

# Combinatoire énumérative et génération aléatoire

Mireille Bousquet-Mélou

*<http://www.labri.fr/~bousquet>*

## La combinatoire énumérative : l'art de *compter*

COMPTER [κ ò t e ]. I. V. tr. 1. Déterminer (une quantité) par le calcul ; spécialt. Établir le nombre de  $\Rightarrow$  chiffrer, dénombrer, nombrer. *Compter les moutons.*

### Problème générique

$\mathcal{A}$  ensemble des moutons

$\mathcal{A}_n$  ensemble des moutons à  $n$  pattes

### Question

évaluer, estimer, caractériser le nombre

$$a_n = |\mathcal{A}_n|.$$

## La combinatoire énumérative : l'art de *compter*

COMPTER [κ ò t e ]. I. V. tr. 1. Déterminer (une quantité) par le calcul ; spécialt. Établir le nombre de  $\Rightarrow$  chiffrer, dénombrer, nombrer. *Compter les moutons.*

### Problème générique

$\mathcal{A}$  ensemble des moutons

$\mathcal{A}_{n,p}$  ensemble des moutons à  $n$  pattes et  $p$  cornes

### Question

évaluer, estimer, caractériser le nombre

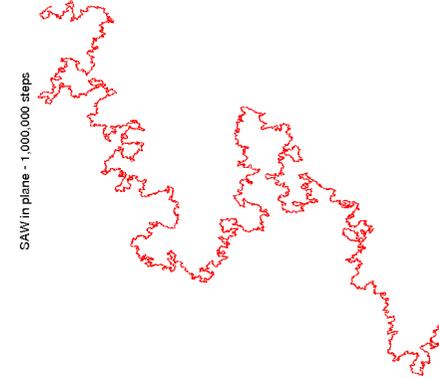
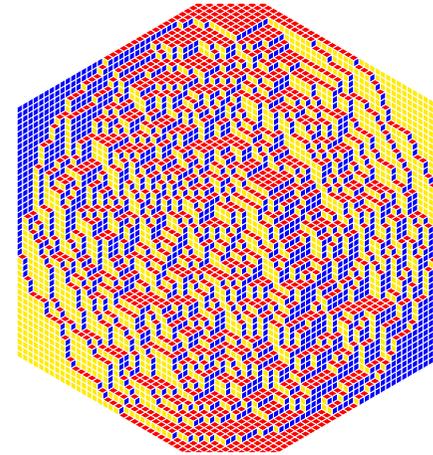
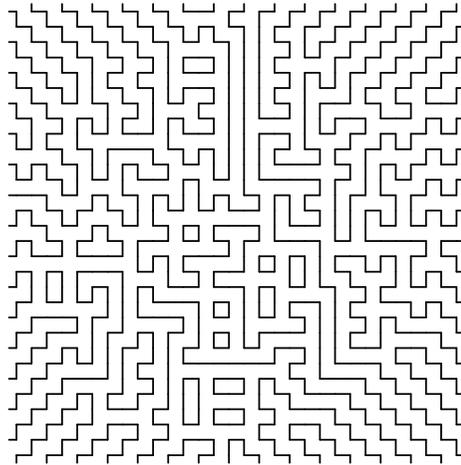
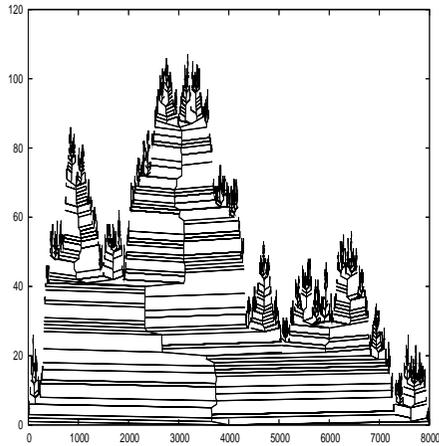
$$a_{n,p} = |\mathcal{A}_{n,p}|.$$

## Le coefficient binomial ou la combinatoire de Monsieur Jourdain

$$\begin{aligned} C_n^p &= \text{nombre de sous-ensembles de } \{1, 2, \dots, n\} \text{ ayant } p \text{ éléments} \\ &= \binom{n}{p} = \frac{n!}{p!(n-p)!} = \frac{n(n-1)\cdots(n-p+1)}{p!} \end{aligned}$$

# Génération aléatoire

Trouver une procédure (rapide) qui tire au hasard de façon uniforme un objet de taille  $n$  fixée.



