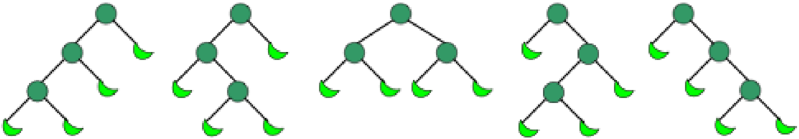
Some tree-valued chains and their bridges

Steven N. Evans

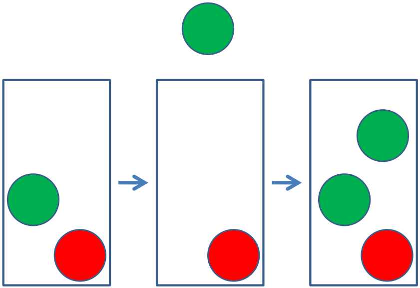
August, 2021



Pólya urn

Write (G_n, R_n) for the numbers of green and red balls added by time n in the classical Pólya urn:

- Start at time 0 with one green and one red ball in an urn.
- At each subsequent point in time, pick a ball uniformly at random from the urn and replace it along with one of the same color.



Running the urn backwards: exchangeability

- Check that if we condition on the event $\{G_{n+1} = g, R_{n+1} = r\}$, then the order in which the g green balls and r red balls appear is uniformly distributed over the $\binom{g+r}{g} = \binom{g+r}{r}$ possibilities.
- In particular,

$$\mathbb{P}\{G_n = g - 1, R_n = r \,|\, G_{n+1} = g, R_{n+1} = r\} = \frac{g}{g+r}$$

and

$$\mathbb{P}\{G_n = g, R_n = r - 1 \,|\, G_{n+1} = g, R_{n+1} = r\} = \frac{r}{g+r}.$$

Running coin-tossing backwards: exchangeability

- Let H_n (resp. T_n) be the number of heads (resp. tails) that appear in n independent tosses of a coin that comes up heads with probability p.
- Check that if we condition on the event $\{H_{n+1} = h, T_{n+1} = t\}$, then the order in which the h heads and t tails appear is uniform over the $\binom{h+t}{h} = \binom{h+t}{t}$ possibilities.
- In particular,

$$\mathbb{P}\{H_n = h-1, T_n = t \,|\, H_{n+1} = h, T_{n+1} = t\} = \frac{h}{h+t}$$

and

$$\mathbb{P}\{H_n = h, T_n = t-1 \,|\, H_{n+1} = h, T_{n+1} = t\} = \frac{t}{h+t}.$$

Pólya's urn backwards is coin-tossing backwards

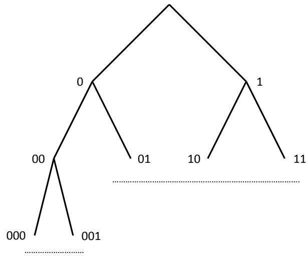
- Pólya's urn run backwards has the same transition dynamics as coin-tossing run backwards.

Binary trees

- Let's start with trees that are rooted, binary (each vertex has 0, 1 or 2 children), and ordered (we distinguish between a left child and a right child – even when there is only one child).
- We can identify such a tree as a subtree of the complete rooted binary tree.

Complete rooted binary tree

- Denote by $\{0,1\}^* := \bigsqcup_{k=0}^{\infty} \{0,1\}^k$ the set of finite words drawn from the alphabet $\{0,1\}$ (with the empty word $\emptyset$ allowed).
- Write a word $(v_1, \ldots, v_\ell) \in \{0,1\}^*$ more simply as $v_1 \ldots v_\ell$.
- Define a directed graph with vertex set $\{0,1\}^*$ by declaring that if $u = u_1 \ldots u_k$ and $v = v_1 \ldots v_\ell$ are two words, then (u,v) is a directed edge (that is, $u \to v$) if and only if $\ell = k+1$ and $u_i = v_i$ for $i = 1, \ldots, k$ (i.e. v is a child of u).
- This directed graph is the complete rooted binary tree (rooted at $\emptyset$).

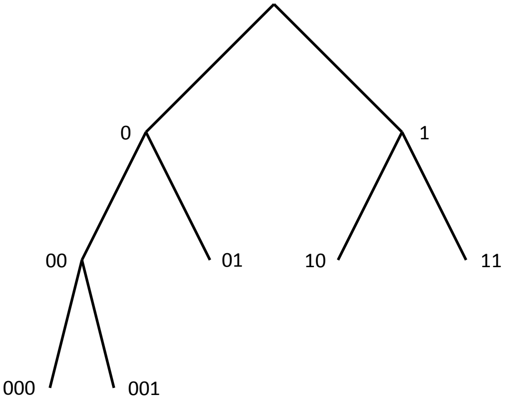


A partial order on the complete rooted binary tree and its "leaves"

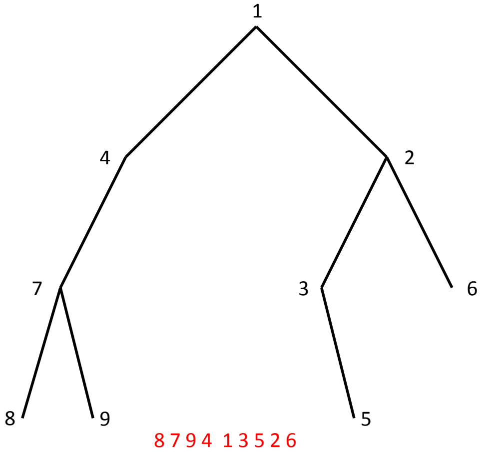
- We can think of $\{0,1\}^\infty$ as the leaves of the complete rooted binary tree.
- Define a partial order on $\{0,1\}^*$ by declaring that $u < v$ if $u \to w_1 \to \cdots \to w_m \to v$ for some words $w_1, \ldots, w_m$ (i.e. v is a descendant of u). This partial order extends to $\{0,1\}^* \sqcup \{0,1\}^\infty$.

Finite rooted binary trees

- A finite rooted binary tree is a non-empty subset $\mathbf{t}$ of $\{0,1\}^*$ with the property that if $v \in \mathbf{t}$ and $u \in \{0,1\}^*$ is such that $u \to v$, then $u \in \mathbf{t}$.
- The vertex $\emptyset$ belongs to any such tree $\mathbf{t}$ and is the root of $\mathbf{t}$.



Storing an ordered listing of $[n]$ as a finite rooted binary tree



One way of building a random finite rooted binary tree

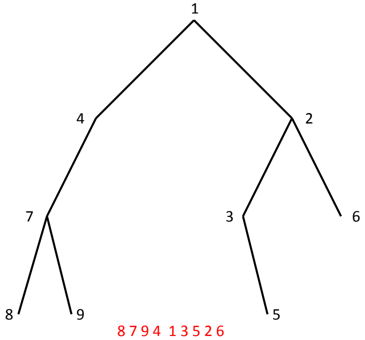


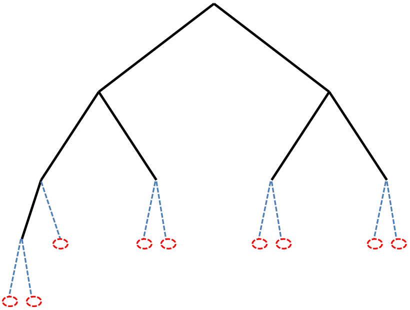
Figure: Finite rooted binary tree built from a realization of the uniform i.i.d.r.v. $U_1, \ldots, U_9$ with $U_8 < U_7 < U_9 < U_4 < U_1 < U_3 < U_5 < U_2 < U_6$.

Binary search tree process

- If we do the construction of the previous slide for successive values of n and only keep track of the resulting the rooted binary trees (i.e. we throw away the labeling of vertices by $[n]$), then we get a Markov chain taking values in the space of finite rooted binary trees called the binary search tree process.

Binary search tree process transition probabilities

- The binary search tree process $\{T_n\}_{n \in \mathbb{N}}$ evolves as follows.
  - $T_1 = \emptyset$
  - Given T_n, a tree with n leaves, there are $n+1$ words of the form $v = v_1 \ldots v_\ell$ such that v is not a vertex of the tree T_n but the word $v_1 \ldots v_{\ell-1}$ is.
  - Pick such a word uniformly at random and adjoin it to produce the tree T_{n+1}.



