Journées ALEA, CIRM, 21st-25th March 2022

Scaling limits of random trees and graphs

Christina Goldschmidt

https://www.stats.ox.ac.uk/~goldschm/ALEAminicourse.html



1. INTRODUCTION: BINARY TREES



Binary trees

- Let \mathbb{T}_n be the set of planted plane binary trees with *n* leaves.
- ▶ Note that every element of \mathbb{T}_n has n-1 internal vertices and 2n-1 edges.



Binary trees

- Let \mathbb{T}_n be the set of planted plane binary trees with *n* leaves.
- ▶ Note that every element of \mathbb{T}_n has n-1 internal vertices and 2n-1 edges.
- Let T^{*}_n be the set of planted plane binary leaf-labelled trees with n labelled leaves.
- ► The root, labelled 0 is, by convention, not a leaf.



 \mathbb{T}_n is the set of planted plane binary trees with *n* leaves. \mathbb{T}_n^* is the set of planted plane binary leaf-labelled trees with *n* labelled leaves.

$$|\mathbb{T}_n| = rac{1}{n} inom{2n-2}{n-1}$$
 (Catalan numbers), $|\mathbb{T}_n^*| = n! |\mathbb{T}_n|.$

$$|\mathbb{T}_n| = rac{1}{n} inom{2n-2}{n-1} \sim rac{4^{n-1}}{n^{3/2}\sqrt{\pi}} \quad ext{as } n o \infty.$$

$$|\mathbb{T}_n| = rac{1}{n} inom{2n-2}{n-1} \sim rac{4^{n-1}}{n^{3/2}\sqrt{\pi}} \quad ext{as } n o \infty.$$

Our first object of interest is a uniform random element of \mathbb{T}_n .

$$|\mathbb{T}_n| = \frac{1}{n} \binom{2n-2}{n-1} \sim \frac{4^{n-1}}{n^{3/2}\sqrt{\pi}} \quad \text{as } n \to \infty.$$

Our first object of interest is a uniform random element of \mathbb{T}_n . It will be more convenient to work with \mathbb{T}_n^* and then ignore the leaf labels.

$$|\mathbb{T}_n| = \frac{1}{n} \binom{2n-2}{n-1} \sim \frac{4^{n-1}}{n^{3/2}\sqrt{\pi}} \quad \text{as } n \to \infty.$$

Our first object of interest is a uniform random element of \mathbb{T}_n . It will be more convenient to work with \mathbb{T}_n^* and then ignore the leaf labels.

Rémy's algorithm recursively constructs a sequence $(T_n)_{n\geq 1}$ of trees such that T_n is uniform on \mathbb{T}_n^* for each n.

- Start from a single edge with endpoints labelled 0 and 1.
- ► At step n ≥ 2, pick an edge uniformly at random, divide it into two edges, insert a new vertex in the middle and attach to that vertex a new edge with a leaf labelled n at its other end, chosen to point in one of the two possible directions each with probability 1/2.















Claim: for each *n*, T_n is a uniform element of \mathbb{T}_n^* .



[J.-L. Rémy, Un procédé itératif de dénombrement d'arbres binaires et son application à leur génération aléatoire, RAIRO. Informatique théorique 19:2 (1985), pp.179–195]

Taking limits

Vague question: what can we say about T_n as $n \to \infty$?

Taking limits

Vague question: what can we say about T_n as $n \to \infty$?

Concrete first question: as $n \to \infty$, how does the distance between 0 and 1 behave?

The total number of edges present at step n is equal to 2n - 1.

Consider the number of edges in the path between 0 and 1:

- If we add our new leaf somewhere along that path, it gets longer by 1.
- If we add our new leaf anywhere else, the length of the path remains the same.

We have an urn process with two colours, say black and white, where each black ball represents an edge in the path between 0 and 1, and each white ball represents an edge elsewhere.

When we pick a black ball, we replace it in the urn together with one black and one white ball.

When we pick a white ball, we replace it in the urn together with two new white balls.

We start with a single black ball. At step n, we always have 2n - 1 balls present.



Let B_n be the number of black balls at step n.

Let B_n be the number of black balls at step n.

We have $B_1 = 1$.

For $n \geq 1$,

$$\mathbb{E}[B_{n+1}|\mathcal{F}_n] = \frac{B_n}{2n-1}(B_n+1) + \frac{2n-1-B_n}{2n-1}B_n$$

Let B_n be the number of black balls at step n.

We have $B_1 = 1$.

For $n \geq 1$,

$$\mathbb{E}[B_{n+1}|\mathcal{F}_n] = \frac{B_n}{2n-1}(B_n+1) + \frac{2n-1-B_n}{2n-1}B_n = \frac{2n}{2n-1}B_n.$$

Let B_n be the number of black balls at step n.

We have $B_1 = 1$.

For $n \geq 1$,

$$\mathbb{E}[B_{n+1}|\mathcal{F}_n] = \frac{B_n}{2n-1}(B_n+1) + \frac{2n-1-B_n}{2n-1}B_n = \frac{2n}{2n-1}B_n.$$

Define a sequence by $b_1 = 1$ and $b_{n+1} = \frac{2^{2n}(n!)^2}{(2n)!}$ for $n \ge 1$. Then

$$b_{n+1}=\frac{2n}{2n-1}b_n$$

Let B_n be the number of black balls at step n.

We have $B_1 = 1$.

For $n \geq 1$,

$$\mathbb{E}[B_{n+1}|\mathcal{F}_n] = \frac{B_n}{2n-1}(B_n+1) + \frac{2n-1-B_n}{2n-1}B_n = \frac{2n}{2n-1}B_n.$$

Define a sequence by $b_1 = 1$ and $b_{n+1} = \frac{2^{2n}(n!)^2}{(2n)!}$ for $n \ge 1$. Then

$$b_{n+1}=\frac{2n}{2n-1}b_n$$

Then we have that

$$\left(\frac{B_n}{b_n}\right)_{n\geq 1}$$
 is a non-negative martingale.

