

# ALEA 2011 : Programme des journées

	Lundi	Mardi	Mercredi	Jeudi	Vendredi
09:00 - 10:00	Ouverture M Bousquet-Mélou +K. Raschel	X Viennot	X Viennot	D Piau	M Mishna H Cheballah C Banderier
10:00 - 11:00					
11:00 - 12:00	M Soria H Tafat	F Bassino P Duchon	B Salvy A Bostan	J Marckert E Gutter	M Régnier M Josuat-Vergès D Gardy
12:00 - 13:00					
13:00 - 14:00					
14:00 - 15:00					HK Hwang M Madritsch V Féray
15:00 - 16:00	K Raschel+ M Bousquet-Mélou	D Piau		C Mailler F Avram G Louchard	
16:00 - 17:00					
17:00 - 18:00	R Aguech O Roussel S Dasse-hartaut H Ennafti	N Pouyanne X Wang T Monteil		B Morcrette M Sablik Boite à idées	
18:00 - 19:00					
19:00 - 20:00	M Bousquet-Mélou +K Raschel	X Viennot		D Piau	

## Lundi 07 mars

09h00 – Ouverture

### Présentation des journées

Rôle et histoire des journées ALEA, grandes thématiques scientifiques ...

09h15 – Mireille Bousquet-mélou (Labri) + K. Raschel

### Chemins dans le quart de plan I

Soit  $S$  un sous-ensemble fini de  $Z^2$ . On s'intéresse au comptage des chemins du plan à pas dans  $S$ , issus de l'origine et confinés au premier quadrant. En particulier, on voudrait connaître la nature de la série génératrice associée : est-elle rationnelle, algébrique, différentiellement finie ?

On présentera deux types d'approche pour résoudre ce problème, l'une basée sur la manipulation de séries formelles (mbm), l'autre utilisant des méthodes plus analytiques (K. Raschel).

11h00 – Michèle Soria (LIP6/UPMC)

### Expressivité et complexité de la génération aléatoire de structures combinatoires sous modèle de Boltzmann

Le modèle de Boltzmann, proposé en 2004 par Duchon, Flajolet, Louchard et Schaeffer, définit une approche générique pour la génération uniforme de structures engendrées par des grammaires combinatoires (à base d'opérateurs Union, Produit, Sequence, Cycle et Ensemble), dont la complexité est linéaire si l'on autorise la génération en taille approchée. Nous examinerons dans cet exposé différents travaux récents visant à enrichir l'expressivité du modèle de Boltzmann et à optimiser la génération effective.

12h00 – Hanane Tafat (CNRS-LIPN Paris13) + Cyril Banderier, Olivier Bodini, Yann Ponty

### Asymptotique et loi limite des automates et des grammaires.

Dans le cadre d'automates ou de grammaires fortement connexe, différents travaux permettent d'obtenir asymptotique et loi limite. (nous illustrons ceci via l'étude de motif dans un automate dans le modèle de Schelling [prix Nobel d'économie 2005]).

La situation se complique singulièrement dans les cas non fortement connexe, et nous donnons alors une classification de phénomènes universaux qui subsistent en matière d'asymptotique/loi limite, ce qui a des applications à la génération aléatoire via la méthode de Boltzmann.

Nous concluons sur des recherches en cours : que peut-il être dit au-delà du théorème de Dromta-Lalley-Woods ? (Tout un zoo de lois limites apparaît alors).

15h00 – Kilian Raschel (LabRI) + M. Bousquet-Mélou

### Chemins dans le quart de plan II

Soit  $S$  un sous-ensemble fini de  $Z^2$ . On s'intéresse au comptage des chemins du plan à pas dans  $S$ , issus de l'origine et confinés au premier quadrant. En particulier, on voudrait connaître la nature de la série génératrice associée : est-elle rationnelle, algébrique, différentiellement finie ? On présentera deux types d'approche pour résoudre ce problème, l'une basée sur la manipulation de séries formelles (mbm), l'autre utilisant des méthodes plus analytiques (K. Raschel).