Pólya urn connection

- For any $u = u_1 \ldots u_m \in \{0,1\}^*$ the sequences

$$\#\{v \in T_n : u_1 \ldots u_m 0 \leqslant v\}$$

  and

$$\#\{v \in T_n : u_1 \ldots u_m 1 \leqslant v\}$$

  evolve like time changes of the numbers of new green and red balls in a classical Pólya urn that starts with 1 green and 1 red ball.
- The binary search tree process is thus an infinite hierarchical system of Pólya urns.

Digital search tree process

- The digital search tree process $\{D_n\}_{n \in \mathbb{N}}$ is a Markov chain taking values in the space of finite binary trees that evolves as follows.
  - $D_1 = \emptyset$
  - Given D_n, a tree with n leaves, there are $n+1$ words of the form $v = v_1 \dots v_\ell$ such that v is not a vertex of the tree D_n but the word $v_1 \dots v_{\ell-1}$ is.
  - Pick such a word v with probability $2^{-\ell} = 2^{-|v|}$ and adjoin it to produce the tree D_{n+1}.

Random walk connection

- For any $u = u_1 \ldots u_m \in \{0,1\}^*$ the sequences

$$\#\{v \in D_n : u_1 \ldots u_m 0 \leqslant v\}$$

and

$$\#\{v \in D_n : u_1 \ldots u_m 1 \leqslant v\}$$

evolve like time changes of the numbers of heads and tails in fair coin-tossing.

- The digital search tree process is thus an infinite hierarchical system of simple random walks.

Binary search tree process versus digital search tree process

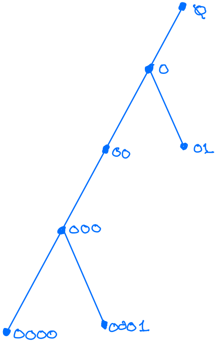
- The binary search tree process run backwards has the same transition dynamics as the digital search tree process run backwards.

Radix sort trees

- Suppose that $z_1, \dots, z_n \in \{0,1\}^\infty$ are distinct infinite binary words.
- For each $i \in [n]$ we may construct a finite binary word y_i that is an initial segment of z_i such that $y_1, \dots, y_n$ are the distinct leaves of a finite rooted binary tree and $y_1, \dots, y_n$ are the minimal length words with this property.
- The resulting finite rooted binary tree $\mathbf{R}(z_1, \dots, z_n)$ is called the radix sort tree defined by the infinite binary words $z_1, \dots, z_n$.
- A depth first search of this tree visits the leaves in an order that coincides with the lexicographic order of the corresponding infinite binary words.

Example of a radix sort tree

Radix sort tree for the words 01...
0000...
0001...



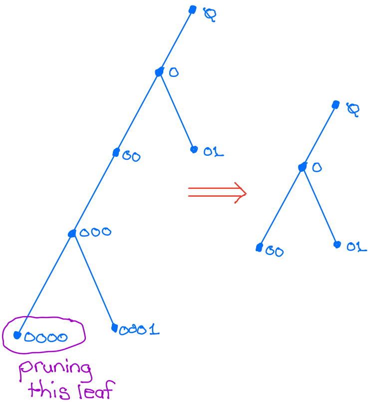
Some notation for radix sort trees

- Denote by $\mathbb{S}$ the class of finite rooted binary trees that can arise as radix sort trees.
- A finite rooted binary tree $\mathbf{s}$ belongs to $\mathbb{S}$ if and only if $\mathbf{s} = \{\emptyset\}$ or $\mathbf{s}$ has at least two leaves and for any leaf $u_1 \dots u_p \in \mathbf{s}$ the sibling word $u_1 \dots u_{p-1}\bar{u}_p$ is also a vertex of $\mathbf{s}$, where we set $\bar{0} := 1$ and $\bar{1} := 0$.
- Write $\mathbb{S}_n$, $n \in \mathbb{N}$, for the elements of $\mathbb{S}$ with n leaves.
- In particular, $\mathbb{S}_1$ contains only the trivial tree with the single vertex $\{\emptyset\}$.

Pruning a leaf from a radix sort tree

- Consider $\mathbf{t} \in \mathbb{S}_{n+1}$ and let $v = v_1 \ldots v_m$ be a leaf of $\mathbf{t}$.
- Suppose first that the sibling $v_1 \ldots v_{m-1}\bar{v}_m$ is not a leaf of $\mathbf{t}$. Let $\kappa(\mathbf{t}, v) \in \mathbb{S}_n$ be the finite rooted binary tree with the same leaf set as $\mathbf{t}$ except that v has been removed.
- On the other hand, suppose that the sibling $v_1 \ldots v_{m-1}\bar{v}_m$ is also a leaf of $\mathbf{t}$. There is a largest $\ell < m$ such that the siblings $v_1 \ldots v_\ell$ and $v_1 \ldots v_{\ell-1}\bar{v}_\ell$ are both vertices of $\mathbf{t}$. In this case, let $\kappa(\mathbf{t}, v) \in \mathbb{S}_n$ be the tree with the same leaf set as $\mathbf{t}$ except that the leaf v and its sibling leaf $v_1 \ldots v_{m-1}\bar{v}_m$ have both been removed and replaced by the single leaf $v_1 \ldots v_\ell$.
- If $\mathbf{t}$ is the radix sort tree for the infinite binary inputs $z_1, \ldots, z_{n+1}$ and y_{n+1} is the leaf of $\mathbf{t}$ corresponding to the input z_{n+1}, then $\kappa(\mathbf{t}, y_{n+1})$ is the radix sort tree for the inputs $z_1, \ldots, z_n$.

Example of pruning a leaf from a radix sort tree



Radix sort processes

- Let $Z_1, Z_2, \ldots$ be i.i.d. $\{0,1\}^\infty$-valued random variables with common distribution some diffuse probability measure ν.
- Set ${}^\nu R_n := \mathbf{R}(Z_1, \ldots, Z_n)$.

The radix sort process is Markov

- The radix sort process $({}^{\nu}R_n)_{n \in \mathbb{N}}$ is Markov. The easiest way to see this is to show that the backwards in time process is Markov.
- For $\mathbf{s} \in \mathbb{S}_n$ and $\mathbf{t} \in \mathbb{S}_{n+1}$, the associated backwards transition probability is

$$\mathbb{P}\{{}^{\nu}R_n = \mathbf{s} \,|\, {}^{\nu}R_{n+1} = \mathbf{t}\} = \frac{1}{n+1}\{v : \mathbf{s} = \kappa(\mathbf{t}, v)\}.$$

- That is, $({}^{\nu}R_n)_{n \in \mathbb{N}}$ evolves backwards in time by picking leaves uniformly at random and pruning them.
- NOTE: The backwards transition probabilities of $({}^{\nu}R_n)_{n \in \mathbb{N}}$ are the same for all ν.

