ALEA 2002

CIRM, Marseille, 18 au 22 février 2002

Résumés des exposés

Jean-Paul Allouche, LRI Orsay. *Complexité palindromique.* (Avec M. Baake, J. Cassaigne et D. Damanik)

Une des manières traditionnelles de mesurer la “complexité” d'une suite infinie sur un alphabet fini est de compter les facteurs (ou blocs) finis distincts de longueur donnée qui apparaissent dans cette suite. Nous introduisons la “complexité palindromique” qui compte, elle, les palindromes distincts de longueur donnée qui apparaissent dans la suite. Après avoir survolé quelques propriétés et étudié quelques familles d'exemples, nous indiquerons des applications en physique (spectre d'opérateurs de Schrödinger discrets unidimensionnels).

Anne-Elisabeth Baert, LaRIA Amiens. Modeles aleatoires de diffusion dans les reseaux mobiles : diagrammes de Voronoi et Triangulation de Delaunay. (Avec D.Seme et L.Thimonier).

Pour couvrir la plus grande aire possible avec un minimum de stations (sans perdre d'espaces, et minimiser les problemes d'interferences entre les cellules), on a recours a un pavage du plan. La modelisation de la diffusion dans les reseaux mobiles utilisait jusqu'ici une maille hexagonale reguliere, ne rendant pas completement compte d'aspects irreguliers de la realite (repartition geographique, densite demographique). Les differentes proprietes de ces reseaux sont examines ici a travers l'algorithme de parcours Breadth First Search (excentricite, diametre et rapport avec l'arbre couvrant associe). Ces notions sont etudies d'abord avec la maille hexagonale reguliere, puis etendus au diagramme de Voronoi et a la triangulation de Delaunay, representations plus realistes de ces reseaux mobiles, a l'aide de proprietes combinatoires et probabilistes des graphes planaires.

Viviane Baladi, IHES. Spectre d'operateurs de transfert et fonctions zeta en presence de points periodiques neutres.

La théorie spectrale des opérateurs de transfert fonctionne bien (trou spectral) dans un cadre hyperbolique. Il est cependant possible de traiter le cas parabolique (point fixe neutre: cf fraction continue de type Farey par opposition à la fraction continue de Gauss), du moins si la dynamique est assez régulière. Après avoir rappelé un résultat de Rugh en dimension un, nous discuterons comment il peut s'étendre á la dimension deux.”

Cyril Banderier, INRIA et Max-Planck-Institut. Etude d'un modele de marches avec multiplicite.

Une famille de marches sur N avec des sauts dont la multiplicite depend de la position est etudiee et resolue (ce modele comprend, par bijection, les animaux diriges). A partir de la serie generatrice (algebrique), on peut etablir l'asymptotique de divers parametres. La resolution des equations fonctionnelles dans le cas general demeure ouverte.

Jeremy Barbay, LRI Orsay. Théorie de l'évolution et les modèles de Bak Sneppen.

Les paleontologues ont observe de brusques variation dans la variete des especes, caracterisees par des extinction massives d'especes, dont la plus connue est celle des dinosaures. Ces variations sont usuellement expliquees par des catastrophes cosmiques (cometes, etc). Bak et Sneppen ont propose une theorie alternative basee sur la dynamique du systeme d'evolution, justifiee par des experimentations. Nous essayons de prouver le comportement de ce modele et de modifier le modele de maniere a le rendre plus plausible. Je pense exposes la problematique du point de vue biologique, physique et modele aleatoire, mes resultats (en collaboration avec Claire Kenyon) exposes a SODA2001, et des resultats plus recents prochainement publies, d'une equipe d'Amsterdam.

Nicolas Bonichon, LaBRI, Bordeaux. Génération aléatoire de p-uplets de chemins de Dyck qui ne se coupent pas

Un p-uplet de chemins de Dyck (resp. Grand Dyck) qui ne se coupent pas est formé de p chemins composés de pas Nord-Est et Sud-Est partant du point (0,0) et arrivant au point (2n, 0). Le i-ème chemin ne doit pas passer sous le (i-1)-ème chemin. Ces chemins ne peuvent pas (resp. peuvent) couper l'axe des abscisses. Ces objets sont en bijection avec de nombreux autres objets combinatoires : les polyominos parallélogrammes, les permutations de baxter, les réaliseurs, ... C'est pourquoi nous nous intéressons à leur génération aléatoire. Je présenterai un algorithme de génération aléatoire uniforme de p chemins de Dyck (ou Grand Dyck) qui ne se coupent pas. Cet algorithme est basé sur celui de Arnold et Sleep qui construit un chemin de Dyck pas à pas.

