Combinatorial Markov chains

R. Grübel
Leibniz Universität Hannover

AofA, Menorca 2013

Combinatorial Markov chains

Combinatorial Markov chains

- A combinatorial family is a set $\mathbb{F}$ with a size function $\phi : \mathbb{F} \to \mathbb{N}$ such that
$$\mathbb{F}_n := \{x \in \mathbb{F} : \phi(x) = n\}$$
is finite for all $n \in \mathbb{N}$. We assume that $\mathbb{F}_1 = \{e\}$.

Combinatorial Markov chains

- A combinatorial family is a set $\mathbb{F}$ with a size function $\phi : \mathbb{F} \to \mathbb{N}$ such that
$$\mathbb{F}_n := \{x \in \mathbb{F} : \phi(x) = n\}$$
is finite for all $n \in \mathbb{N}$. We assume that $\mathbb{F}_1 = \{e\}$.

- A combinatorial Markov chain (CMC) is a Markov chain $X = (X_n)_{n \in \mathbb{N}}$ that is adapted to $\mathbb{F}$:
$$P(X_n \in \mathbb{F}_n) = 1 \quad \text{for all } n \in \mathbb{N}.$$

Combinatorial Markov chains

- A combinatorial family is a set $\mathbb{F}$ with a size function $\phi : \mathbb{F} \to \mathbb{N}$ such that
$$\mathbb{F}_n := \{x \in \mathbb{F} : \phi(x) = n\}$$
is finite for all $n \in \mathbb{N}$. We assume that $\mathbb{F}_1 = \{e\}$.

- A combinatorial Markov chain (CMC) is a Markov chain $X = (X_n)_{n\in\mathbb{N}}$ that is adapted to $\mathbb{F}$:
$$P(X_n \in \mathbb{F}_n) = 1 \quad \text{for all } n \in \mathbb{N}.$$

AofA context: Consider an sequential algorithm that transforms a sequence $(\eta_n)_{n\in\mathbb{N}}$ of input values into a sequence $(x_n)_{n\in\mathbb{N}}$ of combinatorial objects such that

$$x_n \in \mathbb{F}_n, \quad x_{n+1} \text{ depends on } x_n \text{ and } \eta_{n+1} \text{ only.}$$

For i.i.d. input this algorithm generates a CMC.

Growth / Transition structure

Growth / Transition structure

On $\mathbb{F}$ we have a relation '$\hookrightarrow$' with

$$x \hookrightarrow y \implies \phi(x) + 1 = \phi(y).$$

Growth / Transition structure

On $\mathbb{F}$ we have a relation '$\hookrightarrow$' with

$$x \hookrightarrow y \implies \phi(x) + 1 = \phi(y).$$

This turns $\mathbb{F}$ into a directed graph with edges

$$E(\mathbb{F}) = \{(x, y) : x, y \in \mathbb{F}, x \hookrightarrow y\}.$$

Such structures $(\mathbb{F}, \hookrightarrow)$ appear elsewhere, e.g. as Bratelli diagrams.

On $\mathbb{F}$ we have a relation '$\hookrightarrow$' with

$$x \hookrightarrow y \implies \phi(x) + 1 = \phi(y).$$

This turns $\mathbb{F}$ into a directed graph with edges

$$E(\mathbb{F}) = \{(x, y) : x, y \in \mathbb{F}, x \hookrightarrow y\}.$$

Such structures $(\mathbb{F}, \hookrightarrow)$ appear elsewhere, e.g. as Bratelli diagrams.

More assumptions:

- weak connectedness – there is a path from the root to each node,

$$\forall\, x \in \mathbb{F} \;\exists\, y_i \in \mathbb{F}_i,\; i = 1, \ldots, n := \phi(x) :$$
$$y_1 = e,\; y_n = x,\; (y_{i-1}, y_i) \in E(\mathbb{F}) \text{ for } i = 2, \ldots, n$$

- the transition probabilities $p(x, y) = P(X_{n+1} = y | X_n = x)$ are adapted to this structure,

$$p(x, y) > 0 \iff x \hookrightarrow y,$$

Example 1: Random walk, urns, and special numbers

Example 1: Random walk, urns, and special numbers

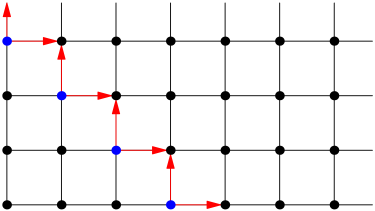
$\mathbb{F} = \mathbb{N}_0 \times \mathbb{N}_0, \;\; \phi(i,j) = i + j + 1,$

$(i,j) \hookrightarrow (i+1,j), \;\; (i,j) \hookrightarrow (i,j+1)$

Example 1: Random walk, urns, and special numbers

$\mathbb{F} = \mathbb{N}_0 \times \mathbb{N}_0, \;\; \phi(i,j) = i + j + 1,$

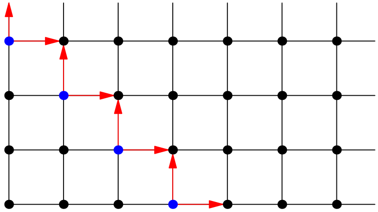
$(i,j) \hookrightarrow (i+1,j), \;\; (i,j) \hookrightarrow (i,j+1)$



Example 1: Random walk, urns, and special numbers

$\mathbb{F} = \mathbb{N}_0 \times \mathbb{N}_0, \;\; \phi(i,j) = i + j + 1,$

$(i,j) \hookrightarrow (i+1,j), \;\; (i,j) \hookrightarrow (i,j+1)$



CMC	$p\big((i,j),(i+1,j)\big)$	$\mathcal{L}(X_n)$
north-east random walk	param. $\theta \in (0,1)$	$\binom{i+j}{i}\,\theta^i(1-\theta)^j$
"	$\theta = 1/2$	$\sim$ binomial coefficients
record walk	$1/(i+j+2)$	$\sim$ Stirling(1) numbers
Friedman urn	$(j+1)/(i+j+2)$	$\sim$ Euler numbers
Pólya urn	$(i+1)/(i+j+2)$	uniform distribution

Example 2: Partitions (of integers)

Example 2: Partitions (of integers)

$$\mathbb{F}_n := \left\{ \lambda = (\lambda_1, \ldots, \lambda_k) \in \mathbb{N}^\star : \sum_{i=1}^k \lambda_i = n,\ \lambda_1 \geq \cdots \geq \lambda_k \right\},$$

$$\mathbb{F}_n \ni \lambda = (\lambda_1, \ldots, \lambda_k) \hookrightarrow \mu = (\mu_1, \ldots, \mu_l) \in \mathbb{F}_{n+1}$$

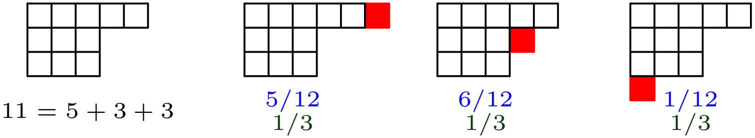
iff $k \leq l$ and $\lambda_j \leq \mu_j$.

Example 2: Partitions (of integers)

$$\mathbb{F}_n := \left\{ \lambda = (\lambda_1, \ldots, \lambda_k) \in \mathbb{N}^\star : \textstyle\sum_{i=1}^{k} \lambda_i = n,\ \lambda_1 \geq \cdots \geq \lambda_k \right\},$$

$$\mathbb{F}_n \ni \lambda = (\lambda_1, \ldots, \lambda_k) \hookrightarrow \mu = (\mu_1, \ldots, \mu_l) \in \mathbb{F}_{n+1}$$

$$\text{iff } k \leq l \text{ and } \lambda_j \leq \mu_j.$$

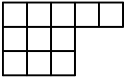


Example 2: Partitions (of integers)

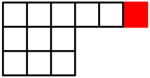
$$\mathbb{F}_n := \left\{ \lambda = (\lambda_1, \dots, \lambda_k) \in \mathbb{N}^\star : \sum_{i=1}^k \lambda_i = n,\ \lambda_1 \geq \cdots \geq \lambda_k \right\},$$

$$\mathbb{F}_n \ni \lambda = (\lambda_1, \dots, \lambda_k) \hookrightarrow \mu = (\mu_1, \dots, \mu_l) \in \mathbb{F}_{n+1}$$

$$\text{iff } k \leq l \text{ and } \lambda_j \leq \mu_j.$$

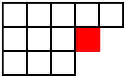


$11 = 5 + 3 + 3$



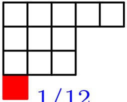
$5/12$

$1/3$



$6/12$

$1/3$



$1/12$

$1/3$

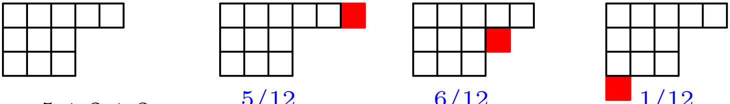
- Thoma: $p(\lambda, \mu) = H(\lambda)/H(\mu)$, hook length formula.