The transition probabilities of the radix sort process

- The radix sort process $({}^{\nu}R_n)_{n \in \mathbb{N}}$ is a Markov chain with the following transition dynamics. Suppose that ${}^{\nu}R_n$ is the finite rooted binary tree $\mathbf{s}$ with leaves $\mathbf{L}(\mathbf{s})$. There are two cases to consider.
  - **Case I.** Suppose that $y = u_1 u_2 \ldots u_{m-1} u_m$ is not a vertex of $\mathbf{s}$ and $u_1 u_2 \ldots u_{m-1} \bar{u}_m$ is a (non-leaf) vertex of $\mathbf{s}$. Let $\mathbf{t}$ be the finite rooted binary tree with leaves $\mathbf{L}(\mathbf{s}) \sqcup \{y\}$. Then

    $$\mathbb{P}\{{}^{\nu}R_{n+1} = \mathbf{t} \mid {}^{\nu}R_n = \mathbf{s}\} = \nu(\tau(y)),$$

    where $\tau(y) := \{z \in \{0,1\}^{\infty} : y < z\}$.
  - **Case II.** Suppose that $y \in \mathbf{L}(\mathbf{s})$ and y', y'' are siblings with $y < y'$ and $y < y''$. Let $\mathbf{t}$ be the finite rooted binary tree with leaves $(\mathbf{L}(\mathbf{s}) \backslash \{y\}) \sqcup \{y', y''\}$. Then

    $$\mathbb{P}\{{}^{\nu}R_{n+1} = \mathbf{t} \mid {}^{\nu}R_n = \mathbf{s}\} = 2 \frac{\nu(\tau(y'))\nu(\tau(y''))}{\nu(\tau(y))}.$$

Finite rooted full binary trees

- A finite rooted full binary tree is a finite rooted binary tree in which every vertex has 0 or 2 children.
- The number of such trees with $n+1$ leaves (and hence $2n+1$ vertices) is the Catalan number $\frac{1}{n+1}\binom{2n}{n}$.
- Rémy's (1985) algorithm generates a sequence of random binary trees $(U_n)_{n\in\mathbb{N}}$ such that U_n is uniformly distributed on the set of finite rooted full binary trees with $n+1$ leaves.

Rémy's algorithm

- Start with U_1 being the unique finite rooted full binary tree $\aleph := \{\emptyset, 0, 1\}$ with 3 vertices.
- Supposing that U_n has been generated, pick a vertex v uniformly at random.
- Cut off the subtree rooted at v and set it aside.
- Attach a copy of the tree $\aleph$ with 3 vertices to the end of the edge that previously led to v.
- Re-attach the subtree rooted at v uniformly at random to one of the two leaves in the copy of $\aleph$.
- Call the two new vertices that have been produced clones of v.

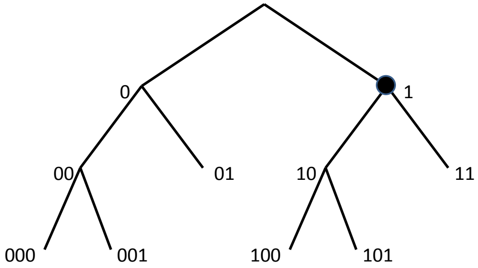


Figure: First step in an iteration of Rémy's algorithm: pick a vertex v uniformly at random.

Example of one iteration of Rémy's algorithm – continued

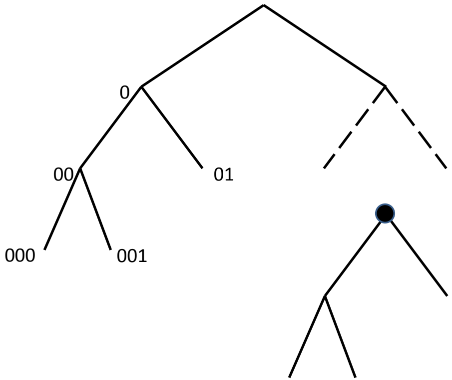


Figure: Second step in an iteration of Rémy's algorithm: cut off the subtree rooted at v and attach a copy of $\aleph$ to the end of the edge that previously led to v.

Example of one iteration of Rémy's algorithm – continued

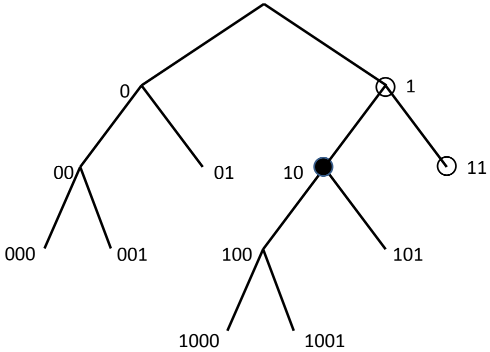


Figure: Third step in an iteration of Rémy's algorithm: re-attach the subtree rooted at v to one of the two leaves of the copy of $\aleph$, and re-label the vertices appropriately. The solid circle is the new location of v and the open circles are the clones of v.

Digression: Embeddings

- An embedding of a finite rooted full binary tree $\mathbf{s}$ into a finite rooted full binary tree $\mathbf{t}$ is a map from the vertex set of $\mathbf{s}$ into the vertex set of $\mathbf{t}$ such that the following hold.
  - The image of a leaf of $\mathbf{s}$ is a leaf of $\mathbf{t}$.
  - If u, v are vertices of $\mathbf{s}$ such that v is below and to the left (resp. right) of u, then the image of v in $\mathbf{t}$ is below and to the left (resp. right) of the image of u in $\mathbf{t}$.
- Write $N(\mathbf{s}, \mathbf{t})$ for the number of embeddings of $\mathbf{s}$ into $\mathbf{t}$.

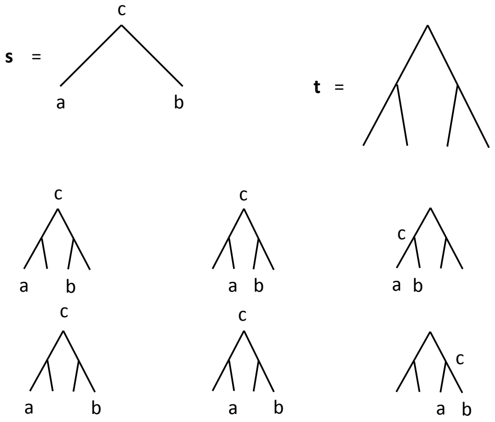


Figure: All the embeddings of the unique finite rooted full binary tree $\mathbf{s} = \aleph$ with 3 vertices into a particular finite rooted full binary tree $\mathbf{t}$ with 7 vertices.

Digression: Embeddings – continued

- Note that an embedding of $\mathbf{s}$ into $\mathbf{t}$ is uniquely determined by the images of the leaves of $\mathbf{s}$, because if x and y are vertices of $\mathbf{s}$, then the image of the most recent common ancestor of x and y in $\mathbf{s}$ must be the most recent common ancestor in $\mathbf{t}$ of the images of x and y.

Transition probabilities for the Rémy chain

- Suppose that $\mathbf{s}$ and $\mathbf{t}$ are two finite rooted full binary trees with, respectively, $m+1$ and $m+n+1$ leaves. Then, the probability that the Rémy chain transitions from $\mathbf{s}$ to $\mathbf{t}$ in n steps is

$$p^n(\mathbf{s},\mathbf{t}) = n! \frac{1}{(2m+1)\times(2m+3)\times\cdots\times(2(m+n)-1)} \frac{1}{2^n} N(\mathbf{s},\mathbf{t}),$$

where $N(\mathbf{s},\mathbf{t})$ is the number of ways of embedding $\mathbf{s}$ into $\mathbf{t}$.

What are the multi-step transition probabilities of Rémy's chain?

- Condition on T_m.
- Say that a vertex of T_{m+n} is a clonal descendant of a vertex $v \in T_m$ if it is v itself, a clone of v, a clone-of-a-clone of v, etc.
- We can decompose T_{m+n} into connected pieces according to clonal descent from the vertices of T_m.

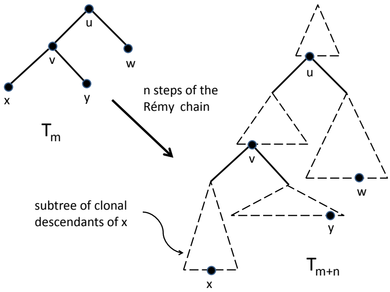


Figure: Decomposition of T_{m+n} via clonal descent from the vertices of T_m.

What are the multi-step transition probabilities? – continued

- The numbers of clonal descendants of the $2m+1$ vertices is the result of n steps in a Polya urn that starts with $2m+1$ balls of different colors and at each stage a ball is chosen uniformly at random and replaced along with two balls of the same color.
- Conditional on the numbers of clonal descendants, the binary trees of clonal descendants are independent and uniformly distributed.
- Conditional on the trees of clonal descendants, the ancestors from T_m are located at independently and uniformly chosen leaves of their respective trees of clonal descendants.

