PATRICIA BRIDGES

STEVEN N. EVANS AND ANTON WAKOLBINGER

Abstract. Given a finite collection of pairwise distinct infinite binary words, there is a minimal length finite prefix for each word such that the resulting collection of finite prefixes is also pairwise distinct. The set of initial segments of these finite prefixes may be identified with the vertices of a finite rooted tree in which each interior vertex has a left child, a right child or both, and the finite prefixes themselves correspond to the leaves of the tree. A depth first search of this radix sort tree produces an ordering of the leaves that agrees with the lexicographic ordering of the corresponding infinite binary words. The radix sort tree stores information that is redundant for the purpose of sorting the input infinite binary words into lexicographic order. Indeed, if one deletes the out-degree 1 vertices in the radix sort tree and “closes up the gaps”, then the resulting PATRICIA tree maintains all the information that is necessary for sorting.

We investigate the PATRICIA chains – the tree-valued Markov chains that arise when successively building the PATRICIA trees for the collection of infinite binary words $Z_1, \ldots, Z_n$, $n = 1, 2, \ldots$, where the source words $Z_1, Z_2, \ldots$ are independent and have a common diffuse distribution on $\{0,1\}^\infty$. It turns out that the PATRICIA chains share a common collection of backward transition probabilities and that these are the same as those of a chain introduced by Rémy for successively generating uniform random binary trees with larger and larger numbers of leaves. This means that the infinite bridges of any PATRICIA chain (that is, the chains obtained by conditioning a PATRICIA chain on its remote future) coincide with the infinite bridges of the Rémy chain. The infinite bridges of the Rémy chain are characterized concretely in Evans, Grübel, and Wakolbinger 2017 and we recall that characterization here while adding some details and clarifications.

1. Introduction

We first fix some notation. Denote by $\{0,1\}^\ast := \bigsqcup_k \{0,1\}_k^k$ the set of finite tuples or words drawn from the alphabet $\{0,1\}$ (with the empty word $\emptyset$ allowed) – the symbol $\bigsqcup$ emphasizes that this is a disjoint union. Write an $\ell$-tuple $v = (v_1, \ldots, v_\ell) \in \{0,1\}^\ast$ more simply as $v_1 \ldots v_\ell$ and set $|v| = \ell$.

Define a partial order on $\{0,1\}^\ast \cup \{0,1\}_\infty$ by declaring that $u < v$ if and only if $u$ is a strict prefix of $v$; that is, $u < v$ if and only if $u = u_1 \ldots u_k$ for some $k \geq 0$, $v$ is of the form $v_1 \ldots v_\ell \in \{0,1\}_k^\ast$ where $k < \ell$ or $v_1 v_2 \ldots \in \{0,1\}_\infty$, and $v_1 \ldots v_k = u_1 \ldots u_k$. The empty word is the unique minimal element for this partial order.

Date: June 19, 2018.

2010 Mathematics Subject Classification. Primary 60J50, secondary 60J10, 68W40.

Key words and phrases. binary tree, tail $\sigma$-field, exchangeability, radix sort, Markov function, didendritic system.

SNE supported in part by NSF grant DMS-1512933, and NIH grant 1R01GM109454-01. AW supported in part by DFG priority program 1590.
A rooted binary tree (or binary tree for short) is a non-empty subset \( t \) of \( \{0,1\}^* \) with the property that if \( v \in t \) and \( u \in \{0,1\}^* \) is such that \( u < v \), then \( u \in t \). The vertex \( \emptyset \) (that is, the empty word) belongs to any such tree \( t \) and is the root of \( t \). For a finite rooted binary tree \( t \) the leaves of \( t \) are the maximal elements of \( t \) for the partial order \(<\). A finite rooted binary tree is uniquely determined by its leaves; it is the smallest rooted binary tree that contains the set of leaves and it consists of the leaves and the elements \( u \in \{0,1\}^* \) such that \( u < v \) for some leaf \( v \).

Suppose now that \( z_1, \ldots, z_n \in \{0,1\}^\infty \) 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 \( y_1, \ldots, y_n \) are distinct leaves of a binary tree and \( y_1, \ldots, y_n \) are the minimal length words with this property. The resulting binary tree is called the radix sort tree defined by the infinite words \( z_1, \ldots, z_n \): a depth first search of this tree visits the leaves in an order that coincides with the lexicographic order of the corresponding infinite 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 out-degree 1 vertices 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 were invented independently in [Mor68, Gwe68]. Note that a PATRICIA tree is a full binary 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 \( v_1 \ldots v_m0 \) and \( v_1 \ldots v_m1 \) are also vertices of the tree.