Martingale limit

 $(B_n/b_n)_{n\geq 1}$ is also bounded in L^2 , hence uniformly integrable, and so it has an almost sure limit by the martingale convergence theorem.

Since

$$b_{n+1} = \frac{2^{2n}(n!)^2}{(2n)!} \sim \sqrt{\pi n},$$

we get that

$$rac{B_n}{\sqrt{2n}}
ightarrow L$$
 a.s. as $n
ightarrow \infty$

for some limit random variable L.

[P. Marchal, A note on the fragmentation of the stable tree, Fifth Colloquium on Mathematics and Computer Science, DMTCS (2008), pp.489–500]

Limiting distribution for the length

It also turns out (using a generating function argument) that the law of B_{n+1} is explicit:

$$\mathbb{P}(B_{n+1} = k) = \frac{k-1}{n} 2^{k-1} \frac{\binom{2n-k}{n-1}}{\binom{2n}{n}}$$

and so

$$\mathbb{P}\left(B_{n+1}=\lfloor x\sqrt{2n}\rfloor\right)\sim \frac{x}{\sqrt{2n}}e^{-x^2/2}, \quad x>0.$$

Limiting distribution for the length

It also turns out (using a generating function argument) that the law of B_{n+1} is explicit:

$$\mathbb{P}(B_{n+1}=k) = \frac{k-1}{n} 2^{k-1} \frac{\binom{2n-k}{n-1}}{\binom{2n}{n}}$$

and so

$$\mathbb{P}\left(B_{n+1}=\lfloor x\sqrt{2n}\rfloor\right)\sim \frac{x}{\sqrt{2n}}e^{-x^2/2},\quad x>0.$$

In other words, we get

$$rac{B_n}{\sqrt{2n}}
ightarrow L$$
 a.s. as $n
ightarrow \infty$,

where the limit *L* has the Rayleigh distribution, with density $xe^{-x^2/2}$ on \mathbb{R}_+ .

[P. Flajolet, P. Dumas and V. Puyhaubert, Some exactly solvable models of urn process theory, Fourth Colloquium on Mathematics and Computer Science: Algorithms, Trees, Combinatorics and Probabilities, DMTCS (2006), pp.59–118]

Consequences

The distance between 0 and 1 varies as $\sqrt{2n}$, with a nice almost sure limit. What can we say about the distances between the other leaves as $n \to \infty$?

Consequences

The distance between 0 and 1 varies as $\sqrt{2n}$, with a nice almost sure limit. What can we say about the distances between the other leaves as $n \to \infty$?

For example, let's think about the distance from 2 to the path between 0 and 1, and the position along that path at which it branches off.



At step 2 of Rémy's algorithm, we always have



At step 2 of Rémy's algorithm, we always have



Each of the three parts here behaves precisely as a little copy of Rémy's algorithm, although the numbers of leaves we add to each copy are dependent.

Each of the three parts here behaves precisely as a little copy of Rémy's algorithm, although the numbers of leaves we add to each copy are dependent.

A useful consequence is that given the three sets of leaves, these three trees are themselves uniform binary plane trees.



The numbers of vertices in each of the three little trees evolve according to a variant of Pólya's urn with three colours, red, green and blue. We start with one ball of each colour. We pick a ball at random and replace it in the urn with **two** more of the same colour. Let R_n , G_n , B_n be the numbers of red, green and blue balls respectively at step n.


It is then standard that

$$\frac{1}{2n+3}(R_n,G_n,B_n)\to (\Delta_1,\Delta_2,\Delta_3) \quad \text{a.s. as } n\to\infty,$$

where $(\Delta_1, \Delta_2, \Delta_3) \sim \text{Dirichlet}(1/2, 1/2, 1/2).$

It is then standard that

$$rac{1}{2n+3}(R_n,G_n,B_n)
ightarrow (\Delta_1,\Delta_2,\Delta_3)$$
 a.s. as $n
ightarrow\infty$,

where $(\Delta_1, \Delta_2, \Delta_3) \sim \text{Dirichlet}(1/2, 1/2, 1/2).$

The Dirichlet distribution with parameters $\alpha_1,\alpha_2,\ldots,\alpha_k>0$ has density

$$\frac{\Gamma(\sum_{i=1}^{k} \alpha_i)}{\prod_{i=1}^{k} \Gamma(\alpha_i)} x_1^{\alpha_1 - 1} \dots x_k^{\alpha_k - 1}$$

with respect to Lebesgue measure on

$$\left\{\mathbf{x}=(x_1,\ldots,x_k)\in\mathbb{R}_+^k:\sum_{i=1}^kx_i=1\right\}.$$

The numbers of leaves in each of the three subtrees are given by

$$N_n^R = (R_n + 1)/2, \quad N_n^G = (G_n + 1)/2, \quad N_n^B = (B_n + 1)/2.$$

So we have

$$\frac{1}{n}(N_n^R,N_n^G,N_n^B) \to (\Delta_1,\Delta_2,\Delta_3) \quad \text{a.s.}$$

The numbers of leaves in each of the three subtrees are given by

$$N_n^R = (R_n + 1)/2, \quad N_n^G = (G_n + 1)/2, \quad N_n^B = (B_n + 1)/2.$$

So we have

$$\frac{1}{n}(N_n^R,N_n^G,N_n^B) \to (\Delta_1,\Delta_2,\Delta_3) \quad \text{a.s.}$$

Writing L_n^R , L_n^G , L_n^B for the lengths of the three paths at step *n*, we see that they look like small copies of the first urn model run for numbers of steps which are approximately $n\Delta_1$, $n\Delta_2$ and $n\Delta_3$.

