Résumés des exposés des journées Aléa 2005

du 7 au 11 mars 2005, CIRM, Luminy, France

Organisateurs : C. Lavault, JM Le Bars et V. Ravelomanana

Analyse de l'algorithme glouton de couplage. Tsiriniaina Andriamampianina

Dans cet exposé nous nous intéressons à l'analyse de la performance en moyenne de l'algorithme glouton pour le problème de couplage maximum. Le problème de couplage maximum est un problème difficile. L'utilisation de l'algorithme glouton donne rapidement un couplage maximal, mais généralement non maximum. Nous généralisons le résultat de Martin Dyer, de Alan Frieze et de Borris Pittel en calculant exactement les fonctions génératrices de l'espérance et de la variance de la variable aléatoire : nombre d'arêtes choisi par l'algorithme quand l'entrée est choisie aléatoirement parmi les graphes connexes ayant L arêtes de plus que de sommets.

Énumération des automates déterministes accessibles et complets sur un alphabet à 2 lettres
Frédérique Bassino et Cyril Nicaud

On s'intéresse aux automates déterministes accessibles (tout état peut êre atteint à partir de l'état initial) et complets (toute lettre peut être lue à partir de chaque état de l'automate). On montre que le nombre $|\mathcal{A}_n|$ d'automates déterministes accessibles et complets à n états sur un alphabet à 2 lettres est $\Theta(n\,2^n\{{2n \atop n}\})$, on donne également des bornes supérieure et inférieure explicites pour $|\mathcal{A}_n|$. La preuve repose sur des bijections qui permettent aussi d'obtenir un algorithme de génération aléatoire des éléments de l'ensemble $\mathcal{A}_n$.

Comptage bijectif des cartes planaires boisées.
Olivier Bernardi

On donne une démonstration bijective d'un résultat de comptage concernant les cartes planaires. On appelle *carte boisée* une carte planaire enracinée dont un arbre couvrant est distingué. Il est connu (Mullin 1966) que le nombre de cartes planaires boisées à n arêtes est $\mathcal{C}_n\mathcal{C}_{n+1}$ où $\mathcal{C}_n = \frac{1}{n+1}\binom{2n}{n}$ est le $n^{\text{ème}}$ nombre de Catalan. Ce résultat de comptage suggère une décomposition des cartes boisées (à n arêtes) en couples d'arbres (à n et $n+1$ arêtes respectivement). Cependant, aucune bijection de ce type n'était connue sur les cartes. Nous comblons cette lacune en proposant une décomposition naturelle des cartes planaires boisées en un arbre et une partition non-croisée.

Mettre à jour des k-d trees presque à peu de coût
N. Broutin, K. Dalal, L. Devroye, E. McLeish

Les k-d trees sont des structures géométriques qui encodent des partitions d'un espace, $[0,1]^2$ par example. Le decoupage est fait recursivement par des hyperplans passant par un point d'un l'ensemble donné P. Elles sont très contraignantes à cause de l'alternance des directions des découpes imposée par la définition. En particulier, il est difficile d'insérer de nouveaux points, si ce n'est dans les feuilles. A. Duch a introduit des structures (relaxed k-d trees) qui relachent la contrainte d'alternance dans le but de rendre la mise à jour plus facile. Cependant, le temps nécéssaire pour les mises à jour n'a pas été étudié. Nous prouvons qu'en moyenne, la mise à jour d'un arbre peut se faire de manière remarquablement rapide sous une hypothèse d'uniformité des données.

Comment trouver une base d'un réseau à partir d'un oracle ?
Guillaume Dabosville

Soient $b_1, ..., b_n$, n vecteurs linéairement indépendants de Z^n. On note $L = L(b_1, ..., b_n)$ l'ensemble des combinaisons linéaires entières des vecteurs $b_1, ..., b_n$. L forme alors ce que l'on appelle un réseau. Un réseau est un ensemble infini de points (ici des points de Z^n). La façon la plus naturelle et la plus utilisée en algorithmique pour représenter un tel objet, est d'en donner une base, ici constituée des vecteurs $b_1, ..., b_n$.

Dans l'exposé nous proposons de répondre à la question suivante : Soit L un réseau de Z^n et O_L une fonction (appelée oracle) qui prend en entrée des points x de Z^n et qui renvoie 1 (répond OUI) si x est un point de L, 0 (NON) sinon. Peut-on lorsque l'on détient un tel oracle associé à L, en déduire une base de L ? L'oracle est en quelque sorte une abstraction de toute représentation possible d'un réseau, partant du principe qu'une représentation permet au minimum de tester l'appartenance d'un point au réseau.

La Tortue de Lyapunov et le Lièvre Dyadique.
Benoît Daireaux, Véronique Maume, **Brigitte Vallée**.