Suppose now that \( Z_1, Z_2, \ldots \) is an infinite i.i.d. sequence of random elements of \( \{0,1\}^\infty \) with common diffuse distribution \( \nu \). Let \( \nu^R_n, n \in \mathbb{N}, \) be the PATRICIA tree constructed from \( Z_1, \ldots, Z_n \). We show in Proposition 3.7 that \( (\nu^R_n)_{n \in \mathbb{N}} \) is a Markov chain which we call a PATRICIA chain. Features of PATRICIA trees for random inputs were first studied in [Knu73] and this topic has since been the subject of quite a large literature (see, for example, [Lec15, MS18, TGL16, Mag15, MKS14, TGL14, FL14, Dev05, KS02, Dev02, Bun01, RJS93, And92, Dev92, Szp90, KPS99]).

Our aim in this paper is to characterize the infinite bridges of the PATRICIA chains; that is, for each \( \nu \) we wish to characterize the family of Markov chains that have the same backward transition probabilities as \( (\nu^R_n)_{n \in \mathbb{N}} \). As we observe in Proposition 3.7, the backward transition probabilities of \( (\nu^R_n)_{n \in \mathbb{N}} \) are the same for all \( \nu \) and this common backward transition mechanism may be described as follows: pick a leaf uniformly at random, delete it and its sibling, and close up the gap in the tree if there is one. It follows that all the PATRICIA chains have the same infinite bridges.

Before we proceed, we need to recall the Rémy chain of [Rémi58] (see, also, [AS95, Mar03, LG99]). The state space of this chain is also the set of full binary trees. Writing \( \bar{N} \) for the full binary tree with two leaves, the Rémy chain evolves forwards in time as follows:

- Pick a vertex \( v \) uniformly at random.
- Cut off the subtree rooted at \( v \) and set it aside.
- Attach a copy of the tree \( \bar{N} \) to the end of the edge that previously led to \( v \).
- Re-attach the subtree that was rooted at \( v \) in uniformly at random to one of the two leaves in the copy of \( \bar{N} \).
The Rémy chain started with the trivial tree consisting of a single vertex has the important feature that after \( n \) steps it is uniformly distributed over the set of full binary trees with \( n \) leaves. The backward transition probabilities of the Rémy chain were identified in [EGW17] and they coincide with those of the PATRICIA chains. It follows that the common family of infinite bridges of the PATRICIA chains is the same as that of the Rémy chain and the latter was determined in [EGW17].

An outline of the remainder of the paper is as follows.

In Section 2 we introduce some notation and recall more formally the construction of a radix sort tree and a PATRICIA tree from a set of infinite binary words. In Section 3 we begin considering the stochastic processes built by applying the radix sort tree and PATRICIA tree constructions to sequences of independent identically distributed infinite binary words with common diffuse distribution \( \nu \). We observe that these processes are Markov, show that all the radix sort chains have the same family of backward transition probabilities, and that the same is true of the PATRICIA chains. We also investigate the issue that different diffuse probabilities \( \nu \) can give rise to the same PATRICIA chain.

Section 4 initiates the study of infinite bridges per se. Since the backward transition probabilities are the same for all \( \nu \) for both the radix sort chains and the PATRICIA chains, we concentrate on the case where \( \nu \) is fair coin-tossing measure. We also recall that the distribution of any infinite bridge for either the radix sort chain or the PATRICIA chain will be a mixture of extremal infinite bridge distributions and that an infinite bridge is extremal if and only if it has an almost surely trivial tail \( \sigma \)-field, so it suffices for the determination of the distributions of general infinite bridges to characterize the extremal ones. We recall from [EW17] that the extremal infinite bridges for the radix sort chain are nothing other than the radix sort chains with general input distribution \( \nu \). By the observation that all the PATRICIA chains have the same backward transition probabilities, it follows that a PATRICIA chain with general input distribution \( \nu \) is an infinite bridge for our “reference” PATRICIA chain with fair coin-tossing input distribution. Because of the fact we have just recalled for the radix sort chain, it is natural to hope that these chains (which, by the Hewitt-Savage zero-one law, clearly have almost surely trivial tail \( \sigma \)-fields) exhaust the collection of extremal infinite PATRICIA bridges. We present an example to show that this is not the case. Rather, to describe the totality of the extremal infinite PATRICIA bridges we need to recall from [EGW17] the description of the extremal infinite bridges for the Rémy chain because, as we have already noted, these two families coincide. This retelling takes up the remainder of the paper.

