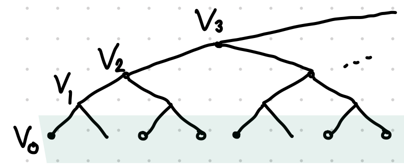
Convergence of recursive equations via numerical analysis

Louigi Addario-Berry, Luc Devroye, Celine Kerriou, Rivka Maclaine Mitchell

Analysis of Algorithms, CIRM, Luminy, France

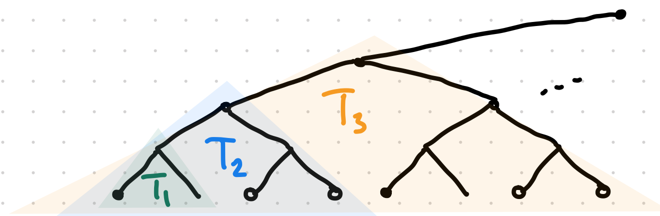
June 24, 2019

T = infinite binary canopy tree



$\mathcal{L}$ = leaves of T

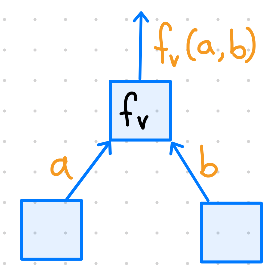
- one-way infinite path $v_0, v_1, v_2, \dots$
- node v_n is the root of a complete binary tree of depth n



T_n = subtree rooted at v_n

$\mathcal{L}_n$ = leaves of T_n

Functions on T:
- input from children
- combination function at nodes
- output to parents



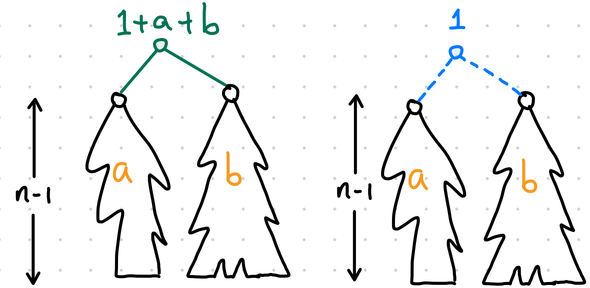
Choose functions $(f_v, v \in T_n \setminus \mathcal{L}_n)$; this turns T_n into a function,

$$x = (x_v, v \in \mathcal{L}_n) \longmapsto T_n(x) \longleftarrow \text{output at root } v_n \text{, on input } x.$$

Either or both of x and $(f_v, v \in T_n \setminus \mathcal{L}_n)$ can be random

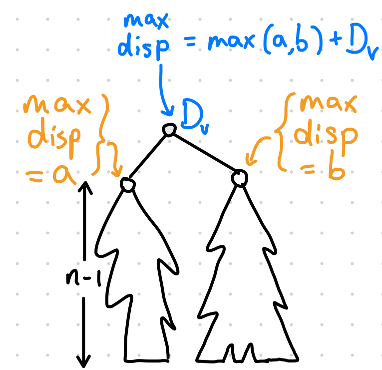
Examples

① $f_v = \begin{cases} (a,b) \longmapsto 1+a+b & \text{with prob. } p \\ (a,b) \longmapsto 1 & \text{with prob. } 1-p \end{cases}$



Then $T_n(\vec{1}) \overset{d}{=}$ # nodes at level $\leq n$ in a Galton-Watson tree

with offspring dist. $\begin{cases} 2 & \text{with probability } p \\ 0 & \text{with probability } 1-p \end{cases}$

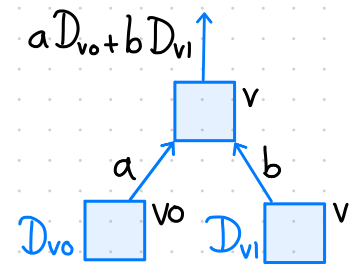


② Let $(D_v, v \in T)$ be IID with law μ, let $f_v(a,b) = \max(a,b) + D_v$

Then $T_n(\vec{0}) \overset{d}{=}$ maximum position in generation n of a binary branching random walk with displacement dist μ

(displacements at vertices)

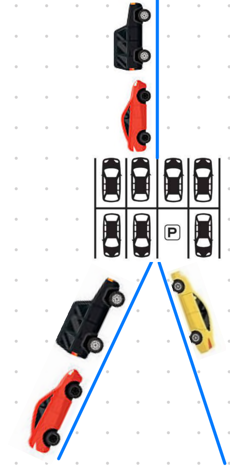
③ Let $(D_v, v \in T)$ be IID with law μ, let $f_v(a,b) = a D_{v0} + b D_{v1}$.