Stephane Boucheron, LRI Orsay. (Avec Gabor Lugosi et Pascal Massart) Inégalités de concentration par la méthode entropique.

Les inégalités de concentration constituent une extension des classiques inégalités exponentielles (Hoeffding, Bennett, Bernstein). Elles permettent de contrôler les fluctuations de variables aléatoires "régulières" dans les espaces produits, c'est-à-dire les quantités qui dépendent de beaucoup de variables aléatoires indépendantes sans trop dépendre de chacune d'elles. La concentration de la mesure est une question qui touche les probabilités, l'analyse, la géométrie, la théorie de l'information. Ses aspects les plus fondamentaux se sont vigoureusement développés durant les 25 dernières années sous l'impulsion de M. Ledoux et M. Talagrand. Les inégalités de concentration ont trouvé des applications en informatique théorique (analyse en moyenne des problèmes d'optimisation, complexité de circuit), en statistique (processus empiriques), en probabilités sur les espaces de Banach ... Cet exposé présentera la méthode entropique popularisée par M. Ledoux. Cette méthode consiste à déduire des inégalités de concentration à partir d'inégalités de Sobolev logarithmiques. Elle permet de retrouver (à peu de frais) certaines des inégalités éléborées par M. Talagrand dans les années 1990. La simplicité de la méthode permet parfois d'obtenir des avancées significatives, notamment en théorie statistique de l'apprentissage et en probabilité combinatoire.

Mireille Bousquet-Melou, LaBRI, Bordeaux. Enumération des chemins confinés dans un quadrant

Tous les problèmes d'énumération de chemins confinés dans un DEMI-PLAN peuvent être décrits par une équation fonctionnelle linéaire, très simple à établir, qui fait intervenir des dérivées discrètes par rapport à UNE variable. Ces modèles sont en fait, par essence, unidimensionnels, et équivalents à l'énumération des chemins sur une demi-droite. La "méthode du noyau", qui remonte aux années 70, permet de résoudre les équations associées. C. Banderier et Ph. Flajolet viennent d'écrire sur le sujet un bel article exhaustif. De façon similaire, tous les problémes d'énumération de chemins confinés dans un QUADRANT peuvent être décrits par une équation fonctionnelle linéaire qui fait intervenir des dérivées discrètes par rapport à DEUX variables. Il n'existe pas de technique permettant de résoudre systématiquement ces équations. Je propose un exposé explorant (partiellement) ce problème. Je présenterai notamment une nouvelle solution d'un modèle où, pour des raisons mystérieuses, la série génératrice est algébrique, un critère garantissant l'holonomie de cette même série, et enfin, un cas très simple où la série n'est pas holonome.

Sylvie Corteel, PRISM, Versailles. Partitions aleatoires dont les riemes differences sont positives ou nulles. (Avec Rod Canfield et Pawel Hiczenko).

Soit $P_r(n)$ l'ensemble des partitions de n dont les $r^{ièmes}$ différences sont positives ou nulles. Soit λ une partition tirée au hasard dans $P_r(n)$ et $d(\lambda)$ une $r^{ième}$ différence positive tirée au hasard. Le but de ce travail est de montrer que la probabilité que $d(\lambda) \geq m$ est $m^{-1/r}$ quand $n \to \infty$. Ce travail est une généralisation d'un résultat sur les partitions ordinaires de Corteel, Savage, Pittel et Witlf et a été motivé par une identité découverte par le logiciel Omega d'Andrews, Paule et Riese. Pour prouver ce résultat, nous utilisons les combinatoires bijective, asymptotique/analytique et probabiliste.

Philippe Duchon, LaBRI, Bordeaux. Generation aleatoire "Boltzmannienne". (Avec Ph. Flajolet, G. Louchard et G. Schaeffer)

Nous proposons d'engendrer des structures combinatoires selon une distribution "Boltzmannienne" - répartie sur l'ensemble de la classe au lieu d'etre concentrée sur les structures d'une taille donnée. Pour des familles de structures construites au moyen d'opérateurs "classiques" (unions, produits, séquences; plus, dans un cadre étiqueté, cycles et multi-ensembles), nous montrons que des générateurs boltzmanniens peuvent être combinés de manière simple, et peuvent généralement être paramétrés pour fournir de manière efficace des structures dont la taille est proche d'une taille "cible" (typiquement, en temps O(n) pour des tailles comprises entre n et C.n).

