Descent-weighted trees and permutations

Stephan Wagner

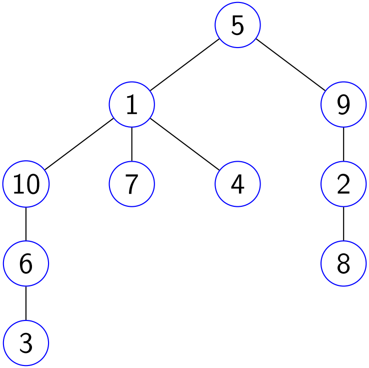
joint work with Paul Thévenin

Uppsala University
Department of Mathematics

November 14, 2022

Rooted labelled trees

In a rooted labelled tree, all vertices have a unique label in $\{1, 2, \ldots, n\}$.



Rooted labelled trees

- Rooted labelled trees form a *simply generated family* of trees.

Rooted labelled trees

- Rooted labelled trees form a *simply generated family* of trees.
- Uniformly random rooted labelled trees are a special case of *conditioned Galton–Watson trees*.

Rooted labelled trees

- Rooted labelled trees form a *simply generated family* of trees.
- Uniformly random rooted labelled trees are a special case of *conditioned Galton–Watson trees*.
- The number of rooted labelled trees with n vertices is n^{n-1}, and many other combinatorial formulas are known.

Rooted labelled trees

- Rooted labelled trees form a *simply generated family* of trees.
- Uniformly random rooted labelled trees are a special case of *conditioned Galton–Watson trees*.
- The number of rooted labelled trees with n vertices is n^{n-1}, and many other combinatorial formulas are known.
- The height and the average distance from the root are of order $\sqrt{n}$.

Recursive trees

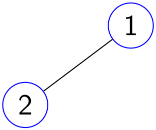
Recursive trees can be obtained by adding vertices step by step.

Recursive trees

Recursive trees can be obtained by adding vertices step by step.

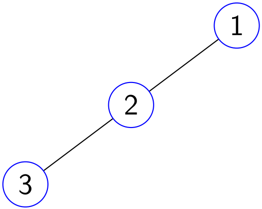
Recursive trees

Recursive trees can be obtained by adding vertices step by step.



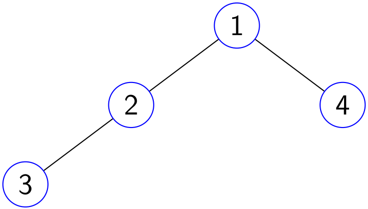
Recursive trees

Recursive trees can be obtained by adding vertices step by step.



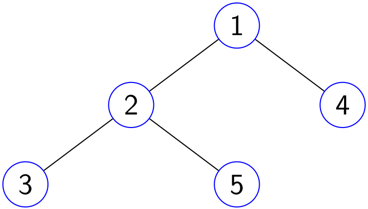
Recursive trees

Recursive trees can be obtained by adding vertices step by step.



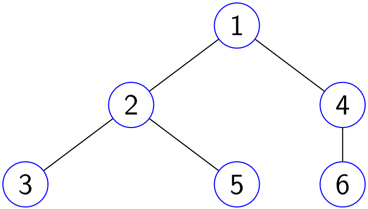
Recursive trees

Recursive trees can be obtained by adding vertices step by step.



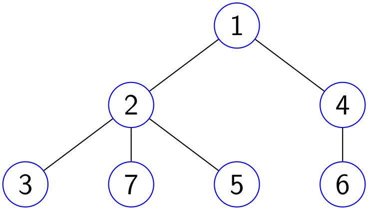
Recursive trees

Recursive trees can be obtained by adding vertices step by step.



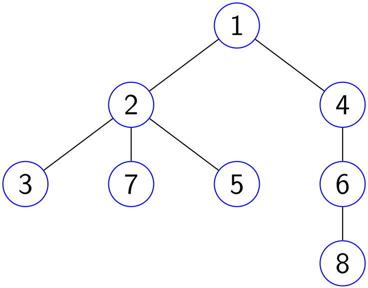
Recursive trees

Recursive trees can be obtained by adding vertices step by step.



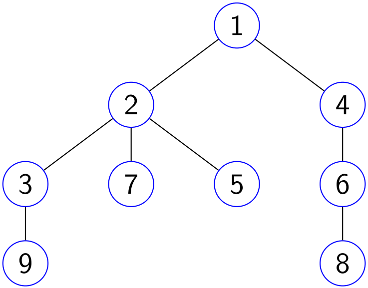
Recursive trees

Recursive trees can be obtained by adding vertices step by step.