A crucial element of the development in [EGW17] was the introduction of the notion of a didendritic system with a given label set. This class of objects is a generalization of the class of finite leaf-labeled full binary trees that includes objects with infinitely many “leaves”. Section 5 contains a review of some facts about didendritic systems. In particular, we show that the class of didendritic systems with finite label sets is in a natural bijective correspondence with the class of finite leaf-labeled full binary trees. This fact was claimed in [EGW17] with a slightly different definition of didendritic system but we show by an example that under the axioms in [EGW17] this assertion is false. We stress that this gap does not illegitimate the further development in [EGW17] which does not depend on the details of the definition of a didendritic system but only on the two facts that
a finite didendritic system is effectively a finite leaf-labeled binary tree and that the class of didentritic systems is closed under projective limits in a natural way – both of which are true for our new definition. We also develop in an explicit manner an alternative description of the class of finite didendritic systems that was somewhat implicit in \cite{EGW17}. This alternative description, which is obtained in Proposition 5.9, is crucial for the later exposition.

We recall in Section 6 that there is a bijective correspondence between infinite PATRICIA (equivalently, Rémy) bridges and random didentric systems with label set $\mathbb{N}$ that are exchangeable in a natural sense. Moreover, the extremal infinite PATRICIA bridges correspond to the exchangeable random didentric systems with label set $\mathbb{N}$ that are ergodic, where ergodicity is equivalent to the property that the random didentritic systems induced on disjoint subsets of $\mathbb{N}$ are independent. In order to characterize the class consisting of all PATRICIA bridges it therefore suffices to determine the class of ergodic exchangeable random didendritic systems. We recall the concrete representation of the latter from \cite{EGW17} in Section 7. The ingredients of this representation are a rooted complete separable $\mathbb{R}$-tree, a diffuse probability measure on the $\mathbb{R}$-tree, and a possibly random mechanism for giving a left-versus-right ordering to the tree built by sampling countably many points from the $\mathbb{R}$-tree according to the given probability measure.

Finally, we note that an alternative concrete representation of extremal infinite Rémy bridges has recently been given in \cite{Ger18} as part of a program that extends the study of Markov chains with Rémy-like transition probabilities to classes of discrete structures other than binary trees.

2. **Radix sort trees and PATRICIA trees**

For a finite rooted binary tree $t$, we write $L(t)$ for the set of leaves of $t$. For $y_1, \ldots, y_m \in \{0,1\}^*$, write

$$T(y_1, \ldots, y_m) := \bigcup_{j=1}^m \{ u \in \{0,1\}^* : u \leq y_j \}$$

for the smallest finite rooted binary tree containing $y_1, \ldots, y_m \in \{0,1\}^*$; the leaves of this tree form a subset of $\{y_1, \ldots, y_m\}$ and this subset is proper if and only if $y_i < y_j$ for some pair $1 \leq i \neq j \leq m$.

A collection $z_1, \ldots, z_n$ of distinct elements of $\{0,1\}^\infty$ determines a finite rooted binary tree in the following manner. For $1 \leq i \leq n$ let $y_i$ be the minimal length prefix of $z_i$ that differs from all prefixes of $z_j, j \neq i$. The words $y_1, \ldots, y_n$ are distinct and, moreover, they are incomparable for the partial order $<$. The **radix sort tree determined by the input** $z_1, \ldots, z_n$ is the finite rooted binary tree $R(z_1, \ldots, z_n)$ with leaves $y_1, \ldots, y_n$; that is,

$$R(z_1, \ldots, z_n) := T(y_1, \ldots, y_n).$$

**Remark 2.1.** Observe that

$$R(z_1, \ldots, z_n) = R(z_{\sigma(1)}, \ldots, z_{\sigma(n)})$$

for any permutation $\sigma$ of $[n]$.

**Notation 2.2.** Denote by $S$ the class of binary trees that can arise as radix sort trees. A binary tree $s$ belongs to $S$ if and only if $s = \{\emptyset\}$ or $s$ has at least two leaves.
and for any leaf \( u_1 \ldots u_p \in S \) the word \( u_1 \ldots u_{p-1}u_p \) is also a vertex of \( S \), where we set \( \bar{0} := 1 \) and \( \bar{1} := 0 \). Write \( S_n, n \in \mathbb{N} \), for the elements of \( S \) with \( n \) leaves.