One-step backward transition probabilities for the Rémy chain

- Note that if $\mathbf{s}, \mathbf{t}$ are finite rooted full binary trees with $n+1$ and $n+2$ leaves, respectively, then, using the expressions for the forward transition probabilities and the Catalan numbers,

$$\mathbb{P}\{T_n = \mathbf{s} \,|\, T_{n+1} = \mathbf{t}\} = \frac{1}{n+2} N(\mathbf{s}, \mathbf{t}).$$

Backward dynamics for the Rémy chain in words

- The Rémy chain evolves one step backward in time as follows.
  - Pick a leaf uniformly at random.
  - Delete the chosen leaf and its sibling (which may or may not be a leaf).
  - If the sibling is not a leaf, then close up the resulting gap by attaching the subtree below the sibling to the parent of the chosen leaf and the sibling.

Backward dynamics for the Rémy chain – continued

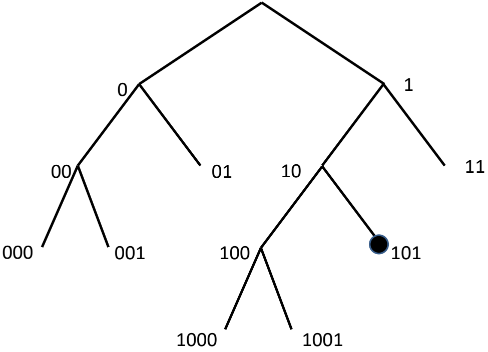


Figure: Pick a leaf uniformly at random.

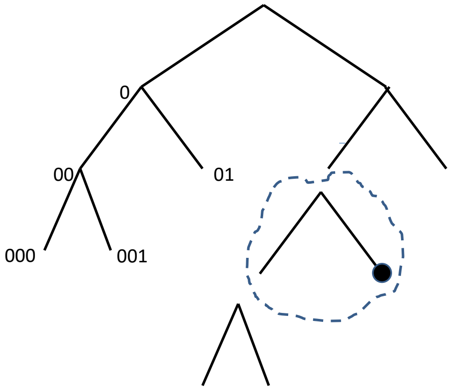


Figure: Delete the chosen leaf and its sibling.

Backward dynamics for the Rémy chain – continued

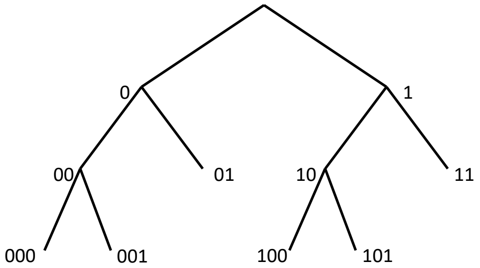


Figure: Close up the gap if there is one.

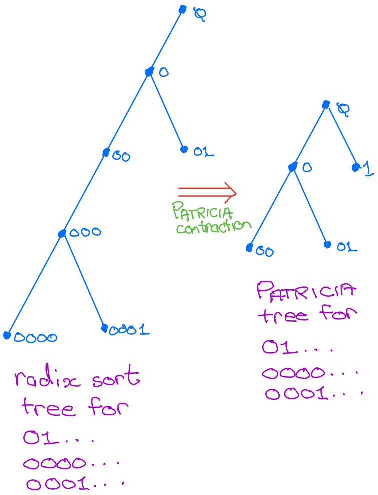
PATRICIA trees

- Recall that a depth first search of a radix sort tree visits the leaves in an order that coincides with the lexicographic order of the corresponding infinite binary words.
- The radix sort tree stores more information than is necessary for the purpose of sorting the infinite binary words into lexicographic order.
- More precisely, if one deletes the vertices with a single child in the radix sort tree and closes up the gaps, then a depth first search of the resulting PATRICIA tree still visits the leaves in an order that coincides with the lexicographic order of the corresponding infinite words.
- PATRICIA is an acronym for "Practical Algorithm To Retrieve Information Coded In Alphanumeric".

PATRICIA trees – continued

- Note that in a PATRICIA tree each non-leaf vertex of the tree has two children; that is, if the finite binary word $v = v_1 \ldots v_m$ is a vertex of the tree that is not a leaf, then both of the words $v_1 \ldots v_m 0$ and $v_1 \ldots v_m 1$ are also vertices of the tree.

Example of a PATRICIA tree



Some notation for PATRICIA trees

- Denote by $\bar{\mathbb{S}}$ the set of finite rooted binary trees which can arise as PATRICIA trees.
- A finite rooted binary tree belongs to $\bar{\mathbb{S}}$ if and only if each vertex has 2 or 0 children.
- Write $\bar{\mathbb{S}}_n$, $n \in \mathbb{N}$, for the elements of $\bar{\mathbb{S}}$ with n leaves.
- In particular, $\bar{\mathbb{S}}_1$ contains only the trivial tree with the single vertex $\{\varnothing\}$.

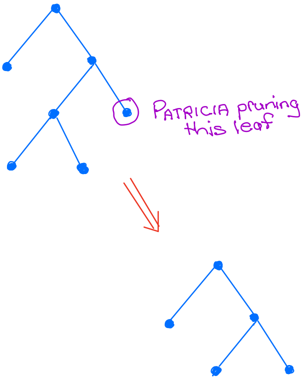
The PATRICIA contraction

- Recall that $\mathbb{S}$ is the class of finite rooted binary trees that can arise as radix sort trees.
- The PATRICIA contraction is the map $\Phi : \mathbb{S} \to \bar{\mathbb{S}}$ that deletes vertices with a single child and closes up the gaps.
- The PATRICIA contraction maps $\mathbb{S}_n$ to $\bar{\mathbb{S}}_n$ for all $n \in \mathbb{N}$.

Pruning a leaf from a PATRICIA tree

- Consider $\bar{\mathbf{t}} \in \bar{\mathbb{S}}_{n+1}$ and let $v = v_1 \ldots v_m$ be a leaf of $\mathbf{t}$.
- The finite rooted binary tree $\bar{\kappa}(\bar{\mathbf{t}}, v) \in \bar{\mathbb{S}}_n$ is obtained by removing $v = v_1 \ldots v_m$ and $v_1 \ldots v_{m-1}\bar{v}_m$ and closing up the gap if there is one (this will be the case if the sibling $v_1 \ldots v_{m-1}\bar{v}_m$ is not a leaf).
- If $\bar{\mathbf{t}}$ is the PATRICIA tree for the infinite binary inputs $z_1, \ldots, z_{n+1}$ and y_{n+1} is the leaf of $\mathbf{t}$ corresponding to the input z_{n+1}, then $\bar{\kappa}(\bar{\mathbf{t}}, y_{n+1})$ is the PATRICIA tree for the inputs $z_1, \ldots, z_n$.

Example of pruning a leaf from a PATRICIA tree



PATRICIA processes

- The PATRICIA process $({}^{\nu}\bar{R}_n)_{n \in \mathbb{N}}$ is obtained by taking ${}^{\nu}\bar{R}_n$, $n \in \mathbb{N}$, to be the PATRICIA tree for the i.i.d. inputs $Z_1, \ldots, Z_n$ with common distribution ν
- Note that ${}^{\nu}\bar{R}_n = \Phi({}^{\nu}R_n)$, $n \in \mathbb{N}$, where we recall that Φ is the PATRICIA contraction.

The simplest radix sort and PATRICIA processes

- Set $\gamma := \pi^{\otimes\infty}$, where $\pi(\{0\}) = \pi(\{1\}) = \frac{1}{2}$; that is, γ is fair coin-tossing measure on $\{0,1\}^\infty$.
- Write $R_n := {}^\gamma R_n$ and $\bar{R}_n := {}^\gamma \bar{R}_n$.

The PATRICIA processes are Markov

- A PATRICIA process $({}^{\nu}\bar{R}_n)_{n\in\mathbb{N}}$ is Markov.
- For $\bar{\mathbf{s}} \in \bar{\mathbb{S}}_n$ and $\bar{\mathbf{t}} \in \bar{\mathbb{S}}_{n+1}$, the associated backward transition probability is

$$\mathbb{P}\{{}^{\nu}\bar{R}_n = \bar{\mathbf{s}} \mid {}^{\nu}\bar{R}_{n+1} = \bar{\mathbf{t}}\} = \frac{1}{n+1}\#\{v : \bar{\mathbf{s}} = \bar{\kappa}(\bar{\mathbf{t}}, v)\}.$$

- That is, $({}^{\nu}\bar{R}_n)_{n\in\mathbb{N}}$ evolves backward in time by picking leaves uniformly at random and pruning them.
- NOTE: The backward transition probabilities of $({}^{\nu}\bar{R}_n)_{n\in\mathbb{N}}$ are the same for all ν.