Applications : tests, expérimentations

# Plan

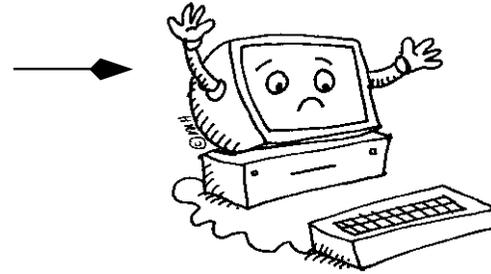
- Compter quoi, et pour quoi ?
- Comment compter ?
- Quelques questions “simples”

## I. Pourquoi compter ?

- **Analyse d'algorithmes** : énumération des entrées selon la taille et le coût de traitement → analyse de la **complexité** de l'algorithme (en moyenne, voire en distribution).

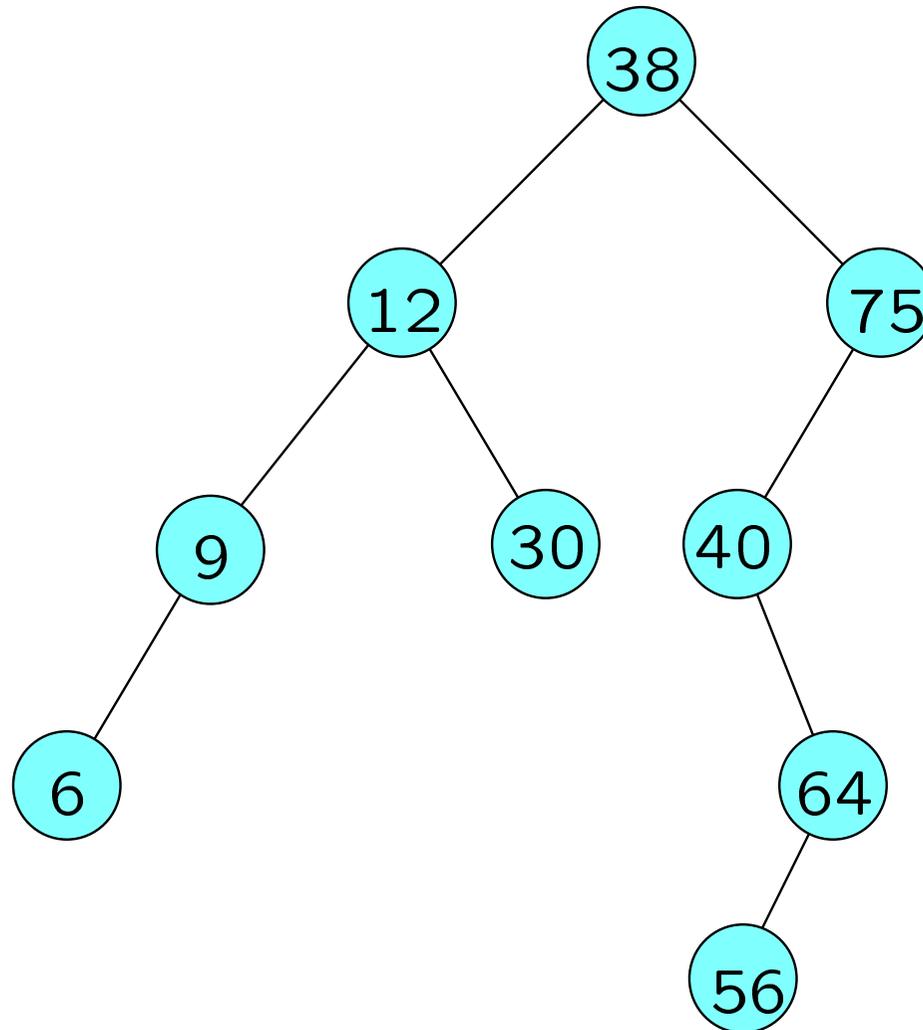
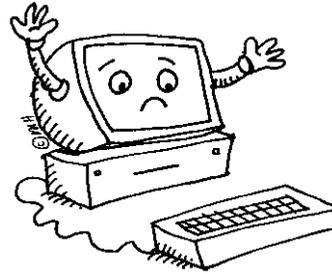
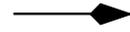
## Stockage d'une liste : les arbres binaires de recherche

56 6 64 9 40 30 75 12 38



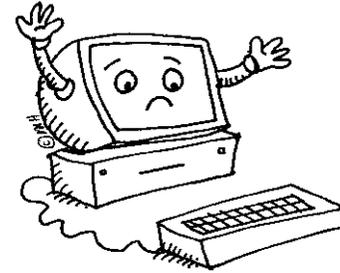
# Stockage d'une liste : les arbres binaires de recherche