One of the reasons for constructing the radix sort tree for a set of inputs \( z_1, \ldots, z_n \) is that if \( y_1, \ldots, y_n \) are the leaves of the tree (where the indexing is such that \( y_i < z_i \) for \( 1 \leq i \leq n \)), then the lexicographic ordering of \( z_1, \ldots, z_n \) is the same as that of \( y_1, \ldots, y_n \); that is, \( \sigma \) is the unique permutation of \( [n] \) such that \( z_{\sigma(i)}, i \in [n] \), are increasing in the lexicographic order if and only if \( \sigma \) is the unique permutation of \( [n] \) such that \( y_{\sigma(i)}, i \in [n] \), are increasing in the lexicographic order. We now describe another procedure for associating a finite set of inputs with the leaves of a finite binary tree that shares this feature but typically uses a finite binary tree with fewer vertices.

**Notation 2.3.** Denote by \( \bar{S} \) the set of finite binary tree for which all vertices have out-degree 2 or 0. That is, \( \bar{S} \) consists of finite binary trees \( t \) such that if \( v \in t \), then either \( v \in L(t) \) or both \( v0 \in t \) and \( v1 \in t \). We call the elements of \( \bar{S} \) full finite binary trees. Write \( S_n, n \in \mathbb{N} \), for the elements of \( \bar{S} \) with \( n \) leaves. Note that \( \#S_n = C_{n-1} \), where \( C_n := \frac{1}{m+1} \binom{2m}{m} \) is the \( m \)th Catalan number, see, for example, [Sta15, Bijective Exercise 7]).

**Definition 2.4.** The PATRICIA contraction is a map \( \Phi : S \to \bar{S} \). It maps \( S_n \) to \( \bar{S}_n, n \geq 0 \). It is defined recursively as follows. For \( n = 0 \), put \( \Phi([\emptyset]) := \emptyset \). Now consider \( s \in S_n \) for \( n \geq 1 \). There is a unique maximal \( m \) and \( u_1 \ldots u_m \in \{0,1\}^m \) such that \( u_1 \ldots u_m < y \) for every leaf \( y \in L(s) \). We have a decomposition

\[
 s = \{\emptyset\} \cup u_1 \cup u_1u_2 \cup \ldots \cup u_1u_2 \ldots u_m \} \sqcup u_1 \ldots u_m 0s(0) \cup u_1 \ldots u_m 1s(1)
\]

for two binary trees \( s(0), s(1) \in S \) that both have fewer leaves than \( s \). Put

\[
 \Phi(s) := \{\emptyset\} \sqcup 0\Phi(s(0)) \sqcup 1\Phi(s(1)).
\]

**Remark 2.5.** The PATRICIA contraction of \( s \in S \) is the unique \( t \in \bar{S} \) with the following properties.

- Writing \( S(s) \) (respectively, \( S(t) \)) for the vertices of \( s \) (respectively, \( t \)) with out-degree 2, there is a bijective correspondence between \( S(s) \) and \( S(t) \) and a bijective correspondence between \( L(s) \) and \( L(t) \).
- Suppose that \( w, x \in s \) and \( y, z \in t \) are such that \( w \) corresponds to \( y \) and \( x \) corresponds to \( z \). Then \( w = u_1 \ldots u_p \) and \( x = u_1 \ldots u_p u_{p+1} \ldots u_q \) for some \( u_1, \ldots, u_q \in \{0,1\} \) with \( u_{p+1} = 0 \) (resp. \( u_{p+1} = 1 \)) if and only if \( y = v_1 \ldots v_m \) and \( z = v_1 \ldots v_m v_{m+1} \ldots v_n \) for some \( v_1, \ldots, v_n \in \{0,1\} \) with \( v_{m+1} = 0 \) (resp. \( v_{m+1} = 1 \)).

3. **Radix sort chains and PATRICIA chains**

**Definition 3.1.** Given a diffuse probability measure \( \nu \) on \( \{0,1\}^\infty \), let \( (Z_n)_{n \in \mathbb{N}} \) be a sequence of independent identically distributed \( \{0,1\}^\infty \)-valued random variables with common distribution \( \nu \). The corresponding radix sort chain \( (R_n)_{n \in \mathbb{N}} \) is defined by \( R_n := R(Z_1, \ldots, Z_n), n \in \mathbb{N} \), and the corresponding PATRICIA chain \( (\tilde{R}_n)_{n \in \mathbb{N}} \) is defined by \( \tilde{R}_n = \Phi(R_n), n \in \mathbb{N} \). Setting \( \gamma := \pi^\infty \), where \( \pi(\{0\}) = \pi(\{1\}) = \frac{1}{2} \), we will write \( R_n := \gamma R_n \) and \( \tilde{R}_n := \gamma \tilde{R}_n \).

The processes \( (R_n)_{n \in \mathbb{N}} \) and \( (\tilde{R}_n)_{n \in \mathbb{N}} \) are indeed Markovian. This observation does not seem to have appeared explicitly in the literature except in [EW17] for
the former process. We will state this formally along with a description of the respective backward transition probabilities, but we first need some notation.

