Most trees are short and fat.

Louigi Addario-Berry

XVI (?)

One-day Meeting
in Combinatorics

Oxford, UK, May 27, 2015
**Plane Trees:**

Rooted; children of each node have left-to-right order.

Height: Greatest distance from any node to the root \( h^T(T) \)

Width: Greatest # nodes on a single level \( \text{wid}(T) \)
Random Plane Trees: Galton-Watson Trees

- Each node has random # of children
- Nodes reproduce independently
- Model Parameter: Reproduction Law/Offspring distribution.

Construction
- \((C_i, i \geq 1)\) independent copies of a random variable \(C\) with \(\sum_{k \geq 0} P(C = k) = 1\).

Rule
- The sequence \((C_i, i \geq 1)\) gives # children of nodes, in breadth-first search order.

Example: \((2, 1, 3, 0, 0, 1, 0, 2, 0, 0, 1, 4, 0, \ldots)\)

Size = # vertices = \(\sigma\)

[possible that tree is infinite, if so then \(\sigma = \infty\)]

Key definition:
Such \(T\) is GW\((C)\)-distributed
Main Results Fix any r.v. \( C \) with \( \sum_{k \geq 0} \mathbb{P}(C=k) = 1 \). Let \( T \) be \( GW(C) \) distributed.

Write \( p_k = \mathbb{P}(C=k) \).

Theorem ("Most trees are short & fat")

There is a universal constant \( \delta > 0 \) s.t.

\[
\mathbb{P}(\text{ht}(T) \geq \frac{k}{1-p_1} \cdot \text{wid}(T)) \leq \exp(-\delta k).
\]

Remarks

1. If \( IE_C > 1 \) then \( \mathbb{P}(\sigma = \infty) > 0 \), and \( \mathbb{P}(\text{ht}(T) = \infty | \sigma = \infty) = 1 \).

   Also, given that \( \sigma < \infty \), the cond. dist. of \( T \) is \( GW(H) \) where \( \mathbb{P}(H=1) = p_1 \). So can assume \( EC \leq 1 \).

2. Write \( T_k = \{\text{nodes at level } k\} \). Easy: \( \mathbb{E}|T_k| = [IE_C]^k \).

   If \( EC < 1 \) then \( \mathbb{P}(\text{ht}(T) \geq k) = \mathbb{P}(|T_k| > 0) \leq \mathbb{E}|T_k| = [IE_C]^k \)

   So \( \mathbb{P}(\text{ht}(T) \geq k \cdot \text{wid}(T)) \leq \exp(-k \cdot \log(\frac{1}{IE_C})) \).

   \( \text{Exp. tails, but } \log(\frac{1}{IE_C}) \text{ isn't } \leq \frac{1}{1-p_1} \).

3. If \( EC = 1 \) \& \( IE[C^2] < \infty \) then \( \mathbb{P}(\text{wid}(T) \geq x) = \Theta(\frac{1}{x}) \),

   \( \mathbb{P}(\text{ht}(T) \geq x) = \Theta(\frac{1}{x}) \), so not trivially true.

   [Actually, when \( EC = 1 \) \& \( IE[C^2] < \infty \), \( \mathbb{P}(\text{ht}(T) \geq \frac{k}{1-p_1} \text{wid}(T)) \geq \exp(-\delta k); \text{thm is tight}] \)
Random Plane Trees: Galton-Watson Trees

- Each node has random # of children
- Nodes reproduce independently
- Model Parameter: Reproduction Law/Offspring distribution.

**Construction**
- \((C_i, i \geq 1)\) independent copies of a random variable \(C\) with \(\sum_{k \geq 0} P(C = k) = 1\).

**Rule**
The sequence \((C_i, i \geq 1)\) gives # children of nodes, in breadth-first search order.

**Example**: \((2, 1, 3, 0, 0, 1, 0, 2, 0, 0, 1, 4, 0, ...\))

**Halting Condition**
Trees with \(n\) nodes have \(n-1\) edges.
For \(n \geq 0\)

\[
\text{nodes discovered by time } n = 1 + \sum_{i=1}^{n} C_i
\]

\[#\text{ vertices explored by time } n = n\]

**Size** = # vertices = \(\sigma\)

\[
= \inf\{n : 1 + \sum_{i=1}^{n} C_i = n\}
\]

Key definition: Such \(T\) is \(GW(C)\)-distributed
Setup

\[ \sum_{j=1}^{i} C_j = \# \text{nodes discovered by time } i. \]

Let \( S_i = \sum_{j=1}^{i} (C_j - 1) \)

= \# nodes in "BFS queue" at time \( i \)

\[ EC < 1 \Rightarrow \sum_{n=1}^{\infty} n(\frac{EC}{1}) < 1. \]

\( \sigma = \inf \{ t : S_t = 0 \} = \text{first time no nodes left to explore} \)

**Prop:** Let \( W(T) = \max(S_i, 0 \leq i < \sigma) \). Then \( \text{wid}(T) \in (W(T)/2, W(T)] \)

**Proof:** During BFS on level \( k \), "exploration queue" \( \in T_k \cup T_{k+1} \) and \( = T_k \) at start of level \( k \).