**16h45** – Rafik Aguech (Maître de conférence/Université de Monastir) + Cyril Banderier

### **Loi limite d'un processus de Moran**

On dispose d'une machine comportant  $m$  composantes, initialement toutes neuves. A chaque unité de temps, l'une des  $m$  composantes, choisie au hasard, est remplacée par une composante neuve avec probabilité  $p$ . Au bout de  $n$  unités de temps, quel est le comportement de  $L_n$  l'âge de la plus vieille composante ? Quelle aura été la durée,  $H_n$ , de vie maximale d'une composante ? Ce problème est lié au problème de Pat Moran en biologie des populations. Dans ce travail on s'intéresse au cas où  $m = 1$  et  $p \in ]0, 1[$ . Grâce à des outils d'analyse complexe (études des singularités de la série génératrice associée à  $H_n$ ) on obtient des expressions explicites de l'espérance et de la variance de  $H_n$ . Ensuite on obtient la loi limite de  $H_n$  : on montre l'existence d'une suite déterministe explicite  $beta_n$  d'ordre  $\ln(n)/-\ln(1-p)$  telle que le processus  $H_n - beta_n$  converge en loi vers une loi double exponentielle.

**17h15** – Olivier Roussel (Doctorant UPMC) + Michèle Soria ; Olivier Bodini

### **Génération aléatoire de structures ordonnées**

Dans le cadre de la génération aléatoires sous le modèle de Boltzmann de structures combinatoires, nous montrerons comment il est possible de prendre en compte des objets et structures dites ordonnées. Il s'agit de structures pouvant être décrites par un jeu d'équations *différentielles*. Ainsi, toutes les familles d'arbres croissants rentrent dans cette catégorie, ainsi que les permutations alternantes, par exemple.

**17h30** – Sandrine Dasse-hartaut (Doctorante) + Pawel Hitczenko

### **Propriétés des tableaux escaliers aléatoires**

Les tableaux escaliers sont des objets combinatoires assez récents liés au modèle de physique statistique PASEP. On présente une approche probabiliste qui permet d'avoir des informations sur la distribution de certains de ses paramètres. L'idée est de s'appuyer sur un paramètre dont dépendent les autres et de décomposer tout tableau de taille  $(n)$  en une colonne de taille  $(n)$  et un tableau de taille  $(n-1)$ .

**17h45** – Hazar Ennafti (étudiante chercheur) + Sana Louhichi

### **Inégalité de Concentration pour des variables aléatoires négativement associées**

Dans cet exposé, on établit une inégalité de concentration qui permet de majorer

$$P(g(X_1, \dots, X_n) - Eg(X_1, \dots, X_n)),$$

où  $\{X_1, \dots, X_n\}$  est une famille des variables aléatoires négativement associées. On applique enfin, le résultat à un processus d'exclusion symétrique possédant une distribution initiale 'Strong Rayleigh'.

**18h30** – Mireille Bousquet-mélou (LabRI) + K. Raschel

### **Exercices**

## **Mardi 08 mars**

**09h00** – Xavier Viennot (LabRI)

### **Processus d'exclusion par approche cellulaire I**

**10h45** – Frédérique Bassino (LIPN)

**Analyse d’algorithmes et automates**

**11h45** – Philippe Duchon (PR/Université Bordeaux 1)

**Génération Boltzmannienne exacte avec oracles inexacts**

Le modèle de Boltzmann est une technique séduisante pour la génération aléatoires d’objets combinatoires, mais il a potentiellement le défaut de reposer sur un ”oracle” chargé d’évaluer diverses séries génératrices en un même point.

Une question naturelle est celle de l’influence, sur la qualité du résultat, de la précision numérique de l’oracle. On apportera à cette question des réponses plus ou moins satisfaisantes et intrigantes.