Example 2: Partitions (of integers)

$$\mathbb{F}_n := \left\{ \lambda = (\lambda_1, \ldots, \lambda_k) \in \mathbb{N}^\star : \sum_{i=1}^k \lambda_i = n,\ \lambda_1 \geq \cdots \geq \lambda_k \right\},$$

$$\mathbb{F}_n \ni \lambda = (\lambda_1, \ldots, \lambda_k) \hookrightarrow \mu = (\mu_1, \ldots, \mu_l) \in \mathbb{F}_{n+1}$$

$$\text{iff } k \leq l \text{ and } \lambda_j \leq \mu_j.$$



$11 = 5 + 3 + 3$ — $5/12$, $1/3$ — $6/12$, $1/3$ — $1/12$, $1/3$

- Thoma: $p(\lambda, \mu) = H(\lambda)/H(\mu)$, hook length formula.
- The restaurant (or cycle) chain:
  - If $\mu_j = \lambda_j + 1$, $p(\lambda, \mu) = \lambda_j \cdot \#\{i \geq j : \lambda_i = \lambda_j\}/(n+1)$.
  - If $\mu = (\lambda_1, \ldots, \lambda_k, 1)$: $p(\lambda, \mu) = 1/(n+1)$

Example 3: Binary trees

Example 3: Binary trees

Let $\mathbb{V} = \{0,1\}^\star$ be the set of nodes, with prefix relation '$\leq$'.

$$\mathbb{F}_n := \big\{x \subset \mathbb{V} :\ \#x = n,\ (u \in x, v \leq u) \Rightarrow v \in x\big\},$$
$$x \hookrightarrow y \ \text{ iff } \#y = \#x + 1 \text{ and } x \subset y.$$

Example 3: Binary trees

Let $\mathbb{V} = \{0,1\}^\star$ be the set of nodes, with prefix relation '$\leq$'.

$$\mathbb{F}_n := \big\{x \subset \mathbb{V} : \#x = n, \, (u \in x, v \leq u) \Rightarrow v \in x\big\},$$
$$x \hookrightarrow y \text{ iff } \#y = \#x + 1 \text{ and } x \subset y.$$

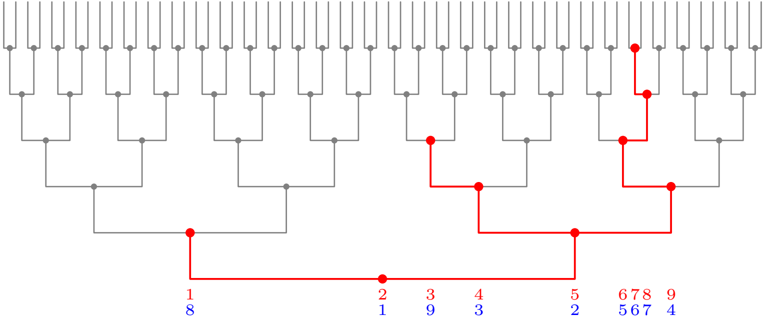
BST chain: $\mathcal{L}(X_{n+1}|X_n = x) = \mathrm{unif}\big(\{y \in \mathbb{F}_{n+1} : x \hookrightarrow y\}\big).$

Example 3: Binary trees

Let $\mathbb{V} = \{0,1\}^\star$ be the set of nodes, with prefix relation '$\leq$'.

$$\mathbb{F}_n := \{x \subset \mathbb{V} : \#x = n, \, (u \in x, v \leq u) \Rightarrow v \in x\},$$
$$x \hookrightarrow y \;\text{ iff } \#y = \#x + 1 \text{ and } x \subset y.$$

BST chain: $\mathcal{L}(X_{n+1} | X_n = x) = \text{unif}\big(\{y \in \mathbb{F}_{n+1} : x \hookrightarrow y\}\big).$



Boundaries: Background

Boundaries: Background

potential theory	Markov processes
Laplace operator	transition probabilities
harmonic functions	martingales
Dirichlet problem	Brownian motion
electrical networks	random walks on graphs

Boundaries: Background

potential theory	Markov processes
Laplace operator	transition probabilities
harmonic functions	martingales
Dirichlet problem	Brownian motion
electrical networks	random walks on graphs

Doob, Hunt, Dynkin, Meyer . . .; an 'icon formula' is

$$h(x) = E_x \phi(B_\tau)$$

Boundaries: Background

potential theory	Markov processes
Laplace operator	transition probabilities
harmonic functions	martingales
Dirichlet problem	Brownian motion
electrical networks	random walks on graphs

Doob, Hunt, Dynkin, Meyer . . .; an ‘icon formula’ is

$$h(x) = E_x \phi(B_\tau)$$

Recent textbook:

W. Woess, Denumerable Markov chains. EMS 2009.

Boundaries: Background

potential theory	Markov processes
Laplace operator	transition probabilities
harmonic functions	martingales
Dirichlet problem	Brownian motion
electrical networks	random walks on graphs

Doob, Hunt, Dynkin, Meyer ...; an ‘icon formula’ is

$$h(x) = E_x \phi(B_\tau)$$

Recent textbook:

W. Woess, Denumerable Markov chains. EMS 2009.

Connection to representation theory, random walks on groups:
Vershik, Kaimanovich, Kerov (book), Gnedin, Okounkov, Woess ...

Doob-Martin compactification: Construction

Doob-Martin compactification: Construction

Let $X = (X_n)_{n \in \mathbb{N}}$ be a CMC on the combinatorial family $\mathbb{F}$.

Doob-Martin compactification: Construction

Let $X = (X_n)_{n \in \mathbb{N}}$ be a CMC on the combinatorial family $\mathbb{F}$.

- Regard $\mathbb{F}$ as a discrete topological space.

Doob-Martin compactification: Construction

Let $X = (X_n)_{n \in \mathbb{N}}$ be a CMC on the combinatorial family $\mathbb{F}$.

- Regard $\mathbb{F}$ as a discrete topological space.
- Define the Martin kernel by

$$K(x, y) := \frac{P(X_n = y | X_m = x)}{P(X_n = y)}, \; x \in \mathbb{F}_m, \; y \in \mathbb{F}_n.$$

Doob-Martin compactification: Construction

Let $X = (X_n)_{n\in\mathbb{N}}$ be a CMC on the combinatorial family $\mathbb{F}$.

- Regard $\mathbb{F}$ as a discrete topological space.
- Define the Martin kernel by

$$K(x,y) := \frac{P(X_n = y | X_m = x)}{P(X_n = y)}, \; x \in \mathbb{F}_m, \; y \in \mathbb{F}_n.$$

- The Doob-Martin compactification $\bar{\mathbb{F}}$ is the minimal compact enlargement of $\mathbb{F}$ that allows for a continuous extension of the (bounded) functions $y \mapsto K(x,y)$, $x \in \mathbb{F}$.

Doob-Martin compactification: Construction

Let $X = (X_n)_{n\in\mathbb{N}}$ be a CMC on the combinatorial family $\mathbb{F}$.

- Regard $\mathbb{F}$ as a discrete topological space.
- Define the Martin kernel by

$$K(x,y) := \frac{P(X_n = y | X_m = x)}{P(X_n = y)}, \ x \in \mathbb{F}_m, \ y \in \mathbb{F}_n.$$

- The Doob-Martin compactification $\bar{\mathbb{F}}$ is the minimal compact enlargement of $\mathbb{F}$ that allows for a continuous extension of the (bounded) functions $y \mapsto K(x,y)$, $x \in \mathbb{F}$.

Then $\partial\mathbb{F} := \bar{\mathbb{F}} \setminus \mathbb{F}$ is the Martin boundary.

Doob-Martin compactification: h-transforms

Doob-Martin compactification: h-transforms

Call $h : \mathbb{F} \to \mathbb{R}$ harmonic if it has the 'mean value property':

$$h(x) = \sum_{y \in \mathbb{F}} p(x, y) h(y) \quad \text{for all } x \in \mathbb{F}.$$

Doob-Martin compactification: h-transforms

Call $h : \mathbb{F} \to \mathbb{R}$ harmonic if it has the 'mean value property':

$$h(x) = \sum_{y \in \mathbb{F}} p(x, y) h(y) \quad \text{for all } x \in \mathbb{F}.$$

Let $\mathbb{H}_+$ resp. $\mathbb{H}_b$ be the set of non-negative resp. bounded harmonic functions.

Doob-Martin compactification: h-transforms

Call $h : \mathbb{F} \to \mathbb{R}$ harmonic if it has the 'mean value property':

$$h(x) = \sum_{y \in \mathbb{F}} p(x, y) h(y) \quad \text{for all } x \in \mathbb{F}.$$

Let $\mathbb{H}_+$ resp. $\mathbb{H}_b$ be the set of non-negative resp. bounded harmonic functions. For $h \in \mathbb{H}_+$, $h \neq 0$,

$$p^h(x, y) := \frac{1}{h(x)} \, p(x, y) \, h(y), \; x, y \in \mathbb{F},$$

defines another transition probability on $\mathbb{F}$.

Doob-Martin compactification: *h*-transforms

Call $h : \mathbb{F} \to \mathbb{R}$ harmonic if it has the 'mean value property':

$$h(x) = \sum_{y \in \mathbb{F}} p(x, y)h(y) \quad \text{for all } x \in \mathbb{F}.$$

Let $\mathbb{H}_+$ resp. $\mathbb{H}_b$ be the set of non-negative resp. bounded harmonic functions. For $h \in \mathbb{H}_+$, $h \neq 0$,

$$p^h(x, y) := \frac{1}{h(x)}\, p(x, y)\, h(y), \; x, y \in \mathbb{F},$$

defines another transition probability on $\mathbb{F}$.

With $X_1^h = e$ and transitions p^h we obtain another CMC $X^h = (X_n^h)_{n \in \mathbb{N}}$, the *h*-transform of X.