Philippe Flajolet, INRIA Rocquencourt. Introduction "light" à l'alea discret et a l'analyse d'algorithmes.

Danièle Gardy, PRISM, Versailles. Sur les arbres ET/OU. (Avec B.Chauvin et B.Gittenberger)

Les arbres et/ou sont des arbres binaires representant des expressions booleennes. On definit la complexite d'une formule f comme la taille minimale d'un arbre et/ou representant f. Il existe dans la litterature des resultats reliant la complexite d'une formule booleenne f a sa probabilite d'apparition, pour une definition "naturelle" de la loi de probabilite sur l'ensemble des formules booleennes. Nous etudions ici un tel resultat, du a Lefman et Slavicky (RSA 97), et montrons comment une utilisation fine de series generatrices permet d'ameliorer sensiblement la majoration de la probabilite d'une formule en fonction de sa complexite.

Florent Gillet, IECN Nancy 1. Etude de la stabilite de quicksort face aux erreurs.

On desire connaitre le comportement de l'algorithme de tri Quicsort lorsque des erreurs surviennent pendant son deroulement. Par erreur, on entend une mauvaise comparaison entre deux nombres. L'algorithme Quicksort avec erreurs ne trie donc plus correctement une liste. On se propose d'etudier le "desordre" I_n engendre par ces erreurs sur la liste finale. Apres avoir donne une equation fonctionnelle verifiee par I_n, on montre, qu'une fois bien renormalise, I_n converge vers une variable aleatoire I caracterisee par des equations fonctionnelles (differentes suivant la probabilite de se tromper lors d'une comparaison). Les methodes utilisees dans cette expose sont proches de celles de Rosler dans l'etude du cout de Quicksort.

Christian Lavault, LIPN, Paris 13. Algorithme et analyse d'un protocole d'élection efficace en energie. (Avec V.Ravelomanana).

Etant donné n stations mobiles (n non connu), nous proposons : 1) un protocole d'élection permettant aux stations de rester inactives un certain temps et d'économiser ainsi l'énergie du système ("Energy Efficient") ; 2) l'analyse en moyenne (complexité, estimation de n, etc.) de cet algorithme.

Guy Louchard, ULB, Bruxelles. Additive Decompositions, random allocations and threshold phenomena. (Avec O. Dubois et J. Mandler).

An *additive decomposition* of a set A of integers is an expression of A as the arithmetic sum of two of its proper subsets. If the smaller of these has k elements, we have a *k-decomposition.* If A is obtained by randomly removing n^α integers from $\{0,...,n-1\}$, decomposability translates into a balls-and-urns problem which we start to investigate by first showing that the number of k-decompositions exhibits a threshold phenomenon as α crosses a k-dependent critical value. We then study in detail the distribution of the number of 2-decompositions.

Régine Marchand, IECN Nancy I. Deforestation (un probleme de Meir et Moon). (Avec Ph. Chassaing).

On s'intéresse à un vieux problème de Meir et Moon, qui nous a été signalé par J.F. Marckert: combien de coupes au hasard faut-il pour détruire complètement un arbre aléatoire de taille n? Si l'arbre était linéaire, $\log n$ coupes seraient suffisantes, mais Meir et Moon ont obtenu un nombre moyen de coupes $\mathcal{O}(\sqrt{n})$. A l'aide d'une bijection entre arbres aléatoires et parking, on obtient, après normalisation par $\sqrt{n}$, que la loi limite du nombre de coupes est celle de

$$X = \int_0^{+\infty} \frac{da}{1+T_a},$$

où T_a désigne le subordinateur stable $1/2$. On peut alors montrer, par un calcul direct, que X suit la loi de Rayleigh

$$\Pr(X \in dx) = x\ e^{-x^2/2}\ dx.$$

Jean-Francois Marckert, LAMA, Versailles . Marches aléatoires branchantes de taille n.