The numbers of leaves in each of the three subtrees are given by

$$N_n^R = (R_n + 1)/2, \quad N_n^G = (G_n + 1)/2, \quad N_n^B = (B_n + 1)/2.$$

So we have

$$\frac{1}{n}(N_n^R,N_n^G,N_n^B) \to (\Delta_1,\Delta_2,\Delta_3) \quad \text{a.s.}$$

Writing L_n^R , L_n^G , L_n^B for the lengths of the three paths at step *n*, we see that they look like small copies of the first urn model run for numbers of steps which are approximately $n\Delta_1$, $n\Delta_2$ and $n\Delta_3$. It follows that

$$\frac{1}{\sqrt{2n}}(L_n^R, L_n^G, L_n^B) \to (\sqrt{\Delta_1}L_1, \sqrt{\Delta_2}L_2, \sqrt{\Delta_3}L_3) \quad \text{a.s.}$$

where L_1, L_2, L_3 are i.i.d. Rayleigh random variables, independent of $(\Delta_1, \Delta_2, \Delta_3)$.

Limiting subtree lengths

An elementary distributional calculation yields that

$$(\sqrt{\Delta_1}L_1, \sqrt{\Delta_2}L_2, \sqrt{\Delta_3}L_3) \stackrel{d}{=} \sqrt{\Gamma_2} \times \text{Dir}(1, 1, 1),$$

where $\Gamma_2 \sim \text{Gamma}(2, 1/2)$ and the two factors are independent.

Limiting subtree lengths

An elementary distributional calculation yields that

$$(\sqrt{\Delta_1}L_1, \sqrt{\Delta_2}L_2, \sqrt{\Delta_3}L_3) \stackrel{d}{=} \sqrt{\Gamma_2} \times \text{Dir}(1, 1, 1),$$

where $\Gamma_2 \sim \text{Gamma}(2, 1/2)$ and the two factors are independent.

More generally, if we consider the subtree spanned by 0 and the leaves labelled $1, 2, \ldots, k$, we get 2k - 1 edges whose lengths are distributed as

$$\sqrt{\Gamma_k} \times \operatorname{Dir}(\underbrace{1,1,\ldots,1}_{2k-1}),$$

where again $\Gamma_k \sim \text{Gamma}(k, 1/2)$ and the two factors are independent.

Limiting subtree lengths

An elementary distributional calculation yields that

$$(\sqrt{\Delta_1}L_1, \sqrt{\Delta_2}L_2, \sqrt{\Delta_3}L_3) \stackrel{d}{=} \sqrt{\Gamma_2} \times \text{Dir}(1, 1, 1),$$

where $\Gamma_2 \sim \text{Gamma}(2, 1/2)$ and the two factors are independent.

More generally, if we consider the subtree spanned by 0 and the leaves labelled $1, 2, \ldots, k$, we get 2k - 1 edges whose lengths are distributed as

$$\sqrt{\Gamma_k} \times \operatorname{Dir}(\underbrace{1, 1, \dots, 1}_{2k-1}),$$

where again $\Gamma_k \sim \text{Gamma}(k, 1/2)$ and the two factors are independent.

(Note that the k = 1 case fits into this pattern, since Rayleigh $\stackrel{d}{=} \sqrt{\Gamma_1}$.)

Take an inhomogeneous Poisson process on \mathbb{R}_+ of intensity *t* at *t*.



Take an inhomogeneous Poisson process on \mathbb{R}_+ of intensity t at t.



A useful way of constructing this is to let $E_1, E_2, ...$ be i.i.d. Exp(1/2) and set $C_i = \sqrt{\sum_{j=1}^{i} E_j}$.

Take an inhomogeneous Poisson process on \mathbb{R}_+ of intensity t at t.



A useful way of constructing this is to let $E_1, E_2, ...$ be i.i.d. Exp(1/2) and set $C_i = \sqrt{\sum_{j=1}^{i} E_j}$.

• Consider the line-segments $[0, C_1)$, $[C_1, C_2)$,....

Take an inhomogeneous Poisson process on \mathbb{R}_+ of intensity t at t.



A useful way of constructing this is to let $E_1, E_2, ...$ be i.i.d. Exp(1/2) and set $C_i = \sqrt{\sum_{j=1}^{i} E_j}$.

• Consider the line-segments $[0, C_1)$, $[C_1, C_2)$,

▶ Start from [0, *C*₁) and proceed inductively.

Take an inhomogeneous Poisson process on \mathbb{R}_+ of intensity t at t.



A useful way of constructing this is to let $E_1, E_2, ...$ be i.i.d. Exp(1/2) and set $C_i = \sqrt{\sum_{j=1}^{i} E_j}$.

- Consider the line-segments $[0, C_1)$, $[C_1, C_2)$,
- ▶ Start from [0, *C*₁) and proceed inductively.
- For i ≥ 2, attach [C_{i−1}, C_i) at a random point chosen uniformly over the existing tree.















Why is this the right limit?

Claim: this gives the almost sure limit of the subtree spanned by 0 and the leaves 1, 2, ..., k in the rescaled version of Rémy's algorithm.

• The tree at step $k \ge 1$ has total length

$$C_k = \sqrt{\sum_{i=1}^k E_i} \stackrel{d}{=} \sqrt{\Gamma_k},$$

where $\Gamma_k \sim \text{Gamma}(k, 1/2)$.

- The combinatorics of the attachment mechanism are exactly the same as in Rémy's algorithm – so the underlying binary leaf-labelled tree has the right distribution.
- ► A calculation shows that the cut-points and attachment points split up the interval [0, C_k) uniformly.