Doob-Martin compactification: Results

Doob-Martin compactification: Results

(reprDM)

The extensions $x \mapsto K(x, \alpha)$, $\alpha \in \partial\mathbb{F}$, are minimal harmonic, and all $h \in \mathbb{H}_+$ with $h(e) = 1$ are mixtures of these:

$$h(x) = \int_{\partial\mathbb{F}} K(x, \alpha)\, \nu_h(d\alpha)$$

Doob-Martin compactification: Results

(reprDM)

The extensions $x \mapsto K(x, \alpha)$, $\alpha \in \partial\mathbb{F}$, are minimal harmonic, and all $h \in \mathbb{H}_+$ with $h(e) = 1$ are mixtures of these:

$$h(x) = \int_{\partial\mathbb{F}} K(x, \alpha)\, \nu_h(d\alpha)$$

(conv)

As $n \to \infty$, $X_n \to X_\infty \in \partial\mathbb{F}$ almost surely, with $\mathcal{L}(X_\infty)$ the measure ν representing $h \equiv 1$.

Doob-Martin compactification: Results

(reprDM)

The extensions $x \mapsto K(x, \alpha)$, $\alpha \in \partial\mathbb{F}$, are minimal harmonic, and all $h \in \mathbb{H}_+$ with $h(e) = 1$ are mixtures of these:

$$h(x) = \int_{\partial\mathbb{F}} K(x, \alpha)\, \nu_h(d\alpha)$$

(conv)

As $n \to \infty$, $X_n \to X_\infty \in \partial\mathbb{F}$ almost surely, with $\mathcal{L}(X_\infty)$ the measure ν representing $h \equiv 1$.

(struct)

The conditional distribution $\mathcal{L}(X | X_\infty = \alpha)$ is equal to the distribution of the h-transform $\mathcal{L}(X^h)$ with $h = K(\cdot, \alpha)$ (in particular, this is the distribution of a Markov chain).

Doob-Martin compactification: The ‘why’

Doob-Martin compactification: The 'why'

In a CMC $(X_n)_{n\in\mathbb{N}}$ with Martin kernel K,

$$E\big[K(x, X_{m-1})\big|X_m, \ldots, X_n\big] \leq K(x, X_m).$$

Doob-Martin compactification: The 'why'

In a CMC $(X_n)_{n\in\mathbb{N}}$ with Martin kernel K,

$$E\big[K(x, X_{m-1})\big|X_m, \dots, X_n\big] \le K(x, X_m).$$

Hence, for each $n \in \mathbb{N}$, $(Y_k^n, \mathcal{F}_k^n)_{k=1,\dots,n}$ with

$$Y_k^n := K(x, X_{n+1-k}), \quad \mathcal{F}_k^n = \sigma(X_{n+1-k}, \dots, X_n)$$

is a supermartingale.

Doob-Martin compactification: The 'why'

In a CMC $(X_n)_{n\in\mathbb{N}}$ with Martin kernel K,

$$E\big[K(x,X_{m-1})\big|X_m,\ldots,X_n\big] \le K(x,X_m).$$

Hence, for each $n\in\mathbb{N}$, $(Y_k^n,\mathcal{F}_k^n)_{k=1,\ldots,n}$ with

$$Y_k^n := K(x,X_{n+1-k}),\quad \mathcal{F}_k^n = \sigma(X_{n+1-k},\ldots,X_n)$$

is a supermartingale. For the number $U_n(a,b)$ of upcrossings of some interval $[a,b]$ this means

$$EU_n[a,b] \le \frac{1}{b-a}E(Y_n^n - a)^- = \frac{1}{b-a}\big(K(x,e)-a\big)^-.$$

It follows that $(K(x,X_n))_{n\in\mathbb{N}}$ converges almost surely.

Poisson boundary

Poisson boundary

A measurable space $(E, \mathcal{B})$ together with a family ν_x, $x \in \mathbb{F}$, of probability measures on $(E, \mathcal{B})$ such that $\nu_x \ll \nu_e$ for all $x \in \mathbb{F}$ 'is' the Poisson boundary of $(X, \mathbb{F})$ if

$$L^\infty(E, \mathcal{B}, \nu_e) \cong \mathbb{H}_b$$

as Banach spaces, with the supremum norm on $\mathbb{H}_b$,

Poisson boundary

A measurable space $(E, \mathcal{B})$ together with a family ν_x, $x \in \mathbb{F}$, of probability measures on $(E, \mathcal{B})$ such that $\nu_x \ll \nu_e$ for all $x \in \mathbb{F}$ 'is' the Poisson boundary of $(X, \mathbb{F})$ if

$$L^\infty(E, \mathcal{B}, \nu_e) \cong \mathbb{H}_b$$

as Banach spaces, with the supremum norm on $\mathbb{H}_b$, via:

For each $\phi \in L^\infty$

(reprP) $$h(x) = \int_E \phi(y)\, \nu_x(dy), \quad x \in \mathbb{F},$$

is in $\mathbb{H}_b$, and for each $h \in \mathbb{H}_b$ there is a unique $\phi \in L^\infty$ such that this representation holds.

Tail boundary

Tail boundary

Let $\mathcal{T} := \bigcap_{n=1}^{\infty} \sigma\big(\{X_m : m \geq n\}\big)$ be the tail σ-field of $(X_n)_{n\in\mathbb{N}}$.

Tail boundary

Let $\mathcal{T} := \bigcap_{n=1}^{\infty} \sigma\big(\{X_m : m \geq n\}\big)$ be the tail σ-field of $(X_n)_{n\in\mathbb{N}}$.
Call $h : \mathbb{N} \times \mathbb{F} \to \mathbb{R}$ space-time harmonic if

$$h(n, x) = \sum_{y\in\mathbb{F}} p(x, y)\, h(n+1, y) \quad \text{for all } x \in \mathbb{F},\ n \in \mathbb{N}.$$

Tail boundary

Let $\mathcal{T} := \bigcap_{n=1}^{\infty} \sigma\big(\{X_m : m \geq n\}\big)$ be the tail σ-field of $(X_n)_{n\in\mathbb{N}}$.
Call $h : \mathbb{N} \times \mathbb{F} \to \mathbb{R}$ space-time harmonic if

$$h(n,x) = \sum_{y\in\mathbb{F}} p(x,y)\, h(n+1,y) \quad \text{for all } x \in \mathbb{F},\ n \in \mathbb{N}.$$

Let $\mathbb{H}_b^{\text{st}}$ be the set of all bounded space-time harmonic functions.

Tail boundary

Let $\mathcal{T} := \bigcap_{n=1}^{\infty} \sigma\big(\{X_m : m \geq n\}\big)$ be the tail σ-field of $(X_n)_{n\in\mathbb{N}}$.
Call $h : \mathbb{N} \times \mathbb{F} \to \mathbb{R}$ space-time harmonic if

$$h(n,x) = \sum_{y\in\mathbb{F}} p(x,y)\, h(n+1,y) \quad \text{for all } x \in \mathbb{F},\ n \in \mathbb{N}.$$

Let $\mathbb{H}_b^{\text{st}}$ be the set of all bounded space-time harmonic functions.
For each $h \in \mathbb{H}_b^{\text{st}}$

$$\Big(h(n,X_n), \sigma\big(\{X_m : m \leq n\}\big)\Big)_{n\in\mathbb{N}}$$

is a bounded martingale, hence $h(n,X_n) \to Y(h)$ a.s., with $Y(h)$ $\mathcal{T}$-mesurable.

Tail boundary

Let $\mathcal{T} := \bigcap_{n=1}^{\infty} \sigma\big(\{X_m : m \ge n\}\big)$ be the tail σ-field of $(X_n)_{n\in\mathbb{N}}$.
Call $h : \mathbb{N} \times \mathbb{F} \to \mathbb{R}$ space-time harmonic if

$$h(n,x) = \sum_{y\in\mathbb{F}} p(x,y)\, h(n+1,y) \quad \text{for all } x \in \mathbb{F},\ n \in \mathbb{N}.$$

Let $\mathbb{H}_b^{\text{st}}$ be the set of all bounded space-time harmonic functions.
For each $h \in \mathbb{H}_b^{\text{st}}$

$$\Big(h(n,X_n), \sigma\big(\{X_m : m \le n\}\big)\Big)_{n\in\mathbb{N}}$$

is a bounded martingale, hence $h(n,X_n) \to Y(h)$ a.s., with $Y(h)$ $\mathcal{T}$-mesurable. This leads to $L^\infty(\Omega,\mathcal{T},P \restriction \mathcal{T}) \cong \mathbb{H}_b^{\text{st}}$ via

(reprT) $$h(n,x) = E\big[Y(h)\big|X_n = x\big].$$

Boundaries: General relations

Boundaries: General relations

- From the topological Martin boundary $\partial\mathbb{F}$ we obtain the measure-theoretic Poisson boundary via

$$E = \partial\mathbb{F}, \quad \mathcal{B} \text{ the Borel subsets of } E,$$

$$\nu_e = \mathcal{L}(X_\infty), \quad \frac{d\nu_x}{d\nu_e}(\alpha) = K(x, \alpha).$$

Boundaries: General relations

- From the topological Martin boundary $\partial\mathbb{F}$ we obtain the measure-theoretic Poisson boundary via

$$E = \partial\mathbb{F}, \quad \mathcal{B} \text{ the Borel subsets of } E,$$

$$\nu_e = \mathcal{L}(X_\infty), \quad \frac{d\nu_x}{d\nu_e}(\alpha) = K(x,\alpha).$$

- The Martin boundary does not change when passing to an h-transform, the Poisson boundary may do so.