This is a smoothing transform; fixed points studied by Durrett & Liggett (1983), many others.

In fact, all these equations have been studied from the perspective of fixed-point equations (sometimes wish to introduce a rescaling or shift).

Examples without a fixed-point theory



④ Derrida-Retaux model / "Parking on trees". Here $f_v(a,b) = \max(a+b-1, 0)$

Question: Large-n behaviour of $T_n(X)$ where $X = (X_v, v \in \mathcal{L}_n)$ IID with some law μ

(answer of course depends on μ)

[Refs: Hu, Mallein, Pain, 1811.08749v2 ; Hu, Shi, 1705.03792 ; Goldschmidt, Przykucki, 1610.08786]

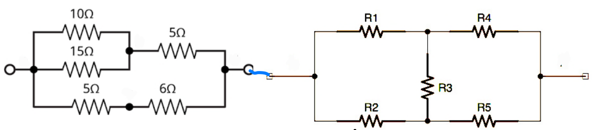
There is exactly one model where this can be analyzed so far.

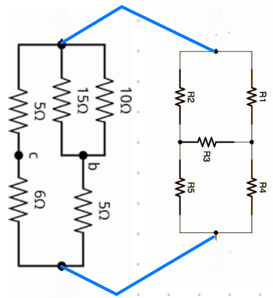
⑤ Random hierarchical lattice.

Series connection Resistance $\to a+b$

Parallel connection Resistance $\to \frac{ab}{a+b}$

$$f_v = \begin{cases} (a,b) \longmapsto a+b & \text{with prob. } p \\ (a,b) \longmapsto \frac{ab}{a+b} & \text{with prob. } 1-p \end{cases}$$





[Ref: Hambly - Jordan 2004. $p > \frac{1}{2} \to T_n(\vec{1})$ grows exponentially; $p < \frac{1}{2} \to T_n(\vec{1})$ decays exp.]

⑥ Pemantle's Min-Plus tree

$$f_v = \begin{cases} (a,b) \longmapsto a+b & \text{with prob. } p \\ (a,b) \longmapsto \min(a,b) & \text{with prob. } 1-p \end{cases}$$

Theorem (A-C) $\dfrac{\log T_n(\vec{0})}{(\pi^2 n/3)^{1/2}} \xrightarrow{d} \text{Beta}(2,1)$

[Ref: Auffinger - Cable: 1709.07849

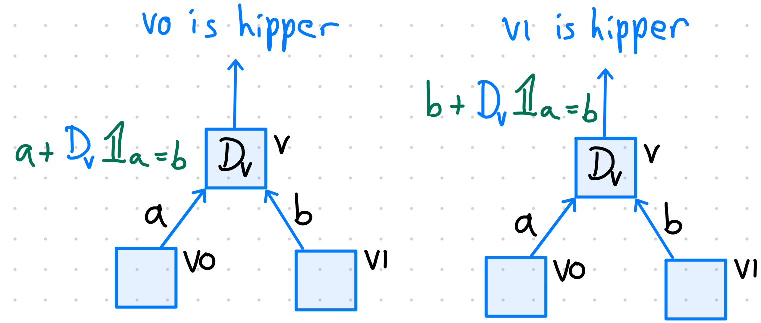
(Open question: universality: what happens for other inputs?)

New model — Hipster random walk

Fix $(D_v, v \in \mathcal{L})$ IID. Let f_v be defined by

$(a,b) \xmapsto{f_v} a + D_v \mathbb{1}_{a=b}$ with prob. $\frac{1}{2}$

$(a,b) \xmapsto{f_v} b + D_v \mathbb{1}_{a=b}$ with prob. $\frac{1}{2}$



Idea Think of time as running up the tree

1. One of v0, v1 is hipper than the other (chosen randomly)
2. If another particle shows up, hipper child takes off.

We will study

- symmetric simple hipster random walk → $D_v = \begin{cases} 1 & \text{w/ prob. } \frac{1}{2} \\ -1 & \text{w/ prob. } \frac{1}{2} \end{cases}$

  SSHRW

- totally asymmetric lazy simple hipster random walk → $D_v = \begin{cases} 1 & \text{w/ prob. } p \\ 0 & \text{w/ prob. } 1-p \end{cases}$ $p \in (0,1)$.

  TALSHRW

Theorem

For SSHRW, $\dfrac{T_n(\vec{0})}{(36n)^{1/3}} \xrightarrow{d} \mathrm{Beta}(2,2) - \frac{1}{2}$.