The Rémy chain and the PATRICIA processes

- The Rémy chain starts with the finite rooted full binary tree $\aleph = \{\emptyset, 0, 1\}$.
- A PATRICIA process starts with finite rooted full binary tree $\{\emptyset\}$.
- However, both processes have the same backward transition probabilities.

What do all these processes have in common?

- We have a Markov chain $(X_n)_{n \in \mathbb{N}_0}$ with a countable state space E.
- The state space is partitioned into disjoint pieces $E_0 \sqcup E_1 \sqcup E_2 \sqcup \ldots$, where $E_0 = \{e\}$ for some distinguished state e.
- The transition probabilities satisfy $p(i, j) = 0$ unless $i \in E_n$ and $j \in E_{n+1}$ for some $n \in \mathbb{N}_0$.
- Consequently, if $X_0 = e$, then $X_n \in E_n$ for all $n \in \mathbb{N}_0$.

Bridges

- A bridge for the type of Markov chain X we are considering is a Markov chain Y with:
  - The same state space as X.
  - Initial state $Y_0 = e$.
  - The same backward transition probabilities as X.

The central question

- WHAT ARE ALL THE BRIDGES FOR A GIVEN MARKOV CHAIN?

The central question sharpened

- A mixture of two bridges is a bridge. So what we really want to know is:
- WHAT ARE ALL THE EXTREMAL BRIDGES FOR A GIVEN MARKOV CHAIN?
- FACT: A bridge is extremal if and only if it has almost surely trivial tail σ-field.

The extremal bridges for the simplest radix sort chain

- Recall the simplest radix sort chain $R := {}^{\gamma}R$, where γ is fair coin-tossing measure.
- We have observed that each chain of the form ${}^{\nu}R$ is a bridge for R.
- We will show that these bridges are extremal and they are the only extremal bridges for R.

A conjecture

- We have seen that the PATRICIA chains ${}^{\nu}\bar{R}$ are bridges for the Rémy chain T or, equivalently the simplest PATRICIA chain $\bar{R}$.
- The above result for the simplest radix sort chain R suggest that all the extremal bridges for T (equivalently, $\bar{R}$) are of the form ${}^{\nu}\bar{R}$.

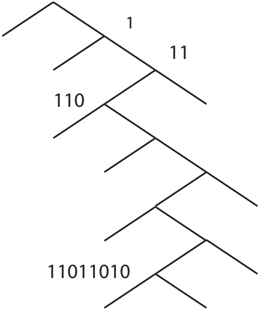


Figure: The value at time n of an extremal Rémy bridge. The tree consists of leaves hanging off a single spine that moves to the left or right according to successive tosses of a fair coin. It is clear that this chain is not of the form ${}^{\nu}\bar{R}$ for any ν.

What are the extremal Rémy bridges?

- What are the extremal bridges for the Rémy chain T (equivalently, the simplest PATRICIA process $\bar{R}$)?
- Write $(T_n^\infty)_{n\in\mathbb{N}}$ for an extremal Rémy bridge.

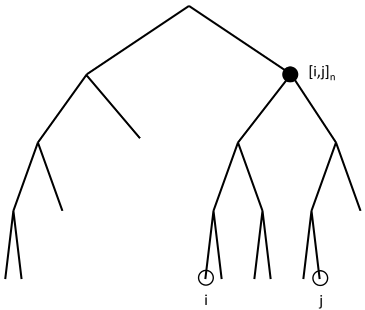
Consistently labeling the leaves in a Rémy bridge

By Kolmogorov's extension theorem, there is a Markov process $(\tilde{T}_n^\infty)_{n\in\mathbb{N}}$ such that for each $n \in \mathbb{N}$ the random element $\tilde{T}_n^\infty$ is a leaf-labeled binary tree with $n+1$ leaves labeled by $[n+1]$ and the following hold.

- The binary tree obtained by removing the labels of $\tilde{T}_n^\infty$ is T_n^∞.
- For every $n \in \mathbb{N}$, the conditional distribution of $\tilde{T}_n^\infty$ given T_n^∞ is uniform over the $(n+1)!$ possible labelings of T_n^∞.
- In going backward from time $n+1$ to time n, $\tilde{T}_{n+1}^\infty$ is transformed into $\tilde{T}_n^\infty$ as follows:
  - The leaf labeled $n+2$ is deleted, along with its sibling.
  - If the sibling of the leaf labeled $n+2$ is also a leaf, then the common parent (which is now a leaf) is assigned the sibling's label.

Most recent common ancestors

- We want to use the labeling and a projective construction to build an infinite binary-tree-like structure for which $\mathbb{N}$ plays the role of the leaves.
- If $i, j \in \mathbb{N}$ are the labels of two leaves T_n^∞ that are represented as the words $u_1 \dots u_k$ and $v_1 \dots v_\ell$ in $\{0,1\}^\star$, then set $[i,j]_n := u_1 \dots u_m = v_1 \dots v_m$, where $m := \max\{h : u_h = v_h\}$.
- That is, $[i,j]_n$ is the most recent common ancestor in T_n^∞ of the leaves labeled i and j.



Interior vertices

- Define an equivalence relation $\equiv$ on the Cartesian product $\mathbb{N} \times \mathbb{N}$ by declaring that $(i', j') \equiv (i'', j'')$ if and only if $[i', j']_n = [i'', j'']_n$ for some (and hence all) n such that $i', j', i'', j'' \in [n+1]$.
- Write $\langle i, j \rangle$ for the equivalence class of the pair (i, j).
- Think of $\langle i, j \rangle$ as the being the most recent common ancestor of the leaves i and j and of such points being interior vertices of a tree-like object.

Ordering equivalence classes – below and to the left

- Define a partial order $<_L$ on the set of equivalence classes by declaring for $(i', j'), (i'', j'') \in \mathbb{N} \times \mathbb{N}$ that $\langle i', j' \rangle <_L \langle i'', j'' \rangle$ if and only if for some (and hence all) n such that $i', j', i'', j'' \in [n+1]$ we have $[i', j']_n = u_1 \ldots u_k$ and $[i'', j'']_n = u_1 \ldots u_k 0 v_1 \ldots v_\ell$ for some $u_1, \ldots, u_k, v_1, \ldots, v_\ell \in \{0, 1\}$.
- Interpret the ordering $\langle i', j' \rangle <_L \langle i'', j'' \rangle$ as the "vertex" $\langle i'', j'' \rangle$ being below and to the left of the "vertex" $\langle i', j' \rangle$.

Ordering equivalence classes – below and to the right

- Similarly, define another partial order $<_R$ by declaring that $\langle i', j' \rangle <_R \langle i'', j'' \rangle$ if and only if for some (and hence all) n such that $i', j', i'', j'' \in [n+1]$ we have $[i', j']_n = u_1 \ldots u_k$ and $[i'', j'']_n = u_1 \ldots u_k 1 v_1 \ldots v_\ell$ for some $u_1, \ldots, u_k, v_1, \ldots, v_\ell \in \{0, 1\}$.
- Interpret the ordering $\langle i', j' \rangle <_R \langle i'', j'' \rangle$ as the "vertex" $\langle i'', j'' \rangle$ being below and to the right of the "vertex" $\langle i', j' \rangle$.

Yet another partial order – below

- Define a third partial order $<$ on the set of equivalence classes of $\mathbb{N} \times \mathbb{N}$ by declaring that $\langle i', j' \rangle < \langle i'', j'' \rangle$ if either $\langle i', j' \rangle <_L \langle i'', j'' \rangle$ or $\langle i', j' \rangle <_R \langle i'', j'' \rangle$.
- Interpret the ordering $\langle i', j' \rangle < \langle i'', j'' \rangle$ as the "vertex" $\langle i'', j'' \rangle$ being below the "vertex" $\langle i', j' \rangle$.