**15h00** – Didier Piau (UJF)

**Sur quelques processus stochastiques en biologie moléculaire I**

On évoquera quelques modèles récents d’évolution de séquences d’ADN qui rendent compte de la dynamique singulière (mais tout à fait bien documentée par les biologistes) du dinucléotide CpG et d’autres observations similaires, de la résolution miraculeuse d’une certaine classe de ces modèles, et de l’extension au forceps de ce miracle à des modèles suffisamment proches des précédents pour qu’un processus de Galton-Watson sous-jacent résumant toute l’affaire reste sous-critique. On procèdera à des rappels de biologie moléculaire. Les notions mathématiques mobilisées concerneront les processus de Markov en temps continu, des variantes de couplages à partir du passé et quelques rudiments de systèmes de particules et de processus de branchement.

**16h45** – Nicolas Pouyanne (Labo math Versailles) + Brigitte Chauvin (Versailles), Peggy Cénac (Dijon), Frédéric Paccaut (Amiens)

**Le peigne riemannien mélange peu**

Les chaînes de Markov à mémoire de longueur variable (en anglais VLMC) constituent une classe de sources probabilistes. Elles fournissent notamment des exemples ”concrets” de sources intermittentes. Il sera question des propriétés de mélange, de l’analyse du trie et de celle du trie des suffixes dans un cas stationnaire simple où l’arbre des contextes est un peigne infini.

**17h15** – Xiaomin Wang (LIP6) + Michèle Soria, Matthieu Latapy

**Deciding the type of a graph from a BFS tree**

La distribution du degré de la topologie d’internet est considérée comme une de ses principales propriétés. Néanmoins, on ne la connaît qu’à travers une procédure de mesure qui donne une estimation biaisée. En première approximation, cette mesure peut être modélisée par un arbre de parcours en largeur (ou BFS, pour Breadth First Search). On explore notre capacité à déduire le type de la distribution du degré (Poisson ou power-law) à partir de connaissances limitées. On construit des procédures qui estiment la distribution du degré d’un graphe à partir de son BFS et on montre expérimentalement (sur des modèles et des données réelles) que cette approche réussit à distinguer le type de distribution entre Poisson et power-law.

**17h30** – Thierry Monteil (LIRMM)

**SAGE : Maths libres !**

Présentation du logiciel libre SAGE

**18h30** – Xavier Viennot (LabRI)

**Exercices**

## **Mercredi 09 mars**

**09h00** – Xavier Viennot (LabRI)

**Processus d'exclusion par approche cellulaire II**

**10h45** – Bruno Salvy (INRIA Rocquencourt) + Carine Pivoteau et Michèle Soria

**Itération de Newton combinatoire et applications à l'aléa**

Nous utilisons et étendons des idées qui proviennent de la théorie des espèces de structures. Nous en tirons d'une part des algorithmes de complexité quasi-optimale pour calculer des suites d'énumération (ce qui donne un prétraitement quasi-optimal à la génération aléatoire par la méthode récursive), d'autre part des oracles numériques pour les générateurs aléatoires dans le modèle de Boltzmann. Il s'agit d'un travail en commun avec Carine Pivoteau et Michèle Soria.

**11h45** – Alin Bostan (CR INRIA Paris-Rocquencourt) + Manuel Kauers

**Algébricité de la série génératrice complète des chemins de Gessel**

Nous montrerons que certaines questions de combinatoire énumérative peuvent être traitées de façon systématique en utilisant une approche de type "mathématiques expérimentales" guidée par des algorithmes modernes du calcul formel. Nous décrivons la découverte et la preuve assistées par ordinateur de l'algébricité de la série génératrice des chemins confinés au quart de plan et utilisant des pas de longueur un vers l'est, l'ouest, le sud-ouest, ou le nord-est.

## **Jeudi 10 mars**

**09h00** – Didier Piau (UJF)

**Sur quelques processus stochastiques en biologie moléculaire II**

On évoquera quelques modèles récents d'évolution de séquences d'ADN qui rendent compte de la dynamique singulière (mais tout à fait bien documentée par les biologistes) du dinucléotide CpG et d'autres observations similaires, de la résolution miraculeuse d'une certaine classe de ces modèles, et de l'extension au forceps de ce miracle à des modèles suffisamment proches des précédents pour qu'un processus de Galton-Watson sous-jacent résumant toute l'affaire reste sous-critique. On procèdera à des rappels de biologie moléculaire. Les notions mathématiques mobilisées concerneront les processus de Markov en temps continu, des variantes de couplages à partir du passé et quelques rudiments de systèmes de particules et de processus de branchement.

