

Action Génération Aléatoire

Olivier Bodini, Maryse Pelletier, Michèle Soria

Alexis Darrasse, Carine Pivoteau

Modèle de Boltzmann

► Outils d'analyse

- Modélisation algorithmique
- Combinatoire analytique
- Résolution de systèmes fonctionnels

► Points forts de la méthode

- Généricité : structures combinatoires complexes
- Efficacité en temps et mémoire

► Notre contribution

- Développements du modèle (oracle, extensions)
- Développements logiciels (générique et dédiés)
- Domaines applicatifs (réseaux, tests)

► Interactions

- ANR – projet GAMMA : P6-P7-Orsay-Marne
- INRIA – Algorithms
- Spécialistes domaines : grands graphes, tests logiciels



Pourquoi ?

- ▶ **Observer Expérimenter → Analyser Conjecturer**
Structures complexes,
visualisation des formes et propriétés limites
- ▶ **Tester → Valider**
Validation de programmes, adéquation de modèles
sans choix humain du test : couverture typique
- ▶ **Modéliser structures massives**
mots, arbres, graphes, pavages...
- ▶ **Large échelle → Efficacité**
structures de grande taille— $10^6 - 10^9$

Génération Aléatoire de Structures Combinatoires

- ▶ Structures Combinatoires Décomposables
Spécifications par **grammaire**
Union, Produit, Sequence, Cycle, Ensemble
Taille : chaque structure a une **taille**
- ▶ Tirage Aléatoire
Distribution **uniforme** sur tous les objets d'une taille donnée

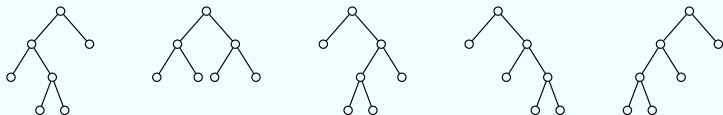
Arbres binaires

$$\mathcal{B} = \mathcal{Z} \cup \mathcal{Z} \times \mathcal{B} \times \mathcal{B}$$

$$\text{taille}(\mathcal{Z}) = 1$$

$$\text{taille}(\mathcal{Z}.B_1.B_2) = 1 + \text{taille}(B_1) + \text{taille}(B_2)$$

Tirer($\mathcal{B}, 5$) :



Spécifications combinatoires et séries génératrices

Système de spécifications $\rightarrow$ Système d'équations fonctionnelles

	specification	ordinary g.f. (unlabelled)	exponential g.f. (labelled)
ε / atom	$1 / \mathcal{Z}$	$1 / z$	$1 / z$
Union	$\mathcal{C} = \mathcal{A} \cup \mathcal{B}$	$C(z) = A(z) + B(z)$	$\hat{C}(z) = A(z) + B(z)$
Product	$\mathcal{C} = \mathcal{A} \times \mathcal{B}$	$C(z) = A(z) \times B(z)$	$\hat{C}(z) = A(z) \times B(z)$
Sequence	$\mathcal{C} = \text{SEQ}(\mathcal{A})$	$C(z) = \frac{1}{1-A(z)}$	$\hat{C}(z) = \frac{1}{1-A(z)}$
Set	$\mathcal{C} = \text{SET}(\mathcal{A})$	$\exp\left(\sum_{k=1}^{\infty} \frac{(-1)^{k-1}}{k} A(z^k)\right)$	$\hat{C}(z) = \exp(A(z))$
Multi-Set	$\mathcal{C} = \text{MSET}(\mathcal{A})$	$\exp\left(\sum_{k=1}^{\infty} \frac{1}{k} A(z^k)\right)$	—
Cycle	$\mathcal{C} = \text{CYC}(\mathcal{A})$	$\sum_{k=1}^{\infty} \frac{1}{k} \log \frac{1}{1-A(z^k)}$	$\hat{C}(z) = \log \frac{1}{1-A(z)}$

Résolution Système d'équations + Algorithmes probabilistes $\implies$
Génération aléatoire Boltzmann

Générateur de Boltzmann

Classe Combinatoire $\mathcal{C}$ $\rightarrow$ Série génératrice $C(z) = \sum C_n z^n$