Properties of the equivalence relation and partial orders

The equivalence relation $\equiv$ and the partial orders $<_L$, $<_R$, and $<$ have the following properties.

(A) For $i, j \in \mathbb{N}$, $(i,j) \equiv (j,i)$.

(B) For distinct $i, j \in \mathbb{N}$, either $\langle i,j \rangle <_L \langle i,i \rangle$ and $\langle i,j \rangle <_R \langle j,j \rangle$, or $\langle i,j \rangle <_R \langle i,i \rangle$ and $\langle i,j \rangle <_L \langle j,j \rangle$.

(C) *"Triplet property"* For distinct i, j, k, exactly one of

$$\langle i,j \rangle = \langle i,k \rangle < \langle j,k \rangle$$

$$\langle j,k \rangle = \langle j,i \rangle < \langle k,i \rangle$$

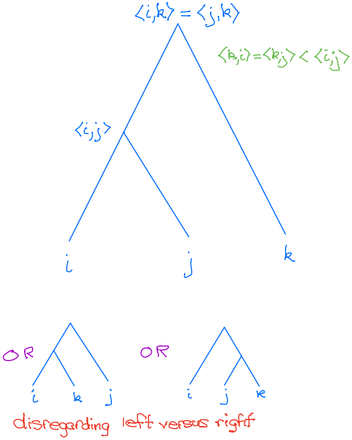
or

$$\langle k,i \rangle = \langle k,j \rangle < \langle i,j \rangle$$

is valid.

(D) For $i, j, k, \ell \in \mathbb{N}$, at most one of the relations $\langle i,j \rangle <_L \langle k,\ell \rangle$ and $\langle i,j \rangle <_R \langle k,\ell \rangle$ can hold and $\langle i,j \rangle < \langle k,\ell \rangle$ if and only if either $\langle i,j \rangle <_L \langle k,\ell \rangle$ or $\langle i,j \rangle <_R \langle k,\ell \rangle$.

(E) Fix $f, g, h, i, j, k \in \mathbb{N}$. If $\langle f,g \rangle <_L \langle h,i \rangle < \langle j,k \rangle$, then $\langle f,g \rangle <_L \langle j,k \rangle$. Similarly, if $\langle f,g \rangle <_R \langle h,i \rangle < \langle j,k \rangle$, then $\langle f,g \rangle <_R \langle j,k \rangle$.


⟨i,k⟩ = ⟨j,k⟩
⟨k,i⟩ = ⟨k,j⟩ < ⟨i,j⟩
⟨i,j⟩
i
j
k
OR
i k j
OR
i j k
disregarding left versus right

Didendritic systems

- Suppose that $\mathcal{N}$ is a set, $\equiv$ is an equivalence relation on $\mathcal{N}$, and each of $<_L, <_R, <$ is a partial order on the resulting collection of equivalence classes. We say that $\mathbf{D} = (\mathcal{N}, \equiv, \langle \cdot, \cdot \rangle, <_L, <_R, <)$ is a didendritic system if the properties (A)-(E) hold with $\mathbb{N}$ replaced by $\mathcal{N}$.
- A didendritic system with a finite label set $\mathcal{N}$ is just a rooted full binary tree with its leaves labeled by $\mathcal{N}$.

Didendritic systems and sequences of leaf-labeled finite rooted full binary trees

- Suppose that $(\tilde{t}_n)_{n \in \mathbb{N}}$ is a sequence of finite rooted full binary trees such that the leaves of $\tilde{t}_n$, $n \in \mathbb{N}$, are labeled by $[n+1]$. Assume that this sequence is is consistent in the sense that $\tilde{t}_n$ is produced from $\tilde{t}_{n+1}$ by:
  - deleting leaf labeled $n+2$ along with its sibling;
  - if the sibling was also a leaf, assigning its label to the common parent (which is now a leaf).

  Then there is a corresponding didendritic system constructed as above for $(\tilde{T}_n^\infty)_{n \in \mathbb{N}}$.
- Moreover, any didendritic system on $\mathbb{N}$ arises in this way for a unique consistent sequence $(\tilde{t}_n)_{n \in \mathbb{N}}$.

Permuting a didendritic systems

- Given a didendritic system $\mathbf{D} = (\mathbb{N}, \equiv, \langle \cdot, \cdot \rangle, <_L, <_R, <)$ and a permutation σ of $\mathbb{N}$, the didendritic system $\mathbf{D}^\sigma = (\equiv^\sigma, \langle \cdot, \cdot \rangle^\sigma, <_L^\sigma, <_R^\sigma)$ is defined by
  - $(i', j') \equiv^\sigma (i'', j'')$ if and only if $(\sigma(i'), \sigma(j')) \equiv (\sigma(i''), \sigma(j''))$,
  - $\langle i, j \rangle^\sigma$ is the equivalence class of the pair (i, j) for the equivalence relation $\equiv^\sigma$,
  - $\langle h, i \rangle^\sigma <_L^\sigma \langle j, k \rangle^\sigma$ if and only if $\langle \sigma(h), \sigma(i) \rangle <_L \langle \sigma(j), \sigma(k) \rangle$,
  - $\langle h, i \rangle^\sigma <_R^\sigma \langle j, k \rangle^\sigma$ if and only if $\langle \sigma(h), \sigma(i) \rangle <_R \langle \sigma(j), \sigma(k) \rangle$.

Exchangeable random didendritic systems

- A random didendritic system $\mathbf{D} = (\mathbb{N}, \equiv, \langle \cdot, \cdot \rangle, <_L, <_R, <)$ is exchangeable if for each permutation σ of $\mathbb{N}$ such that $\sigma(i) = i$ for all but finitely many $i \in \mathbb{N}$ the random didendritic system $\mathbf{D}^\sigma$ has the same distribution as $\mathbf{D}$.

Didendritic systems and Rémy bridges

- The random didendritic system corresponding to the labeled version of a Rémy bridge is exchangeable.
- Conversely, the sequence of random leaf-labeled finite rooted full binary trees produced from an exchangeable random didendritic system is the labeled version of a Rémy bridge.

Ergodic exchangeable random didendritic systems

- An exchangeable random didendritic system $\mathbf{D}$ is ergodic if

$$\mathbb{P}(\{\mathbf{D} \in A\} \triangle \{\mathbf{D}^\sigma \in A\}) = 0$$

for some measurable set A for all permutations σ of $\mathbb{N}$ with $\sigma(i) = i$ for all but finitely many $i \in \mathbb{N}$ implies that

$$\mathbb{P}\{\mathbf{D} \in A\} \in \{0, 1\}.$$

- Any exchangeable random didendritic system is a mixture of ergodic exchangeable random didendritic systems.
- The tail σ-field of a Rémy bridge is almost surely trivial (equivalently, the Rémy bridge is extremal) if and only if the corresponding exchangeable random didendritic system is ergodic.

Constructing a distance

- Consider an extremal Rémy bridge and the corresponding ergodic exchangeable random didendritic system. By de Finetti and the strong law of large numbers,

$$d(i,j) := \lim_{n \to \infty} \frac{1}{n} \sum_{p=1}^{n} \mathbb{1}\{\langle i,j \rangle \leqslant p\}$$

  exists for each pair $i, j \in \mathbb{N}$.
- Almost surely, d is an ultrametric on $\mathbb{N}$. That is, almost surely the following hold.
  - For all $i, j \in \mathbb{N}$, $d(i,j) \geqslant 0$, and $d(i,j) = 0$ if and only if $i = j$.
  - For all $i, j \in \mathbb{N}$, $d(i,j) = d(j,i)$.
  - For all $i, j, k \in \mathbb{N}$, $d(i,k) \leqslant d(i,j) \vee d(j,k)$.
- *A fortiori*, d is almost surely a metric on $\mathbb{N}$.