The line-breaking definition of the Brownian CRT

- Start from $[0, C_1)$ and proceed inductively.
- For i ≥ 1, sample B_i uniformly from [0, C_i) and attach [C_i, C_{i+1}) at the corresponding point of the tree constructed so far (this is a point chosen uniformly at random over the existing tree).

The line-breaking definition of the Brownian CRT

- Start from $[0, C_1)$ and proceed inductively.
- For i ≥ 1, sample B_i uniformly from [0, C_i) and attach [C_i, C_{i+1}) at the corresponding point of the tree constructed so far (this is a point chosen uniformly at random over the existing tree).

Now take the union of all the branches, thought of as a path metric space, and then take its completion.

This procedure gives (somewhat informally expressed) definition of the Brownian continuum random tree (CRT) which is a key object in this minicourse.

The line-breaking definition of the Brownian CRT



[Picture by Igor Kortchemski]

The scaling limit of the uniform binary plane tree

In the next section, we will make sense of the following statement.

Theorem. (Marchal (2003), Curien and Haas (2013)) As $n \to \infty$, $\frac{1}{\sqrt{2n}}T_n \to \mathcal{T}$ a.s.

where \mathcal{T} is the Brownian CRT.

The scaling limit of the uniform binary plane tree

In the next section, we will make sense of the following statement.

Theorem. (Marchal (2003), Curien and Haas (2013)) As $n \to \infty$, 1 $T \to T$ as

$$rac{1}{\sqrt{2n}}T_n o \mathcal{T}$$
 a.s.

where \mathcal{T} is the Brownian CRT.

We need to know what sort of objects we're really dealing with, and what is the topology in which the convergence occurs. (Also: why is \mathcal{T} Brownian?!)

Before we do that, let's record an immediate consequence of the theorem.

Self-similarity of the Brownian CRT

Recall: we split our uniform binary plane tree into three little uniform binary plane trees of random sizes.



This property passes to the limit, and so the Brownian CRT can be split into three randomly rescaled Brownian CRTs. In particular, the Brownian CRT is a random fractal.

2. **R-TREES AND CONVERGENCE**

Key reference:

Jean-François Le Gall, **Random trees and applications**, *Probability Surveys* **2** (2005) pp.245-311.



We want a continuous notion of a tree. We don't really care about vertices: the important aspects are the shape of the tree and the distances. So it makes sense to think in terms of metric spaces.

\mathbb{R} -trees

Definition. A compact metric space (\mathcal{T}, d) is an \mathbb{R} -tree if for all $x, y \in \mathcal{T}$,

► There exists a unique shortest path [[x, y]] from x to y (of length d(x, y)).

▶ The only non-self-intersecting path from x to y is [[x, y]].

\mathbb{R} -trees

Definition. A compact metric space (\mathcal{T}, d) is an \mathbb{R} -tree if for all $x, y \in \mathcal{T}$,

- ► There exists a unique shortest path [[x, y]] from x to y (of length d(x, y)). (There is a unique isometric map f_{x,y} from [0, d(x, y)] into T such that f(0) = x and f(d(x, y)) = y. We write f_{x,y}([0, d(x, y)]) = [[x, y]].)
- ► The only non-self-intersecting path from x to y is [[x, y]]. (If g is a continuous injective map from [0, 1] into T, such that g(0) = x and g(1) = y, then g([0, 1]) = [[x, y]].)

\mathbb{R} -trees

Definition. A compact metric space (\mathcal{T}, d) is an \mathbb{R} -tree if for all $x, y \in \mathcal{T}$,

- ► There exists a unique shortest path [[x, y]] from x to y (of length d(x, y)). (There is a unique isometric map f_{x,y} from [0, d(x, y)] into T such that f(0) = x and f(d(x, y)) = y. We write f_{x,y}([0, d(x, y)]) = [[x, y]].)
- ► The only non-self-intersecting path from x to y is [[x, y]]. (If g is a continuous injective map from [0, 1] into T, such that g(0) = x and g(1) = y, then g([0, 1]) = [[x, y]].)

An element $v \in \mathcal{T}$ is called a vertex.

A rooted \mathbb{R} -tree has a distinguished vertex ρ called the root. The height of a vertex v is its distance $d(\rho, v)$ from the root. A leaf is a vertex v such that $v \notin [[\rho, w]]$ for any $w \neq v$.

Coding \mathbb{R} -trees

Let $h: [0,1] \to \mathbb{R}^+$ be an excursion, that is a continuous function such that h(0) = h(1) = 0 and h(x) > 0 for $x \in (0,1)$.



Coding \mathbb{R} -trees

Now put glue on the underside of the excursion and push the two sides together...



Coding \mathbb{R} -trees

Now put glue on the underside of the excursion and push the two sides together...



$\mathsf{Coding}\ \mathbb{R}\text{-}\mathsf{trees}$

Now put glue on the underside of the excursion and push the two sides together...












Formally, use h to define a distance:

$$d_h(x,y) = h(x) + h(y) - 2 \inf_{x \wedge y \le z \le x \lor y} h(z).$$



Coding \mathbb{R} -trees

Let $y \sim y'$ if $d_h(y,y') = 0$ and take the quotient $\mathcal{T}_h = [0,1]/\sim$.



Theorem. For any excursion h, (\mathcal{T}_h, d_h) is an \mathbb{R} -tree.

Write $\pi_h : [0,1] \to \mathcal{T}_h$ for the projection map.

We will often root \mathcal{T}_h at $\rho = \pi_h(0) = \pi_h(1)$.

We will want to be able to sample random points in our trees. There is a natural "uniform" measure μ_h which is the push-forward of the Lebesgue measure on [0, 1] onto \mathcal{T}_h .

We will want to be able to sample random points in our trees. There is a natural "uniform" measure μ_h which is the push-forward of the Lebesgue measure on [0, 1] onto \mathcal{T}_h .