On superpose sur chaque branche d'un arbre simple, une marche aléatoire. Ainsi deux marches quelconques sont dépendantes l'une de l'autre, ayant des ancêtres communs. On étudie alors des fonctionnelles de l'ensemble de ces marches. Les méthodes utilisent des propriétés des records d'une marche aléatoire associée ainsi qu'une étude des types des ancêtres de chaque noeud dans les arbres simples. Par exemple, des résultats de concentration uniforme sur le nombre de k-ancêtres d'un noeud u (ancêtres de u ayant k enfants), k, j-ancêtres (ancêtres de u ayant k enfants, dont u est un descendant du j ème) sont établis. Dans une deuxième partie ("si le temps le permet") on évoquera le fait que correctement renormalisée, la marche aléatoire branchante converge vers le serpent Brownien.

Michel Nguyen-The, LIX, Ecole Polytechnique. Distribution de la taille d'arbres simplifiés et réduits.

On représente des expressions algébriques ou logiques par des arbres m-aires étiquetés aux feuilles. Ces expressions peuvent se simplifier (loi idempotente) ou se réduire (loi nilpotente), ce qui se traduit sur les arbres de la manière suivante : l'opération de simplification se définit récursivement de telle sorte qu'un arbre formé d'une seule feuille se simplifie en lui-même, et qu'un arbre de sous-arbres identiques se simplifie en ce sous-arbre identique ; l'opération de réduction se définit de manière quasi similaire, en ajoutant une étiquette particulière e, de telle sorte qu'un arbre formé d'une seule feuille se réduise en lui-même, et qu'un arbre de sous-arbres identiques se simplifie en e. Dans sa thèse en 1987, María Inés Fernández Camacho a étudié la moyenne et la variance de la taille de l'arbre simplifié et de l'arbre réduit d'un arbre m-aire aléatoire tiré de manière uniforme au sein de l'ensemble des arbres m-aires d'une taille donnée. On montrera que la technologie actuelle en analyse de singularité, à base du théorème des fonctions implicites et du théorème des quasi-puissances, permet d'établir que la distribution limite de la taille des arbres simplifiés et réduits suit une loi gaussienne.

Pierre Nicodeme, Statistique et Genomes, Evry. Statistique de motifs rapide et comparaison de protéomes.

La plupart des motifs de protéines ont une probabilité d'apparition faible et une probabilité d'auto-corrélation encore plus faible. Nous considérons la sous-classe de motifs sans auto-corrélation qui est une bonne approximation de la plupart des motifs. Cette classe se prête bien à l'analyse combinatoire. Dans le même esprit que ce qui a été réalisé par Mireille Régnier et Wojciech Szpankowski pour un mot à occurrences rares, nous montrons une loi asymptotique Poisson pour les motifs à apparitions rares. Dans le cas de motifs à apparitions fréquentes, nous utilisons l'approche générale proposée par Bruno Salvy, Philippe Flajolet et par l'orateur. Nous appliquons ces résultats à des comparaisons à grande échelle de protéomes (un protéome est l'ensemble des protéines fabriquées par une espèce).

Mireille Régnier, INRIA Rocquencourt. Comptage de mots: grandes deviations. (Avec A.Denise et M.Vandenbogaert)

One studies the distribution of words in random texts that are generated by a Bernoulli or a Markov model. First, a single pattern H_1 is given, and one studies the tail distribution of the number of occurrences of H_1. One establishes large deviation results. Large deviation is also used to compute the distribution of a second word H_2 conditioned by overrepresentation of H_1. The formulae also hold when H_2 is a set of patterns. Some related algorithmic and complexity issues are discussed. The approach is illustrated with typical examples of applications on biological data.

Gilles Schaeffer, LORIA, Nancy. Cartes aléatoires et serpent brownien. (Avec Ph. Chassaing).

Dans ce travail, un lien suprenant est établi entre les quadrangulations aléatoires et une variante de serpent brownien considérée par Aldous (ISE). Ce lien permet de montrer que le rayon r_n d'une quadrangulation à n faces converge, après normalisation, vers la largeur $r = R - L$ du support du processus ISE de dimension un. Plus précisément,

$$n^{-1/4} r_n \xrightarrow{law} (8/9)^{1/4}\, r.$$

Les outils combinatoires sont d'une part un codage à l'aide d'arbres bien étiquetés à la mode Cori-Vauquelin, et la conjugaison d'arbres, chère à l'orateur, qui permettra de lier les arbres bien étiquetés aux arbres plongés considéré par Aldous. En probabilité, nous irons chercher un nouveau résultat, intéressant en soi, à savoir la convergence faible du codage d'un arbre plongé aléatoire par deux marches $(e^{(n)}, \hat{W}^{(n)})$ vers la description de ISE en termes du serpent brownien $(e, \hat{W})$.