The Strange Logic of Galton-Watson Trees

Joel Spencer
joint with Moumanti Podder

Courant Institute of Mathematical Sciences
New York University

MSRNE Colloquium
December 2, 2015
Rarely does a mathematical problem convey so much of the flavour of its time, colonialism and male supremacy hand in hand, as well as the underlying concern for a diminished fertility of noble families, paving the way for the crowds from the genetically dubious lower classes.

– Peter Jagers on the Galton-Watson process
Dead white men

Figure: Francis Galton (1822-1911)

Figure: Henry William Watson (1827-1903)
The Poisson Distribution

Parameter $\lambda$, the mean. Poisson $P = P_\lambda$

$$\Pr[P = i] = e^{-\lambda} \frac{\lambda^i}{i!} \text{ for } i = 0, 1, 2, \ldots$$
Parameter $\lambda$, the mean. Poisson $P = P_{\lambda}$

$$\Pr[P = i] = e^{-\lambda} \frac{\lambda^i}{i!} \text{ for } i = 0, 1, 2, \ldots$$

$\text{BIN}[n, \frac{\lambda}{n}] \rightarrow P_{\lambda}$
The Poisson Distribution

Parameter $\lambda$, the mean. Poisson $P = P_\lambda$

$$\Pr[P = i] = e^{-\lambda} \frac{\lambda^i}{i!} \text{ for } i = 0, 1, 2, \ldots$$

$\text{BIN}[n, \frac{\lambda}{n}] \rightarrow P_\lambda$

**The Amazing Property:** Let Eve have $P_\lambda$ children. Suppose each child, independently, is of type $j$ with probability $p_j$. ($1 \leq j \leq M$) Then, equivalently, Eve has $P_{p_j \lambda}$ children of type $j$ and the number of children of different types are mutually independent.
TIME TO TOSS
Alon/JS The Probabilistic Method, Edition 1
Alon/JS The Probabilistic Method, Edition 2
Alon/JS The Probabilistic Method, Edition 3
TIME TO BUY
Alon/JS The Probabilistic Method, Edition FOUR!
Available at Fine Bookstores Everywhere!
Parameter $\lambda$. Random rooted tree $T = T_\lambda$. 
Eve (root) has $P$ children 
Each child has $P$ children 
Grandchildren, greatgrandchildren, etc. 
May be finite or infinite.
Fictitious Continuation

\[ X_1, X_2, \ldots \text{ mutually independent, } X_i \sim \text{Pois}(\lambda) \]

\( i \)-th node has \( X_i \) children.

Example: 2, 0, 1, 0, 3, 2, \ldots

Eve has Anna and Barbara \((X_1 = 2)\)
Fictitious Continuation

\[X_1, X_2, \ldots \text{ mutually independent, } X_i \sim \text{Pois}(\lambda)\]

\(i\)-th node has \(X_i\) children.

Example: 2, 0, 1, 0, 3, 2, \ldots
Eve has Anna and Barbara \((X_1 = 2)\)
Anna has no children \((X_2 = 0)\)
Fictitious Continuation

$X_1, X_2, \ldots$ mutually independent, $X_i \sim \text{Pois}(\lambda)$
i-th node has $X_i$ children.
Example: 2, 0, 1, 0, 3, 2, \ldots
Eve has Anna and Barbara ($X_1 = 2$)
Anna has no children ($X_2 = 0$)
Barbara has Carol ($X_3 = 1$)
Fictitious Continuation

\[ X_1, X_2, \ldots \text{ mutually independent, } X_i \sim \text{Pois}(\lambda) \]

i-th node has \( X_i \) children.

Example: 2, 0, 1, 0, 3, 2, \ldots

Eve has Anna and Barbara \((X_1 = 2)\)
Anna has no children \((X_2 = 0)\)
Barbara has Carol \((X_3 = 1)\)
Carol has no children \((X_4 = 0)\)
Fictitious Continuation

\[ X_1, X_2, \ldots \] mutually independent, \[ X_i \sim \text{Pois}(\lambda) \]

\( i \)-th node has \( X_i \) children.