56 6 64 9 40 30 75 12 38



# Arbres binaires de recherche

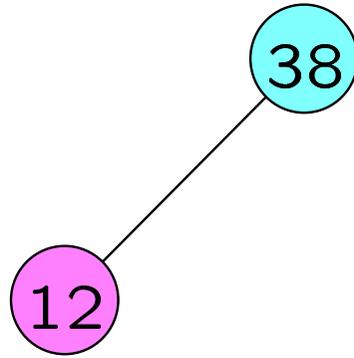
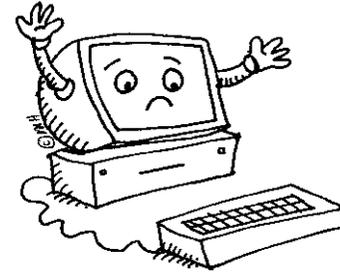
56 6 64 9 40 30 75 12 38 →



12 → (38)

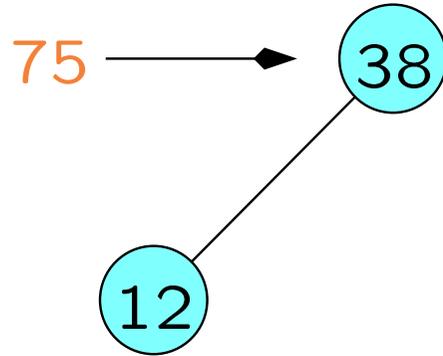
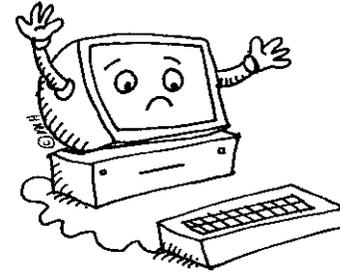
# Arbres binaires de recherche

56 6 64 9 40 30 75 12 38 →



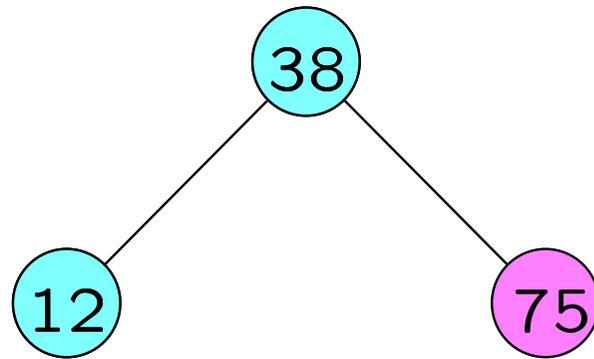
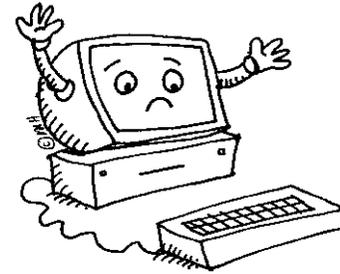
# Arbres binaires de recherche

56 6 64 9 40 30 75 12 38 →



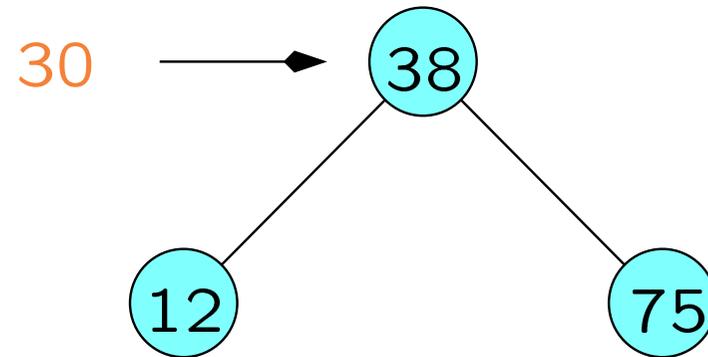
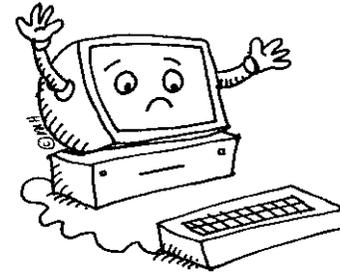
# Arbres binaires de recherche

56 6 64 9 40 30 75 12 38 →



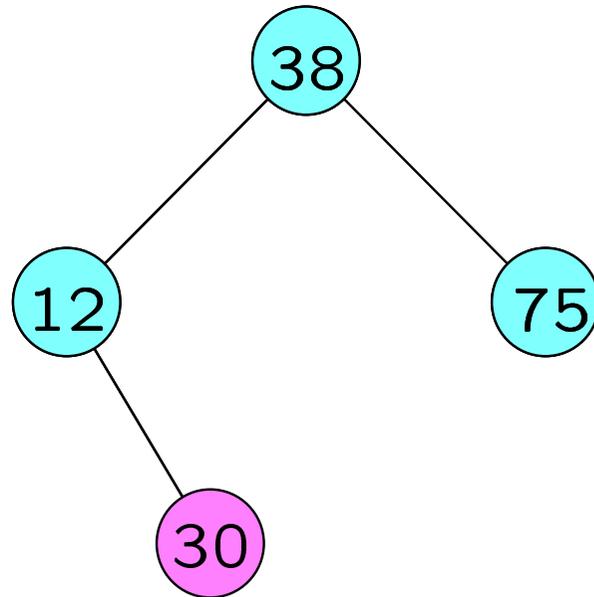
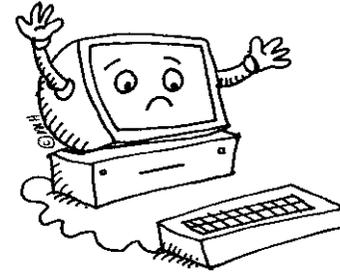
# Arbres binaires de recherche