Nous étudions un algorithme de pgcd dirigés par les bits les moins significatifs des entiers, l'algorithme LSB, et effectuons une analyse en moyenne précise de ses principaux paramètres (nombre d'itérations, nombre de shifts, etc ...) Cette analyse est basée sur une étude précise des systèmes dynamiques qui constituent l'extension continue de l'algorithme. Il apparaît qu'il est plus efficace de travailler ici une double extension, à la fois 2-adique et réelle. Ceci nous amène dans le cadre des produits de matrices aléatoires, et notre principal résultat s'exprime donc à l'aide de l'exposant de Lyapunov γ de l'ensemble des matrices relatives à l'algorithme. Ce dernier peut finalement être vu comme une course entre un Lièvre dyadique, d'une vitesse de 2 bits par étape, et une Tortue de Lyapunov d'une vitesse de $\gamma \sim 0.05$ bits par étape. Même si la Tortue part avant le Lièvre, celui-ci la rattrape et l'algorithme s'arrête.

Représentation succincte de triangulations avec un bord
Luca Castelli Aleardi

Nous considérons le problème de représenter des structures géometriques de manière compacte en gardant une implémentation efficace des opérations de navigation. Pour le cas des triangulations planaires à m faces, nous proposons une représentation compacte de l'information combinatoire qui améliore à 2.175 bits par triangle le coût asymptotique en espace et qui permet la navigation entre triangles adjacents en temps constant.

Pour les triangulations à m faces d'une surface de genre g, notre représentation nécessite asymptotiquement de $36(g-1)\lg m$ bits supplémentaires. La structure permet aussi l'accès en temps constant des informations associées aux sommets, notamment leurs coordonnées, cependant nous ne traitons pas ici la compression de cette information géometrique.

Graphes aléatoires réguliers et diffusion efficace
Philippe Duchon et Nicolas Hanusse

Nous proposons, pour le problème de la diffusion de rumeurs, deux protocoles de type "push and pull", efficaces aussi bien en temps total qu'en nombre de transmissions de rumeurs. L'analyse est basée sur des propriétés de modèles de graphes aléatoires réguliers (connexité et faible diamètre).

Élection Uniforme dans les Polyominoïdes
A. El Hibaoui, N. Saheb-Djahromi et A. Zemmari

Dans cet exposé, nous présentons une famille de graphes appelés polyominoïdes. Nous introduisons ensuite un algorithme distribué d'élection uniforme dans les polyominoïdes. Le processus d'élection est un processus d'éliminations qui enlève du polyominoïde les sommets actifs un par un jusqu'à le réduire à un sommet unique, appelé sommet élu. En utilisant seulement un calcul local et une seule passe, l'algorithme affecte une durée de vie à tout sommet actif. Cette durée de vie est une variable aléatoire exponentielle de paramètre défini.

Le processus d'éliminations est modélisé par un processus markovien dans le temps continu. Notre algorithme est totalement équitable dans la mesure où tous les sommets ont la même probabilité d'être élus.

Automates cellulaires asynchrones
Nazim Fatès et Nicolas Schabanel
Travail réalisé en collaboration avec Michel Morvan et Éric Thierry

Les automates cellulaires sont très utilisés pour simuler des phénomènes "réels". Or, la plupart des phénomènes réels est très robuste, en particulier à l'asynchronisme latent (il semble peu probable qu'un système biologique soit synchrone). Il apparaît pourtant que le comportement d'un automate cellulaire varie énormément du régime synchrone à un régime asynchrone. Par exemple, le diagramme espace-temps du jeu de la vie 2D "converge" vers la forme bien connue en synchrone (avec des clignotants et des gliders), alors qu'il prend l'aspect de labyrinthes en régime asynchrone. Nous assistons donc à des transitions de phases entre ces régimes, différentes pour chaque automate. Nous présenterons ici différentes simulations illustrant les différents comportements que nous avons observés sur les automates cellulaires élémentaires (i.e. à deux états), puis les résultats théoriques que nous avons obtenus, caractérisant complètement le comportement totalement asynchrone des 64 automates élémentaires doublement quiescents (dont les états 0 et 1 sont stables). Ces résultats reposent sur l'utilisation de couplage avec des variants de type Martingales.

Analyse de la profondeur dans un arbre des suffixes sous modèle markovien
Julien Fayolle et Mark D. Ward

La profondeur dans un arbre des suffixes est liée, entre autres en bio-informatique, à la recherche de chaînes se répétant. En utilisant des méthodes classiques d'analyse combinatoire et les résultats antérieurs de Jacquet et Szpankowski, on montrera qu'asymptotiquement la profondeur a un comportement moyen similaire à celui des arbres digitaux, soit en $\log n/h + O(1)$. On se limitera par simplicité à un modèle markovien d'ordre 1.