**10h45** – Jean-françois Marckert (LaBRI) + Nicolas Bonichon

### **Navigation dans le plan**

On se donne  $S$ , un ensemble fini de points inclus dans un certain domaine du plan  $D$ . Ces points sont vus comme un ensemble d'étapes possibles pour un voyageur. Le voyageur partant d'un point  $s$  et allant à un point  $t$ , doit s'arrêter "au premier point" de l'ensemble  $t$  union  $S$ , situé "dans la direction de  $t$ ".

Nous examinons alors la distance parcourue par le voyageur, et la comparons à la distance Euclidienne entre  $s$  et  $t$ , lorsque l'ensemble  $S$  est tiré aléatoirement (et à une taille tendant vers l'infini), et pour diverses notions de "premiers points" et de "dans la direction de  $t$ ".

**11h45** – Emmanuel Guitter (IPhT-CEA Saclay) + Jérémie Bouttier

### **Cartes planaires et fractions continues**

Je montrerai qu'il existe un lien surprenant entre les deux problèmes combinatoires suivants : (1) compter des cartes planaires avec un bord de longueur fixée et (2) compter des cartes planaires avec deux points marqués à distance prescrite. La solution à ces deux problèmes est codée dans le développement d'une même quantité de deux manières : en série pour (1) et en fractions continues pour (2). Je montrerai alors comment utiliser une solution connue de (1) pour résoudre (2) et dériver ainsi certaines formules exactes pour la fonction à deux points des cartes planaires.

**15h00** – Cécile Mailler (Université de Versailles) + Brigitte Chauvin, Danièle Gardy

### **Le modèle des arbres croissants pour la représentation de fonctions booléennes.**

On définit, via leur représentation arborescente, une loi de probabilité sur l'ensemble des fonctions Booléennes à  $k$  variables. Cette nouvelle loi, que l'on nomme *loi des arbres croissants*, est inspirée du modèle de croissance des arbres binaires de recherche. On l'étudie dans différents systèmes logiques et on la compare à des distributions déjà étudiées : celle des arbres de Catalan, celle des arbres de Galton-Watson, et celle des arbres équilibrés.

**15h30** – Florin Avram (Universite de Pau)

### **Sur les probabilités stationnaires des files d'attente avec ressaies**

Les fonctions génératrices des probabilités stationnaires pour les files d'attente multiserveur avec ressaies sont holonomiques. Pour un ou deux serveurs, il s'agit des fonctions hypergéométriques bien connues, mais pour plusieurs serveurs elles appartiennent à une classe encore peu étudiée des fonctions hypergéométriques de type Okubo, ayant une singularité irrégulière en 0. Pour plusieurs serveurs, l'identification de l'unique fonction analytique autour de cette singularité semble être un problème ouvert.

**15h45** – Guy Louchard (Universite Libre de Bruxelles) + C. Martinez and H. Prodinger

**The Asymmetric Leader Election Algorithm with swedish stopping**

At the last ANALCO 2011, a talk was presented by C. Martinez (in collaboration with G. Louchard and H. Prodinger) about the mean analysis of the Swedish leader election protocol. The goal is to select one among  $n > 0$  players, by proceeding through a number of rounds. If there is only one player remaining, the protocol stops and the player is declared the leader. Otherwise, all remaining players flip a biased coin, with probability  $q$  the player survives to the next round, with probability  $p = 1-q$  the player loses (is killed) and plays no further.. unless all players lose in a given round (null round) , so all them play again. In the classical leader election protocol, any number of null rounds may take place, and with probability 1 some player will ultimately be elected. In the Swedish leader election protocol there is a maximum number of consecutive null rounds, and if the threshold is attained the protocol fails without declaring a leader. The analysis was based on an analytic treatment of some recurrences. In the present talk, we present a more probabilistic analysis, based on an urn model. We obtain asymptotic distributions and the first two moments.