**Notation 3.2.** Consider \( t \in S_{n+1} \) and let \( v = v_1 \ldots v_m \) be a leaf of \( t \).

Suppose first that \( v_1 \ldots v_{m-1} \bar{v}_m \) is not a leaf of \( t \). Let \( \kappa(t, v) \in S_n \) be the tree \( t \setminus \{v\} \). That is, \( \kappa(t, v) \) is the tree with the same leaf set as \( t \) except that \( v \) has been removed.

On the other hand, suppose that \( v_1 \ldots v_{m-1} \bar{v}_m \) is also a leaf of \( t \). There is a largest \( \ell < m \) such that \( v_1 \ldots v_{\ell} \) and \( v_1 \ldots v_{\ell-1} \bar{v}_m \) are both vertices of \( t \). In this case let \( \kappa(t, v) \in S_n \) be the tree \( t \setminus \{v_1 \ldots v_p : \ell + 1 \leq p \leq m \} \cup \{v_1 \ldots v_{m-1} \bar{v}_m\} \). That is, \( \kappa(t, v) \) is the tree with the same leaf set as \( t \) 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 \( v_1 \ldots v_{\ell} \).

**Remark 3.3.** If \( t = R(z_1, \ldots, z_{n+1}) \) for distinct \( z_1, \ldots, z_{n+1} \) and \( y_{n+1} \) is the leaf of \( t \) corresponding to the input \( z_{n+1} \), then \( \kappa(t, y_{n+1}) = R(z_1, \ldots, z_n) \).

**Notation 3.4.** Consider \( \bar{t} \in \bar{S}_{n+1} \) and let \( v = v_1 \ldots v_m \) be a leaf of \( \bar{t} \). Let \( \bar{\kappa}(\bar{t}, v) \in \bar{S}_n \) be the tree \( (\bar{t} \setminus \{w \in \bar{t} : v_1 \ldots v_{m-1} < w\}) \cup \{w_1 \ldots w_m : w \in \bar{t}, v_1 \ldots v_{m-1} w_m \leq w_1 \ldots w_p\} \).

**Remark 3.5.** If \( \bar{t} = \Phi \circ R(z_1, \ldots, z_{n+1}) \) for distinct \( z_1, \ldots, z_{n+1} \) and \( y_{n+1} \) is the leaf of \( \bar{t} \) corresponding to the input \( z_{n+1} \), then \( \bar{\kappa}(\bar{t}, y_{n+1}) = \Phi \circ R(z_1, \ldots, z_n) \).

**Remark 3.6.** Note that \( \bar{\kappa}(\bar{t}, v) \) is the tree obtained from \( \bar{t} \) by deleting \( v \) and its sibling and closing up the resulting gap if there is one (there will be a gap if and only if the sibling of \( v \) is not a leaf). This operation is the same as one that appears in the backward transition of the Rémy chain, and indeed, as part (ii) of the next proposition shows, the common backward transition probabilities of the PATRICIA chains are the same as that of the Rémy chain described at the beginning of [EGW17, Section 4].

**Proposition 3.7.**

(i) The process \( (\nu R_n)_{n \in \mathbb{N}} \) is Markov. For \( s \in S_n \) such that \( \mathbb{P}\{\nu R_n = s\} > 0 \) and \( t \in S_{n+1} \) such that \( \mathbb{P}\{\nu R_{n+1} = t\} > 0 \) the associated backward transition probability is

\[
\mathbb{P}\left\{\nu R_n = s \mid \nu R_{n+1} = t\right\} = \begin{cases} \frac{1}{n+1}, & \text{if } s = \kappa(t, v) \text{ for some } v \in L(t), \\ 0, & \text{otherwise.} \end{cases}
\]

(ii) The process \( (\nu \bar{R}_n)_{n \in \mathbb{N}} \) is Markov. For \( \bar{s} \in \bar{S}_n \) such that \( \mathbb{P}\{\nu \bar{R}_n = \bar{s}\} > 0 \) and \( \bar{t} \in S_{n+1} \) such that \( \mathbb{P}\{\nu \bar{R}_{n+1} = \bar{t}\} > 0 \) the associated backward transition probability is

\[
\mathbb{P}\left\{\nu \bar{R}_n = \bar{s} \mid \nu \bar{R}_{n+1} = \bar{t}\right\} = \begin{cases} \frac{1}{n+1}, & \text{if } \bar{s} = \bar{\kappa}(\bar{t}, v) \text{ for some } v \in L(\bar{t}), \\ 0, & \text{otherwise.} \end{cases}
\]

