Local limits of random trees

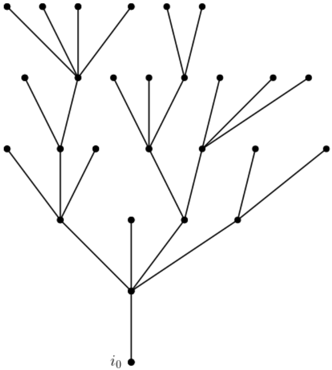
Meltem Ünel, August 2022

Journées de Rentrée des Masters

Overview

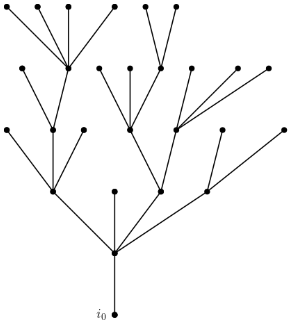
1. Preliminaries
   - Trees: roots, labels and planarity
   - Metric space of planar rooted trees
2. Limits
   - Local vs Scalar: the idea
   - (Weak) convergence of probability measures
   - Generating functionology and transfers
   - UIPT: what does it look like?
3. Horizons?

Trees



- Connected graphs without cycles,
- Fundamental objects in graph theory, combinatorics and probability,
- ...also for data structures and algorithms in computer science,
- Physicists like them for multiple reasons!

Construction: a basic classification



1. **Rooted** / unrooted
   - A vertex is marked,
   - can be described by generations,
   - possibility to construct them recursively.

2. **Planar** / non-planar
   - Trees **are** planar graphs BUT might have different embeddings!
   - Planar means count all embeddings $\rightarrow$

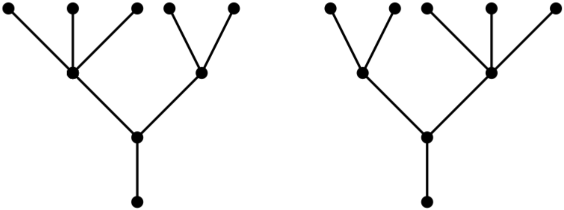
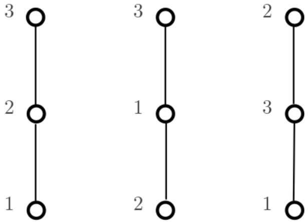


Figure: Two different embeddings of a tree

3. Labelled / **unlabelled** There is much lattitude in choosing labels. Here is a simple example:



Attention!

When we say **rooted planar tree**, we mean a rooted tree where the children of each vertex is ordered from left to right. This induces a **natural embedding** into the half-plane.

From now on, these are the trees we are concerned with.

Recall: *We want to do probability!*

Let's start by constructing the metric space of rooted planar trees...

Metric space of rooted planar trees

- $\mathcal{T}_{\text{fin}} = \bigcup_{N=1}^{\infty} \mathcal{T}_N \quad , \quad \mathcal{T} = \mathcal{T}_{\text{fin}} \cup \mathcal{T}_{\infty}$,
- Denote the vertex set at graph distance s from the root by $D_s(T)$,
- Subgraph $B_R(T)$ of T spanned by $V(B_R(T)) = \bigcup_{s=0}^{R} D_s(T)$,
- For $T, T' \in \mathcal{T}$

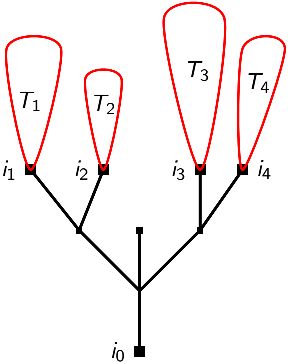
$$\text{dist}(T, T') = \inf\{\frac{1}{R} \mid R \in \mathbb{N},\, B_R(T) = B_R(T')\}\,,$$

- $(\mathcal{T}, \text{dist})$ is separable and complete.

Metric space of rooted planar trees

The ball of radius $\frac{1}{R}$ around $T_0 \in \mathcal{T}$

$$\mathcal{B}_{\frac{1}{R}}(T_0) = \{ T \in \mathcal{T} \mid B_R(T) = B_R(T_0) \} .$$



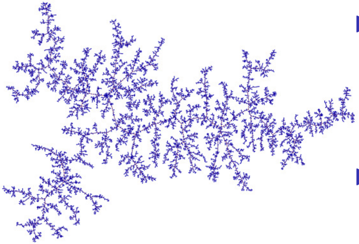
Local vs Scaling

Given some probabilistic structure on $\mathcal{T}_N$, we are interested in the asymptotic behavior of a "typical element" of $\mathcal{T}_N$ as $N \to \infty$.