To pick a point of \mathcal{T}_h according to μ_h , we simply sample $U \sim U[0, 1]$ and then take our point to be $\pi_h(U)$.

We will want to be able to sample random points in our trees. There is a natural "uniform" measure μ_h which is the push-forward of the Lebesgue measure on [0, 1] onto \mathcal{T}_h .

To pick a point of \mathcal{T}_h according to μ_h , we simply sample $U \sim U[0, 1]$ and then take our point to be $\pi_h(U)$.

We will typically think of our continuous trees as triples $(\mathcal{T}_h, d_h, \mu_h)$.

Let \mathbb{M} be the space of compact metric spaces endowed with a Borel probability measure, up to measure-preserving isometry.

We will define a metric d_{GHP} , the Gromov-Hausdorff-Prokhorov distance on \mathbb{M} .

Suppose that (X, d) and (X', d') are compact metric spaces.



Suppose that (X, d) and (X', d') are compact metric spaces.

A correspondence *R* is a subset of $X \times X'$ such that for every $x \in X$, there exists $x' \in X'$ with $(x, x') \in R$ and vice versa.



Suppose that (X, d) and (X', d') are compact metric spaces.

A correspondence *R* is a subset of $X \times X'$ such that for every $x \in X$, there exists $x' \in X'$ with $(x, x') \in R$ and vice versa.



Topological considerations The distortion of R is

$$dis(R) = \sup\{|d(x,y) - d'(x',y')| : (x,x'), (y,y') \in R\}.$$



Suppose that μ is a Borel probability measure on (X, d) and that μ' is a Borel probability measure on (X', d').

A measure ν on $X \times X'$ is a coupling of μ and μ' if $\nu(\cdot, X') = \mu(\cdot)$ and $\nu(X, \cdot) = \mu'(\cdot)$.

Suppose that μ is a Borel probability measure on (X, d) and that μ' is a Borel probability measure on (X', d').

A measure ν on $X \times X'$ is a coupling of μ and μ' if $\nu(\cdot, X') = \mu(\cdot)$ and $\nu(X, \cdot) = \mu'(\cdot)$.

Idea: find a correspondence and a coupling such that the correspondence has small distortion and the coupling "lines up" well with the correspondence i.e. if $(V, V') \sim \nu$ then $\mathbb{P}((V, V') \in R) = \nu(R)$ is close to 1.

The Gromov-Hausdorff-Prokhorov distance between (X, d, μ) and (X', d', μ') is defined to be

$$\mathsf{d}_{\mathsf{GHP}}((X,d,\mu),(X',d',\mu')) = \frac{1}{2}\inf_{R,\nu} \mathsf{max}\{\mathsf{dis}(R),\nu(R^c)\}.$$

The Gromov-Hausdorff-Prokhorov distance between (X, d, μ) and (X', d', μ') is defined to be

$$\mathsf{d}_{\mathsf{GHP}}((X,d,\mu),(X',d',\mu')) = \frac{1}{2}\inf_{R,\nu}\max\{\mathsf{dis}(R),\nu(R^c)\}.$$

Theorem. $(\mathbb{M}, \mathsf{d}_{\mathsf{GHP}})$ is a complete separable metric space.

[S. Evans, J. Pitman and A. Winter, Rayleigh processes, real trees, and root growth with re-grafting, *Probability Theory and Related Fields* 134 (2006) pp.81-126.]

[R. Abraham, J.-F. Delmas and P. Hoscheit, A note on the Gromov-Hausdorff-Prokhorov distance between (locally) compact metric measure spaces, *Electronic Journal of Probability* 18 (2013), no. 14.]

The Brownian CRT

Definition. The Brownian continuum random tree is $(\mathcal{T}_{2e}, d_{2e}, \mu_{2e})$, where *e* is a standard Brownian excursion.



[Pictures by Igor Kortchemski]

A planar ordering



Observe that the excursion comes with more information than the the tree: if s < t and $\pi_{2e}(s)$ and $\pi_{2e}(t)$ are leaves, it is natural to think of $\pi_{2e}(s)$ appearing to the left of $\pi_{2e}(t)$ (c.f. Rémy's algorithm).

Discrete trees as metric spaces

We want to think of $(T_n, n \ge 1)$ as metric spaces.

The vertices of T_n (labelled and unlabelled) come equipped with a natural metric: the graph distance d_n .



We sometimes write aT_n for the metric space (T_n, ad_n) given by the vertices of T_n with the graph distance scaled by a.

Uniform measure



We will endow T_n with μ_n , the measure which puts mass 1/(2n) on each of the 2n vertices.

Convergence

Theorem. As $n \to \infty$,

$$\left(T_n, \frac{d_n}{\sqrt{2n}}, \mu_n\right)
ightarrow \left(\mathcal{T}_{2e}, d_{2e}, \mu_{2e}\right)$$
 a.s.

with respect to the Gromov-Hausdorff-Prokhorov topology.

[P. Marchal, **Constructing a sequence of random walks strongly converging to Brownian motion**, *Discrete Mathematics and Theoretical Computer Science*, 2003, pp.181–190.]

[N. Curien & B. Haas, The stable trees are nested, Probability Theory and Related Fields 157, 2013, pp.847-883.]

There is a well-known bijection between planted binary plane trees with n leaves and lattice excursions with 2n steps.



There is a well-known bijection between planted binary plane trees with n leaves and lattice excursions with 2n steps.



There is a well-known bijection between planted binary plane trees with n leaves and lattice excursions with 2n steps.



There is a well-known bijection between planted binary plane trees with n leaves and lattice excursions with 2n steps.



There is a well-known bijection between planted binary plane trees with n leaves and lattice excursions with 2n steps.



There is a well-known bijection between planted binary plane trees with n leaves and lattice excursions with 2n steps.