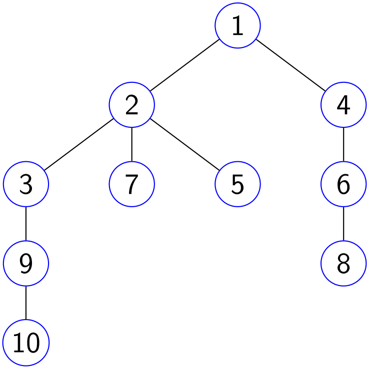
Recursive trees

Recursive trees can be obtained by adding vertices step by step.



Recursive trees

Recursive trees can be obtained by adding vertices step by step.



Recursive trees

- Recursive trees are special case of a family of *increasing* trees.

Recursive trees

- Recursive trees are special case of a family of *increasing* trees.
- The number of recursive trees with n vertices is $(n-1)!$, and there are indeed many nice combinatorial connections to permutations.

Recursive trees

- Recursive trees are special case of a family of *increasing* trees.
- The number of recursive trees with n vertices is $(n-1)!$, and there are indeed many nice combinatorial connections to permutations.
- The height and the average distance from the root are of order $\log n$.

Motivation

We would like a model of random trees that *interpolates* between uniformly random rooted labelled trees and recursive trees.

Motivation

We would like a model of random trees that *interpolates* between uniformly random rooted labelled trees and recursive trees.

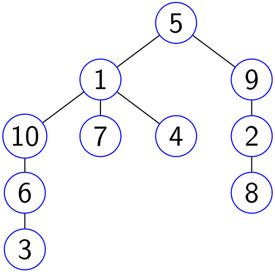
This is achieved by defining a *weight* based on *descents*.

Descents in rooted labelled trees

A *descent* is a pair of adjacent vertices labelled i (parent) and j (child) respectively such that $j < i$.

Descents in rooted labelled trees

A *descent* is a pair of adjacent vertices labelled i (parent) and j (child) respectively such that $j < i$.



The example has four descents: $(5, 1)$, $(9, 2)$, $(6, 3)$ and $(10, 6)$.

The model

Let q be a positive real number. We consider random rooted labelled trees with n vertices whose probabilities are *proportional* to $q^{\text{number of descents}}$. The parameter q is allowed to depend on n.

The model

Let q be a positive real number. We consider random rooted labelled trees with n vertices whose probabilities are *proportional* to $q^{\text{number of descents}}$. The parameter q is allowed to depend on n.

- The q-enumeration of labelled trees by descents goes back to Eğecioğlu and Remmel (1986).

The model

Let q be a positive real number. We consider random rooted labelled trees with n vertices whose probabilities are *proportional* to $q^{\text{number of descents}}$. The parameter q is allowed to depend on n.

- The q-enumeration of labelled trees by descents goes back to Eğecioğlu and Remmel (1986).
- Note that we obtain uniformly random rooted labelled trees for $q = 1$, random recursive trees as $q \to 0$ and random recursive trees with labels reversed as $q \to \infty$.

The model

Let q be a positive real number. We consider random rooted labelled trees with n vertices whose probabilities are *proportional* to $q^{\text{number of descents}}$. The parameter q is allowed to depend on n.

- The q-enumeration of labelled trees by descents goes back to Eğecioğlu and Remmel (1986).
- Note that we obtain uniformly random rooted labelled trees for $q = 1$, random recursive trees as $q \to 0$ and random recursive trees with labels reversed as $q \to \infty$.
- Replacing q by $1/q$ amounts to reversing all labels. It is therefore enough to consider $q \leq 1$.

Descents in permutations

As a related model, consider random *permutations* of $\{1, 2, \ldots, n\}$ whose probabilities are proportional to $q^{\text{number of descents}}$.

Descents in permutations

As a related model, consider random *permutations* of $\{1, 2, \ldots, n\}$ whose probabilities are proportional to $q^{\text{number of descents}}$.

$$582469713$$

The example has three descents: 82, 97 and 71.

Descents in permutations

As a related model, consider random *permutations* of $\{1, 2, \ldots, n\}$ whose probabilities are proportional to $q^{\text{number of descents}}$.

582469713

The example has three descents: 82, 97 and 71.

- Permutations of $\{1, 2, \ldots, n\}$ with k descents are counted by the *Eulerian numbers* $\left\langle {n \atop k} \right\rangle$.

Descents in permutations

As a related model, consider random *permutations* of $\{1, 2, \ldots, n\}$ whose probabilities are proportional to $q^{\text{number of descents}}$.

582469713

The example has three descents: 82, 97 and 71.