56 6 64 9 40 30 75 12 38 →



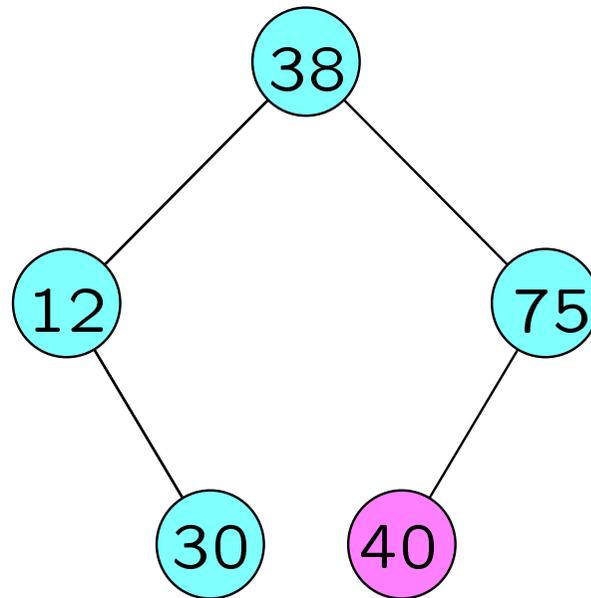
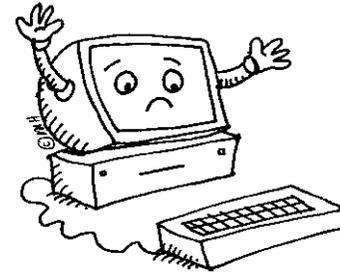
# Arbres binaires de recherche

56 6 64 9 40 30 75 12 38 →



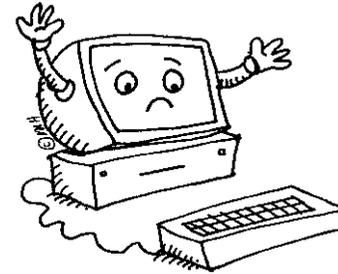
# Arbres binaires de recherche

56 6 64 9 40 30 75 12 38 →



## Arbres binaires de recherche

56 6 64 9 40 30 75 12 38 →



Combien faut-il faire de comparaisons, en moyenne, pour construire l'arbre associé à une liste de  $n$  entrées ?

Soit  $C(n)$  ce nombre. Alors :

$$C(n) = n - 1 + \frac{2}{n} \sum_{i=0}^{n-1} C(i),$$

dont on déduit que, pour  $n$  grand,

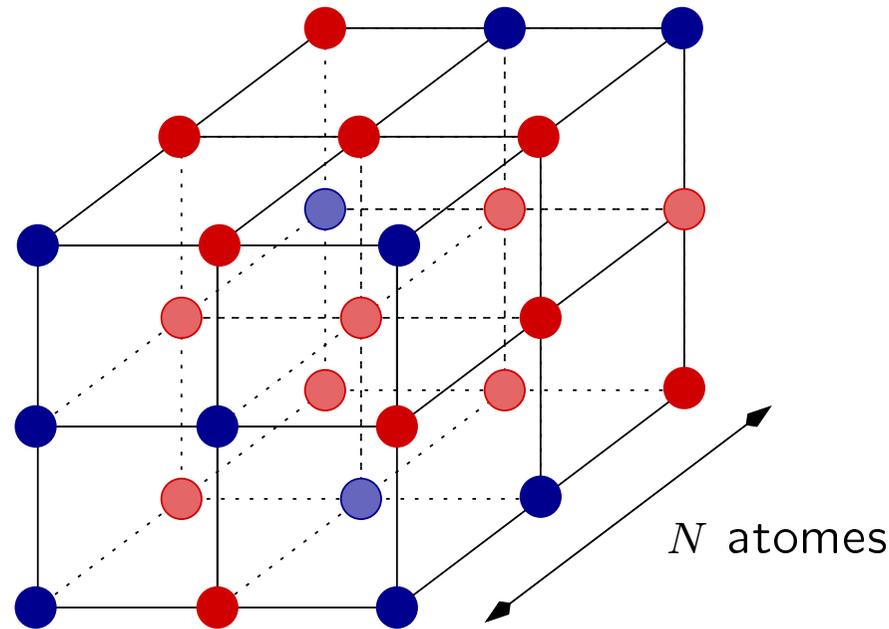
$$C(n) \sim 2n \log n.$$

## Pourquoi compter ?

- **Analyse d'algorithmes** : énumération des entrées selon la taille et le coût de traitement → complexité de l'algorithme
- **Physique statistique** : modèles discrets décrivant les interactions entre composants microscopiques → prédiction du comportement macroscopique

