Boltzmann Sampling and Properties of Trees

Michèle Soria

LIP6-UPMC
Paris 6 University
Same distribution & same properties

Subtrees of a large uniform random tree

Random trees independently generated with a singular Boltzmann sampler
Same distribution & same properties

Subtrees of a large uniform random tree

Random trees independently generated with a singular Boltzmann sampler

Michèle Soria
Same distribution & same properties

Subtree of a large uniform random tree

Random trees independently generated with a singular Boltzmann sampler
Illustration: size in ternary trees

Ternary trees: \[ T(z) = 1 + zT^3(z) \]

\[
T(z) = 3/2 - \sqrt{3}/2 \sqrt{1 - z/\rho} + ... \quad \rho = 4/27
\]

Singular Boltzmann sampler \equiv \text{branch with probability } 1/3

\[
\Pr(|A| = k) = \frac{T_k \rho^k}{T(\rho)} \sim Ck^{-3/2}
\]

Power law
Boltzmann generation
P. Duchon, P. Flajolet, G. Louchard, and G. Schaeffer

**Boltzmann model parameter** $x$

Combinatorial class $\mathcal{A}$, g. f.
$$A(z) = \sum a_n z^n$$

$\forall A \in \mathcal{A}, \ Pr_x(A) = \frac{x^{|A|}}{A(x)}$

- uniform random sampling
- but result size is random:
  $$\Pr_x(|A| = n) = \frac{a_n x^n}{A(x)}$$

- Probabilistic algorithm
- combinatorial constr.
- linear time

<table>
<thead>
<tr>
<th>construction</th>
<th>sampler with parameter $x$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$C = I$</td>
<td>$\Gamma C(x) := \varepsilon$</td>
</tr>
<tr>
<td>$C = Z$</td>
<td>$\Gamma C(x) := z$</td>
</tr>
<tr>
<td>$C = \mathcal{A} + \mathcal{B}$</td>
<td>$\Gamma C(x) := \text{Bern} \frac{A(x)}{C(x)} \implies \Gamma A(x) \mid \Gamma B(x)$</td>
</tr>
<tr>
<td>$C = \mathcal{A} \times \mathcal{B}$</td>
<td>$\Gamma C(x) := \langle \Gamma A(x) ; \Gamma B(x) \rangle$</td>
</tr>
<tr>
<td>$C = \text{SEQ}(\mathcal{A})$</td>
<td>$\Gamma C(x) := \text{Geom} A(x) \implies \Gamma A(x)$</td>
</tr>
<tr>
<td>...</td>
<td>...</td>
</tr>
</tbody>
</table>
Tree sampling

### Recursive structures

\[ T(z) = z \Phi(T(z)), \quad \rho > 0 \]

- \[ T(z) = \tau - h \sqrt{1 - z/\rho} + g(1 - z/\rho) + O(1 - z/\rho)^{3/2} \]
- \[ T_n \sim \frac{h}{2\sqrt{\pi}} \rho^{-n} n^{-3/2} \]
- derivative \[ T'(z) = \frac{h}{2\rho \sqrt{1 - z/\rho}} + \ldots \]
- \[ T'_n \sim n T_n / \rho \]
- bivariate g.f. \[ T(z, u) = \sum T_{n,k} u^k z^n \]

### Singular Boltzmann free generation

- \( \forall A \in T \), \( \Pr_{\rho}(A) = \frac{\rho^{|A|}}{T(\rho)} \) \( \mathbb{E}(A) \) infinite
- parameter \( \Omega : T \to \mathbb{N} \)
- \[ \Pr(\Omega = k) = \sum_n \frac{T_{n,k}}{T_n} \times \frac{T_n \rho^n}{T(\rho)} = \frac{[u^k]T(\rho, u)}{T(\rho)} \]
Atomic probability of trees

**Proposition**

\( T \) simple family of trees; \( T(z) = z\phi(T(z)) \), \( Rdc \rho \).