Boundaries: General relations

- From the topological Martin boundary $\partial\mathbb{F}$ we obtain the measure-theoretic Poisson boundary via

$$E = \partial\mathbb{F}, \quad \mathcal{B} \text{ the Borel subsets of } E,$$

$$\nu_e = \mathcal{L}(X_\infty), \quad \frac{d\nu_x}{d\nu_e}(\alpha) = K(x,\alpha).$$

- The Martin boundary does not change when passing to an h-transform, the Poisson boundary may do so.
- The tail boundary is the Poisson boundary of the space-time chain $(X_n^{\mathrm{st}})_{n\in\mathbb{N}}$, with $X_n^{\mathrm{st}} = (n, X_n)$ for all $n \in \mathbb{N}$.

Boundaries: General relations

- From the topological Martin boundary $\partial\mathbb{F}$ we obtain the measure-theoretic Poisson boundary via

$$E = \partial\mathbb{F}, \quad \mathcal{B} \text{ the Borel subsets of } E,$$

$$\nu_e = \mathcal{L}(X_\infty), \quad \frac{d\nu_x}{d\nu_e}(\alpha) = K(x,\alpha).$$

- The Martin boundary does not change when passing to an h-transform, the Poisson boundary may do so.
- The tail boundary is the Poisson boundary of the space-time chain $(X_n^{\text{st}})_{n\in\mathbb{N}}$, with $X_n^{\text{st}} = (n, X_n)$ for all $n \in \mathbb{N}$.
- CMCs have the space-time property, i.e. $n = \phi(X_n)$, hence tail boundary and Poisson boundary coincide.

Boundaries for the Pólya urn (Blackwell and Kendall, 1964)

Boundaries for the Pólya urn (Blackwell and Kendall, 1964)

Path counting leads to

$$K\big((i,j),(k,l)\big) \;=\; \frac{\binom{k+l-i-j}{k-i}}{\binom{k+l}{k}}\,\frac{(i+j+1)!}{i!\,j!}, \;\; i \le k, \; j \le l.$$

Boundaries for the Pólya urn (Blackwell and Kendall, 1964)

Path counting leads to

$$K\big((i,j),(k,l)\big) \;=\; \frac{\binom{k+l-i-j}{k-i}}{\binom{k+l}{k}}\,\frac{(i+j+1)!}{i!\,j!}, \;\; i \le k,\; j \le l.$$

Inspection shows that $\big(K((i,j),(k_n,l_n))\big)_{n\in\mathbb{N}}$ with $k_n + l_n \to \infty$ converges for all $(i,j) \in \mathbb{F}$ iff $\frac{k_n}{k_n+l_n} \to \alpha \in [0,1]$,

Boundaries for the Pólya urn (Blackwell and Kendall, 1964)

Path counting leads to

$$K\big((i,j),(k,l)\big) \;=\; \frac{\binom{k+l-i-j}{k-i}}{\binom{k+l}{k}}\,\frac{(i+j+1)!}{i!\,j!},\;\; i \le k,\; j \le l.$$

Inspection shows that $\big(K((i,j),(k_n,l_n))\big)_{n\in\mathbb{N}}$ with $k_n + l_n \to \infty$ converges for all $(i,j) \in \mathbb{F}$ iff $\frac{k_n}{k_n+l_n} \to \alpha \in [0,1]$, and then

- $K\big((i,j),\alpha\big) = \dfrac{(i+j+1)!}{i!\,j!}\,\alpha^i\,(1-\alpha)^j, \quad 0 \le \alpha \le 1,$

Boundaries for the Pólya urn (Blackwell and Kendall, 1964)

Path counting leads to

$$K\big((i,j),(k,l)\big) \;=\; \frac{\binom{k+l-i-j}{k-i}}{\binom{k+l}{k}}\,\frac{(i+j+1)!}{i!\,j!}, \;\; i \le k,\; j \le l.$$

Inspection shows that $\big(K((i,j),(k_n,l_n))\big)_{n\in\mathbb{N}}$ with $k_n + l_n \to \infty$ converges for all $(i,j) \in \mathbb{F}$ iff $\frac{k_n}{k_n+l_n} \to \alpha \in [0,1]$, and then

- $K\big((i,j),\alpha\big) = \dfrac{(i+j+1)!}{i!\,j!}\,\alpha^i\,(1-\alpha)^j, \quad 0 \le \alpha \le 1,$
- $\partial\mathbb{F} \cong [0,1],$

Boundaries for the Pólya urn (Blackwell and Kendall, 1964)

Path counting leads to

$$K\big((i,j),(k,l)\big) \;=\; \frac{\binom{k+l-i-j}{k-i}}{\binom{k+l}{k}}\,\frac{(i+j+1)!}{i!\,j!},\;\; i \le k,\; j \le l.$$

Inspection shows that $\big(K((i,j),(k_n,l_n))\big)_{n\in\mathbb{N}}$ with $k_n + l_n \to \infty$ converges for all $(i,j) \in \mathbb{F}$ iff $\frac{k_n}{k_n+l_n} \to \alpha \in [0,1]$, and then

- $K\big((i,j),\alpha\big) = \dfrac{(i+j+1)!}{i!\,j!}\,\alpha^i\,(1-\alpha)^j, \quad 0 \le \alpha \le 1,$
- $\partial\mathbb{F} \cong [0,1]$,
- $\mathcal{L}(X_\infty) = \mathcal{L}(U, 1-U)$ with $\mathcal{L}(U) = \mathrm{unif}(0,1)$,

Boundaries for the Pólya urn (Blackwell and Kendall, 1964)

Path counting leads to

$$K\big((i,j),(k,l)\big) \;=\; \frac{\binom{k+l-i-j}{k-i}}{\binom{k+l}{k}}\,\frac{(i+j+1)!}{i!\,j!},\;\; i \le k,\; j \le l.$$

Inspection shows that $\big(K((i,j),(k_n,l_n))\big)_{n\in\mathbb{N}}$ with $k_n + l_n \to \infty$ converges for all $(i,j) \in \mathbb{F}$ iff $\frac{k_n}{k_n+l_n} \to \alpha \in [0,1]$, and then

- $K\big((i,j),\alpha\big) = \dfrac{(i+j+1)!}{i!\,j!}\,\alpha^i\,(1-\alpha)^j, \quad 0 \le \alpha \le 1,$
- $\partial\mathbb{F} \cong [0,1]$,
- $\mathcal{L}(X_\infty) = \mathcal{L}(U, 1-U)$ with $\mathcal{L}(U) = \mathrm{unif}(0,1)$,
- the h-transform with $h = K(\cdot,\alpha)$ is the north-east random walk with parameter $\theta = \alpha$.

Boundaries for the Pólya urn (Blackwell and Kendall, 1964)

Path counting leads to

$$K\big((i,j),(k,l)\big) \;=\; \frac{\binom{k+l-i-j}{k-i}}{\binom{k+l}{k}}\,\frac{(i+j+1)!}{i!\,j!}, \;\; i \le k,\; j \le l.$$

Inspection shows that $\big(K((i,j),(k_n,l_n))\big)_{n\in\mathbb{N}}$ with $k_n + l_n \to \infty$ converges for all $(i,j) \in \mathbb{F}$ iff $\frac{k_n}{k_n+l_n} \to \alpha \in [0,1]$, and then

- $K\big((i,j),\alpha\big) = \dfrac{(i+j+1)!}{i!\,j!}\,\alpha^i\,(1-\alpha)^j, \quad 0 \le \alpha \le 1,$
- $\partial\mathbb{F} \cong [0,1]$,
- $\mathcal{L}(X_\infty) = \mathcal{L}(U, 1-U)$ with $\mathcal{L}(U) = \text{unif}(0,1)$,
- the h-transform with $h = K(\cdot,\alpha)$ is the north-east random walk with parameter $\theta = \alpha$.

Note: The Poisson boundary for the NE random walk is trivial.

Boundaries for the BST chain (Evans, Gr., Wakolbinger, 2012)

Boundaries for the BST chain (Evans, Gr., Wakolbinger, 2012)

For a tree $x \subset \mathbb{V} = \{0,1\}^\star$ let

$$x(u) := \{v \in \mathbb{V} : \ u + v \in x\}$$

be the subtree rooted at u.

Boundaries for the BST chain (Evans, Gr., Wakolbinger, 2012)

For a tree $x \subset \mathbb{V} = \{0,1\}^\star$ let

$$x(u) := \{v \in \mathbb{V} : \ u + v \in x\}$$

be the subtree rooted at u. Then

- $\partial\mathbb{F}$ 'is' the space of probability measures μ on $\{0,1\}^\infty$,

Boundaries for the BST chain (Evans, Gr., Wakolbinger, 2012)

For a tree $x \subset \mathbb{V} = \{0,1\}^\star$ let

$$x(u) := \{v \in \mathbb{V} : \ u + v \in x\}$$

be the subtree rooted at u. Then

- $\partial\mathbb{F}$ 'is' the space of probability measures μ on $\{0,1\}^\infty$,
- $x_n \to \mu \in \partial\mathbb{F}$ iff $\#x_n \to \infty$ and, for all $u \in \mathbb{V}$,

$$\frac{\#x(u)}{\#x} \to \mu(A_u), \ A_u := \{v \in \{0,1\}^\infty : \ u \leq v\},$$

Boundaries for the BST chain (Evans, Gr., Wakolbinger, 2012)

For a tree $x \subset \mathbb{V} = \{0,1\}^\star$ let

$$x(u) := \{v \in \mathbb{V} : \ u + v \in x\}$$

be the subtree rooted at u. Then

- $\partial\mathbb{F}$ 'is' the space of probability measures μ on $\{0,1\}^\infty$,
- $x_n \to \mu \in \partial\mathbb{F}$ iff $\#x_n \to \infty$ and, for all $u \in \mathbb{V}$,

$$\frac{\#x(u)}{\#x} \to \mu(A_u), \ A_u := \{v \in \{0,1\}^\infty : \ u \le v\},$$

- $\mathcal{L}(X_\infty)$ 'can be given',

Boundaries for the BST chain (Evans, Gr., Wakolbinger, 2012)

For a tree $x \subset \mathbb{V} = \{0,1\}^\star$ let

$$x(u) := \{v \in \mathbb{V} : \; u + v \in x\}$$

be the subtree rooted at u. Then

- $\partial\mathbb{F}$ 'is' the space of probability measures μ on $\{0,1\}^\infty$,
- $x_n \to \mu \in \partial\mathbb{F}$ iff $\#x_n \to \infty$ and, for all $u \in \mathbb{V}$,

$$\frac{\#x(u)}{\#x} \to \mu(A_u), \; A_u := \{v \in \{0,1\}^\infty : \; u \le v\},$$

- $\mathcal{L}(X_\infty)$ 'can be given',
- the h-transform with $h = K(\cdot,\mu)$ is the DST chain with parameter μ (generalizes the classical digital search tree, which has $\mu(A_u) = 2^{-|u|}$).