Example: 2, 0, 1, 0, 3, 2, …

Eve has Anna and Barbara \((X_1 = 2)\)

Anna has no children \((X_2 = 0)\)

Barbara has Carol \((X_3 = 1)\)

Carol has no children \((X_4 = 0)\)

Process STOPS.
$X_1, X_2, \ldots$ mutually independent, $X_i \sim \text{Pois}(\lambda)$

$i$-th node has $X_i$ children.

Example: $2, 0, 1, 0, 3, 2, \ldots$

Eve has Anna and Barbara ($X_1 = 2$)
Anna has no children ($X_2 = 0$)
Barbara has Carol ($X_3 = 1$)
Carol has no children ($X_4 = 0$)
Process STOPS.

Fictitious Continuation (convenient!)
Danielle (no parent) has Florence, Gabrielle, Harriet
Never Ends
One Grandchild

A: Some node has one child with one child.

**Finite** State Space \( \Sigma = \{\bullet, \bullet, \bullet\} \)

Count: 0, 1, \( \omega \) (\( \omega \) means \( \geq 2 \).)

\( \bullet \): Yes; \( \bullet \): One Child, not \( \bullet \); \( \bullet \): All Else

Node state determined by count of each state of children.

\[
(1, -, -) \Rightarrow \bullet \\
(\omega, -, -) \Rightarrow \bullet \\
(0, 1, 0) \Rightarrow \bullet \\
(0, 0, 1) \Rightarrow \bullet
\]

\( x = \Pr[\bullet], \ y = \Pr[\bullet], \ z = \Pr[\bullet] = 1 - x - y \)

\( x = 1 - e^{-x\lambda} + y\lambda e^{-\lambda}, \ y = z\lambda e^{-\lambda} \)

Solution \( x = f_A(\lambda) \) unique.
Figure: Probability $p(\lambda)$ of having no node with one child with one child.
Immortality

$B$: The tree is infinite. ☐: Yes; ☐: No

Count 0, $\omega$.

$(\omega, -) \Rightarrow ☐$

$(0, -) \Rightarrow ☐$

$x = \Pr[\bullet]$

$x = 1 - e^{-x\lambda}$

Solution $x = f_B(\lambda)$ not unique when $\lambda > 1$. 
Probability of Immortality, as a function of $\lambda$

**Figure:** Probability $\rho(\lambda)$ of having an infinite tree, $0 \leq \lambda \leq 3$. 
Draconian Fecundity

If 0 or 1 Living Children: Death!

\( C: \) Root Survives. •: Yes; •: No

Count 0, 1, \( \omega \).

\[(\omega, -) \Rightarrow \bullet \]
\[(1, -) \Rightarrow \bullet \]
\[(0, -) \Rightarrow \bullet \]

\[ x = \Pr[\bullet] \]

\[ x = 1 - e^{-x\lambda} - x\lambda e^{-x\lambda} \]

Solution \( x = f_C(\lambda) \) **not unique** when \( \lambda \geq \lambda_0 \).
Draconian Fecundity, various values of $\lambda$

Figure: When $\lambda = 2.7$

Figure: When $\lambda = 3.9$
Draconian Fecundity, function of $\lambda$

Figure: Probability $x(\lambda)$ of non-empty 2-core containing root.
The First Order World

Constant Symbol: root
Equality: \( x = y \),
Parent: \( \pi(x) = y \) (\( y \) is parent of \( x \), binary predicate),
Variable Symbols \( x, y, z \ldots \),
Boolean \( \lor, \land, \neg, \rightarrow, \leftrightarrow \), etc,
Quantification \( \forall_x, \exists_y \) over vertices only.

\[ A: \exists_x \exists_y \exists_z \pi(y) = x \land \pi(z) = y \]
\[ \land [\forall_w \pi(w) = y \rightarrow z = w] \land [\forall_v \pi(v) = x \rightarrow y = v]. \]

\( B \) not expressible in First Order Language.
Quantifier depth \( qf \): Nesting of Quantifiers.
\( qf(A) = 4. \)
Locality

$k$ fixed. $\Sigma = \Sigma_k = \text{all } \equiv_k$ Ehrenfeucht classes.
$\text{RAD} = \text{RAD}(k) = 4^k$ (not best possible)
$\text{BALL}[v, \text{RAD}]$ of finitely many balltypes.