Génération aléatoire de graphes planaires en utilisant la méthode de Boltzmann
Eric Fusy

La méthode de Boltzmann est un processus tres général de génération aléatoire. Comme la méthode récursive, elle s'applique aux structures combinatoires ayant une décomposition récursive. Nous utilisons une décomposition bien connue des graphes planaires par degré croissant de connectivité et en déduisons un générateur de Boltzmann très efficace pour la génération aléatoire de graphes planaires étiquetés.

Limite Locale pour des Algorithmes Euclidiens
Aïcha Hachemi

On s'interesse au comportement asymptotique d'une fonction coût quelconque associée à des algorithmes euclidiens rapides, et on obtient un théorème de la limite locale.

Peut-on augmenter tout graphe en un petit monde ?
Emmanuelle Lebhar

Kleinberg a montré en 2000 qu'en plus des propriétés statistiques de graphe (diamètre, degré, clustering, ...), un petit-monde peut aussi être vu comme un graphe dans lequel le routage se fait de facilement, et efficacement, malgré l'absence de connaissance globale. Précisément, dans une grille augmentée d'arcs aléatoires (choisis non uniformément), un chemin court de longueur moyenne polylogarithmique peut être trouvé en utilisant l'algorithme glouton et une connaissance locale des noeuds. On appelle un tel graphe un petit monde navigable, puisque des chemins courts existent et peuvent être suivis avec une connaissance partielle du réseau. Nous montrons qu'il existe une large classe de graphes qui peuvent être augmentés, avec un arc par noeud, en petits mondes navigables.

Lois normales pour la complexité en bits d'algorithmes euclidiens
Loïck Lhote

Nous étudions deux algorithmes euclidiens agissant sur les polynômes ou les entiers : l'algorithme d'Euclide classique qui calcule uniquement le pgcd de deux éléments [deux polynômes ou deux entiers] et, l'algorithme d'Euclide étendu qui calcule en plus les coefficients de Bezout. Nous montrons

que les complexités en bits de ces algorithmes suivent des lois limites gaussiennes sauf pour l'algorithme d'Euclide classique sur les entiers. Nous donnons aussi l'ordre de grandeur des espérances et variances.

Les preuves de ces résultats se basent d'une part sur la décomposition des coûts en plusieurs familles coûts simples et d'autre part, sur le fait que les coûts à croissance modérée suivent une loi limite gaussienne (Baladi et Vallée). Ce dernier résultat est fondamental mais ne s'applique pas à la complexité en bits qui n'est pas à croissance modérée.

Les analyses dans le cas polynômial et dans le cas entier peuvent s'effectuer dans le cadre commun des systèmes dynamiques. Les séries génératrices s'expriment alors en fonction d'opérateurs, les pôles en fonction de valeurs propres, etc...

Asymptotic Analysis of a Leader Election Algorithm

Guy Louchard et Christian Lavault

Itai and Rodeh showed that, on the average, the communication of a leader election algorithm takes no more than Ln bits, where $L \simeq 2.441716$. We give a precise asymptotic analysis of the average number of rounds $M(n)$ required by the algorithm, proving for example that $L = \lim_{n\to\infty} M(n) = 2.441715879\ldots$ Accurate asymptotic expressions of the second moment $M^{\{}(2)\}(n)$ of the discrete random variable at hand and of the probability distribution, as well as the generalization to all moments are given. Corresponding asymptotic expansions for large n are provided. We also analyze the optimal probability of asking to be elected. Finally, our numerical results show that both computations fit in perfectly.

Tresse aléatoire à trois brins et raie manta

Jean Mairesse

On considère le groupe de tresses à trois brins $B_3 = \langle a, b | aba = bab \rangle$. On forme une tresse aléatoire en choisissant à chaque étape, de façon indépendante, un des générateurs $\{a, a-1, b, b-1\}$ suivant une loi de probabilité fixée. Ce modèle a été proposé par des physiciens comme pertinent pour l'étude de la formation de polymères. On veut déterminer la vitesse asymptotique à laquelle se forme la tresse aléatoire (la vitesse de fuite de la marche aléatoire). Pour ce faire, on détermine explicitement la "mesure harmonique" (qui donne la "direction" de la fuite vers l'infini) de la marche aléatoire sur le groupe B3 quotienté par son centre. Cette mesure harmonique est obtenue en composant une mesure Markovienne multiplicative avec une transduction rationnelle. Les techniques développées s'appliquent également à l'étude des marches aléatoires sur d'autres groupes : produits amalgamés de groupes finis, et extensions HNN de groupes finis.