- ▶ Modèle Boltzmann de paramètre x :
 - ▶ Génération d'un objet γ de $\mathcal{C}$: $\Pr_x(\gamma) = \frac{x^{|\gamma|}}{C(x)}$
 - ▶ Taille objet = variable aléatoire : $\Pr_x(|\gamma| = n) = \frac{C_n x^n}{C(x)}$
- ▶ Propriétés
 - ▶ Génération uniforme
 - ▶ Complexité linéaire pour génération libre
 - ▶ Linéarité conservée pour taille approchée
- ▶ Algorithme probabiliste

construction	sampler
$\mathcal{C} = \emptyset$ or $\mathcal{Z}$	$\Gamma C(x) := \varepsilon$ or atom
$\mathcal{C} = \mathcal{A} + \mathcal{B}$	$\Gamma C(x) := \text{Bern} \frac{A(x)}{C(x)} \longrightarrow \Gamma A(x) \mid \Gamma B(x)$
$\mathcal{C} = \mathcal{A} \times \mathcal{B}$	$\Gamma C(x) := \langle \Gamma A(x) ; \Gamma B(x) \rangle$
$\mathcal{C} = \text{SEQ}(\mathcal{A})$	$\Gamma C(x) := \text{Geom} A(x) \implies \Gamma A(x)$

$$+ \frac{1}{2}B(z)^2 + \frac{1}{2}B(z^2)$$

$$+ \mathcal{B} \times \mathcal{B}$$

$$+ B(z)^2$$

$$\text{Bern}\left(\frac{x}{B(x)}\right) = 1$$

$$\text{Bern}(x/B(x))$$

$$\text{Bern}\left(\frac{B^2(x)}{B^2(x)+B(z^2)}\right) = 1$$

$$\langle \Gamma B(x), \Gamma B(x) \rangle$$

$$\langle \Gamma B(x), \Gamma B(x) \rangle$$

$$b \leftarrow \Gamma B(x^2)$$

$$\langle b, b \rangle$$

GenL(exp)

If terminal(exp) Then $x \rightarrow \boxed{\text{exp}}$

Else

If +?(exp) Then **** Unions ****

$x \rightarrow$ If bern($fg_x(\text{filsG}(\text{exp}))/fg_x(\text{exp})$) Then

GenL(filsG(exp))

Else

GenL(filsD(exp))

If x?(exp) Then **** Produits ****

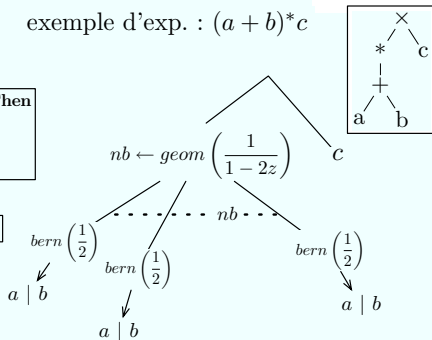
$x \rightarrow$ GenL(filsG(exp)), GenL(filsD(exp))

If *(exp) Then **** Sequences ****

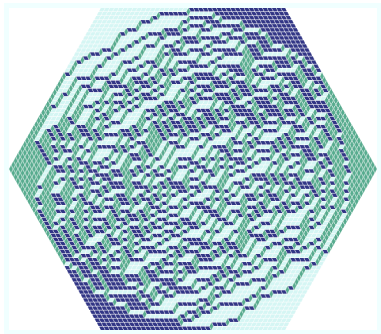
$x \rightarrow$ nb $\leftarrow$ geom($fg_x(\text{exp})$)
repeat nb GenL(fils(exp))

End

exemple d'exp. : $(a + b)^*c$

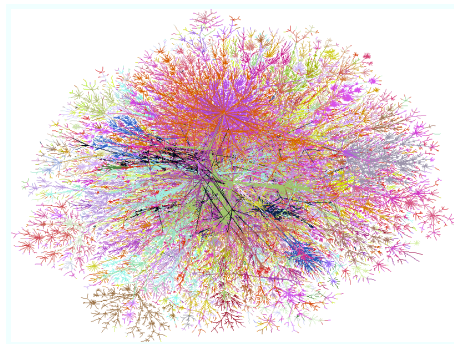


Exemples d'applications



Phénomènes de transition de phase
Modèles de physique statistique

Partitions planes



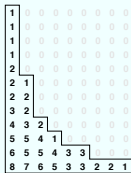
Modèles de réseaux d'interaction
Distribution de paramètres clés

Réseaux apolloniens

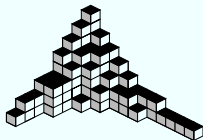
Partitions planes

$$\mathcal{P} = \text{MSET}(\mathcal{Z} \times \text{SEQ}(\mathcal{Z})^2)$$

$$P(z) = \prod_{i,j \geq 0} \frac{1}{1 - z^{i+j+1}} = \prod_{r \geq 1} (1 - z^r)^{-r}$$



