity of semi-regular polynomial systems. In *MEGA'05* (2005). Eighth International Symposium on Effective Methods in Algebraic Geometry, Porto Conte, Alghero, Sardinia (Italy), May 27th – June 1st.

[81] Boeva, V., Clément, J., Régnier, M., and Vandenbogaert, M. Assessing the significance of sets of words. In *Combinatorial Pattern Matching 05* (2005), vol. 3537 of *Lecture Notes in Computer Science*, Springer Verlag, pp. 358–370. In Proceedings CPM'05, Jeju Island, Korea.

[82] Bostan, Alin, Cluzeau, Thomas, and Salvy, Bruno. Fast algorithms for polynomial solutions of linear differential equations. In *ISSAC'05* (New York, 2005), M. Kauers, Ed., ACM Press, pp. 45–52. Proceedings of the 2005 International Symposium on Symbolic and Algebraic Computation, July 2005, Beijing, China.

[83] Bostan, A., González-Vega, L., Perdry, H., and Schost, É. From Newton sums to coefficients: complexity issues in characteristic p. In *MEGA'05* (2005). Eighth International Symposium on Effective Methods in Algebraic Geometry, Porto Conte, Alghero, Sardinia (Italy), May 27th – June 1st.

[84] Fayolle, Julien, and Ward, Mark Daniel. Analysis of the average depth in a suffix tree under a markov model. In *2005 International Conference on Analysis of Algorithms* (2005), C. Martínez, Ed., vol. AD, Discrete Mathematics & Theoretical Computer Science, pp. 95–104.

[85] Fusy, Éric. Quadratic exact size and linear approximate size random generation of planar graphs. In *2005 International Conference on Analysis of Algorithms* (2005), C. Martínez, Ed., vol. AD, Discrete Mathematics & Theoretical Computer Science, pp. 125–138.

[86] Fusy, Éric, Poulalhon, Dominique, and Schaeffer, Gilles. Dissections and trees, with applications to optimal mesh encoding and to random sampling. In *SODA* (2005), pp. 690–699. Proceedings of the Sixteenth Annual ACM-SIAM Symposium on Discrete Algorithms, SODA 2005, Vancouver, British Columbia, Canada, January 23-25, 2005.

[87] Giroire, Frédéric. Order statistics and estimating cardinalities of massive data sets. In *2005 International Conference on Analysis of Algorithms* (2005), C. Martínez, Ed., vol. AD, Discrete Mathematics & Theoretical Computer Science, pp. 157–166.

[88] Salvy, Bruno. D-finiteness: Algorithms and applications. In *ISSAC'05* (2005), M. Kauers, Ed., ACM Press, pp. 2–3. Invited talk. Proceedings of the 2005 International Symposium on Symbolic and Algebraic Computation, Beijing, July 2005.

[89] Amini, O., Bermond, J.-C., Giroire, F., Huc, F., and Pérennes, S. Design of minimal fault tolerant networks: Asymptotic bounds. In *Huitièmes Rencontres Francophones sur les Aspects Algorithmiques des Télécommunications (AlgoTel'06)* (Trégastel, France, May 2006), pp. 37–40.

[90] Bodini, Olivier, Fusy, Éric, and Pivoteau, Carine. Random sampling of plane partitions. In *Gascom 2006* (Dijon, France, 2006), R. Pinzani and V. Vajnovszki, Eds., LE2I, pp. 124–135.

[91] Bostan, Alin, Chyzak, Frédéric, Cluzeau, Thomas, and Salvy, Bruno. Low complexity algorithms for linear recurrences. In *ISSAC'06* (2006), J.-G. Dumas, Ed., ACM Press, pp. 31–38.

[92] Flajolet, Philippe. The ubiquitous digital tree. In *STACS 2006* (2006), B. Durand and W. Thomas, Eds., vol. 3884 of *Lecture Notes in Computer Science*, pp. 1–22. Proceedings of 23rd Annual Symposium on Theoretical Aspects of Computer Science, Marseille, February 2006.

[93] Flajolet, Philippe, Dumas, Philippe, and Puyhaubert, Vincent. Some exactly solvable models of urn process theory. In *Fourth Colloquium on Mathematics and Computer Science* (2006), P. Chassaing, Ed., vol. AG, Discrete Mathematics & Theoretical Computer Science, pp. 59–118.

[94] Fusy, Éric. Transversal structures on triangulations, with application to straight-line drawing. In *Graph Drawing 2005* (2006), P. Healy and N. S. Nikolov, Eds., vol. 3843 of *Lecture Notes in Computer Science*, Springer, pp. 177–188. Limerick, Ireland, September 12–14, 2005.

[95] Giroire, F. Directions to use probabilistic algorithms for cardinality for DNA analysis. In *Journées Ouvertes Biologie Informatique Mathématiques (JOBIM 2006)* (July 2006), pp. 3–5.

[96] Bodirsky, Manuel, Fusy, Éric, Kang, Mihyun, and Vigerske, Stefan. An unbiased pointing operator for unlabeled structures, with applications to counting and sampling. In *18th ACM-SIAM Symposium on Discrete Algorithms, New Orleans, January 2007* (To appear).

[97] Bostan, Alin, Chyzak, Frédéric, Ollivier, François, Salvy, Bruno, Schost, Éric, and Sedoglavic, Alexandre. Fast computation of power series solutions of systems of differential equations. In *18th ACM-SIAM Symposium on Discrete Algorithms, New Orleans, January 2007* (To appear).

[98] Flajolet, Philippe. Analytic combinatorics: a guided tour. Invited talk at the *18th ACM-SIAM Symposium on Discrete Algorithms, New Orleans, January 2007* (To appear).

[99] Fusy, Éric. Straight-line drawing of quadrangulations. In *Graph Drawing 2006* (To appear), Springer-Verlag. Karlsruhe, Germany, September 18–20, 2006.

6.5 Internal Reports

[100] Chyzak, Frédéric, and Paule, Peter. Computer algebra. Methodological chapter for the *Digital Library of Mathematical Functions*, D. Lozier, F. Olver, C. Clark, and R. Boisvert, Eds., 2005. 40 pages. Under the process of validation by NIST.

[101] Bodirsky, Manuel, Fusy, Éric, Kang, Mihyun, and Vigerske, Stefan. Enumeration of unlabeled outerplanar graphs. Submitted to the *Journal of Combinatorial Theory, Series A*, February 2006.

[102] Bostan, Alin, Morain, François, Salvy, Bruno, and Schost, Éric. Fast algorithms for computing isogenies between elliptic curves. Tech. Rep. 91441, Inria, 2006. Submitted to *Mathematics of Computation*.

[103] Chyzak, Frédéric, Drmota, Michael, Klausner, Thomas, and Kok, Gerard. The distribution of patterns in random trees. 30 pages. Submitted to *Combinatorics, Probability & Computing*, May 2006.

[104] Flajolet, Philippe, Fusy, Éric, Gourdon, Xavier, Panario, Daniel, and Pouyanne, Nicolas. A hybrid of Darboux's method and singularity analysis in combinatorial asymptotics. Submitted to the *Electronic Journal of Combinatorics*, June 2006.

[105] Gomez, Claude, and Salvy, Bruno. Calcul formel. *Les Techniques de l'Ingénieur Dossier AF1460* (To appear). 34 pages.