- Permutations of $\{1, 2, \ldots, n\}$ with k descents are counted by the *Eulerian numbers* $\left\langle {n \atop k} \right\rangle$.
- The *Eulerian polynomial* is

$$\sum_{k=0}^{n-1} \left\langle {n \atop k} \right\rangle q^k = (1-q)^{n+1} \sum_{m=1}^{\infty} m^n q^{m-1}.$$

Descents in permutations

As a related model, consider random *permutations* of $\{1, 2, \ldots, n\}$ whose probabilities are proportional to $q^{\text{number of descents}}$.

582469713

The example has three descents: 82, 97 and 71.

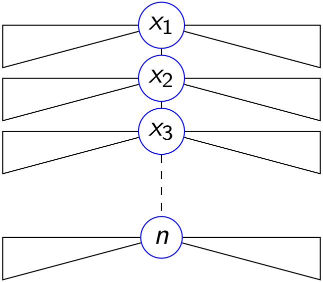
- Permutations of $\{1, 2, \ldots, n\}$ with k descents are counted by the *Eulerian numbers* $\left\langle {n \atop k} \right\rangle$.
- The *Eulerian polynomial* is

$$\sum_{k=0}^{n-1} \left\langle {n \atop k} \right\rangle q^k = (1-q)^{n+1} \sum_{m=1}^{\infty} m^n q^{m-1}.$$

- This model is very similar to *Mallows permutations* (number of inversions) and *Ewens permutations* (number of cycles).

A connection between permutations and trees

Consider the path from the root to vertex n. The labels of the vertices on this path follow the descent-biased permutation model.



Permutations: phases and local limits

1. If $q2^n \to 0$, then the permutation is with high probability the identity.

Permutations: phases and local limits

1. If $q2^n \to 0$, then the permutation is with high probability the identity.
2. If $\log(1/q) \sim cn$ as $n \to \infty$, where $0 < c \leq \log 2$ is constant, then there exists a nonnegative integer k_c such that, with high probability, the number of descents is k_c or $k_c + 1$. The first r elements $(\pi_1, \ldots, \pi_r)$ converge in distribution to $(X_1, X_1 + X_2, \ldots, X_1 + \cdots + X_r)$, where the X_1 are geometrically distributed random variables.

Permutations: phases and local limits

1. If $q2^n \to 0$, then the permutation is with high probability the identity.
2. If $\log(1/q) \sim cn$ as $n \to \infty$, where $0 < c \leq \log 2$ is constant, then there exists a nonnegative integer k_c such that, with high probability, the number of descents is k_c or $k_c + 1$. The first r elements $(\pi_1, \ldots, \pi_r)$ converge in distribution to $(X_1, X_1 + X_2, \ldots, X_1 + \cdots + X_r)$, where the X_1 are geometrically distributed random variables.
3. If $q \to 0$ and $\log(1/q) = o(n)$, then

$$\frac{\log(1/q)}{n}(\pi_1, \ldots, \pi_r) \xrightarrow{d} (E_1, E_1 + E_2, \ldots, E_1 + \cdots + E_r),$$

where the E_i are i.i.d. $\mathrm{Exp}(1)$-variables.

Permutations: phases and local limits

4. If q is constant, then define the following Markov process: X_1 has density $\frac{\log(1/q)}{1-q} q^x$ on $[0,1]$, and for all $j \geq 1$, X_{j+1} has density

$$\frac{q^{x-X_j}}{\int_0^{X_j} q^z dz + \int_{X_j}^1 q^{z+1} dz} \left(q + (1-q)\mathbb{1}_{x \geq X_j}\right),$$

also on $[0,1]$. Then

$$\frac{1}{n}(\pi_1, \ldots, \pi_r) \xrightarrow{d} (X_1, X_2, \ldots, X_r).$$

Some proof ideas

In the *degenerate case* ($\log(1/q) \sim cn$), we can use direct counting: the number of permutations of $\{1, 2, \ldots, n\}$ with k descents is asymptotically equal to $(k+1)^n$, so the total weight of $\sim (k+1)^n q^k$ is maximal for k maximizing $\log(k+1) - ck$.

Some proof ideas

In the *degenerate case* ($\log(1/q) \sim cn$), we can use direct counting: the number of permutations of $\{1, 2, \ldots, n\}$ with k descents is asymptotically equal to $(k+1)^n$, so the total weight of $\sim (k+1)^n q^k$ is maximal for k maximizing $\log(k+1) - ck$.

In particular, if $q2^n \to 0$, then the weight of the identity permutation is greater than that of all others combined.