**Idea:** \[ \text{ht}(T) = \frac{\text{ht}(T)}{k=1} 1 = \sum_{k=1}^{\infty} \frac{1}{T_k}. \]

When \( v_i \in T_k \) then \( S_i \approx |T_k| \) so perhaps

\[ \text{ht}(T) \approx \sum_{k=1}^{\infty} \sum_{v_i \in T_k} \frac{1}{S_i} = \sum_{i=1}^{\infty} \frac{1}{S_i} = H(T) \]

[False; consider a star with \( n \) leaves. But...]

**Prop:** \[ \text{ht}(T) \leq 3H(T). \]

**Corollary** Suffices to prove \[ P(H(T) \geq \frac{k}{1-p}, W(T)) \leq e^{-\delta k}, \]

thm. follows.
\[ W(\sigma) = \max(S_i, 0 \leq i < \sigma) \quad H(\sigma) = \frac{1}{\sum_{i=1}^{\sigma} \frac{1}{S_i}} \quad \text{Aim: } \mathbb{P}(H(\sigma) > \frac{k}{1-p}, W(\sigma)) \leq e^{-\delta k} \]

**Decomposition into scales.**

When \( S_i \approx 2^k \), takes about \( 2^k \) steps for \( H(i) \) to increase by 1

\[ H(i+2^k) - H(i) = \frac{\sum_{j=i+1}^{i+2^k} \frac{1}{S_j}}{\sum_{j=i+1}^{i+2^k} \frac{1}{S_j}} \approx 2^k \cdot \frac{1}{S_i} \approx 1. \]

**Definition**

- \( T_0 = 0 = \text{initial time} \)
- \( L_0 = 0 = \log_2 S_0 = \text{initial scale} \)
- \( T_{n+1} = \min\{t \geq T_n : S_t \notin [2^{L_{n-1}}, 2^{L_{n+1}}]\} \)
- \( L_{n+1} = \sup\{k : 2^k \leq S_{T_{n+1}}\} \)

**Diagram:**

- \( L_{n+1} > L_n + 2 \)
- \( L_{n+1} = L_n - 2 \)

**Note:**

- \( S_{T_n} < 2^{L_n} \) so \( \mathbb{P}(L_{n+1} < L_n) = \mathbb{P}(\text{Hit } 2^{L_n-1} \text{ before } 2^{L_{n+1}^2}) > \frac{1}{2} \)
\( W(\sigma) = \max \{ S_i, 0 \leq i < \sigma \} \quad H(\sigma) = \sum_{i=1}^{\sigma} \frac{1}{S_i} \quad \text{Aim: } \Pr \{ H(\sigma) \geq \frac{k}{1-p} \} \cdot W(\sigma) \leq e^{-\delta k} \\

\text{Number of visits to a scale}

Let \( \Lambda(t) = \) current scale at time \( t = \sum_n \) (with \( n \) s.t. \( \tau_n \leq t < \tau_{n+1} \))

Let \( N(l) = \#\{ t < \sigma : \Lambda(t) = l \} \)

Then

\[
H(\sigma) = \sum_{l \geq 0} \sum_{\{ t < \sigma : \Lambda(t) = l \}} \frac{1}{S_t} \leq \sum_{l \geq 0} N(l) \cdot 2^{-(l-1)}
\]

\( N(l) = N(l,1) + N(l,2) + \ldots + N(l,M(l)) \)

\( N(l,i) = \) Duration of \( i \) th visit to scale \( l \)

\( M(l) = \# \text{ visits to scale } l = \#\{ i : L_i = l \} \)

**Fact:** Given that \( M(l) \neq 0 \), \( M(l) \) dominated by sum of 2 \( \text{Geom}(\frac{1}{2}) \) r.v.s. \( \Rightarrow \Pr \{ M(l) > k \mid M(l) > 0 \} \leq 2^{-k/2} \).

**Proof:** visits to scale \( l \) entail upcrossings of \( [2^{l-1}, 2^{-l}] \) or of \( [2^{l+1}, 2^{l+2}] \). Both are hard since walk has non-positive drift.