For TALSHRW $\dfrac{T(\vec{0})}{(4(1-p)n)^{1/2}} \xrightarrow{d} \mathrm{Beta}(2,1)$

Note Result for TALSHRW very similar to that of Auffinger-Cable.
Recall Auffinger-Cable:

$$f_v = \begin{cases} (a,b) \longmapsto a+b & \text{with prob. } p \\ (a,b) \longmapsto \min(a,b) & \text{with prob. } 1-p \end{cases}$$

Theorem (A-C) $\dfrac{\log T_n(\vec{0})}{(\pi^2 n/3)^{1/2}} \xrightarrow{d} \mathrm{Beta}(2,1)$

Intuition Suppose $T_n(\vec{0})$ is growing on a (stretched) exponential scale.

Write L, R for values at children of root of T_n

If $|\log L - \log R|$ small then $\begin{cases} L+R \approx 2L \\ \min(L,R) \approx L \end{cases}$ $\left.\begin{array}{l} \log(L+R) \approx \log(L)+1 \\ \min(\log L, \log R) \approx \log L \end{array}\right\}$ This is the common value plus a $\{0,1\}$-valued increment

If $|\log L - \log R|$ big then $\begin{cases} L+R \approx \max(L,R) \\ \min(L,R) = \min(L,R) \end{cases}$ $\left.\begin{array}{l} \log(L+R) \approx \max(\log L, \log R) \\ \log(\min(L,R)) = \min(\log L, \log R) \end{array}\right\}$ This is just log(value of a random child)

Similar intuition should work for the hierarchical lattice:

$$f_v = \begin{cases} (a,b) \longmapsto a+b & \text{with prob. } p \\ (a,b) \longmapsto \dfrac{ab}{a+b} & \text{with prob. } 1-p \end{cases}$$

<u>Intuition</u>: Suppose $\vec{T}_n(\vec{0})$ is growing on a (stretched) exponential scale.

Write L, R for values at children of root.

If $|\log L - \log R|$ small then $\begin{cases} L+R \approx 2L & \log(L+R) \approx \log(L)+1 \\ \frac{LR}{L+R} \approx L/2 & \log\left(\frac{LR}{L+R}\right) \approx \log(L)-1 \end{cases}$ This is the common value plus a $\{-1,1\}$-valued increment

If $|\log L - \log R|$ big then $\begin{cases} L+R \approx \max(L,R) & \log(L+R) \approx \max(\log L, \log R) \\ \frac{LR}{L+R} \approx \min(L,R) & \log\left(\frac{LR}{L+R}\right) = \min(\log L, \log R) \end{cases}$ This is just log(value of a random child)

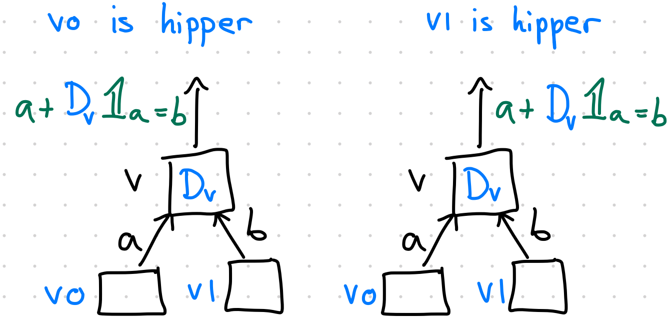
Motivates the following conjecture: in the random hierarchical lattice with $p=\frac{1}{2}$, $\exists\, c>0$ s.t.

$$\frac{\log T_n(\bar{0})}{(c\,n)^{1/3}} - \frac{1}{2} \xrightarrow{d} \mathrm{Beta}(2,2)$$

Theorem (Totally asymmetric lazy) $\dfrac{T(\vec{0})}{(2n)^{1/2}} \xrightarrow{d} \mathrm{Beta}(2,1)$

Proof Idea

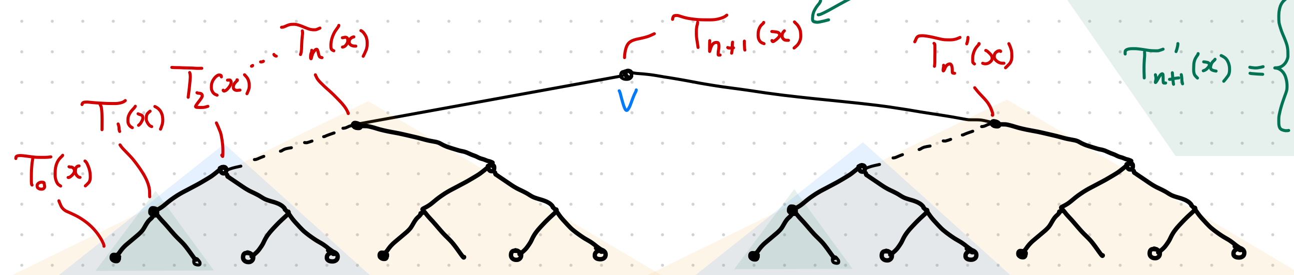
Original dynamics:



$D_v \sim \mathrm{Bernoulli}(\frac{1}{2})$

By symmetry, can assume left child is always chosen.