From boundaries to AofA

From boundaries to AofA

Rough idea: Instead of functionals $\Phi_n(X_n)$ consider X_n directly.

From boundaries to AofA

Rough idea: Instead of functionals $\Phi_n(X_n)$ consider X_n directly. Boundary theory may give $X_n \to X_\infty$ a.s., where X_∞ generates the tail σ-field $\mathcal{T}$.

Rough idea: Instead of functionals $\Phi_n(X_n)$ consider X_n directly. Boundary theory may give $X_n \to X_\infty$ a.s., where X_∞ generates the tail σ-field $\mathcal{T}$.

Suppose that $\Phi_n(X_n) \to Z$ a.s., with $E|Z| < \infty$. Then

$$Z_n := E[Z|\mathcal{F}_n] = \Psi_n(X_n) \to Z \text{ a.s.}$$

by Lévy's martingale convergence theorem.

From boundaries to AofA

Rough idea: Instead of functionals $\Phi_n(X_n)$ consider X_n directly. Boundary theory may give $X_n \to X_\infty$ a.s., where X_∞ generates the tail σ-field $\mathcal{T}$.

Suppose that $\Phi_n(X_n) \to Z$ a.s., with $E|Z| < \infty$. Then

$$Z_n := E[Z|\mathcal{F}_n] = \Psi_n(X_n) \to Z \text{ a.s.}$$

by Lévy's martingale convergence theorem.

As Z is $\mathcal{T}$-measurable, we have $Z = \Phi(X_\infty)$ for some Φ.

Rough idea: Instead of functionals $\Phi_n(X_n)$ consider X_n directly. Boundary theory may give $X_n \to X_\infty$ a.s., where X_∞ generates the tail σ-field $\mathcal{T}$.

Suppose that $\Phi_n(X_n) \to Z$ a.s., with $E|Z| < \infty$. Then

$$Z_n := E[Z|\mathcal{F}_n] = \Psi_n(X_n) \ \to \ Z \ \text{ a.s.}$$

by Lévy's martingale convergence theorem.

As Z is $\mathcal{T}$-measurable, we have $Z = \Phi(X_\infty)$ for some Φ.

Boundary theory approach:

- Try to guess Φ,
- work out Ψ_n,
- prove $\Phi_n - \Psi_n \to 0$.

Boundary theory approach: Examples

Boundary theory approach: Examples

Throughout, $(X_n)_{n\in\mathbb{N}}$ is the BST chain.

Throughout, $(X_n)_{n\in\mathbb{N}}$ is the BST chain.

(1) Internal path length: BTA recovers Régnier's result.
Essentially, it 'replaces inspiration by perspiration'.
Provides representation of the limit.

Boundary theory approach: Examples

Throughout, $(X_n)_{n\in\mathbb{N}}$ is the BST chain.

(1) Internal path length: BTA recovers Régnier's result.
Essentially, it 'replaces inspiration by perspiration'.
Provides representation of the limit.

(2) Wiener index: BTA works well.

Boundary theory approach: Examples

Throughout, $(X_n)_{n \in \mathbb{N}}$ is the BST chain.

(1) Internal path length: BTA recovers Régnier's result.
Essentially, it 'replaces inspiration by perspiration'.
Provides representation of the limit.

(2) Wiener index: BTA works well.

(3) Height and fill level: As Z is a constant, BTA is useless.

Boundary theory approach: Examples

Throughout, $(X_n)_{n \in \mathbb{N}}$ is the BST chain.

(1) Internal path length: BTA recovers Régnier's result.
Essentially, it 'replaces inspiration by perspiration'.
Provides representation of the limit.

(2) Wiener index: BTA works well.

(3) Height and fill level: As Z is a constant, BTA is useless.

(4) Profiles: BTA 'provides insight'.

Boundary theory approach: Examples

Throughout, $(X_n)_{n \in \mathbb{N}}$ is the BST chain.

(1) Internal path length: BTA recovers Régnier's result.
Essentially, it 'replaces inspiration by perspiration'.
Provides representation of the limit.

(2) Wiener index: BTA works well.

(3) Height and fill level: As Z is a constant, BTA is useless.

(4) Profiles: BTA 'provides insight'.

(5) BTA may lead to interesting functionals
(silhouette $\rightarrow$ metric silhouette).

Internal path length (re...revisited)

Internal path length (re...revisited)

$$\mathsf{IPL}(X_n) := \sum_{u \in X_n} |u|, \quad a_n := E\,\mathsf{IPL}(X_n) = 2(n+1)H_n - 4n.$$

Internal path length (re...revisited)

$$\mathrm{IPL}(X_n) := \sum_{u \in X_n} |u|, \quad a_n := E\,\mathrm{IPL}(X_n) = 2(n+1)H_n - 4n.$$

- Rewrite IPL in terms of subtree sizes,

$$\mathrm{IPL}(X_n) = \sum_{u \in X_n} |X_n(u)| - n.$$

Internal path length (re…revisited)

$$\mathrm{IPL}(X_n) := \sum_{u\in X_n} |u|, \quad a_n := E\,\mathrm{IPL}(X_n) = 2(n+1)H_n - 4n.$$

- Rewrite IPL in terms of subtree sizes,

$$\mathrm{IPL}(X_n) = \sum_{u\in X_n} |X_n(u)| - n.$$

- Guess the limit functional,

$$\Phi_P(X_\infty) = \sum_{u\in\mathbb{V}} X_\infty(A_u)C(\xi_u),$$

with

$$\xi_u = \frac{X_\infty(A_{u0})}{X_\infty(A_u)}, \; C(s) = 1 + 2s\log(s) + 2(1-s)\log(1-s).$$

Internal path length (re. . .revisited)

$$\mathrm{IPL}(X_n) := \sum_{u\in X_n} |u|, \quad a_n := E\,\mathrm{IPL}(X_n) = 2(n+1)H_n - 4n.$$

- Rewrite IPL in terms of subtree sizes,

$$\mathrm{IPL}(X_n) = \sum_{u\in X_n} |X_n(u)| - n.$$

- Guess the limit functional,

$$\Phi_P(X_\infty) = \sum_{u\in\mathbb{V}} X_\infty(A_u) C(\xi_u),$$

  with

$$\xi_u = \frac{X_\infty(A_{u0})}{X_\infty(A_u)}, \; C(s) = 1 + 2s\log(s) + 2(1-s)\log(1-s).$$

- Calculate $E[X_\infty(A_u)|\mathcal{F}_n]$ and $E[\xi_u|\mathcal{F}_n]$ to obtain the martingale projection of $\Phi_P(X_\infty)$.

Internal path length (re...revisited)

$$\mathrm{IPL}(X_n) := \sum_{u\in X_n} |u|, \quad a_n := E\,\mathrm{IPL}(X_n) = 2(n+1)H_n - 4n.$$

- Rewrite IPL in terms of subtree sizes,
$$\mathrm{IPL}(X_n) = \sum_{u\in X_n} |X_n(u)| - n.$$
- Guess the limit functional,
$$\Phi_P(X_\infty) = \sum_{u\in\mathbb{V}} X_\infty(A_u)C(\xi_u),$$
with
$$\xi_u = \tfrac{X_\infty(A_{u0})}{X_\infty(A_u)},\ \ C(s) = 1 + 2s\log(s) + 2(1-s)\log(1-s).$$
- Calculate $E[X_\infty(A_u)|\mathcal{F}_n]$ and $E[\xi_u|\mathcal{F}_n]$ to obtain the martingale projection of $\Phi_P(X_\infty)$.

Theorem *As $n\to\infty$, $\dfrac{IPL(X_n)-a_n}{n+1} \to \Phi_P(X_\infty)$ a.s. and in L^2.*

Internal path length (re. . .revisited)

$$\mathrm{IPL}(X_n) := \sum_{u\in X_n} |u|, \quad a_n := E\,\mathrm{IPL}(X_n) = 2(n+1)H_n - 4n.$$

- Rewrite IPL in terms of subtree sizes,
$$\mathrm{IPL}(X_n) = \sum_{u\in X_n} |X_n(u)| - n.$$
- Guess the limit functional,
$$\Phi_P(X_\infty) = \sum_{u\in\mathbb{V}} X_\infty(A_u)C(\xi_u),$$
with
$$\xi_u = \frac{X_\infty(A_{u0})}{X_\infty(A_u)},\ \ C(s) = 1 + 2s\log(s) + 2(1-s)\log(1-s).$$
- Calculate $E[X_\infty(A_u)|\mathcal{F}_n]$ and $E[\xi_u|\mathcal{F}_n]$ to obtain the martingale projection of $\Phi_P(X_\infty)$.