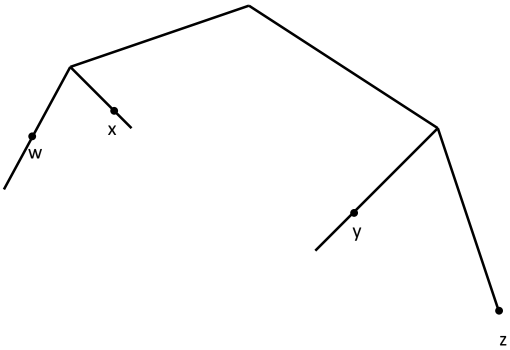
- A segment in a metric space (X, d) is the image of an isometry $\alpha : [a, b] \to X$. The endpoints of the segment are $\alpha(a)$ and $\alpha(b)$.
- A metric space (X, d) is geodesic if for all $x, y \in X$ there is a segment in X with endpoints $\{x, y\}$.
- An $\mathbb{R}$-tree is a metric space (X, d) with the following properties.
  - The space (X, d) is geodesic.
  - If two segments of (X, d) intersect in a single point, which is an endpoint of both, then their union is a segment.
- Fact: If (X, d) is an $\mathbb{R}$-tree, then for all $x, y \in X$ there is a unique segment in X with endpoints $\{x, y\}$.

Constructing an $\mathbb{R}$-tree

- The most recent common ancestor of $\langle h, i\rangle$ and $\langle j, k\rangle$ is of the form $\langle \ell, m\rangle$, where $\ell \in \{h, i\}$ and $m \in \{j, k\}$.
- In terms of the metric d, ℓ and m are any such pair for which $d(\ell, m) = d(h, j) \vee d(h, k) \vee d(i, j) \vee d(i, k)$.
- We therefore extend the metric by setting

$$\begin{aligned} d(\langle h, i\rangle, \langle j, k\rangle) &= \frac{1}{2}([d(\ell, m) - d(h, i)] + [d(\ell, m) - d(j, k)]) \\ &= d(h, j) \vee d(h, k) \vee d(i, j) \vee d(i, k) - \frac{1}{2}(d(h, i) + d(j, k)). \end{aligned}$$

- This extension is indeed a metric and the four point condition holds; that is for equivalence classes w, x, y, z at least one of the following sets of conditions holds
  - $d(w,x) + d(y,z) \leqslant d(w,y) + d(x,z) = d(w,z) + d(x,y)$,
  - $d(w,z) + d(x,y) \leqslant d(w,x) + d(y,z) = d(w,y) + d(x,z)$,
  - $d(w,y) + d(x,z) \leqslant d(w,z) + d(x,y) = d(w,x) + d(y,z)$.

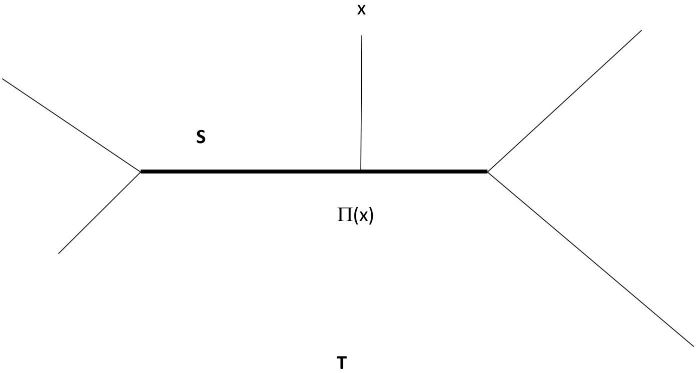


Constructing an $\mathbb{R}$-tree – continued

- Because the four point condition holds, we can embed $\{\langle i, j \rangle : i, j \in \mathbb{N}\}$ in a distance-preserving manner into a minimal, complete $\mathbb{R}$-tree $\mathbf{T}$ with a root ρ in such a way that $\langle i, j \rangle < \langle k, \ell \rangle$ if and only if $\langle i, j \rangle$ is on the geodesic segment from ρ to $\langle k, \ell \rangle$.
- That is, the natural partial order on $(\mathbf{T}, \rho)$ extends the partial order $<$ on the embedded set $\{\langle i, j \rangle : i, j \in \mathbb{N}\}$.

Core and projections

- The core $\mathbf{S}$ of $\mathbf{T}$ is the smallest closed subtree containing $\{\langle i,j\rangle : i,j \in \mathbb{N},\ i \neq j\}$.
- The projection $\Pi(x)$ of $x \in \mathbf{T}$ onto $\mathbf{S}$ is the unique point in $\mathbf{S}$ closest to x.



Sampling measure

- Put $\xi_k := \Pi(k) = \Pi(\langle k, k\rangle) \in \mathbf{S}$.
- The equivalence relation $\equiv$ and the partial order $<$ can be reconstructed from $(\xi_k)_{k\in\mathbb{N}}$.
- By de Finetti and ergodicity, $(\xi_k)_{k\in\mathbb{N}}$ is an i.i.d. sequence with common distribution μ, where μ is a diffuse probability measure on $\mathbf{S}$.

Distinguishing left from right

- Any didendritic system $(\equiv, \langle\cdot,\cdot\rangle, <_L, <_R)$ is uniquely determined by the equivalence relation $\equiv$, the partial order $<$, and a determination for each pair of distinct labeled leaves $i, j \in \mathbb{N}$ whether

$$\langle i,j\rangle <_L i \quad \text{and} \quad \langle i,j\rangle <_R j$$
$$\text{or}$$
$$\langle i,j\rangle <_L j \quad \text{and} \quad \langle i,j\rangle <_R i.$$

- For an ergodic exchangeable random didendritic system, define J_{ij}, $i, j \in \mathbb{N}$, $i \neq j$, by $J_{ij} = 0$ (resp. $J_{ij} = 1$) if the former (resp. latter) alternative holds.
- The array J is jointly exchangeable and ergodic.

Distinguishing left from right – continued

- By the Aldous-Hoover-Kallenberg theory of jointly exchangeable arrays, we may suppose that on some extension of our underlying probability space there exist i.i.d. random variables $(U_i)_{i \in \mathbb{N}}$, and $(U_{ij})_{i,j \in \mathbb{N},\, i<j}$ that are uniform on $[0,1]$ and a function F such that

$$J_{ij} = F(\xi_i, U_i, \xi_j, U_j, U_{ij}),$$

  where $U_{ij} = U_{ji}$ for $i > j$.
- With some extra work, we can show that

$$J_{ij} = G(\xi_i, U_i, \xi_j, U_j)$$

  for a suitable function G.

Where have we got to?

- Any Rémy bridge is a mixture of extremal Rémy bridges (i.e. ones with trivial tail σ-fields).
- There is a bijection between extremal Rémy bridges and ergodic exchangeable random didendritic systems.
- Any ergodic exchangeable random didendritic system is determined by
  - a complete, separable $\mathbb{R}$-tree $\mathbf{S}$,
  - a distinguished root $\rho \in \mathbf{S}$,
  - a diffuse sampling probability measure μ on $\mathbf{S}$,
  - a "left-vs-right" function $G : \mathbf{S} \times [0,1] \times \mathbf{S} \times [0,1] \to \{0,1\}$.
- The ensemble $(\mathbf{S}, \rho, \mu, G)$ has to satisfy certain obvious consistency conditions (e.g. for $\mu^{\otimes 3}$-a.e. $(x,y,z) \in \mathbf{S}^3$, two of the three geodesic segments $[\rho, x] \cap [\rho, y]$, $[\rho, x] \cap [\rho, z]$, $[\rho, y] \cap [\rho, z]$ are equal and these two are strictly contained in the third).
- Conversely, any ensemble $(\mathbf{S}, \rho, \mu, G)$ that satisfies the consistency conditions gives rise to an ergodic exchangeable random didendritic system and hence to an extremal Rémy bridge.