The distribution of subtrees in a uniform random tree of size \( n \) is asymptotically equivalent to the atomic distribution of trees in a singular Boltzmann sampling:

\[
\forall A \in T, \quad \frac{1}{nT_n} \times \text{occ}(A, T_n) \xrightarrow{n \to \infty} \text{Pr}(A, \Gamma_{\rho})
\]

Proof.

\[ \text{Occ}_A(z) = \sum \text{occ}(A, T) z^{|T|} = \sum \text{occ}(A, T_n) z^n = \frac{zT'(z)}{T(z)} z^{|A|} \]

Thus the probability that \( A \) appears as a subtree in a tree of size \( n \) is

\[
\text{Pr}(A, n) = \frac{1}{nT_n} [z^n] \frac{zT'(z)}{T(z)} z^{|A|} \xrightarrow{n \to \infty} \frac{\rho^{|A|}}{T(\rho)}
\]
Atomic probability of trees

Proposition

\( T \) simple family of trees; \( T(z) = z\phi(T(z)), \text{Rdc } \rho. \)

The distribution of subtrees in a uniform random tree of size \( n \) is asymptotically equivalent to the atomic distribution of trees in a singular Boltzmann sampling:

\[
\forall A \in \mathcal{T}, \quad \frac{1}{nT_n} \times \text{occ}(A, \mathcal{T}_n) \rightarrow_{n \to \infty} \Pr(A, \Gamma_\rho)
\]

Proof.

\( \text{Occ}_A(z) \equiv \sum \text{occ}(A, T) z^{|T|} = \sum \text{occ}(A, \mathcal{T}_n) z^n = \frac{zT'(z)}{T(z)} z^{|A|} \)

Thus the probability that \( A \) appears as a subtrees in a tree of size \( n \) is

\[
\Pr(A, n) = \frac{1}{nT_n} [z^n] \frac{zT'(z)}{T(z)} z^{|A|} \rightarrow_{n \to \infty} \frac{\rho^{|A|}}{T(\rho)}
\]
properties of subtrees

\[ \text{Occ}_A(z) = \frac{zT'(z)}{T(z)} |A| \]
Parameters of trees

Corollary

\(T\) simple family of trees, \(\Omega: T \rightarrow \mathbb{N}\) parameter.

Distribution of \(\Omega\) on subtrees of a uniform random tree size \(n\) asymptotically equivalent to distrib. of \(\Omega\) on trees generated by singular Boltzmann sampler.

Proof.

\[
\Lambda_k(z) = \sum_{A;\Omega(A)=k} z^{|A|} = \sum_n T_{n,k}z^n = [u^k]T(z,u), \quad \Lambda_k(z) \xrightarrow{z\to\rho} \Lambda_k(\rho) < \infty
\]

\[
S\Lambda_k(z) = \sum_{0cc(\Omega_k,T)} z^{|T|} = \sum S_{n,k}z^n = \frac{zT'(z)}{T(z)}\Lambda_k(z)
\]

Probability random subtree parameter \(k\), in a random tree of size \(n\):

\[
\Pr(\Omega_k, n) = \frac{1}{nT_n}[z^n]S\Lambda_k(z) \xrightarrow{n\to\infty} \frac{S\Lambda_k(\rho)}{T(\rho)}
\]
**Corollary**

A simple family of trees, $\Omega: T \to \mathbb{N}$ parameter.

Distribution of $\Omega$ on subtrees of a uniform random tree size $n$ asymptotically equivalent to distrib. of $\Omega$ on trees generated by singular Boltzmann sampler.

Proof.

$$\Lambda_k(z) = \sum_{A; \Omega(A) = k} z^{|A|} = \sum_n T_{n,k} z^n = [u^k] T(z, u), \quad \Lambda_k(z) \rightarrow z \rightarrow \rho \Lambda_k(\rho) < \infty$$

$$S\Lambda_k(z) = \sum_{0cc(\Omega_k, T)} z^{|T|} = \sum S_{n,k} z^n = \frac{zT'(z)}{T(z)} \Lambda_k(z)$$

Probability random subtree parameter $k$, in a random tree of size $n$ :

$$\Pr(\Omega_k, n) = \frac{1}{nT_n} [z^n] S\Lambda_k(z) \rightarrow_{n \to \infty} \frac{S\Lambda_k(\rho)}{T(\rho)}$$
Random Appollonian Networks Structures $\leftrightarrow$ Ternary trees

Degree of a vertex in a RANS $\leftrightarrow$ size of binary subtrees
Degrees in RANS

Choose a random vertex in a random RANS of size $n$; degree?
Choose a random node in a random TT of size $n$; size of BT?

Theorem

The limiting distribution of degrees in RANS follows a power law with an exponential cutoff: $\Pr(D = k) \sim C \beta^k k^{-3/2}, \beta = \frac{8}{9}

Simulation

- Root degree in 100,000 trees generated with free Singular Boltzmann
- Root degree in all subtrees of a uniform tree of size 100,000