Theorem *As* $n\to\infty$, $\dfrac{IPL(X_n) - a_n}{n+1} \to \Phi_P(X_\infty)$ *a.s. and in* L^2.

Régnier 1989, Rösler 1991,$\cdots$, Bindjeme and Fill 2012, Gr. 2012

The Wiener index: Convergence in distribution

The Wiener index: Convergence in distribution

The Wiener index of a graph $G = (V, E)$ is the sum of all node distances,

$$\mathrm{WI}(G) := \frac{1}{2} \sum_{(u,v) \in V \times V} d(u, v).$$

The Wiener index: Convergence in distribution

The Wiener index of a graph $G = (V, E)$ is the sum of all node distances,

$$\mathrm{WI}(G) := \frac{1}{2} \sum_{(u,v)\in V\times V} d(u, v).$$

Neininger (2002): For the BST sequence $(X_n)_{n\in\mathbb{N}}$,

$$W_n := \frac{1}{n^2}\mathrm{WI}(X_n) - 2\log n$$

converges in distribution as $n \to \infty$.

The Wiener index: Convergence in distribution

The Wiener index of a graph $G = (V, E)$ is the sum of all node distances,

$$\mathrm{WI}(G) := \frac{1}{2} \sum_{(u,v)\in V\times V} d(u, v).$$

Neininger (2002): For the BST sequence $(X_n)_{n\in\mathbb{N}}$,

$$W_n := \frac{1}{n^2}\mathrm{WI}(X_n) - 2\log n$$

converges in distribution as $n \to \infty$.

Proof uses the contraction method (Rösler, Rüschendorf, Neininger).

The Wiener index: BTA

The Wiener index: BTA

- Rewrite Wiener index in terms of subtree sizes,

$$\mathsf{WI}(X_n) = n \sum_{u \in X_n} |X_n(u)| - \sum_{u \in X_n} |X_n(u)|^2.$$

The Wiener index: BTA

- Rewrite Wiener index in terms of subtree sizes,

$$\mathrm{WI}(X_n) = n \sum_{u \in X_n} |X_n(u)| - \sum_{u \in X_n} |X_n(u)|^2.$$

- Guess the limit functional Φ_W,

$$\Phi_W(X_\infty) = 2\gamma - 3 + \Phi_P(X_\infty) - Z_\infty, \ Z_\infty := \sum_{u \in \mathbb{V}} X_\infty(A_u)^2.$$

The Wiener index: BTA

- Rewrite Wiener index in terms of subtree sizes,

$$\mathsf{WI}(X_n) = n \sum_{u \in X_n} |X_n(u)| - \sum_{u \in X_n} |X_n(u)|^2.$$

- Guess the limit functional Φ_W,

$$\Phi_W(X_\infty) = 2\gamma - 3 + \Phi_P(X_\infty) - Z_\infty, \; Z_\infty := \sum_{u \in \mathbb{V}} X_\infty(A_u)^2.$$

- Show that $EZ_\infty < \infty$ and calculate $E[X_\infty(A_u)^2 | \mathcal{F}_n]$ to obtain the martingale projection of $\Phi_W(X_\infty)$.

The Wiener index: BTA

- Rewrite Wiener index in terms of subtree sizes,

$$\mathrm{WI}(X_n) = n \sum_{u \in X_n} |X_n(u)| - \sum_{u \in X_n} |X_n(u)|^2.$$

- Guess the limit functional Φ_W,

$$\Phi_W(X_\infty) = 2\gamma - 3 + \Phi_P(X_\infty) - Z_\infty, \; Z_\infty := \sum_{u \in \mathbb{V}} X_\infty(A_u)^2.$$

- Show that $EZ_\infty < \infty$ and calculate $E[X_\infty(A_u)^2 | \mathcal{F}_n]$ to obtain the martingale projection of $\Phi_W(X_\infty)$.

Theorem *As $n \to \infty$, $W_n \to \Phi_W(X_\infty)$ a.s. and in L^2.*

Profiles (from Régnier to Jabbour-Hattab)

Profiles (from Régnier to Jabbour-Hattab)

Count the external nodes of x at height k:

$$\Psi_k(x) := \#\{u \in \mathbb{V} : |u| = k, u \notin x, \bar{u} \in x\}.$$

Profiles (from Régnier to Jabbour-Hattab)

Count the external nodes of x at height k:

$$\Psi_k(x) := \#\{u \in \mathbb{V} : |u| = k, u \notin x, \bar{u} \in x\}.$$

Chauvin, Drmota, Jabbour-Hattab (2001) and Chauvin, Klein, Marckert, Roualt (2005) used the martingales

$$M_n(z) := \prod_{j=1}^{n-1} \frac{j+1}{j+2z} \sum_{k=1}^{\infty} \Psi_k(X_n)\, z^k, \; z \in \mathbb{C},$$

to obtain a.s. asymptotics for $k \mapsto \Psi_k(X_n)$ as $n \to \infty$.

Profiles (from Régnier to Jabbour-Hattab)

Count the external nodes of x at height k:

$$\Psi_k(x) := \#\{u \in \mathbb{V} : |u| = k, u \notin x, \bar{u} \in x\}.$$

Chauvin, Drmota, Jabbour-Hattab (2001) and Chauvin, Klein, Marckert, Roualt (2005) used the martingales

$$M_n(z) := \prod_{j=1}^{n-1} \frac{j+1}{j+2z} \sum_{k=1}^{\infty} \Psi_k(X_n) z^k, \; z \in \mathbb{C},$$

to obtain a.s. asymptotics for $k \mapsto \Psi_k(X_n)$ as $n \to \infty$.

- $M_n(z) > 0$ if $z \in [0, \infty)$,

Profiles (from Régnier to Jabbour-Hattab)

Count the external nodes of x at height k:

$$\Psi_k(x) := \#\big\{u \in \mathbb{V} :\ |u| = k, u \notin x, \bar{u} \in x\big\}.$$

Chauvin, Drmota, Jabbour-Hattab (2001) and Chauvin, Klein, Marckert, Roualt (2005) used the martingales

$$M_n(z) \ := \ \prod_{j=1}^{n-1} \frac{j+1}{j+2z} \ \sum_{k=1}^{\infty} \Psi_k(X_n)\, z^k,\ \ z \in \mathbb{C},$$

to obtain a.s. asymptotics for $k \mapsto \Psi_k(X_n)$ as $n \to \infty$.

- $M_n(z) > 0$ if $z \in [0, \infty)$,
- $M_n(z) = h(X_n)$ for some harmonic function $h : \mathbb{F} \to (0, \infty)$,

Profiles (from Régnier to Jabbour-Hattab)

Count the external nodes of x at height k:

$$\Psi_k(x) := \#\{u \in \mathbb{V} : |u| = k, u \notin x, \bar{u} \in x\}.$$

Chauvin, Drmota, Jabbour-Hattab (2001) and Chauvin, Klein, Marckert, Roualt (2005) used the martingales

$$M_n(z) := \prod_{j=1}^{n-1} \frac{j+1}{j+2z} \sum_{k=1}^{\infty} \Psi_k(X_n)\, z^k, \; z \in \mathbb{C},$$

to obtain a.s. asymptotics for $k \mapsto \Psi_k(X_n)$ as $n \to \infty$.

- $M_n(z) > 0$ if $z \in [0, \infty)$,
- $M_n(z) = h(X_n)$ for some harmonic function $h : \mathbb{F} \to (0, \infty)$,
- the corresponding h-transform has a simple interpretation via a marked spine construction used by Lyons, Pemantle and Peres and others in the Bienaymé-Galton-Watson context,

Profiles (from Régnier to Jabbour-Hattab)

Count the external nodes of x at height k:

$$\Psi_k(x) := \#\{u \in \mathbb{V} : |u| = k, u \notin x, \bar{u} \in x\}.$$

Chauvin, Drmota, Jabbour-Hattab (2001) and Chauvin, Klein, Marckert, Roualt (2005) used the martingales

$$M_n(z) := \prod_{j=1}^{n-1} \frac{j+1}{j+2z} \sum_{k=1}^{\infty} \Psi_k(X_n)\, z^k, \; z \in \mathbb{C},$$

to obtain a.s. asymptotics for $k \mapsto \Psi_k(X_n)$ as $n \to \infty$.

- $M_n(z) > 0$ if $z \in [0, \infty)$,
- $M_n(z) = h(X_n)$ for some harmonic function $h : \mathbb{F} \to (0, \infty)$,
- the corresponding h-transform has a simple interpretation via a marked spine construction used by Lyons, Pemantle and Peres and others in the Bienaymé-Galton-Watson context,
- $(2z)^{-1} M_\infty(z) = dP^h/dP$ for $z \in (c_-, c_+)$.

Profiles (from Régnier to Jabbour-Hattab)

Count the external nodes of x at height k:

$$\Psi_k(x) := \#\{u \in \mathbb{V} : |u| = k, u \notin x, \bar{u} \in x\}.$$

Chauvin, Drmota, Jabbour-Hattab (2001) and Chauvin, Klein, Marckert, Roualt (2005) used the martingales

$$M_n(z) := \prod_{j=1}^{n-1} \frac{j+1}{j+2z} \sum_{k=1}^{\infty} \Psi_k(X_n)\, z^k, \; z \in \mathbb{C},$$

to obtain a.s. asymptotics for $k \mapsto \Psi_k(X_n)$ as $n \to \infty$.