**Proof.** Suppose that \( s \in S_n \) and \( t \in S_{n+1} \) is such that \( \mathbb{P}\{\nu R_{n+1} = t\} > 0 \). It is clear from Remark 2.1 and the assumption that \( (Z_n)_{n \in \mathbb{N}} \) are independent that \( \mathbb{P}\{\nu R_n = s \mid \nu R_{n+1} = t, \nu R_{n+2} \ldots\} \) is \( \frac{1}{n+1} \) if \( s = \kappa(t, v) \) for some \( v \in L(t) \) and 0 otherwise. In particular, the time-reversal of \( (\nu R_n)_{n \in \mathbb{N}} \) is Markovian and hence the same is true of \( (\nu R_n)_{n \in \mathbb{N}} \) itself. This establishes (i).

The proof of (ii) is similar. \( \square \)
Notation 3.8. For $y \in \{0,1\}^*$, let $\tau(y) := \{z \in \{0,1\}^\infty : y < z\}$; that is, $\tau(y)$ is the set of infinite extensions of the finite word $y$.

Remark 3.9. An alternative route to establishing the Markov property of $(\nu R_n)_{n \in \mathbb{N}}$ is to note that the random set $\{Z_1, \ldots, Z_n\}$ is conditionally independent of $\nu R_1, \ldots, \nu R_{n-1}$ given $\nu R_n$. The conditional distribution of the random set $\{Z_1, \ldots, Z_n\}$ given $(\nu R_n = s)$, where $L(s) = \{y_1, \ldots, y_n\}$, is that of the random set $\{\tilde{Z}_1, \ldots, \tilde{Z}_n\}$, where $\tilde{Z}_1, \ldots, \tilde{Z}_n$ are independent and $\tilde{Z}_k$ is distributed according to $\nu(\cdot \cap \tau(y_k))/\nu(\tau(y_k))$ (that is, according to $\nu$ conditioned on $\tau(y_k)$) for $1 \leq k \leq n$. Thus,

$$\mathbb{P}(\nu R_{n+1} = t \mid \nu R_1, \ldots, \nu R_{n-1}, \nu R_n = s) = \mathbb{P}(R(\tilde{Z}_1, \ldots, \tilde{Z}_n, Z_{n+1}) = t),$$

where $\tilde{Z}_1, \ldots, \tilde{Z}_n$ are constructed to be independent of $Z_{n+1}$. This observation leads readily to an explicit calculation of the (forward) transition probabilities of $(\nu R_n)_{n \in \mathbb{N}}$ which are independent and $\mathbb{P}(\nu R_{n+1} = t) > 0$ and $\tilde{t} = \Phi(t)$, and $s \in S_n$, then $\mathbb{P}(\Phi(\nu R_n) = s \mid \nu R_{n+1} = t)$ is equal for $s$ of the form $\bar{s}(t, v)$ when $v \in L(t)$ and 0 otherwise. The fact that this conditional probability is the same for all $t \in S_{n+1}$ such that $\mathbb{P}(\nu R_{n+1} = t) > 0$ and $\tilde{t} = \Phi(t)$ gives that the time-reversal of $(\nu R_n)_{n \in \mathbb{N}}$ is Markov and hence the same is true of the process $(\bar{\nu} R_n)_{n \in \mathbb{N}}$ itself. Moreover, this same observation leads readily to the claimed backward transition probabilities.