## Un modèle de cristal ferromagnétique [Lenz] : le modèle d'Ising

À une échelle microscopique...



**But :** Calculer la somme suivante, portant sur tous les coloriage possibles des sommets :

$$Z(N) = \sum_{\text{conf.}} x^{\bullet-\bullet},$$

où  $x$  est un nombre dépendant de la température et du matériau.

## Pourquoi compter ?

- **Analyse d'algorithmes** : énumération des entrées selon la taille et le coût de traitement → complexité de l'algorithme
- **Physique statistique** : modèles discrets décrivant les interactions entre composants microscopiques → prédiction du comportement macroscopique
- **Probabilités** : les probas discrètes sont (en grande partie...) de la combinatoire énumérative !  
Possibilité d'approcher des phénomènes continus par des modèles discrets.

## Pourquoi compter ?

- **Analyse d'algorithmes** : énumération des entrées selon la taille et le coût de traitement → analyse de l'algorithme
- **Physique statistique** : modèles discrets décrivant les interactions entre composants microscopiques → prédiction du comportement macroscopique
- **Probabilités** : les probas discrètes sont (en grande partie...) de la combinatoire énumérative !  
Possibilité d'approcher des phénomènes continus par des modèles discrets.
- Bioinformatique, Algèbre, etc.
- **Pour élucider la structure des objets**

## II. Comment compter ?

- Compter les objets

- Génération aléatoire

- Dénombrement selon plusieurs paramètres

Comprendre la structure  
d'objets discrets

- Prédire les propriétés statistiques des objets de grande taille

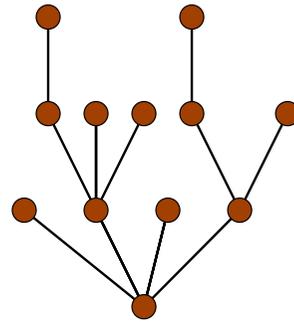
Objets aléatoires

Résultats asymptotiques

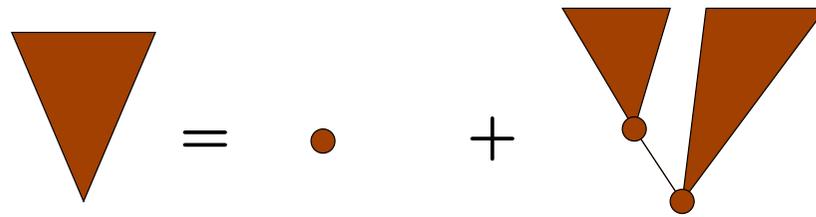
- Codage (compact, efficace)

Thème Graphes

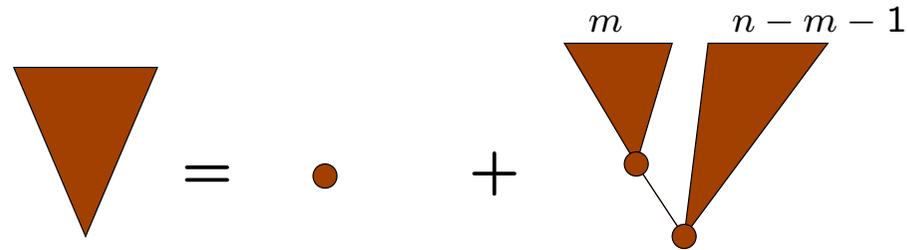
## Exemple : les arbres plans



## Les arbres plans : structure



## Les arbres plans : comptage (# arêtes)



$$a_n = \sum_{m=0}^{n-1} a_m a_{n-m-1}$$

- Série génératrice :