There is a well-known bijection between planted binary plane trees with n leaves and lattice excursions with 2n steps.



There is a well-known bijection between planted binary plane trees with n leaves and lattice excursions with 2n steps.



There is a well-known bijection between planted binary plane trees with n leaves and lattice excursions with 2n steps.



There is a well-known bijection between planted binary plane trees with n leaves and lattice excursions with 2n steps.



There is a well-known bijection between planted binary plane trees with n leaves and lattice excursions with 2n steps.


There is a well-known bijection between planted binary plane trees with n leaves and lattice excursions with 2n steps.



There is a well-known bijection between planted binary plane trees with n leaves and lattice excursions with 2n steps.



There is a well-known bijection between planted binary plane trees with n leaves and lattice excursions with 2n steps.



There is a well-known bijection between planted binary plane trees with n leaves and lattice excursions with 2n steps.



There is a well-known bijection between planted binary plane trees with n leaves and lattice excursions with 2n steps.



There is a well-known bijection between planted binary plane trees with n leaves and lattice excursions with 2n steps.



There is a well-known bijection between planted binary plane trees with n leaves and lattice excursions with 2n steps.



To go back the other way, it's easy to recover the tree:



To go back the other way, it's easy to recover the tree:





Since our trees are uniform, so are the lattice excursions. In other words, they are excursions of simple random walk away from 0.



Since our trees are uniform, so are the lattice excursions. In other words, they are excursions of simple random walk away from 0. So (at least in distribution), it's clear that, suitably rescaled, they should converge to a Brownian excursion.

















Let $(E_n)_{n\geq 1}$ be the sequence of lattice excursions.

Theorem. (Marchal (2003)) As $n \to \infty$, we have

$$\frac{1}{\sqrt{2n}}(E_n(\lfloor 2nt \rfloor), 0 \le t \le 1) \to (e(t), 0 \le t \le 1)$$

uniformly on [0, 1], almost surely.

This is not quite enough to conclude that the trees converge in the GHP sense. The embedding of the tree in the excursion distorts distances.



This is not quite enough to conclude that the trees converge in the GHP sense. The embedding of the tree in the excursion distorts distances.



This is not quite enough to conclude that the trees converge in the GHP sense. The embedding of the tree in the excursion distorts distances.



This is not quite enough to conclude that the trees converge in the GHP sense. The embedding of the tree in the excursion distorts distances.



Write $H_n(k)$ for the distance from the root to the vertex visited at time k. Then

$$H_n(k) = \left| \left\{ 0 \le i \le k-1 : E_n(i) = \min_{i \le j \le k} E_n(k) \right\} \right|.$$

This is not quite enough to conclude that the trees converge in the GHP sense. The embedding of the tree in the excursion distorts distances.



Write $H_n(k)$ for the distance from the root to the vertex visited at time k. Then

$$H_n(k) = \left| \left\{ 0 \le i \le k-1 : E_n(i) = \min_{i \le j \le k} E_n(k) \right\} \right|.$$

It turns out that $H_n(k) \approx 2E_n(k)$.

Theorem. As $n \to \infty$,

$$\frac{1}{\sqrt{2n}}(H_n(\lfloor 2nt \rfloor), 0 \le t \le 1) \to (2e(t), 0 \le t \le 1)$$

uniformly on [0, 1], almost surely.

[J.-F. Marckert & A. Mokkadem, The depth first processes of Galton-Watson trees converge to the same Brownian excursion, Annals of Probability, 31(3), pp.1655–1678, 2003.]

Let's call the vertices be $v_0, v_1, \ldots, v_{2n-1}$ in the order we visit them, where v_0 is the root.

Let's call the vertices be $v_0, v_1, \ldots, v_{2n-1}$ in the order we visit them, where v_0 is the root.

By definition,

 $d_n(v_0,v_k)=H_n(k).$

Let's call the vertices be $v_0, v_1, \ldots, v_{2n-1}$ in the order we visit them, where v_0 is the root.

By definition,

$$d_n(v_0,v_k)=H_n(k).$$

More generally, for $0 \le i < j \le 2n - 1$, write $v_i \land v_j$ for the most recent common ancestor of v_i and v_j in the tree. Then

$$d_n(v_i, v_j) = d_n(v_0, v_i) + d_n(v_0, v_j) - 2d_n(v_0, v_i \wedge v_j).$$

$$d_n(v_0, v_i \wedge v_j) = \begin{cases} \min_{i \le k \le j} H_n(k) - 1 & \text{if } v_i \text{ not an ancestor of } v_j \\ \min_{i \le k \le j} H_n(k) = H_n(i) & \text{if } v_i \text{ an ancestor of } v_j. \end{cases}$$



So

$$\left|d_n(v_0, v_i \wedge v_j) - \min_{i \leq k \leq j} H_n(k)\right| \leq 1.$$

A correspondence

Define a correspondence R_n between $\{v_0, v_1, \ldots, v_{2n-1}\}$ and [0, 1] by declaring $(v_i, s) \in R_n$ if $i = \lfloor 2ns \rfloor$.

Endow [0,1] with the pseudo-metric d_{2e} . We will bound dis (R_n) .

A correspondence

Define a correspondence R_n between $\{v_0, v_1, \ldots, v_{2n-1}\}$ and [0, 1] by declaring $(v_i, s) \in R_n$ if $i = \lfloor 2ns \rfloor$.

Endow [0,1] with the pseudo-metric d_{2e} . We will bound dis (R_n) .