In other words: We pick a tree *uniformly at random* in $\mathcal{T}_N$ and study its *scaling* or *local* limit.

• Scaling Limit

Choose an object T at random inside $\mathcal{T}_N$. When suitably scaled, T approaches a continuous structure as $N \to \infty$.



- In other words

$$\frac{T_N}{2\sqrt{N}} \xrightarrow[N\to\infty]{(d)} \mathcal{T}_{CRT},$$

- Check [arXiv:math/0511515] by Le Gall to learn more about it!

Local vs Scaling

• Local Limit

Different in spirit!

- No scaling of distances,
- Study the convergence of arbitrarily large but **finite** neighborhoods of the root,
- The limiting object is still discrete!

This limit provides more **local** information: limiting behavior of the degrees of vertices, average volume of balls around the root...

This is the limit we will concentrate on in the rest of the talk!

Convergence of probability measures

A sequence μ_N converges weakly to μ on $\mathcal{T}$

$$\int_{\mathcal{T}} f d\mu_N \to \int_{\mathcal{T}} f d\mu$$

as $N \to \infty$ for all bounded continuous functions f on $\mathcal{T}$.

Some facts:

- Any ball in $\mathcal{T}$ is both open and closed,
- Two balls are either disjoint or one is contained in the other,
- $\mathcal{T}_{\text{fin}}$ is a countable dense subset of $\mathcal{T}$.

Our algoritm

Lemma (See Theorems 2.1, 2.2 and 6.5 in [Billingsley])

Suppose μ_N, $N \in \mathbb{N}$ *is a sequence of probability measures on a metric space* M *and* U *is a family of both open and closed subsets of* M *such that*

1. *any finite intersection of sets in* U *belongs to* U,
2. *any open subset of* M *may be written as a finite or countable union of sets from* U,
3. *the sequence* $\mu_N(A)$ *is convergent for all* $A \in U$.

Then the sequence μ_N, $N \in \mathbb{N}$, *is weakly convergent provided it is tight.*

The tightness means that for each $\epsilon > 0$ there exists a compact subset C such that

$$\mu_N(M \setminus C) < \epsilon \quad \text{for all } N.$$

Our algorithm

The set $U = \{\mathcal{B}_{\frac{1}{R}}(T) \mid R \in \mathbb{N}, T \in \mathcal{T}_{\text{fin}}\}$ satisfies (1) and (2) above.
So, we have an algorithm:

1. Show tighness,

2. Show convergence on balls.

Recall: we want to start with the **uniform** distribution on $\mathcal{T}_N$:

$$\mu_N(T) = \frac{1}{C_N}, \quad T \in \mathcal{T}_N\,.$$

To be able to calculate the mass of balls $\mu_N(\mathcal{B}_{\frac{1}{R}}(T))$, we need to understand the asymptotic behavior of C_N, for N large.

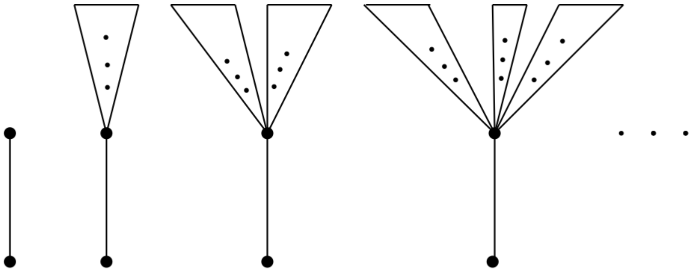
Generating functionology

Recall: we want asymptotics of C_N. Let's encode it!

$$Z(g) = \sum_{N=1}^{\infty} C_N g^N \quad , \quad g \in \mathbb{C} \, .$$

Hence we map a *combinatorial* problem onto an *analytic* one!

- *A tree is a root with a sequence of trees grafted on it.*



Generating functionology

$$Z(g) = g(1 + Z(g) + Z(g)^2 + \dots) = \frac{g}{1 - Z(g)}$$

Solution is simple

$$Z(g) = \frac{1 - \sqrt{1 - 4g}}{2} .$$

Notation: $C_N = [g^N]Z(g) = A^N \theta(N)$.

Principles of Coefficient Asymptotics:

- The *location* of a function's singularities dictates the exponential growth of its coefficients A^N,
- The *nature* of a function's singularities determines the associated **sub**exponential factor $\theta(N)$.