For inputs $x = (x_v, v \in \mathcal{L})$, useful notation: $T_n(x) := T_n((x_v, v \in \mathcal{L}_n))$



$$T_{n+1}'(x) = \begin{cases} T_n(x) & \text{if } T_n(x) \neq T_n'(x) \\ T_n(x) + D_v & \text{if } T_n(x) = T_n'(x) \end{cases}$$

Proof Idea (Totally asymmetric case)

Let $p_n(k) = \mathbb{P}(T_n(\vec{0}) = k)$

Then $p_{n+1}(k) = \underbrace{p_n(k)(1-p_n(k)) + \frac{1}{2}p_n(k-1)^2 + \frac{1}{2}p_n(k)^2}$

$p_n(k)(1-p_n(k)) \leftarrow$ left child $= k$, right child $\neq k$

$\frac{1}{2}p_n(k-1)^2 \leftarrow$ both $= k-1$, make a step

$\frac{1}{2}p_n(k)^2 \leftarrow$ both $= k$, be lazy

Rearranging gives $\boxed{p_{n+1}(k) - p_n(k) = \frac{1}{2}\left(p_n(k)^2 - p_n(k-1)^2\right)}$

This is a discretization of the inviscid Burgers' equation $\frac{\partial}{\partial t}u(x,t) = -\frac{1}{2}\frac{\partial}{\partial x}\left(u(x,t)^2\right)$

$\downarrow$

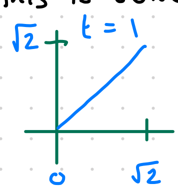
$u_t = -\frac{1}{2}(u^2)_x = -u\,u_x$

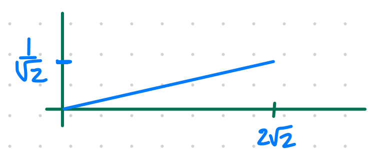
So we are trying to solve the (measure-valued) initial-value problem

$$\begin{cases} u_t = -u\,u_x, & t \geq 0,\ x \in \mathbb{R} \\ u_0(x) = \delta_0(x) = \mathbb{1}_{[x=0]} & \text{(Dirac mass at } 0) \end{cases}$$

Ignoring space-time points of discontinuity, this is solved* by $u: \mathbb{R} \times [0,\infty) \to \mathbb{R}$ given by

$$u(x,t) = \begin{cases} x/t, & 0 \leq x < \sqrt{2t} \\ 0 & \text{otherwise} \end{cases}$$





$t = 4$

Note $u(t,x)$ is always a prob. dist.: the density of a scaled Beta(2,1).

* But solution is not unique!

First step Start Burgers' from a smooth initial condition of the form $u_0(x) = \frac{x}{t_0} \mathbb{1}_{0 \le x \le \sqrt{2t_0}}$

(think of t_0 as small).

Probabilistically what does this mean?

u_0 is density of $\sqrt{2t_0} \cdot B$ where $B \sim \mathrm{Beta}(2,1)$

Fix $M > 0$ and define $u_j^0(M) = M \int_{j/M}^{(j+1)/M} u_0(x)\,dx$. for $j \ge 0$ s.t. $\frac{j}{M} \le \sqrt{2t_0}$

Then $\sum_j u_j^0(M) = 1$, so $(u_j^0(M), j \ge 0)$ defines a probability distribution on $\{0, 1, \dots \lfloor M \cdot \sqrt{2t_0} \rfloor\}$

Let $X^M = (X_v^M, v \in \mathbb{Z})$ be vector of IIDs with $\mathbb{P}(X_v^M = j) = u_j^0(M)$ (discretization of u_0 at mesh size $\frac{1}{M}$)

$T_n(X^M)$ is value of TALSHRW when initial distribution is $\frac{1}{M}$-mesh discretization of $\sqrt{2t_0} \cdot B$.

Lemma We have $\mathbb{P}(T_n(X^M) = j) = \frac{1}{M} \cdot u_j^n(M)$, where $(u_j^n(M))_{n \ge 0, j \ge 0}$ is defined by the recurrence $M \cdot u_j^{n+1} = M \cdot u_j^n - \frac{1}{2}\left((u_j^n)^2 - (u_{j-1}^n)^2\right)$.