Remark 3.10. Once part (i) of Proposition 3.7 is known, an alternative derivation of part (ii) is to use the Markov property of the time-reversal of $(\nu R_n)_{n \in \mathbb{N}}$ and apply Dynkin’s classical criterion for the composition of a function with a Markov process to be Markovian (see, for instance, [RP81, p. 575]). Specifically, one checks that if $\bar{t} \in \bar{S}_{n+1}$ is such that $\mathbb{P}(\nu R_{n+1} = \bar{t}) > 0$, $\tilde{t} \in S_{n+1}$ is such that $\mathbb{P}(\nu R_{n+1} = \tilde{t}) > 0$ and $\bar{t} = \Phi(\tilde{t})$, and $s \in S_n$, then $\mathbb{P}(\tilde{\Phi}(\nu R_n) = s \mid \nu R_{n+1} = \bar{t})$ is equal for $s$ of the form $\bar{s}(\bar{t}, v)$ when $v \in L(\tilde{t})$ and 0 otherwise. The fact that this conditional probability is the same for all $t \in S_{n+1}$ such that $\mathbb{P}(\nu R_{n+1} = t) > 0$ and $\tilde{t} = \Phi(t)$ gives that the time-reversal of $(\nu R_n)_{n \in \mathbb{N}}$ is Markov and hence the same is true of the process $(\bar{\nu} R_n)_{n \in \mathbb{N}}$ itself. Interestingly, Dynkin’s criterion does not hold for the forwards in time processes $(\nu R_n)_{n \in \mathbb{N}}$ and the function $\Phi$, as the following example with $\nu = \gamma$ shows. Define trees $s', s'', \bar{t}$ by $L(s') = \{00, 01, 1\}$ $L(s'') = \{000, 001, 1\}$, and $L(\bar{t}) = \{000, 001, 01\}$. Put $\bar{s} := \Phi(s') = \Phi(s'') = s'$. If $R_3 = s'$, then the only way that $\Phi(R_4)$ can be $\bar{t}$ is if $Z_4$ is of the form 000.; thus, $\mathbb{P}(\Phi(R_4) = \bar{t} \mid R_3 = s') = \frac{1}{8}$. Similarly, if $R_3 = s''$, then the only way that $\Phi(R_4)$ can be $\bar{t}$ is if $Z_4$ is of the form 000.; or 011.; thus $\mathbb{P}(\Phi(R_4) = \bar{t} \mid R_3 = s'') = \frac{1}{8} + \frac{1}{8}$. Since these two conditional probabilities are different, Dynkin’s criterion does not hold.

Before we proceed, we should record the observation that two different diffuse probability measures $\nu'$ and $\nu''$ on $\{0,1\}^\infty$ may result in one and the same distribution of the PATRICIA chains $(\nu' R_n)$ and $(\nu'' R_n)$. As a trivial example, consider $\nu' = \gamma$ and $\nu'' = \delta_0 \otimes \gamma$ – we may couple the two radix sort chains together so that $\nu'' R_n = 0^{\nu'} R_n$ for all $n \in \mathbb{N}$, in which case $\Phi(\nu' R_n) = \Phi(\nu'' R_n)$ for all $n \in \mathbb{N}$.

Declare that two probability measures $\nu'$ and $\nu''$ are PATRICIA-equivalent when $(\nu' R_n)_{n \in \mathbb{N}}$ and $(\nu'' R_n)_{n \in \mathbb{N}}$ have the same distribution. For later use (see Example 4.5 below) we will construct for each probability measure $\nu$ a PATRICIA-equivalent probability measure $\bar{\nu}$ with the property that $\nu'$ and $\nu''$ are PATRICIA-equivalent if and only if $\bar{\nu}' = \bar{\nu}''$. 
Fix $\nu$. We can write $\nu$ as $\delta_{u_1} \otimes \cdots \otimes \delta_{u_k} \otimes \nu^{\emptyset}$ for some finite, possibly empty, maximal sequence $u_1 \ldots u_k$ and some probability measure $\nu^{\emptyset}$ on $\{0,1\}^{\infty}$. Set $\nu_1 := \nu^{\emptyset}$. Note by the assumed maximality of $u_1 \ldots u_k$ that $\nu_1(\tau(0)) > 0$ and $\nu_1(\tau(1)) > 0$.

Suppose for some $n \in \mathbb{N}$ that we have built diffuse probability measures $\nu_1, \ldots, \nu_n$ PATRICIA–equivalent to $\nu$ such that $\nu_n(\tau(y)) > 0$ for every $y \in \{0,1\}^n$ and for $1 \leq m \leq n$ that $\nu_m$ and $\nu_n$ agree on sets of the form $\tau(y)$, $y \in \{0,1\}^m$. Conditioning the probability measure $\nu_n$ on the set $\tau(y)$, $y \in \{0,1\}^n$, gives a probability measure of the form $\delta_{y_1} \otimes \cdots \otimes \delta_{y_n} \otimes \delta_{u_1} \otimes \cdots \otimes \delta_{u_k} \otimes \nu^{\emptyset}$ for some finite, possibly empty, maximal sequence $u_1 \ldots u_k$ (we re-use notation from above and our notation doesn’t record the dependence of this sequence on $y$) and some probability measure $\nu^{\emptyset}$ on $\{0,1\}^{\infty}$. Note that $\nu^{\emptyset}(\tau(y)) > 0$ and $\nu^{\emptyset}(\tau(y)) > 0$.