**Theorem**

*If for all classes $\tau$ there exist $v_1, \ldots, v_k$*

1. $\text{BALL}[v_i, \text{RAD}]$ of balltype $\tau$
2. $v_i$ far $(2 \cdot \text{RAD})$ from root
3. $v_i$ far $(2 \cdot \text{RAD})$ from each other

*then class $\sigma \in \Sigma_k$ is determined by the balltype of $\text{BALL}[\text{root}, \text{RAD}]$*
Universality

Let $\text{UNIV} = \text{UNIV}_k$ be a rooted tree. Suppose for all balltypes $\tau$ there exist $v_1, \ldots, v_k$

1. $\text{BALL}[v_i, \text{RAD}]$ of balltype $\tau$
2. $v_i$ far $(2 \cdot \text{RAD})$ from each other

Suppose $\exists v \in T$ far from root so that $T(v) \cong \text{UNIV}$. Then Ehrenfeucht value $\sigma$ of $T$ determined by the balltype of $\text{BALL}[\text{root}, \text{RAD}]$.

**Theorem**

There exists a finite $\text{UNIV} = \text{UNIV}_k$

We fix one such $\text{UNIV}_k$ of depth $D = D_k$
Tree Automaton

Definition

1. **Finite** state space $\Sigma$,
2. Accepted States $\Gamma \subset \Sigma$
3. State of $v$ determined by tree rooted at $v$
4. $k \geq 1$. We set $C = \{0, 1, \ldots, k - 1, \omega\}$,

such that the state of a node is determined by the “number” $n_\sigma \in C$ of children of each state $\sigma \in \Sigma$. A *recursive* if $A$ holds iff root has accepted state.
Theorem

*First Order* $A$ are recursive.

*Outline:* $k = qf(A)$. $\Sigma$ all $\equiv_k$ Ehrenfeucht classes.

Duplicator wins Ehrenfeucht Game.

*Remark:* $|\Sigma|$ grows like $\text{TOWER}(k)$.

*Monadic second* order properties are recursive.
Solution as Fixed Point

\[ D = \text{set of distributions } \bar{x} \text{ on } \Sigma. \ (D \subset R^{|\Sigma|}). \]

\[ SOL(\lambda) \in D \text{ is distribution of state of } T_\lambda. \]

Map \( \psi_\lambda : D \to D. \)
Solution as Fixed Point

\[ D = \text{set of distributions } \tilde{x} \text{ on } \Sigma. \quad (D \subset R^{|\Sigma|}). \]

\[ \text{SOL}(\lambda) \in D \text{ is distribution of state of } T_\lambda. \]

Map \( \psi_\lambda : D \rightarrow D. \)

1. \( v \) has Poisson mean \( \lambda \) children.
Solution as Fixed Point

\[ D = \text{set of distributions } \tilde{x} \text{ on } \Sigma. \quad (D \subset R^{|\Sigma|}). \]

\[ \text{SOL}(\lambda) \in D \text{ is distribution of state of } T_\lambda. \]

Map \( \psi_\lambda : D \to D. \)

1. \( v \) has Poisson mean \( \lambda \) children.
2. Each child has \( \sigma \in \Sigma \) i.i.d. with distribution \( \tilde{x}. \)
Solution as Fixed Point

\[ D = \text{set of distributions } \vec{x} \text{ on } \Sigma. \quad (D \subset R^{|\Sigma|}). \]

\[ \text{SOL}(\lambda) \in D \text{ is distribution of state of } T_\lambda. \]

Map \( \psi_\lambda : D \to D. \)

1. \( v \) has Poisson mean \( \lambda \) children.
2. Each child has \( \sigma \in \Sigma \) i.i.d. with distribution \( \vec{x} \).
3. OR: \( v \) has Poisson mean \( x_\sigma \lambda \) children in state \( \sigma \).
Solution as Fixed Point

\[ D = \text{set of distributions } \vec{x} \text{ on } \Sigma. \ (D \subset R^{|\Sigma|}). \]

\[ \text{SOL}(\lambda) \in D \text{ is distribution of state of } T_\lambda. \]

Map \( \psi_\lambda : D \rightarrow D. \)