$$A(t) = \sum_{n \geq 0} a_n t^n$$

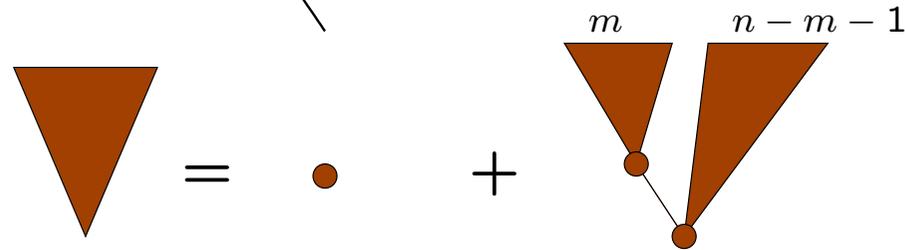
$$\Rightarrow A(t) = 1 + tA(t)^2$$

## Les arbres plans : comptage (# arêtes)

- Comptage

$$A(t) = \frac{1 - \sqrt{1 - 4t}}{2t}$$

$$a_n = \frac{1}{n+1} \binom{2n}{n}$$



$$a_n = \sum_{m=0}^{n-1} a_m a_{n-m-1}$$

- Série génératrice :

$$A(t) = \sum_{n \geq 0} a_n t^n$$

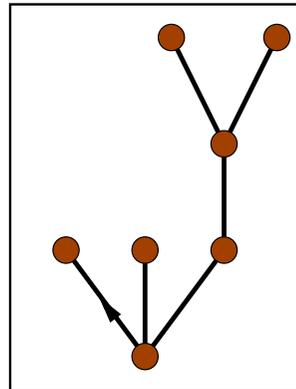
$$\Rightarrow A(t) = 1 + tA(t)^2$$

# Les arbres plans : codage

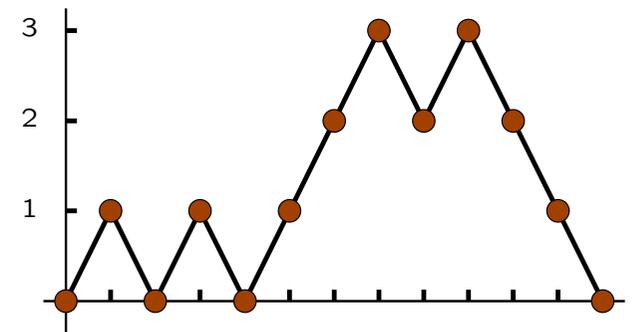
- Comptage

$$A(t) = \frac{1 - \sqrt{1 - 4t}}{2t}$$

$$a_n = \frac{1}{n+1} \binom{2n}{n}$$



- Codage



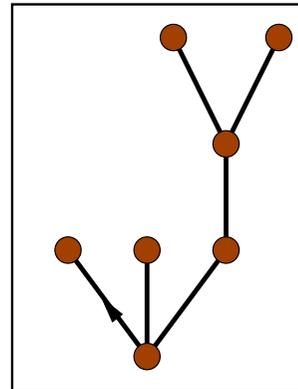
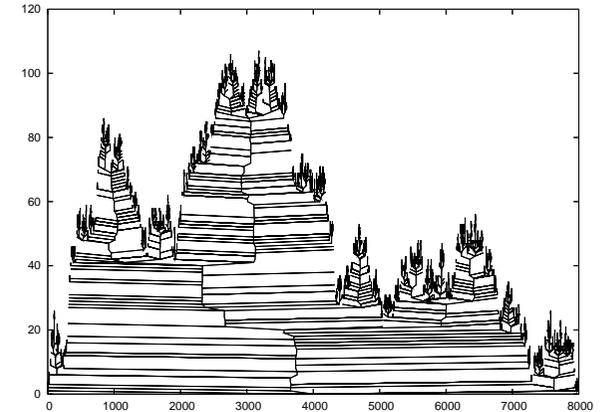
# Les arbres plans : génération aléatoire