**16h45** – Basile Morcrette (INRIA Rocquencourt - LIP6) + Philippe Flajolet

**Urnes de Polya : analyse d’une classe algébrique**

Je considère une classe d’urnes de Polya à 2 couleurs, équilibrées et additives dont la fonction génératrice est algébrique. Par des méthodes de combinatoire analytique, et notamment une méthode de col faisant intervenir des cols ”coalescents”, on obtient des résultats sur la distribution asymptotique de l’urne, avec loi normale ainsi que loi locale limite, et grandes déviations, propriétés jusqu’alors inaccessibles par les méthodes classiques probabilistes. Ce travail poursuit l’étude des urnes de Polya du point de vue de la combinatoire analytique, travail amorcé par Flajolet-Dumas-Puyhaubert en 2006.

**17h15** – Mathieu Sablik (LATP)

**Quel comportement typique peut avoir un automate cellulaire ?**

On peut se demander quelles mesures peuvent être obtenues comme valeurs d’adhérence de la suite des itérés d’une mesure par un automate cellulaire. Un résultat surprenant est que pour chaque mesures calculables, et seulement celles-ci, il existe un automate cellulaire tel que la suite des itérés par une mesure de Bernoulli, et plus généralement une mesure ergodique de support total, converge vers la mesure initialement choisi. On peut même construire un automate cellulaire ”universel” dans le sens que suivant la mesure initiale choisie, les valeurs d’adhérences de la suite des itérés par l’automate de la mesure initiale pourra être n’importe quel ensemble de mesures atteignable par un automate cellulaire.

**17h30** – Boite à idées

**Boite à idées**

Problèmes ouverts et/ou amusants ...

**18h30** – Didier Piau (UJF)

**Exercices**

## Vendredi 11 mars

**09h00** – Marni Mishna (SFU) + Sam Johnson (SFU) et Steve Melczer (SFU)

**Secrets from my step set : avancées sur l'énumération de chemins dans le 1/4 de plan.**

Nous présentons des travaux en cours avec Sam Johnson (SFU) et Steve Melczer (SFU) sur les modèles de marches sur réseaux. Les modèles qui nous intéressent sont les chemins à petits pas dans le quart de plan. Nos résultats comprennent une amélioration des méthodes pour calculer leurs séries génératrices, et une explication combinatoire des formules asymptotiques obtenues automatiquement par Bostin et Kauers. Une élément clé de notre étude réside dans l'utilisation de la dérive des pas permis.

**09h30** – Hayat Cheballah (LIPN - Paris 13) + Philippe Biane

**Gog, Magog et Schutzenberger**

trouver une bijection explicite entre matrices à signes alternants et partitions planes totalement symétriques autoccomplémentaires est une question ouverte importante en combinatoire. On propose une approche de ce problème qui passe par les triangles Gog et Magog de Zeilberger et fait intervenir de façon cruciale l'involution de Schutzenberger.

**10h00** – Cyril Banderier (CNRS/Paris 13) + Pawel Hitczenko

**m enfants jouent à faire des pavages par des dominos...**

$m$  enfants jouent à faire des pavages d'un rectangle  $2 \times n$  par des dominos. Quel est l'ordre de grandeur (pour  $n$  grand) de la probabilité qu'ils aient tous utilisé le même nombre de dominos horizontaux?  $1/(n * m)$ ,  $1/n^m$ ,  $1/exp(nm)$ , ... ?

Quelle est la probabilité que  $m$  cartes de taille  $n$  aient toutes une plus grande composante connexe de même taille ?

Prenez deux jeux de 52 cartes (et on prendra  $52 := n$ ), quelle est la probabilité que la longueur de la plus longue sous-suite croissante de l'un soit la même que la longueur de la plus longue sous-suite croissante de l'autre ?