\( W(\sigma) = \max (S_i, 0 \leq i < \sigma) \quad H(\sigma) = \sum_{i=1}^{\sigma} \frac{1}{S_i} \quad \text{Aim: } P(H(\sigma) \geq \frac{k}{1-p}, W(\sigma)) \leq e^{-\delta k} \)

\[
H(\sigma) \leq \sum_{\lambda \geq 0} N(\lambda) \cdot 2^{-(\lambda-1)}
= \sum_{\lambda \geq 0} \sum_{i=1}^{M(\lambda)} N(\lambda, i) \cdot 2^{-(\lambda-1)}
N(\lambda) = \text{time spent at scale } \lambda
N(\lambda, i) = \text{Duration of } i\text{'th visit to scale } \lambda

\text{Thm (Lévy; Doeblin; Kolmogorov; Rogozin; Le Cam; Esséen; Kesten):}
With \( p = \max p_i \), have
\[
\max_k P(S_n = k) \leq \frac{Cp}{\sqrt{n(1-p)}} \quad C > 0 \text{ universal.}

"Any random walk spreads out over } \geq \sqrt{n} \text{ values by time } n." \text{ Here } \sqrt{n} \approx 2^l

\text{Corollary: } P(N(l,i) > \frac{64C^2p^2}{1-p} \cdot m \cdot 4^l) \leq 2^{-m}

Recall that \( N(l) = \sum_{i=1}^{M(l)} N(l,i) \) and that \( P(M(l) \geq k) \leq 2^{-k/2} \)

It then follows that \( P(\frac{N(l)}{2^{l-1}} > C' m \cdot 2^l) \leq 2^{-m} \)
\[ W(\sigma) = \max(S_i, 0 \leq i < \sigma) \quad H(\sigma) = \sum_{i=1}^{\sigma} \frac{1}{S_i} \quad \text{Aim: } \mathbb{P}(H(\sigma) > \frac{k}{1-p}, W(\sigma)) < e^{-\delta k} \]

\[
H(\sigma) \leq \sum_{l=0}^{\infty} N(l) \cdot 2^{-(l-1)} \quad \mathbb{P}\left( \frac{N(l)}{2^{l-1}} > C' m \cdot 2^l \right) < 2^{-m}
\]

Wrapping up

Note that if \( N(l) > 0 \) then \( W(\sigma) > 2^{l-1} \)

So with \( l^* = \max\{l : N(l) > 0\} \),

\[
\mathbb{P}(H(\sigma) > C' m \cdot W(\sigma))
\]

\[
\leq \mathbb{P}\left( \sum_{l=0}^{l^*} \frac{N(l)}{2^{l-1}} > C' m \cdot 2^{l^*} \right)
\]

\[
\leq \sum_{l=0}^{l^*} \mathbb{P}\left( \frac{N(l)}{2^{l-1}} > C' m \cdot 2^{(l^*+l)/2} \right) \quad \left[ \sum_{l=0}^{l^*} 2^{(l^*+l)/2} \leq 2^{l^*} \frac{1}{\sqrt{2}-1} \right]
\]

\[
\leq \sum_{l=0}^{l^*} 2^{-m} \cdot 2^{(l^*-l)/2} < 2^{1-m}
\]

Behind the curtain: Use Markov property and non-positive drift to “fix” bogus \( \leq \) above.
Remarks

- Claimed theorem with dependence only on $p_i$ proved it with dependence on $p = \max p_i$.

Fix: non-trivial; requires modifying "dispersion" bound for our setting.

(Idea: If subcritical then $p_{\text{max}} = p_0$ or $p_i$; if $p_0$ close to 1 then either very subcritical or make large jumps.)

- Stronger results if add info. about tails of degrees. Ex: If $\Pr(C \geq k) = \Theta(t^{-\alpha})$, $\alpha \in (1,2)$,

  
  then $\Pr(H(\sigma) > C \cdot m \cdot W(\sigma)^{\alpha-1}) \leq 2^{-\delta m}$

- Conjecture: All this works even conditional on size of tree: $\Pr(H(\sigma) > C \cdot m \cdot W(\sigma) \mid \sigma = n) \leq 2^{-\delta m}$

- Conjecture: Binary trees are the tallest & slimmest.

Let $G_{Wn}(C) = [G_{W}(C) \mid \sigma = n]$, $ht_n(C) = \text{ht}(G_{Wn}(C))$, $\text{wid}_n(C) = \text{wid}(G_{Wn}(C))$

Assuming $p_i = 0$, to stochastically maximize $ht_n(C)$

Take

$P_0 = P_2 = \frac{1}{2}$. 

Thank you!