- $M_n(z) > 0$ if $z \in [0, \infty)$,
- $M_n(z) = h(X_n)$ for some harmonic function $h : \mathbb{F} \to (0, \infty)$,
- the corresponding h-transform has a simple interpretation via a marked spine construction used by Lyons, Pemantle and Peres and others in the Bienaymé-Galton-Watson context,
- $(2z)^{-1} M_\infty(z) = dP^h/dP$ for $z \in (c_-, c_+)$.

BTA 'explains' the martingale and gives an interpretation of the limit, but does not (easily) lead to a new proof of the known results.

From AofA to boundaries

From AofA to boundaries

Folklore: In specific cases the state space compactification can generally be obtained directly by using the algebraic or geometric structure.

From AofA to boundaries

Folklore: In specific cases the state space compactification can generally be obtained directly by using the algebraic or geometric structure.

Here: The BST chain X is generated by an algorithm that gives X as a deterministic function of the input sequence $\eta = (\eta_i)_{i \in \mathbb{N}}$.

From AofA to boundaries

Folklore: In specific cases the state space compactification can generally be obtained directly by using the algebraic or geometric structure.

Here: The BST chain X is generated by an algorithm that gives X as a deterministic function of the input sequence $\eta = (\eta_i)_{i\in\mathbb{N}}$.

In particular, X_∞ must be a function of η: For each $u \in \mathbb{V}$, let

$$\tau_u := \inf\{n \in \mathbb{N} : X_n \ni u\},$$

be the entrance time of the node and let

$$\eta_{(i:0)} := 0 < \eta_{(i:1)} < \cdots < \eta_{(i:i)} < \eta_{(i:i+1)} := 1$$

be the order statistics associated with $\eta_1, \ldots, \eta_i$, $i = \tau_u - 1$.

From AofA to boundaries

Folklore: In specific cases the state space compactification can generally be obtained directly by using the algebraic or geometric structure.

Here: The BST chain X is generated by an algorithm that gives X as a deterministic function of the input sequence $\eta = (\eta_i)_{i \in \mathbb{N}}$.

In particular, X_∞ must be a function of η: For each $u \in \mathbb{V}$, let

$$\tau_u := \inf\{n \in \mathbb{N} : X_n \ni u\},$$

be the entrance time of the node and let

$$\eta_{(i:0)} := 0 < \eta_{(i:1)} < \cdots < \eta_{(i:i)} < \eta_{(i:i+1)} := 1$$

be the order statistics associated with $\eta_1, \ldots, \eta_i$, $i = \tau_u - 1$.

Then
$$X_\infty(A_u) = \eta_{(i:j+1)} - \eta_{(i:j)} \quad \text{with} \quad \eta_{(i:j)} < \eta_{i+1} < \eta_{(i:j+1)}.$$

From AofA to boundaries

Folklore: In specific cases the state space compactification can generally be obtained directly by using the algebraic or geometric structure.

Here: The BST chain X is generated by an algorithm that gives X as a deterministic function of the input sequence $\eta = (\eta_i)_{i\in\mathbb{N}}$.

In particular, X_∞ must be a function of η: For each $u \in \mathbb{V}$, let

$$\tau_u := \inf\{n \in \mathbb{N} : X_n \ni u\},$$

be the entrance time of the node and let

$$\eta_{(i:0)} := 0 < \eta_{(i:1)} < \cdots < \eta_{(i:i)} < \eta_{(i:i+1)} := 1$$

be the order statistics associated with $\eta_1, \ldots, \eta_i$, $i = \tau_u - 1$.

Then
$$X_\infty(A_u) = \eta_{(i:j+1)} - \eta_{(i:j)} \ \text{ with } \ \eta_{(i:j)} < \eta_{i+1} < \eta_{(i:j+1)}.$$

Conclusion: Add ‘algorithmic’ to the above structural elements.

From AofA to boundaries

Folklore: In specific cases the state space compactification can generally be obtained directly by using the algebraic or geometric structure.

Here: The BST chain X is generated by an algorithm that gives X as a deterministic function of the input sequence $\eta = (\eta_i)_{i \in \mathbb{N}}$.

In particular, X_∞ must be a function of η: For each $u \in \mathbb{V}$, let

$$\tau_u := \inf\{n \in \mathbb{N} : X_n \ni u\},$$

be the entrance time of the node and let

$$\eta_{(i:0)} := 0 < \eta_{(i:1)} < \cdots < \eta_{(i:i)} < \eta_{(i:i+1)} := 1$$

be the order statistics associated with $\eta_1, \ldots, \eta_i$, $i = \tau_u - 1$.

Then $$X_\infty(A_u) = \eta_{(i:j+1)} - \eta_{(i:j)} \quad \text{with} \quad \eta_{(i:j)} < \eta_{i+1} < \eta_{(i:j+1)}.$$

Conclusion: Add 'algorithmic' to the above structural elements.

Also: A representation may lead to convergence in stronger topologies.

Metric binary trees: Generalities

Metric binary trees: Generalities

Turn $x \subset \mathbb{V}$ into a metric space (x, d):

- assign a positive value $w(u)$ to each non-root node u of x,

Metric binary trees: Generalities

Turn $x \subset \mathbb{V}$ into a metric space (x, d):

- assign a positive value $w(u)$ to each non-root node u of x,
- for $u = (u_1, \ldots, u_k) \in x$, $u \neq \emptyset$, regard $w(u)$ as the distance $d(u, \bar{u})$ of u and its direct ancestor $\bar{u} := (u_1, \ldots, u_{k-1})$,

Metric binary trees: Generalities

Turn $x \subset \mathbb{V}$ into a metric space (x, d):

- assign a positive value $w(u)$ to each non-root node u of x,
- for $u = (u_1, \dots, u_k) \in x$, $u \neq \emptyset$, regard $w(u)$ as the distance $d(u, \bar{u})$ of u and its direct ancestor $\bar{u} := (u_1, \dots, u_{k-1})$,
- for $u, v \in x$ let $d(u, v)$ be the sum of the distances along the unique path from u to v.

Metric binary trees: Generalities

Turn $x \subset \mathbb{V}$ into a metric space (x, d) :

- assign a positive value $w(u)$ to each non-root node u of x,
- for $u = (u_1, \ldots, u_k) \in x$, $u \neq \emptyset$, regard $w(u)$ as the distance $d(u, \bar{u})$ of u and its direct ancestor $\bar{u} := (u_1, \ldots, u_{k-1})$,
- for $u, v \in x$ let $d(u, v)$ be the sum of the distances along the unique path from u to v.

Special cases:

- $w \equiv 1$ leads to the canonical graph distance,

Metric binary trees: Generalities

Turn $x \subset \mathbb{V}$ into a metric space (x, d):

- assign a positive value $w(u)$ to each non-root node u of x,
- for $u = (u_1, \ldots, u_k) \in x$, $u \neq \emptyset$, regard $w(u)$ as the distance $d(u, \bar{u})$ of u and its direct ancestor $\bar{u} := (u_1, \ldots, u_{k-1})$,
- for $u, v \in x$ let $d(u, v)$ be the sum of the distances along the unique path from u to v.

Special cases:

- $w \equiv 1$ leads to the canonical graph distance,
- with $w(u) = w_x(u) := \#x(u)/\#x$ we obtain the (normalized) subtree size metric.

Metric binary trees: Generalities

Turn $x \subset \mathbb{V}$ into a metric space (x, d):

- assign a positive value $w(u)$ to each non-root node u of x,
- for $u = (u_1, \ldots, u_k) \in x$, $u \neq \emptyset$, regard $w(u)$ as the distance $d(u, \bar{u})$ of u and its direct ancestor $\bar{u} := (u_1, \ldots, u_{k-1})$,
- for $u, v \in x$ let $d(u, v)$ be the sum of the distances along the unique path from u to v.

Special cases:

- $w \equiv 1$ leads to the canonical graph distance,
- with $w(u) = w_x(u) := \#x(u)/\#x$ we obtain the (normalized) subtree size metric.
- For $\rho > 0$, define the weighted subtree size metric via $w_\rho(u) = w_{\rho,x}(u) := \rho^{|u|}\, w_x(u)$.

Metric binary trees: Generalities

Turn $x \subset \mathbb{V}$ into a metric space (x, d):

- assign a positive value $w(u)$ to each non-root node u of x,
- for $u = (u_1, \ldots, u_k) \in x$, $u \neq \emptyset$, regard $w(u)$ as the distance $d(u, \bar{u})$ of u and its direct ancestor $\bar{u} := (u_1, \ldots, u_{k-1})$,
- for $u, v \in x$ let $d(u, v)$ be the sum of the distances along the unique path from u to v.

Special cases:

- $w \equiv 1$ leads to the canonical graph distance,
- with $w(u) = w_x(u) := \#x(u)/\#x$ we obtain the (normalized) subtree size metric.
- For $\rho > 0$, define the weighted subtree size metric via $w_\rho(u) = w_{\rho,x}(u) := \rho^{|u|}\, w_x(u)$.

Note: The first one is local, the others are global.