- Comptage

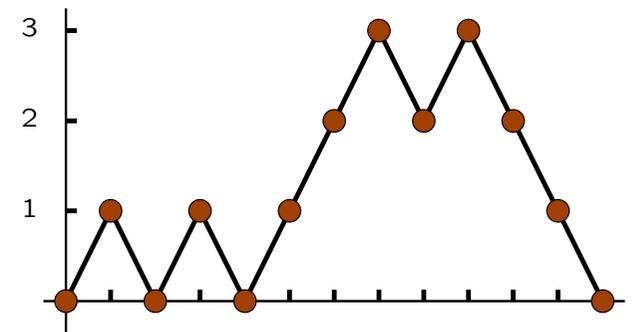
$$A(t) = \frac{1 - \sqrt{1 - 4t}}{2t}$$

$$a_n = \frac{1}{n+1} \binom{2n}{n}$$

- Génération aléatoire



- Codage



# La hauteur des (grands) arbres plans

- Comptage

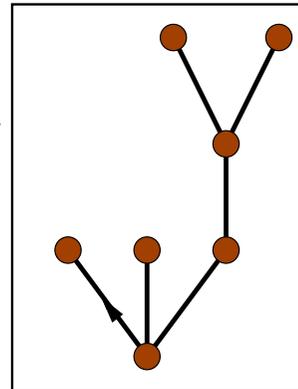
$$A(t) = \frac{1 - \sqrt{1 - 4t}}{2t}$$

$$a_n = \frac{1}{n+1} \binom{2n}{n}$$

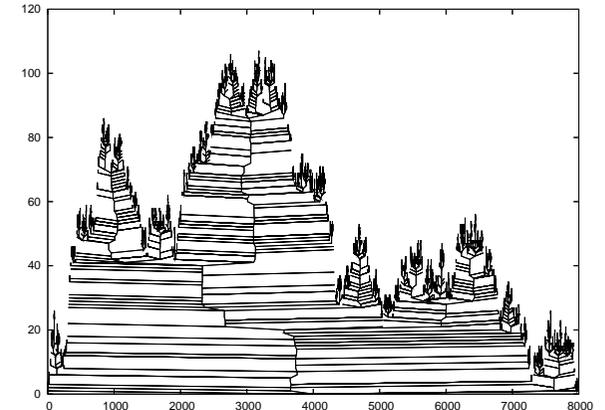
- Arbres de hauteur au plus  $j$

$$A_j(t) = A(t) \frac{1 - Q(t)^{j+1}}{1 - Q(t)^j}$$

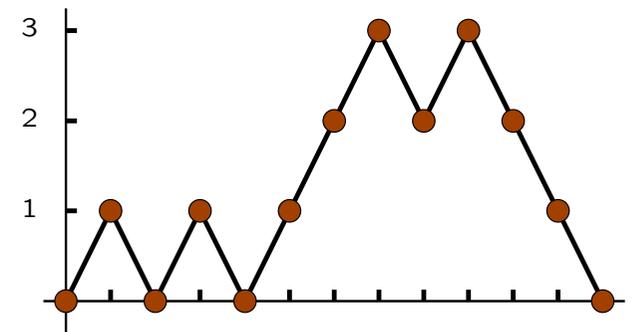
avec  $Q(t) = A(t) - 1$



- Génération aléatoire



- Codage



# La hauteur des (grands) arbres plans

- Comptage

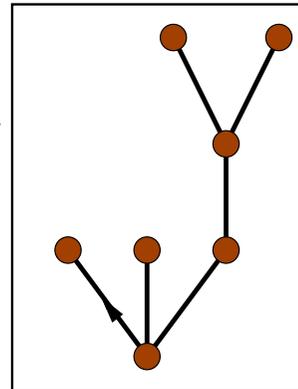
$$A(t) = \frac{1 - \sqrt{1 - 4t}}{2t}$$

$$a_n = \frac{1}{n+1} \binom{2n}{n}$$

- Arbres de hauteur au plus  $j$

$$A_j(t) = A(t) \frac{1 - Q(t)^{j+1}}{1 - Q(t)^j}$$

avec  $Q(t) = A(t) - 1$



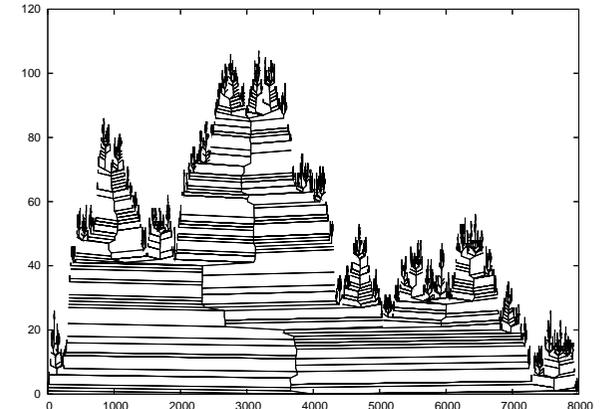
- La hauteur des grands arbres

$$\mathbb{E}(H_n) \sim \sqrt{\pi} \sqrt{n}$$

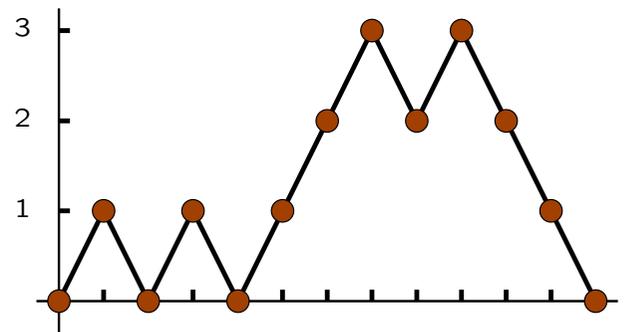
$$\mathbb{P}(H_n \leq x\sqrt{n}) \rightarrow F(x)$$