Proof Easy induction $\square$

Second step Convergence of the fine-mesh approximation.

The spatial mesh is $\frac{1}{M}$. We take a temporal mesh of $\frac{1}{M^2}$.

$$U_M(t,x) = u^{\lfloor tM^2 \rfloor}_{\lfloor xM \rfloor}(M) = \mathbb{P}\left(T_{\lfloor tM^2 \rfloor}(X^M) = \lfloor xM \rfloor\right) \quad \text{for } t, x \geq 0.$$

Call U_M a $\frac{1}{M}$-fine mesh approximation of Burgers' equation

Theorem (Evje & Karlsen, 2000)

From a bounded variation initial condition, the $\frac{1}{M}$-fine mesh approximation converges to the BV entropy solution u of Burgers' equation almost everywhere on $\mathbb{R} \times [0,\infty)$, and for any compact $C \subset \mathbb{R} \times [0,\infty)$, $\int_C |U^M(x,t) - u(x,t)|\, dx\, dt \longrightarrow 0$.

- BV $\longrightarrow$ Bounded variation
- BV entropy solution $\rightarrow$ The correct solution of our problem (verifying this takes some work)

Conclusion $U_M \rightarrow u$ defined by $u(t,x) = \frac{x}{t+t_0} \mathbb{1}_{0 \leq x \leq \sqrt{2(t+t_0)}}$

Implication for TALSHRW

Corollary For $\varepsilon > 0$ small, if $U = \text{Unif}[1-\varepsilon, 1+\varepsilon]$ is independent of X, then as $M \to \infty$,

$$\frac{T_{\lfloor U M^2 \rfloor}(X^M)}{\sqrt{2(t_0+U)}\, M} \xrightarrow{d} \text{Beta}(2,1).$$

Proof: For any compact $C \subset \mathbb{R} \times [0,\infty)$,

$$\iint_C \left| \mathbb{P}\left(T_{\lfloor tM^2 \rfloor}(X^M) = \lfloor xM \rfloor\right) - \frac{x}{t+t_0} \mathbb{1}_{0 \le x \le \sqrt{2(t+t_0)}} \right| dx\, dt \longrightarrow 0$$

Taking $C = \{(x,t) : |t-1| \le \varepsilon, 0 \le x \le a\sqrt{2(t+t_0)}\}$, this yields by the triangle inequality that

$$\int_{[1-\varepsilon,1+\varepsilon]} \left| \mathbb{P}\left(T_{\lfloor tM^2 \rfloor}(X^M) \le a\sqrt{2(t+t_0)}\, M\right) - \int_0^a \frac{x}{t+t_0}\, dx \right| \cdot \frac{1}{2\varepsilon}\, dt \longrightarrow 0 \qquad \text{as } M \to \infty$$

(There are "discretization errors" coming from the floors, but it's easy to see these tend to 0 as $M \to \infty$.)

Since U has density $\frac{1}{2\varepsilon} \mathbb{1}_{|t-1| \le \varepsilon}$, the result follows. $\square$

Last step stochastic domination.

Proposition

If $x = (x_v, v \in \mathcal{L})$ and $y = (y_v, v \in \mathcal{L})$ are such that $x_v \in \mathbb{Z}$, $y_v \in \mathbb{Z}$ and $x_v \leq y_v$ for all $v \in \mathcal{L}$,

then $T_n(x) \preccurlyeq_{st} T_n(y)$ for all $n \geq 1$.

Proof: A straightforward induction. ■

Corollary 1 For all $n, M \in \mathbb{N}$, $T_n(X^M) - \lfloor \sqrt{2 t_0 M} \rfloor \preccurlyeq_{st} T_n(\vec{0}) \preccurlyeq_{st} T_n(\vec{0})$ ■

Allows us to compare all-0 input to random input with $o(M)$ error (recall $t_0 > 0$ is fixed but arbitrarily small).

Corollary 2 For all $M \in \mathbb{N}$, $T_{(1-\varepsilon)M^2}(X^M) \preccurlyeq_{st} T_{UM^2}(X^M) \preccurlyeq_{st} T_{(1+\varepsilon)M^2}(X^M)$

Allows us to compare fixed time near M^2 to random time UM^2.

Since $\frac{1}{\sqrt{2(t_0+U)}\, M} T_{UM^2}(X_M) \xrightarrow{d} \mathrm{Beta}(2,1)$, corollaries yield that $\frac{T_n(X^M)}{\sqrt{2}\, M} \xrightarrow{d} \mathrm{Beta}(2,1)$.


CIRM
HIPSTERS