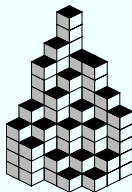
12 7 4 3 2 2 1 1 1



plane partition

$$\mathcal{P}_{a,b} = \text{MSET}(\mathcal{Z} \times \text{SEQ}_{<a}(\mathcal{Z}) \times \text{SEQ}_{<b}(\mathcal{Z}))$$

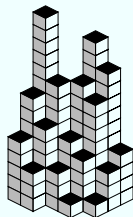
$$P_{a,b}(z) = \prod_{\substack{0 \leq i < a \\ 0 \leq j < b}} \frac{1}{1 - z^{i+j+1}}$$



boxed

$$\mathcal{P}_{a,b,c} = \prod_{(i,j) \in \lambda(a,b,c)} \text{SEQ}(\mathcal{Z}^{h(i,j)})$$

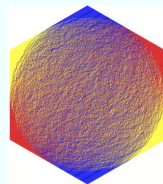
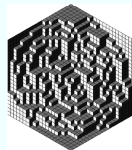
$$P_{a,b,c}(z) = \prod_{(i,j) \in \lambda(a,b,c)} \frac{1}{1 - z^{h(i,j)}}$$



skew

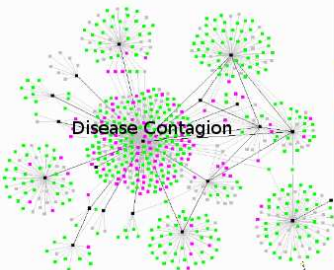
$$P_{a,b,c}(z) = \sum_{n \geq 0} p_n z^n$$

$$p_n = \prod_{i=1}^a \prod_{j=1}^b \prod_{k=1}^c \frac{i+j+k-1}{i+j+k-2}$$



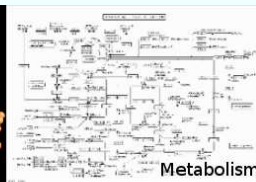
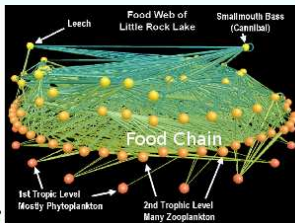
tiling

Réseaux d'Interaction



Disease Contagion

Mean Distance

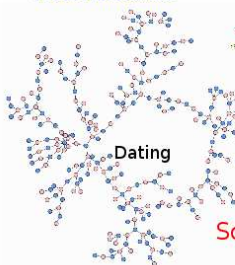


Metabolism

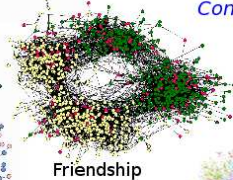
Clustering

Biology

Connectivity



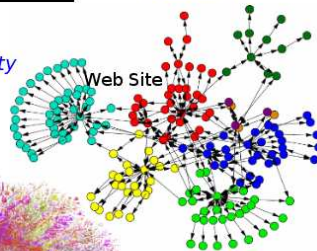
Dating



Friendship

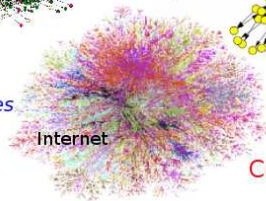
Number of Edges

Sociology



Web Site

Degree Distribution



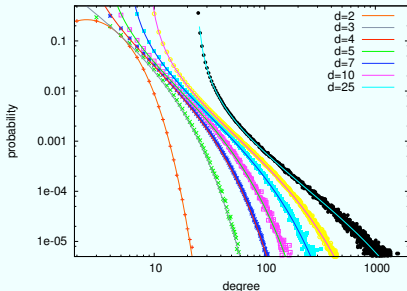
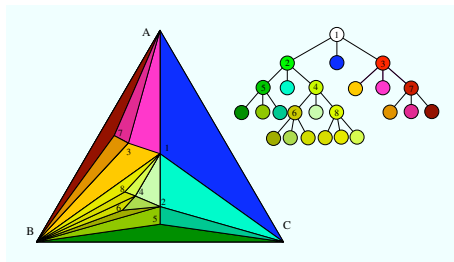
Internet

Computer Science

Réseaux apolloniens

Bijection avec les arbres ternaires

- Paramètres clés des réseaux
 - degré d'un noeud
 - distance entre deux noeuds
- Transfert sur les arbres ternaires
 - taille de certains sous-arbres



Boltzman et analyses de distribution

- séries génératrices bivariées
- indépendance des appels récursifs

Arbre généré Boltzmann singulier

≡

sous-arbre d'un arbre aléatoire