Metric binary trees: Pictures

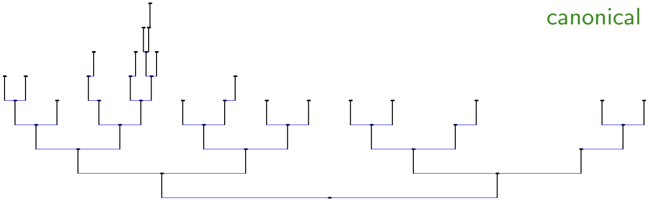
Metric binary trees: Pictures

Map $u = (u_1, \ldots, u_n) \in \mathbb{V}$ to $t = 2^{-1} + \sum_{i=1}^{n} (2u_i - 1)2^{-i-1} \in (0, 1)$, vertical distance is $d(u, \emptyset)$.

Metric binary trees: Pictures

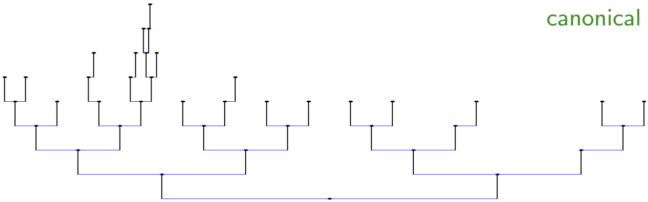
Map $u = (u_1, \ldots, u_n) \in \mathbb{V}$ to $t = 2^{-1} + \sum_{i=1}^{n}(2u_i - 1)2^{-i-1} \in (0, 1)$, vertical distance is $d(u, \emptyset)$.

canonical distance

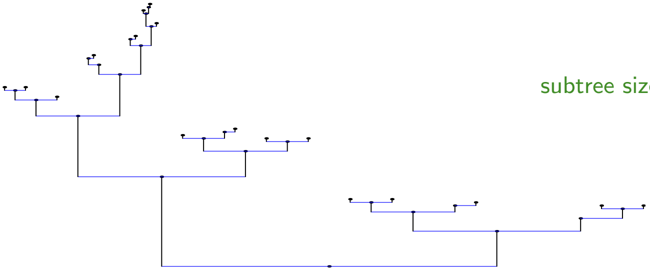


Metric binary trees: Pictures

Map $u = (u_1, \ldots, u_n) \in \mathbb{V}$ to $t = 2^{-1} + \sum_{i=1}^{n}(2u_i - 1)2^{-i-1} \in (0,1)$, vertical distance is $d(u, \emptyset)$.



canonical distance



subtree size metric

Metric binary trees: Convergence

Any measure μ on $\{0,1\}^\infty$ defines a metric on $\mathbb{V}$ via

$$d_\infty(\bar{u}, u) = \mu(A_u) \quad \text{for all } u \in \mathbb{V},\ u \neq \emptyset,$$

(recall that $A_u = \{v \in \{0,1\}^\infty :\ u \leq v\}$).

Metric binary trees: Convergence

Any measure μ on $\{0,1\}^\infty$ defines a metric on $\mathbb{V}$ via

$$d_\infty(\bar{u}, u) = \mu(A_u) \text{ for all } u \in \mathbb{V},\ u \neq \emptyset,$$

(recall that $A_u = \{v \in \{0,1\}^\infty : u \leq v\}$).

Convergence of trees $\leftrightarrow$ convergence of metric spaces

Metric binary trees: Convergence

Any measure μ on $\{0,1\}^\infty$ defines a metric on $\mathbb{V}$ via

$$d_\infty(\bar{u}, u) = \mu(A_u) \text{ for all } u \in \mathbb{V},\ u \neq \emptyset,$$

(recall that $A_u = \{v \in \{0,1\}^\infty : u \leq v\}$).

Convergence of trees $\leftrightarrow$ convergence of metric spaces

Rôle model:
Continuum random tree, Brownian excursion, Gromov convergence.

Metric binary trees: Convergence

Any measure μ on $\{0,1\}^\infty$ defines a metric on $\mathbb{V}$ via

$$d_\infty(\bar{u}, u) = \mu(A_u) \quad \text{for all } u \in \mathbb{V},\ u \neq \emptyset,$$

(recall that $A_u = \{v \in \{0,1\}^\infty : \ u \leq v\}$).

Convergence of trees $\leftrightarrow$ convergence of metric spaces

Rôle model:
Continuum random tree, Brownian excursion, Gromov convergence.

In view of $X_n \uparrow \mathbb{V}$ we can work with

$$(x_n, d_n) \to (\mathbb{V}, d_\infty)$$

meaning that

$$\lim_{n\to\infty} d_n(u,v) = d_\infty(u,v) \quad \text{for all } u, v \in \mathbb{V}.$$

Metric binary trees: A phase transition

Theorem

Let $\rho_0 = 1.26107\cdots$ *be the smaller one of the two roots of the equation* $2e\log(\rho) = \rho$, $\rho > 0$.

Theorem

Let $\rho_0 = 1.26107\cdots$ *be the smaller one of the two roots of the equation* $2e\log(\rho) = \rho$, $\rho > 0$.

(a) *For* $\rho < \rho_0$, *the metric space* $(\mathbb{V}, d_{X_\infty,\rho})$ *is compact w.p. 1.*

Metric binary trees: A phase transition

Theorem

Let $\rho_0 = 1.26107\cdots$ *be the smaller one of the two roots of the equation* $2e\log(\rho) = \rho$*,* $\rho > 0$*.*

(a) *For* $\rho < \rho_0$*, the metric space* $(\mathbb{V}, d_{X_\infty,\rho})$ *is compact w.p. 1.*

(b) *For* $\rho > \rho_0$*,* $(\mathbb{V}, d_{X_\infty,\rho})$ *has infinite diameter w.p. 1.*

Theorem

Let $\rho_0 = 1.26107\cdots$ be the smaller one of the two roots of the equation $2e\log(\rho) = \rho$, $\rho > 0$.

(a) *For $\rho < \rho_0$, the metric space $(\mathbb{V}, d_{X_\infty,\rho})$ is compact w.p. 1.*

(b) *For $\rho > \rho_0$, $(\mathbb{V}, d_{X_\infty,\rho})$ has infinite diameter w.p. 1.*

(c) *For $\rho < \rho_0$, the metric spaces $(X_n, d_{X_n,\rho})$ converge uniformly to $(\mathbb{V}, d_{X_\infty,\rho})$ as $n \to \infty$ in the sense of*

$$\sup_{u,v\in X_n} \left| d_{X_n,\rho}(u,v) - d_{X_\infty,\rho}(u,v) \right| \;\to\; 0 \;\; \textit{a. s. and in mean.}$$

Theorem

Let $\rho_0 = 1.26107\cdots$ be the smaller one of the two roots of the equation $2e\log(\rho) = \rho$, $\rho > 0$.

(a) *For $\rho < \rho_0$, the metric space $(\mathbb{V}, d_{X_\infty,\rho})$ is compact w.p. 1.*

(b) *For $\rho > \rho_0$, $(\mathbb{V}, d_{X_\infty,\rho})$ has infinite diameter w.p. 1.*

(c) *For $\rho < \rho_0$, the metric spaces $(X_n, d_{X_n,\rho})$ converge uniformly to $(\mathbb{V}, d_{X_\infty,\rho})$ as $n \to \infty$ in the sense of*

$$\sup_{u,v\in X_n} \left| d_{X_n,\rho}(u,v) - d_{X_\infty,\rho}(u,v) \right| \to 0 \text{ a. s. and in mean.}$$

(d) *For $\rho > \rho_0$, and with $d_{X_n,\rho}(\bar{u}, u) := 0$ for $u \notin X_n$,*

$$\sup_{u,v\in \mathbb{V}} \left| d_{X_n,\rho}(u,v) - d_{X_\infty,\rho}(u,v) \right| = \infty \text{ w.p. 1.}$$

Outlook

Outlook

- $\mathbb{F}$: Ulam-Harris trees, CMC: random recursive trees:

  Roughly: Nesting leads from the Pólya urn to the BST chain, and from the Chinese restaurant to RRT, see EGW 2012.

Outlook

- $\mathbb{F}$: Ulam-Harris trees, CMC: random recursive trees:

  Roughly: Nesting leads from the Pólya urn to the BST chain, and from the Chinese restaurant to RRT, see EGW 2012.

- Graph asymptotics?

  CMC style growth, see e.g. Janson and Severini 2012

Outlook

- $\mathbb{F}$: Ulam-Harris trees, CMC: random recursive trees:
  Roughly: Nesting leads from the Pólya urn to the BST chain, and from the Chinese restaurant to RRT, see EGW 2012.

- Graph asymptotics?
  CMC style growth, see e.g. Janson and Severini 2012

- Second order results via (struct)?
  IPL: Convergence in distribution in Neininger 2013+.

Outlook

- $\mathbb{F}$: Ulam-Harris trees, CMC: random recursive trees:
  Roughly: Nesting leads from the Pólya urn to the BST chain, and from the Chinese restaurant to RRT, see EGW 2012.
- Graph asymptotics?
  CMC style growth, see e.g. Janson and Severini 2012
- Second order results via (struct)?
  IPL: Convergence in distribution in Neininger 2013+.
- Specific questions.
  Is the continuum random tree the boundary of the Rémy chain?

Outlook

- $\mathbb{F}$: Ulam-Harris trees, CMC: random recursive trees:
  Roughly: Nesting leads from the Pólya urn to the BST chain, and from the Chinese restaurant to RRT, see EGW 2012.
- Graph asymptotics?
  CMC style growth, see e.g. Janson and Severini 2012
- Second order results via (struct)?
  IPL: Convergence in distribution in Neininger 2013+.
- Specific questions.
  Is the continuum random tree the boundary of the Rémy chain?
- Functorial properties of the boundary constructions.