avec  $F(x) = \sum_{k \in \mathbb{Z}} e^{-k^2 x^2} (1 - 2k^2 x^2)$

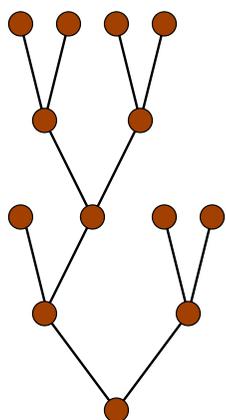
- Génération aléatoire



- Codage



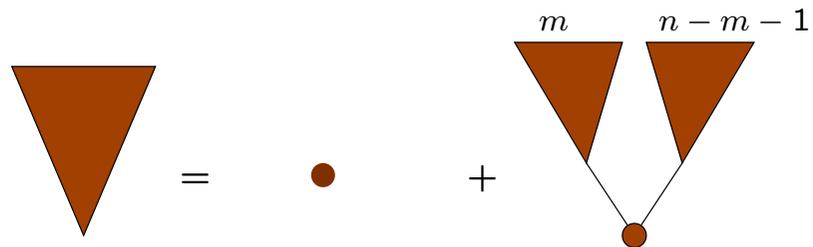
# Une structure peut en cacher une autre !



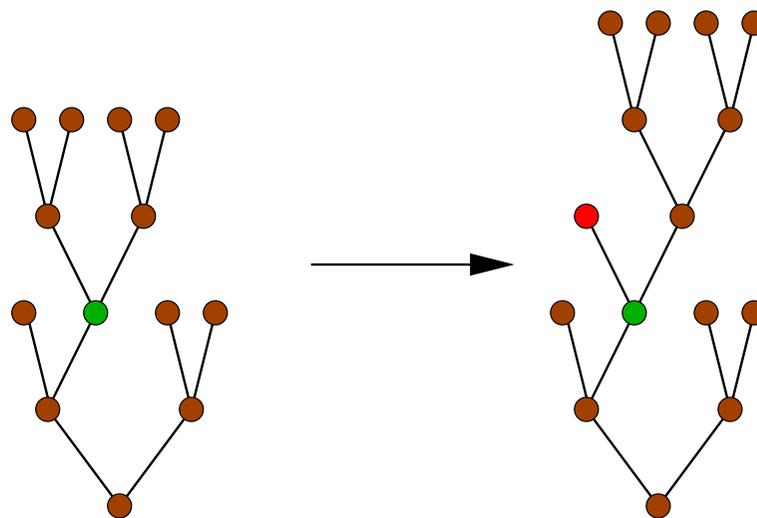
Arbre binaire

Taille  $n$  :

$n + 1$  sommets internes

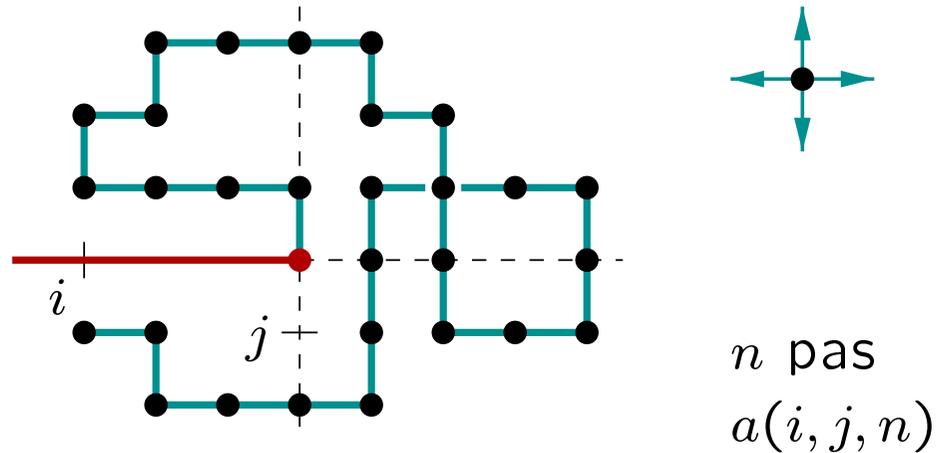


$$a(n) = \sum_{m=0}^{n-1} a(m)a(n-1-m)$$



$$2(2n + 1)a(n) = (n + 2)a(n + 1)$$

### III. Une petite question : Chemins dans un plan incisé



- Soit  $a(i, j; n)$  le nombre de chemins de longueur  $n$  issus de  $(0, 0)$ , finissant en  $(i, j)$ , et évitant la demi-droite rouge.

- Pourquoi a-t-on

$$a(1, 0; 2n + 1) = C_{2n+1}, \quad a(0, 1; 2n + 1) = 4^n C_n, \quad a(-1, 1; 2n) = C_{2n} \dots$$

avec  $C_n = \frac{1}{n+1} \binom{2n}{n} = n^{\text{ème}}$  nombre de Catalan