1. \( \nu \) has Poisson mean \( \lambda \) children.
2. Each child has \( \sigma \in \Sigma \) i.i.d. with distribution \( \vec{x} \).
3. OR: \( \nu \) has Poisson mean \( x_\sigma \lambda \) children in state \( \sigma \).
4. State \( \tau \) of \( \nu \) determined.
Solution as Fixed Point

\[ D = \text{set of distributions } \tilde{x} \text{ on } \Sigma. \quad (D \subset R^{\mid \Sigma \mid}). \]

\[ \text{SOL}(\lambda) \in D \text{ is distribution of state of } T_{\lambda}. \]

Map \( \psi_{\lambda} : D \to D. \)

1. \( v \) has Poisson mean \( \lambda \) children.
2. Each child has \( \sigma \in \Sigma \) i.i.d. with distribution \( \tilde{x}. \)
3. OR: \( v \) has Poisson mean \( x_\sigma \lambda \) children in state \( \sigma \).
4. State \( \tau \) of \( v \) determined.
5. \( \psi_{\lambda}(\tilde{x}) \) is induced distribution \( \tilde{y} \) on \( v \).
Solution as Fixed Point

\[ D = \text{set of distributions } \vec{x} \text{ on } \Sigma. \ (D \subset R^{|\Sigma|}). \]

\[ \text{SOL}(\lambda) \in D \text{ is distribution of state of } T_\lambda. \]

Map \( \psi_\lambda : D \rightarrow D. \)

1. \( \nu \) has Poisson mean \( \lambda \) children.
2. Each child has \( \sigma \in \Sigma \) i.i.d. with distribution \( \vec{x}. \)
3. OR: \( \nu \) has Poisson mean \( \chi_\sigma \lambda \) children in state \( \sigma. \)
4. State \( \tau \) of \( \nu \) determined.
5. \( \psi_\lambda(\vec{x}) \) is induced distribution \( \vec{y} \) on \( \nu. \)

\( \text{SOL}(\lambda) \) is a fixed point of \( \psi_\lambda. \)
Results

When $A$ is a first order property

1. $\Psi_\lambda : D \to D$ is contracting,
Results

When \( A \) is a first order property

1. \( \psi_\lambda : D \rightarrow D \) is contracting,
2. The fixed point of \( \psi_\lambda \) is unique,
When $A$ is a first order property

1. $\psi_\lambda : D \to D$ is contracting,
2. The fixed point of $\psi_\lambda$ is unique,
3. $SOL(\lambda)$ is a real analytic function.
Results

When $A$ is a first order property

1. $\psi_\lambda : D \rightarrow D$ is contracting,
2. The fixed point of $\psi_\lambda$ is unique,
3. $SOL(\lambda)$ is a real analytic function.
4. Conditioning on $T_\lambda$ infinite, $A$ depends only on $BALL[root, RAD]$. 
Rapidly Determined

$T_\lambda$ given by $X_1, X_2, \ldots$ (fictitious continuation). Quite Surely: Exponentially small failure probability.

**Definition**

A (property or function) is *rapidly determined* if quite surely $A$ is tautologically determined by $X_1, \ldots, X_s$. 
$T_\lambda$ given by $X_1, X_2, \ldots$ (fictitious continuation).
Quite Surely: Exponentially small failure probability.

**Definition**

A (property or function) is *rapidly determined* if quite surely $A$ is tautologically determined by $X_1, \ldots, X_s$.

**Theorem**

*First Order* $A$ are rapidly determined. For each $k, \lambda$, $\Sigma = \Sigma_k$ rapidly determined.
One Grandchild

\(I_i: i\) has one child with one child. (fictitious continuation).
\[ Y = \sum_{i \leq s} I_i, (\lambda \epsilon < 1 - \epsilon). \]

\[ E[Y] = s \epsilon (\lambda e^{-\lambda})^2. \]

Martingale \(Y_0, Y_1, \ldots, Y_s: Y_i = E[Y|X_1, \ldots, X_i].\)
Quite surely \(Y_s = Y.\)

Lipschitz: \(|Y_i - Y_{i-1}| \leq 2.\) Only \(I_i, I_{\pi(i)}\) affected by \(X_i.\)
Azuma: Quite surely \(Y \neq 0.\)

1. Either early end (\(A\) determined),
2. or no early end (no fiction, \(A\) quite surely true).

Therefore: \(A\) is rapidly determined.
Unique Fixed Point

To find $\vec{y} = \Psi^s(\vec{x})$:

1. $\nu$ generates $T_\lambda$ to generation $s$.
2. Each $w$ at generation $s$ given state with distribution $\vec{x}$.
3. State $\tau$ of $\nu$ determined.
4. $\Psi^s(\vec{x})$ is induced distribution $\vec{y}$ on $\nu$.