Suppose that $(v_i, s), (v_j, t) \in R_n$ with $s \leq t$. Then

$$\begin{aligned} |d_n(v_i, v_j) - d_{2e}(s, t)| \\ \leq \left| \frac{1}{\sqrt{2n}} \left(H_n(\lfloor 2ns \rfloor) + H_n(\lfloor 2nt \rfloor) - 2\min_{s \le u \le t} H_n(\lfloor 2nu \rfloor) \right) \\ - \left(2e(s) + 2e(t) - 4\min_{s \le u \le t} e(u) \right) \right| + \frac{2}{\sqrt{2n}}. \end{aligned}$$

A correspondence

Define a correspondence R_n between $\{v_0, v_1, \ldots, v_{2n-1}\}$ and [0, 1] by declaring $(v_i, s) \in R_n$ if $i = \lfloor 2ns \rfloor$.

Endow [0,1] with the pseudo-metric d_{2e} . We will bound dis (R_n) .

Suppose that $(v_i, s), (v_j, t) \in R_n$ with $s \leq t$. Then

$$\begin{aligned} |d_n(v_i, v_j) - d_{2e}(s, t)| \\ \leq \left| \frac{1}{\sqrt{2n}} \left(H_n(\lfloor 2ns \rfloor) + H_n(\lfloor 2nt \rfloor) - 2\min_{s \leq u \leq t} H_n(\lfloor 2nu \rfloor) \right) \right. \\ \left. - \left(2e(s) + 2e(t) - 4\min_{s \leq u \leq t} e(u) \right) \right| + \frac{2}{\sqrt{2n}}. \end{aligned}$$

The right-hand side converges to 0 uniformly in $s, t \in [0, 1]$. So

$$\operatorname{dis}(R_n) \to 0$$
 a.s.

A coupling

Recall that μ_n is the measure which puts mass 1/(2n) on each of the vertices $v_0, v_1, \ldots, v_{2n-1}$. Then we may couple μ_n and μ_{2e} by taking $U \sim U[0, 1]$ and taking ν to be the law of the pair

$$(v_{\lfloor 2nU \rfloor}, \pi_{2e}(U)).$$

This is precisely the natural coupling ν_n induced by the correspondence R_n , and so $\nu_n(R_n^c) = 0$.

GHP convergence

But then

$$egin{aligned} &d_{\mathsf{GHP}}\left(\left(\mathcal{T}_n, rac{d_n}{\sqrt{2n}}, \mu_n
ight), \left(\mathcal{T}_{2e}, d_{2e}, \mu_{2e}
ight)
ight)\ &\leq rac{1}{2}\max\left\{\mathsf{dis}(R_n),
u_n(R_n^c)
ight\} o 0, \end{aligned}$$

almost surely as $n \to \infty$.

3. UNIVERSALITY

Key reference:

Jean-François Le Gall, Random trees and applications, *Probability Surveys* **2** (2005) pp.245-311.


Branching processes

A Bienaymé-Galton-Watson (BGW) branching process $(Z_n)_{n\geq 0}$ describes the size of a population which evolves as follows:

- Start with a single individual.
- ► This individual has a number of children distributed according to the offspring distribution p, where p(k) gives the probability of k children, k ≥ 0.
- Each child reproduces as an independent copy of the original individual.

 Z_n gives the number of individuals in generation n (in particular, $Z_0 = 1$).

BGW trees

A BGW tree is the family tree arising from a BGW branching process. We will think of this as a rooted ordered tree.

A BGW tree is the family tree arising from a BGW branching process. We will think of this as a rooted ordered tree.

Consider the case where the offspring distribution p is critical i.e.

$$\sum_{k=1}^{\infty} kp(k) = 1.$$

This ensures, in particular, that the resulting tree, T, is finite.

Combinatorial trees in disguise

Let T be a BGW tree with offspring distribution p and total progeny N.

- If p(0) = 1/2 and p(2) = 1/2 then, conditional on N = 2n − 1, the tree is uniform on the set of rooted (unplanted!) plane binary trees with n leaves.
- If p(k) = 2^{-k-1}, k ≥ 0 (i.e. Geometric(1/2) offspring distribution) then conditional on N = n, the tree is uniform on the set of plane trees with n vertices.
- If p(k) = e⁻¹/k!, k ≥ 0 (i.e. Poisson(1) offspring distribution) then conditional on N = n, if we assign the vertices labels chosen uniformly at random from {1, 2, ..., n} and then forget the ordering and the root, we obtain a labelled tree T̃ which is uniform on the set of possibilities.

The last example will be particularly important in Lecture 2.

A universal scaling limit

Let T_n be the family tree of a BGW process with critical offspring distribution of variance $\sigma^2 \in (0, \infty)$, conditioned to have total progeny n. Let d_n be the graph distance on T_n and let μ_n be the uniform measure on the vertices.

Theorem. (Aldous (1993), Le Gall (2005)) As $n \to \infty$,

$$\left(T_n, \frac{\sigma}{\sqrt{n}}d_n, \mu_n\right) \xrightarrow{d} (\mathcal{T}_{2e}, d_{2e}, \mu_{2e}),$$

where convergence is in the Gromov-Hausdorff-Prokhorov sense.

As we have seen, it is convenient to encode our trees in terms of discrete functions which are easier to manipulate.

We will do this is two different ways:

- the height function
- the depth-first walk (or Łukasiewicz path).

Suppose that our tree has *n* vertices. Let them be $v_0, v_1, \ldots, v_{n-1}$, listed in depth-first order.

Suppose that our tree has *n* vertices. Let them be $v_0, v_1, \ldots, v_{n-1}$, listed in depth-first order.

Then the height function is defined by

$$H(k)=d_n(v_0,v_k), \quad 0\leq k\leq n-1.$$

















We can easily recover the tree from its height function.

Let c(v) be the number of children of v, and that $v_0, v_1, \ldots, v_{n-1}$ is a list of the vertices in depth-first order.

Let c(v) be the number of children of v, and that $v_0, v_1, \ldots, v_{n-1}$ is a list of the vertices in depth-first order.