Some proof ideas

In the non-degenerate cases, we can make use of generating functions and the method of moments.

Some proof ideas

In the non-degenerate cases, we can make use of generating functions and the method of moments.

If $P_n^k(q)$ is the weighted number of permutations of $\{1, 2 \ldots, n\}$ whose first element is k, and $S(x, y) = \sum_{n,k} P_n^k(q) \frac{x^n y^k}{n!}$, then

$$\frac{\partial}{\partial x} S(x, y) = (1-q) \frac{y}{1-y} \left(\frac{1}{e^{(q-1)x} - q} - \frac{qy}{e^{(q-1)xy} - q} - S(x, y) \right).$$

This can be used to analyze the moments of the first element.

Trees: root degree and local limit

Trees: root degree and local limit

For constant q, the root degree has a discrete limit distribution (if $q \to 0$, it goes to infinity), with probabilities given by

$$p_k = \frac{q^{1/(1-q)}}{1-q} \frac{1-q^k}{k!} \Big(\frac{\log(1/q)}{1-q} \Big)^k.$$

Trees: root degree and local limit

For constant q, the root degree has a discrete limit distribution (if $q \to 0$, it goes to infinity), with probabilities given by

$$p_k = \frac{q^{1/(1-q)}}{1-q} \frac{1-q^k}{k!} \Big(\frac{\log(1/q)}{1-q} \Big)^k.$$

Moreover, we have a local limit, i.e., the distribution of the neighbourhood of radius r around the root converges for every fixed r.

Trees: root component

By *root component*, we mean the largest subtree containing the root that forms an increasing tree. Let $R_n(q)$ be the size of this component.

Trees: root component

By *root component*, we mean the largest subtree containing the root that forms an increasing tree. Let $R_n(q)$ be the size of this component.

- For fixed q, $R_n(q)$ converges weakly to a geometric random variable $\mathrm{Geom}(q^{1/(1-q)})$.

Trees: root component

By *root component*, we mean the largest subtree containing the root that forms an increasing tree. Let $R_n(q)$ be the size of this component.

- For fixed q, $R_n(q)$ converges weakly to a geometric random variable $\mathrm{Geom}(q^{1/(1-q)})$.
- If $q \to 0$, but $qn \to \infty$, scaling with q gives a limit: $qR_n(q) \xrightarrow{d} \mathrm{Exp}(1)$.

Trees: distances

Recall that the average distance from the root in uniformly random rooted labelled trees is of order $\Theta(\sqrt{n})$, while it is $\Theta(\log n)$ for random recursive trees. Our model interpolates in the following way:

Trees: distances

Recall that the average distance from the root in uniformly random rooted labelled trees is of order $\Theta(\sqrt{n})$, while it is $\Theta(\log n)$ for random recursive trees. Our model interpolates in the following way:

If q is fixed (and probably if $qn \to \infty$), the average distance of a random vertex from the root is asymptotically equal to $\frac{\log(1/q)}{1-q}\sqrt{\pi qn/2}$.

Some further directions

- "Mesoscopic" limit of permutations: if one considers descent-weighted permutations in windows of size $\Theta(\log(1/q_n))$, the number of descents is of constant order, and one observes a "diagonal pattern".

Some further directions

- "Mesoscopic" limit of permutations: if one considers descent-weighted permutations in windows of size $\Theta(\log(1/q_n))$, the number of descents is of constant order, and one observes a "diagonal pattern".
- Scaling of random trees: for fixed q, descent-weighted trees appear to converge to the continuum random tree after suitable scaling. What happens as $q \to 0$?

Some further directions

- "Mesoscopic" limit of permutations: if one considers descent-weighted permutations in windows of size $\Theta(\log(1/q_n))$, the number of descents is of constant order, and one observes a "diagonal pattern".
- Scaling of random trees: for fixed q, descent-weighted trees appear to converge to the continuum random tree after suitable scaling. What happens as $q \to 0$?
- Further properties and statistics of random trees: what can one say about distributions? Are quantities such as the average height or the average number of leaves generally monotone in q?

Some further directions

- "Mesoscopic" limit of permutations: if one considers descent-weighted permutations in windows of size $\Theta(\log(1/q_n))$, the number of descents is of constant order, and one observes a "diagonal pattern".
- Scaling of random trees: for fixed q, descent-weighted trees appear to converge to the continuum random tree after suitable scaling. What happens as $q \to 0$?
- Further properties and statistics of random trees: what can one say about distributions? Are quantities such as the average height or the average number of leaves generally monotone in q?
- Changing the weight: instead of descents, one could also use *inversions* in trees.