Suppose $\Sigma$ quite surely determined.
$\Psi^s(\vec{x})$ is quite surely independent of $\vec{x}$.
$\lim_s \Psi^s(\vec{x}) = \vec{z}$, independent of $\vec{x}$.
$\Psi^s$ has unique fixed point.
Contraction I

Total Variation $TV[\vec{x}, \vec{y}] = \epsilon$.
Couple: $\Pr[\vec{x} \neq \vec{y}] \leq \epsilon$.
w nodes $v_1, \ldots, v_w$ at level $s$
$TV[\Psi^s(\vec{x}), \Psi^s(\vec{y})] \leq w\epsilon$. 
Total Variation $TV[\vec{x}, \vec{y}] = \epsilon$.
Couple: $\Pr[\vec{x} \neq \vec{y}] \leq \epsilon$.
$w$ nodes $v_1, \ldots, v_w$ at level $s$
$TV[\psi^s(\vec{x}), \psi^s(\vec{y})] \leq w\epsilon$.
Two Stage Process
$s = s_0 + D$, $s_0 > 4^k$, large, $D$ depth of $UNIV_k$
Random $T_0$ to depth $s_0$ with leaves $v_1, \ldots, v_t$.
Each $v_i$ generates GW for $D$ more levels.
Define $\psi^*(\vec{x})$ dependent on $T_0$. 

1. $T_0$ fixed at depth $s_0$ with leaves $v_1, \ldots, v_t$
Define $\Psi^*(\vec{x})$ dependent on $T_0$.

1. $T_0$ fixed at depth $s_0$ with leaves $v_1, \ldots, v_t$
2. Each $v_i$ generates GW for $D$ more levels
Contraction II

Define $\psi^*(\vec{x})$ dependent on $T_0$.

1. $T_0$ fixed at depth $s_0$ with leaves $v_1, \ldots, v_t$
2. Each $v_i$ generates GW for $D$ more levels
3. Nodes $w_1, \ldots, w_y$ at level $s = s_0 + D$
4. Each $w_j$ given state $\sigma \in \Sigma$ with distribution $\vec{x}$
Contraction II

Define $\Psi^*(\vec{x})$ dependent on $T_0$.

1. $T_0$ fixed at depth $s_0$ with leaves $v_1, \ldots, v_t$
2. Each $v_i$ generates GW for $D$ more levels
3. Nodes $w_1, \ldots, w_y$ at level $s = s_0 + D$
4. Each $w_j$ given state $\sigma \in \Sigma$ with distribution $\vec{x}$
5. Determines state $\tau \in \Sigma$ for root
6. $\Psi^*(\vec{x})$ is distribution of $\tau$
Fix $T_0$ to depth $s_0$ with leaves $v_1, \ldots, v_t$.

$Y$ nodes at depth $s = s_0 + D$. $E[Y] = t \lambda^D = O(t)$. 

$\Psi^*(\vec{x})$: Values at depth $s$ determine value at root.

$GOOD$: Some $T(v_i) \equiv UNIV$. $BAD = \neg GOOD$.

$\Pr[T(v_i) \equiv UNIV] = \epsilon_1 > 0$

$\Pr[BAD] \leq (1 - \epsilon_1)^t$

If $GOOD$, $\Psi^*(\vec{x})$ independent of $\vec{x}$.

$E[TV[\Psi^*(\vec{x}), \Psi^*(\vec{y})]] \leq E[Y \chi(BAD)\epsilon]$.