Define

$$X(0) = 0,$$

 $X(i) = \sum_{j=0}^{i-1} (c(v_j) - 1), ext{ for } 1 \le i \le n.$

Let c(v) be the number of children of v, and that $v_0, v_1, \ldots, v_{n-1}$ is a list of the vertices in depth-first order.

Define

$$X(0) = 0,$$

 $X(i) = \sum_{j=0}^{i-1} (c(v_j) - 1), ext{ for } 1 \le i \le n.$

In other words,

$$X(i+1) = X(i) + c(v_i) - 1, \quad 0 \le i \le n-1.$$



















Proposition. For $0 \le i \le n-1$,

$$H(i) = \# \left\{ 0 \le j \le i - 1 : X(j) = \min_{j \le k \le i} X(k) \right\}.$$

Recall that p is a distribution on \mathbb{Z}_+ such that $\sum_{k=1}^{\infty} kp(k) = 1$.

Recall that p is a distribution on \mathbb{Z}_+ such that $\sum_{k=1}^{\infty} kp(k) = 1$.

Proposition. Let $(R(k), k \ge 0)$ be a random walk with initial value 0 and step distribution $\nu(k) = p(k+1), k \ge -1$. Set

$$M = \inf\{k \ge 0 : R(k) = -1\}.$$

Recall that p is a distribution on \mathbb{Z}_+ such that $\sum_{k=1}^{\infty} kp(k) = 1$.

Proposition. Let $(R(k), k \ge 0)$ be a random walk with initial value 0 and step distribution $\nu(k) = p(k+1), k \ge -1$. Set

$$M = \inf\{k \ge 0 : R(k) = -1\}.$$

Now suppose that T is a BGW tree with offspring distribution p and total progeny N. Then

$$(X(k), 0 \le k \le N) \stackrel{d}{=} (R(k), 0 \le k \le M).$$

Recall that p is a distribution on \mathbb{Z}_+ such that $\sum_{k=1}^{\infty} kp(k) = 1$.

Proposition. Let $(R(k), k \ge 0)$ be a random walk with initial value 0 and step distribution $\nu(k) = p(k+1), k \ge -1$. Set

$$M = \inf\{k \ge 0 : R(k) = -1\}.$$

Now suppose that T is a BGW tree with offspring distribution p and total progeny N. Then

$$(X(k), 0 \le k \le N) \stackrel{d}{=} (R(k), 0 \le k \le M).$$

[Careful proof: see Le Gall (2005).]

BGW trees conditioned on their total progeny

Suppose now that we have offspring variance $\sigma^2 := \sum_{k=1}^{\infty} (k-1)^2 \rho(k) \in (0,\infty).$

The depth-first walk X is a random walk with step mean 0 and variance σ^2 , stopped at the first time it hits -1. The underlying random walk has a Brownian motion as its scaling limit, by Donsker's theorem.

The total progeny N is equal to $\inf\{k \ge 0 : X(k) = -1\}$. We want to condition on the event $\{N = n\}$.

BGW trees conditioned on their total progeny

Suppose now that we have offspring variance $\sigma^2 := \sum_{k=1}^{\infty} (k-1)^2 \rho(k) \in (0,\infty).$

The depth-first walk X is a random walk with step mean 0 and variance σ^2 , stopped at the first time it hits -1. The underlying random walk has a Brownian motion as its scaling limit, by Donsker's theorem.

The total progeny N is equal to $\inf\{k \ge 0 : X(k) = -1\}$. We want to condition on the event $\{N = n\}$.

Standing assumption: $\mathbb{P}(N = n) > 0$ for all *n* sufficiently large.

BGW trees conditioned on their total progeny

Write $(X_n(k), 0 \le k \le n)$ for the depth-first walk conditioned on $\{N = n\}$. Then there is a conditional version of Donsker's theorem.

Theorem. As $n \to \infty$,

$$\frac{1}{\sigma\sqrt{n}}(X_n(\lfloor nt \rfloor), 0 \le t \le 1) \xrightarrow{d} (e(t), 0 \le t \le 1),$$

where $(e(t), 0 \le t \le 1)$ is a standard Brownian excursion.

[W.D. Kaigh, An invariance principle for random walk conditioned by a late return to zero, Annals of Probability 4, 1976, pp.115-121.]

Height process

Let $(H_n(i), 0 \le i \le n)$ be the height process of a critical BGW tree with offspring variance $\sigma^2 \in (0, \infty)$, conditioned to have total progeny n, so that

$$H_n(i) = \# \left\{ 0 \le j \le i-1 : X_n(j) = \min_{j \le k \le i} X_n(k) \right\}.$$

Theorem. As $n \to \infty$,

$$\frac{\sigma}{\sqrt{n}}\left(H_n(\lfloor nt \rfloor), 0 \leq t \leq 1\right) \stackrel{d}{\rightarrow} 2\left(e(t), 0 \leq t \leq 1\right)\right),$$

where $(e(t), 0 \le t \le 1)$ is a standard Brownian excursion.
Convergence to the Brownian CRT

The convergence

$$\left(T_n, \frac{\sigma}{\sqrt{n}}d_n, \mu_n\right) \xrightarrow{d} (\mathcal{T}_{2e}, d_{2e}, \mu_{2e}),$$

now follows by applying Skorokhod's theorem (in order to work on a probability space where the height process converges almost surely) and then using the same proof that we had in the case of binary trees.

Universality

The universality class of the Brownian CRT is, in fact, even larger. Some other examples of trees (and graphs!) with the Brownian CRT as their scaling limit are:

- uniform unordered unlabelled rooted trees
- uniform unordered unlabelled unrooted trees
- critical multi-type BGW trees
- random trees with a prescribed degree sequence satisfying certain conditions
- random dissections
- random graphs from subcritical classes.