**10h45** – Mireille Régnier (DR Inria-Saclay) + Saad SHEIKH

**Combinatorics of Overlapping Words**

Overlap words, also named clumps turn out to have interesting combinatorial properties that we explore in this paper. We first introduce a specific subset, the so-called canonic clumps that are shown to be the seed for word enumeration. We demonstrate a combinatorial characterization underlining the importance of these clumps. Then, we define an automaton AutoClump that allows for the generation of all clumps. This Aho-Corasick like automaton relies on a data structure called the overlap graph. Finally, we weight this automaton with generating functions and show how this allows for an efficient computation of probabilities, expectations and distributions for sets of words. Such results are important for word counting, notably to estimate the waiting time of a given set of words.

**11h15** – Matthieu Josuat-vergès (Université de Vienne)

**Des serpents dans les groupes de Coxeter**

Les permutations alternantes sont celles où les descentes sont exactement les entiers impairs. En fait, si on définit une "classe de descente" comme étant les permutations avec un ensemble de descente donnée, les permutations alternantes forme la plus grosse classe de descente. En se basant sur cette idée, les permutations alternantes ont été généralisées dans d'autres groupes, notamment les groupes de permutations signées. Les objets qui apparaissent ont été appelés serpents par Vladimir Arnol'd. Le but de ce travail est de donner une définition combinatoire explicite des serpents dans les groupe de Coxeter (qui généralisent les permutations signées).

**11h45** – Danièle Gardy (PRISM, UVSQ) + Bernhard Gittenberger, Olivier Bodini

**Lambda-termes de hauteur unaire bornée**

Nous nous intéressons à l'énumération asymptotique de lambda-termes de taille donnée, lorsque la hauteur unaire (nombre d'abstractions imbriquées) est bornée. Notre approche relève de la combinatoire analytique ; la fonction génératrice s'exprime en termes de racines carrées imbriquées, et la singularité dominante se révèle avoir un comportement surprenant. L'analyse asymptotique du nombre de tels lambda-termes met en évidence le rôle de la borne sur la hauteur unaire. Nous nous intéressons également à la génération aléatoire de lambda-termes, et mettons en évidence les difficultés rencontrées par la génération de Boltzmann sur ces structures.

**14h00** – HK Hwang

**Asymptotic estimates for k-dominant skylines of random samples**

Skylines emerged as a useful notion in database queries for selecting representative groups in multivariate data samples for further decision making, multi-objective optimization or data processing, and the k-dominant skylines were naturally introduced to resolve the abundance of skylines when the dimensionality grows or when the coordinates are negatively correlated. We show that the number of k-dominant skylines is asymptotically zero for large samples when  $k \ll d$  under two reasonable (continuous) probability assumptions of the input points,  $d$  being the dimensionality. This suggests that practical use of such skylines under similar data conditions has to be more cautious. On the other hand, we show an asymptotic linearity property for the k-dominant skylines under a categorical model, which explains to some extent the difference between theoretical and empirical estimates.

**14h30** – Manfred Madritsch (Université Technique de Graz) + Stephan Wagner

**Partitions avec des entiers à chiffres manquants**

Hwang a montré un théorème central limite pour les  $\Lambda$ -partitions où  $\Lambda$  est une suite croissante tendant vers l'infini qui satisfait certaines conditions. Une de ces conditions est que la série de Dirichlet associée possède un unique pôle d'ordre 1 à l'abscisse de convergence. Dans cet exposé nous montrons que cette condition peut être relâchée et donnons des exemples provenant de l'analyse des entiers ayant des chiffres manquants dans leur représentation en base  $b$ .

14h45 – Valentin Féray (CR LaBRI, Bordeaux) + Amarpreet Rattan, University of London

### **Produits de long cycles dans le groupe symétrique**

Les produits de cycles complets (de longueur  $n$ ) par des cycles un peu plus courts (de longueur  $n-a$ ) ont de jolies propriétés énumératives. Une manière (relativement simple) de les prouver est de partir du cas  $a=1$  (qui correspond à une permutation impaire aléatoire) et de modifier légèrement notre objet pour obtenir un objet de taille  $n+1$  correspondant à  $a=2$ , et ainsi de suite... Je ne sais malheureusement pas comment adapter cette construction aux produits d'un  $n$ -cycle par une permutation de type  $n-2,2$  (par exemple) bien que les propriétés énumératives subsistent.