Example

Recall the Rémy bridge whose value at time n is a finite rooted full binary tree consisting of $n+1$ vertices along a single spinal path that has n leaves coming off to the left of right according to tosses of a fair coin.
Here we may take

- the complete separable $\mathbb{R}$-tree $\mathbf{S}$ to be $[0,1]$ equipped with the usual metric,
- the root ρ to be the point $0 \in [0,1]$,
- the diffuse sampling probability measure μ to be Lebesgue measure on $[0,1]$,
- the "left-vs-right" function $G : \mathbf{S} \times [0,1] \times \mathbf{S} \times [0,1] \to \{0,1\}$ to be given by

$$G(x,u,y,v) = \begin{cases} 1, & \text{if } x < y \text{ and } u < \frac{1}{2}, \\ 0, & \text{if } x < y \text{ and } u > \frac{1}{2}, \\ 1, & \text{if } y < x \text{ and } v < \frac{1}{2}, \\ 0, & \text{if } y < x \text{ and } v > \frac{1}{2}, \\ 0, & \text{otherwise.} \end{cases}$$

Why are things much simpler for the radix sort chain?

- This seems to be a subtle point that depends on monotonicity that is present in the radix sort chain which is not present in the Rémy/PATRICIA chain.

Labeling the leaves of a radix sort bridge

- Suppose that $(R_n^\infty)_{n \in \mathbb{N}}$ is a radix sort bridge.
- By Kolmogorov's extension theorem we may suppose that there is a Markov process $(\tilde{R}_n^\infty)_{n \in \mathbb{N}}$ such that for each $n \in \mathbb{N}$ the random element $\tilde{R}_n^\infty$ is a leaf-labeled rooted binary tree with n leaves labeled by $[n]$ and the following hold.
  - The rooted binary tree obtained by removing the labels of $\tilde{R}_n^\infty$ is R_n^∞.
  - For every $n \in \mathbb{N}$, the conditional distribution of $\tilde{R}_n^\infty$ given R_n^∞ is uniform over the $n!$ possible labelings of R_n^∞.
  - In going backward from time $n+1$ to time n, $\tilde{R}_{n+1}^\infty$ is transformed into $\tilde{R}_n^\infty$ by pruning the leaf labeled $n+1$.

Limit labels

- Given $i \in [n]$, let $\langle i \rangle_n \in \{0,1\}^\star$ be the leaf of R_n^∞ labeled i in $\tilde{R}_n^\infty$.
- Observe that $\langle i \rangle_i \leqslant \langle i \rangle_{i+1} \leqslant \ldots$ and so $\langle i \rangle_\infty = \lim_{n\to\infty} \langle i \rangle_n \in \{0,1\}^\star \sqcup \{0,1\}^\infty$ is well-defined.
- For distinct $i, j \in \mathbb{N}$, the most recent common ancestor $\langle i \rangle_n \wedge \langle j \rangle_n$ is the same for all $n \geqslant i \wedge j$ and coincides with $\langle i \rangle_\infty \wedge \langle j \rangle_\infty$.

CLAIM: The sequence $(\langle i \rangle_\infty)_{i \in \mathbb{N}}$ is exchangeable.

PROOF: It is clear by construction that $(\langle i \rangle_n)_{i \in [n]}$ is (finitely) exchangeable and the claim follows upon taking limits as $n \to \infty$.

Tail triviality

CLAIM: The tail σ-field of $(R_n^\infty)_{n\in\mathbb{N}}$ is $\mathbb{P}$-a.s. trivial if and only if $(\langle i \rangle_\infty)_{i\in\mathbb{N}}$ is an independent identically distributed sequence.

PROOF: The bijective correspondence between the distributions of the bridges $(R_n^\infty)_{n\in\mathbb{N}}$ and the distributions of their labeled versions $(\tilde{R}_n^\infty)_{n\in\mathbb{N}}$ is compatible with convex combinations, and hence preserves extremality.
Therefore the tail σ-field of the bridge $(R_n^\infty)_{n\in\mathbb{N}}$ is $\mathbb{P}$-a.s. trivial if and only if the exchangeable sequence $(\langle i \rangle_\infty)_{i\in\mathbb{N}}$ is ergodic.
A well-known consequence of de Finetti's theorem is that an exchangeable sequence is ergodic if and only if it is independent and identically distributed.

Diffuseness

CLAIM: If $(\langle i \rangle_\infty)_{i \in \mathbb{N}}$ is independent and identically distributed with common distribution ν, then ν is concentrated on $\{0,1\}^\infty$ and diffuse.

PROOF: For any $u \in \{0,1\}^\star$, the sequence $(\mathbb{1}\{u = \langle k \rangle_\infty\})_{k \in \mathbb{N}}$ is independent and identically distributed, and hence $\#\{k \in \mathbb{N} : u = \langle k \rangle_\infty\} = 0$ $\mathbb{P}$-a.s. or $\#\{k \in \mathbb{N} : u = \langle k \rangle_\infty\} = \infty$ $\mathbb{P}$-a.s.
Now, if $\mathbb{P}\{\langle i \rangle_\infty \in \{0,1\}^\star\} > 0$ there would be a $u \in \{0,1\}^\star$ such that with positive probability $\langle i \rangle_n = \langle i \rangle_\infty = u$ for all n sufficiently large.
Then, on the event $\{\langle i \rangle_\infty = u\}$ we would have $\#\{k \in \mathbb{N} : \langle k \rangle_\infty = u\} = 1$, because $\langle j \rangle_\infty \neq \langle i \rangle_\infty$ for $j \neq i$ when $\langle i \rangle_\infty \in \{0,1\}^\star$.
This shows that $\mathbb{P}\{\langle i \rangle_\infty \in \{0,1\}^\star\} = 0$.

Diffuseness – continued

We therefore have that $(\langle k \rangle_\infty)_{k \in \mathbb{N}}$ is an independent identically distributed sequence of $\{0,1\}^\infty$-valued random variables.
Because $\langle i \rangle_\infty \wedge \langle j \rangle_\infty = \langle i \rangle_n \wedge \langle j \rangle_n \in \{0,1\}^\star$ for all $n \geqslant i \vee j$ $\mathbb{P}$-a.s. when $i \neq j$, it follows that $\langle i \rangle_\infty \neq \langle j \rangle_\infty$ $\mathbb{P}$-a.s. for $i \neq j$ and the common distribution of $(\langle k \rangle_\infty)_{k \in \mathbb{N}}$ is diffuse.

Identification of the extremal bridges

CLAIM: The tail σ-field of $(R_n^\infty)_{n \in \mathbb{N}}$ is $\mathbb{P}$-a.s. trivial if and only if $(R_n^\infty)_{n \in \mathbb{N}}$ has the same distribution as $({}^\nu R_n)_{n \in \mathbb{N}}$ for some diffuse probability measure ν on $\{0,1\}^\infty$.

PROOF: We have already seen that when ν is a diffuse probability measure on $\{0,1\}^\infty$ the process $({}^\nu R_n)_{n \in \mathbb{N}}$ is a bridge which, by the Hewitt-Savage zero-one law, has a trivial tail σ-field.

Identification of the extremal bridges – continued

Conversely, suppose that the bridge $(R_n^\infty)_{n \in \mathbb{N}}$ has a trivial tail σ-field. Let ν be the common diffuse distribution of the independent, identically distributed sequence of $\{0,1\}^\infty$-valued random variables $(\langle i \rangle_\infty)_{i \in \mathbb{N}}$. It is clear that $R_n^\infty = \mathbf{R}(\langle 1 \rangle_\infty, \ldots, \langle n \rangle_\infty)$, $n \in \mathbb{N}$, and so $(R_n^\infty)_{n \in \mathbb{N}}$ has the same distribution as $({}^\nu R_n)_{n \in \mathbb{N}}$.

[]

Steven N. Evans, Rudolf Grübel, and Anton Wakolbinger, *Trickle-down processes and their boundaries*, Electron. J. Probab. **17** (2012), no. 1, 58. MR 2869248

______, *Doob-Martin boundary of Rémy's tree growth chain*, Ann. Probab. **45** (2017), no. 1, 225–277. MR 3601650

Steven N. Evans and Anton Wakolbinger, *Radix sort trees in the large*, Electron. Commun. Probab. **22** (2017), Paper No. 68, 13. MR 3734107

______, *Patricia bridges*, Genealogies of Interacting Particle Systems (M. Birkner, R. Sun, and J. Swart, eds.), Lecture Note Series, vol. 38, Institute for Mathematical Sciences, National University of Singapore, World Scientific, 2020, pp. 233–267.