$Y, BAD$ possibly correlated by large deviations: $\leq [c_1(t \lambda^D)(1 - \epsilon_1)^t] \epsilon$

Any $T_0$: $\leq K\epsilon$. 
Double randomness:

\[ E[TV[\Psi^s(\vec{x}), \Psi^s(\vec{y})]] \leq E_{T_0} TV[\Psi^*(\vec{x}), \Psi^*(\vec{y})] \]

Select \( s_0 \) so that, with failure probability \( \leq \frac{1}{2^K} \), \( T_0 \) determines Ehrenfeucht value. In that case \( \Psi^*(\vec{x}) \) is independent of \( \vec{x} \). Otherwise the expansion is by at most a fact \( K \). Thus

\[ TV[\Psi^s(\vec{x}), \Psi^s(\vec{y})] \leq \frac{1}{2^K} K \epsilon \leq \epsilon/2 \]
\( T^\inf_\lambda \): Conditioned on \( T \) infinite. (\( \lambda > 1 \).)

\textit{AS}: Almost Sure Theory.

\( T \): Schema on \( k \): There exists \( v \), at least \( 2 \cdot \text{RAD}(k) \) from root, with \( T(v) \cong \text{UNIV}_k \).

**Theorem**

\( T \) generates \( \text{AS} \).

\( B_i, 1 \leq i \leq M = M(k) \), give balltype of \( \text{BALL}[\text{root}, \text{RAD}] \).

**Theorem**

\( T + B_i \) \( k \)-complete
Corollary
\[ \text{Pr}[A] \text{ is nice function of } \lambda \]

Pr[A] = 1 iff \( \mathcal{T} + B_i \models A \) for all \( i \)
In \( \mathcal{T} \), precisely one of \( B_i \) hold
Therefore \( \mathcal{T} \models A \)

Corollary
\[ \text{Pr}[A] \text{ sum of Pr}[B_i] \text{ with } \mathcal{T} + B_i \models A. \]
MSO: Quantify over sets of vertices.
EMSO: Set quantification existential and at start
Tree Automaton $\implies$ MSO (nonunique)
MSO $\implies$ Tree Automaton (unique)
Tree Automaton $\implies$ EMSO (unique)
Tree Automaton $\implies$ Equation $\Psi(\vec{x}) = \vec{x}$ (unique)
Equation $\implies$ Solution $a = \sum_{\sigma \in \Gamma} x_{\sigma}$ (nonunique)
The Strange Logic of Galton–Watson Trees

Joel Spencer
joint with Moumanti Podder

Probability Lower Bound

Tree Automaton \Rightarrow \text{Equation } \Psi(\vec{x}) = \vec{x} \Rightarrow \text{Solution } a.
Tree Automaton \Rightarrow \text{EMSO A}

\textbf{Theorem}

\[ \Pr[A] \geq a \]

\( n \)-generation \( v \) given state by \( \vec{x} \).

\( \text{YES}_n(T) \): Induced root state accepted.

\[ \Pr[\text{YES}_n(T)] = a \]

\[ \Pr[\Pr[\text{YES}_n(T)] \geq \epsilon] \geq a - \epsilon \]

\( \text{Yes}(T) \): \text{YES}_n(T) for infinitely many \( n \)

\[ \Pr[\text{Yes}(T)] \geq a - \epsilon \]
Axiom of Choice!

Yes($T$): For infinitely many $n$ can give states to first $n$ generations.

**Compactness** $\Rightarrow$ can give states to $T \Rightarrow A$

**BUT** largest soln not always probability!

Green,Red: Green iff *exactly* one green child

$x = x\lambda e^{-x\lambda}$

$x = \frac{\ln \lambda}{\lambda}$

$\neq$ probability $T$ is infinite
Open Question I

Which tree automaton have EMSO $A$ with $\Pr[A]$ equal maximal solution $a$. 
States Green, Red for rooted trees. 
$T(v)$ Green iff $v$ has children $w_1, w_2$ (maybe more) with $T(w_i)$ Green.
Examples: Draconian Fecundity, False

**Question:** Is there such a $D$ with $\Pr[D]$ the smaller nonzero solution to system of equations.
**Conjecture:** No.
Open Question III

Which Tree Automaton have unique solutions to their corresponding system of equations? Which Tree Automaton have corresponding properties unique up to probability zero?
Any new possibility that existence acquires, even the least likely, transforms everything about existence. – from Slowness by Milan Kundera