Inégalités de concentration et estimation de probabilités conditionelles

Véronique Maume

Pour des processus faiblement dépendants, on montre des inégalités de concentration qui conduisent à des estimateurs de certaines probabilités conditionelles. Dans le cadre des mesures de Gibbs et de certains systèmes dynamiques, ces estimations permettent de construire un estimateur consistant de la fonction de potentiel.

L'algorithme du commerce simulé appliqué à la résolution du problème de tournées des véhicules (VRPTW)

Abderrazzak Nejeoui

Nous présentons une heuristique améliorée pour le problème de routage des véhicules. L'heuristique trouve un échange complexe de clients pour améliorer une solution initiale. Notre approche est modulaire, par conséquent elle est facilement adaptable aux différentes versions du problème avec contraintes additives, comme *fenêtres de temps, backhauls...*

Profiles moyens, des tries aux arbres-suffixes

Pierre Nicodème

Park et Szpankowski ont récemment étudié le profile externe asymptotique des tries binaires aléatoires. En utilisant les mêmes méthodes, transformée de Mellin et méthode de col, nous étudions le profile interne de ces tries. Dans le modèle Poisson de paramètre z, nous observons, pour z grand, une inversion de phase du taux de saturation en noeuds internes à la profondeur $k = 2\log z/(\log(1/p) + \log(1/q))$, où p et q sont les probabilités d'apparition de 1 et de 0 dans les clés. Par une approche combinatoire intuitive et l'utilisation de partitions de mots, Julien Fayolle a comparé l'espérance de certains paramètres des tries et des arbres-suffixes. Nous appuyant sur cette approche, nous étudions le nombre des noeuds internes à profondeur k d'un trie qui correspondent à la classe des mots périodiques de longueur k. La comparaison des transformées de Mellin correspondant à différentes classes de mots permet de prouver que les profiles internes asymptotiques moyens des tries et des arbres-suffixes sont équivalents.

Génération aléatoire de mots d'un langage non-algébrique : les chemins culminants
Yann Ponty

Les chemins culminants sont des chemins positifs composés de pas (1,+a) et (1,-b) cheminant dans le quart de plan à partir de (0,0) jusqu'à un point (n,h) "culminant", c'est à dire d'ordonnée jamais précédemment dépassée ni atteinte. Quand l'ordonnée finale est fixée à l'avance, ce langage est rationnel et on dispose alors d'un algorithme de génération aléatoire linéaire sur la taille du chemin à engendrer. Quand le langage est considéré "dans toute sa généralité", alors il n'est ni rationnel, ni algébrique ... On propose alors un algorithme cubique en complexité arithmétique pour la génération uniforme dans le cas général (+a,-b), et on étudie les complexités d'algorithmes "par rejet" pour des valeurs particulières de a et b. En particulier, quand a=1 et b=1, l'asymptotique de la complexité de génération est obtenue par une méthode de combinatoire analytique.

Lois discrètes aléatoires et analyse bayésienne de quelques heuristiques
Christian Paroissin

Les lois discrètes aléatoires (random discrete distribution) sont utilisées pour proposer une analyse probabiliste de certaines heuristiques, en particulier move-to-front et move-to-root. Partant des résultats déterministes connus, nous obtenons diverses caractérisations de ces heuristiques (lois, moments, lois asymptotiques, comparaisons des couts moyens, ...). Nous étudierons plus spécialement deux cas proposés par Kingman (1975) : le cas de la loi de Dirichlet et le cas de lois stables.

Grandes urnes de Polya-Eggenberger
Nicolas Pouyanne

On dira qu'une urne de Polya-Eggenberger est grande lorsque sa matrice de remplacement a des grandes valeurs propres, dans un sens que l'on définira. Dans ces conditions, une fois renormalisée, on montre que sa composition admet toujours une asymptotique presque sûre, oscillante ou non selon le spectre de la matrice de remplacement. On précisera également une manière d'aborder le calcul des moments des variables aléatoires limites en toute dimension et sans hypothèse d'irréductibilité.

La méthode est nouvelle et de nature algébrique. On donnera de nombreux exemples.

La loi convergence de la longueur de facteur droite standard d'un mot aléatoire de Lyndon
Elahe Zohoorian-Azad

Considérez l ensemble des mots finis sur un alphabet totalement ordonné avec $q(> 1)$ lettres. Nous montrons que la loi de la longueur du facteur droit standard d un mot de Lyndon aléatoire de taille n, divisé par n, converge en

$$f(x) = 1/2[1_{\{1\}}(x) + 1_{[0,1)}(x)dx],$$

quand n tend vers l infini. La convergence de tous les moments suit.