A simple transfer

Theorem (Standard function scale, Theorem VI.1 in [AC])

Let α be an arbitrary complex number in $\mathbb{C} \setminus \mathbb{Z}_{\leq 0}$. The coefficient of x^N in

$$Z(x) = (1-x)^{-\alpha}$$

admits for large N a complete asymptotic expansion in descending powers of N

$$[x^N]Z(x) \sim \frac{N^{\alpha-1}}{\Gamma(\alpha)}\left(1 + \mathcal{O}(\frac{1}{N})\right).$$

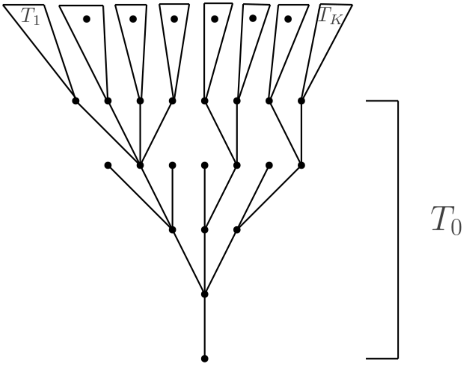
Our case: $\alpha = \frac{1}{2},\ x = 4g\ :$

$$[g^N]Z(g) \sim 4^N N^{-\frac{3}{2}}(1 + \mathcal{O}(\frac{1}{N})).$$

Back to the algorithm

- Tightness: once a suitable compact subset is found, relatively easy.
- Convergence on balls: finer estimations needed.

$$\mu_N(\mathcal{B}_{\frac{1}{R}}(T_0)) = C_N^{-1} \sum_{N_1+\cdots+N_K=N+K-N_0} \prod_{i=1}^{K} C_{N_i} .$$



Limiting measure μ

After a couple of pages of calculation, we get

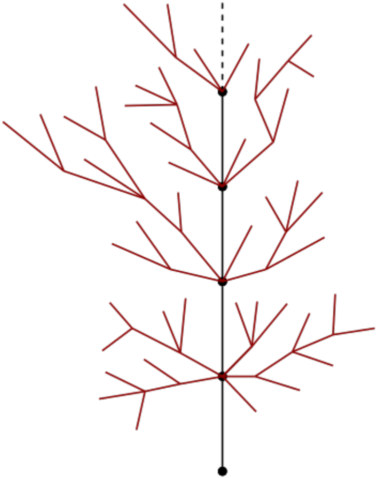
$$\mu_N(\mathcal{B}_{\frac{1}{R}}(T_0)) \xrightarrow{N\to\infty} 2K \cdot 2^K \cdot 4^{-N_0}\,.$$

Hence, we have proven the existence of a limiting measure μ and we also know the mass of balls $\mu(\mathcal{B}_{\frac{1}{R}}(T_0))$!

This measure is supported on $\mathcal{T}_\infty$ and is called **Uniform Infinite Planar Tree** (UIPT).

Let's take a look at its properties...

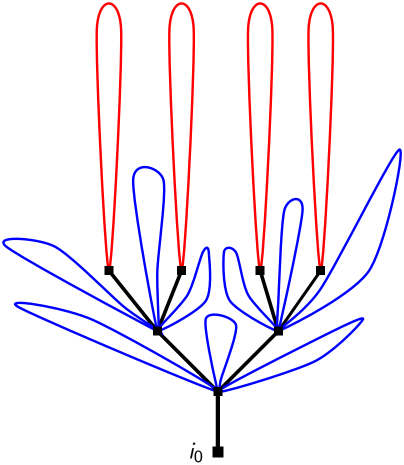
UIPT



- μ is supported on single-spine trees,
- The subbranches are all finite and independently distributed (they are in fact critical BGW trees, if you are familiar with them).

Horizons

What is we start with a sequence of probability measures ν_N that *are not uniform?*



See: [arXiv:2112.06570]

$$\nu_N^{(\lambda)}(T) = \frac{e^{-\lambda h(T)}}{W_N}, \ \lambda \in \mathbb{R},$$

$$W_N := \sum_{T \in \mathcal{T}_N} e^{-\lambda h(T)} .$$

$$\nu_N^{(\lambda)} \to \nu^{(\lambda)}, \ N \to \infty,$$

$\lambda > 0 \Rightarrow$ a new limit!

Merci et bon courage!

Bibliography I

P. Billingsley.
Convergence of probability measures.
John Wiley & Sons, 2013.

B. Durhuus.
Probabilistic aspects of infinite trees and surfaces.
Acta Physica Polonica Series B, 34(10):4795–4812, 2003.

P. Flajolet and R. Sedgewick.
Analytic combinatorics.
Cambridge University press, 2009.

J. F. Le Gall.
Random trees and applications.
Probability surveys, 2:245–311, 2005.