Put $\nu_{n+1} := \sum_{y \in \{0,1\}^n} \nu_n(\tau(y)) \delta_{y_1} \otimes \cdots \otimes \delta_{y_n} \otimes \nu^{\emptyset}$. It is clear that the probability measures $\nu_{n+1}$ and $\nu$ are PATRICIA–equivalent and that for $1 \leq m \leq n + 1$ the probability measures $\nu_m$ and $\nu_{n+1}$ agree on sets of the form $\tau(y)$, $y \in \{0,1\}^m$.

There is thus a unique diffuse probability measure $\bar{\nu}$ such that the probability measures $\nu_m$ and $\bar{\nu}$ agree on sets of the form $\tau(y)$, $y \in \{0,1\}^m$, for all $m \in \mathbb{N}$.

It is not difficult to see that $\nu$ and $\bar{\nu}$ belong to the same PATRICIA–equivalence class: we can couple together two i.i.d. sequences of inputs distributed according to $\nu$ and $\bar{\nu}$ so that for each $n \in \mathbb{N}$ the tree $\bar{\nu} R_n$ is obtained from the tree $\nu R_n$ by a deterministic operation that removes certain vertices with out-degree 1, and hence $\Phi(\nu R_n) = \Phi(\bar{\nu} R_n)$ for all $n \in \mathbb{N}$.

We leave to the reader the proof of the fact that $\nu$ and $\nu^{\emptyset}$ are PATRICIA–equivalent, then $\bar{\nu} = \nu^{\emptyset}$.

4. Infinite bridges

An infinite bridge for $(\nu R_n)_{n \in \mathbb{N}}$ (resp. $(\bar{\nu} R_n)_{n \in \mathbb{N}}$) is a Markov chain $(\nu R_n^\infty)_{n \in \mathbb{N}}$ (resp. $(\bar{\nu} R_n^\infty)_{n \in \mathbb{N}}$) such that $\nu R_1^\infty = \emptyset$ (resp. $\bar{\nu} R_1^\infty = \emptyset$) and

$$\mathbb{P}(\nu R_n^\infty = s \mid \nu R_{n+1}^\infty = t) = \mathbb{P}(\nu R_n = s \mid \nu R_{n+1} = t)$$

for $s \in S_n$, $t \in S_{n+1}$, $n \in \mathbb{N}$ (resp.

$$\mathbb{P}(\bar{\nu} R_n^\infty = s \mid \bar{\nu} R_{n+1}^\infty = \bar{t}) = \mathbb{P}(\bar{\nu} R_n = s \mid \bar{\nu} R_{n+1} = \bar{t})$$

for $s \in S_n$, $\bar{t} \in S_{n+1}$, $n \in \mathbb{N}$). The reason we use this terminology is the following. A finite bridge for $(\nu R_n)_{n \in \mathbb{N}}$ with end-point $t \in S_m$ for some $m \in \mathbb{N}$ is the Markov process $(\nu R_n^k)_{n \in [m]}$ obtained by conditioning $(\nu R_n)_{n \in [m]}$ on the event $\{\nu R_m = t\}$. A finite bridge has the same backward transition probabilities as $(\nu R_n)_{n \in \mathbb{N}}$, and, writing $M(t)$ for the number of leaves of the binary tree $t$, one way to produce an infinite bridge is to find a sequence of trees $(t_k)_{k \in \mathbb{N}}$ such that $M(t_k) \to \infty$ and the finite-dimensional distributions of $(\nu R_k)_{k \in [M(t_k)]}$ converge as $k \to \infty$: the limiting finite-dimensional distributions will be consistent and hence define a stochastic process $(\nu R_n^\infty)_{n \in \mathbb{N}}$ that is an infinite bridge for $(\nu R_n)_{n \in \mathbb{N}}$. Similar remarks hold for the finite bridges of $(\nu R_n)_{n \in \mathbb{N}}$. We call the infinite bridges for the PATRICIA chains infinite PATRICIA bridges.

It follows from Proposition 3.7 that for an arbitrary diffuse probability measure $\nu$ any infinite bridge for $(\nu R_n)_{n \in \mathbb{N}}$ is an infinite bridge for $(R_n)_{n \in \mathbb{N}}$. Conversely, an infinite bridge $(R_n^\infty)_{n \in \mathbb{N}}$ for $(R_n)_{n \in \mathbb{N}}$ is a bridge for $(\nu R_n)_{n \in \mathbb{N}}$ provided $\{t \in S_n : \mathbb{P}(R_n^\infty = t) > 0\} \subseteq \{t \in S_n : \mathbb{P}(\nu R_n = t) > 0\}$ for all $n \in \mathbb{N}$. We may therefore restrict our attention to infinite bridges for $(R_n)_{n \in \mathbb{N}}$. Similar observations give
that in investigating the infinite bridges of the processes \((\nu R_n)_{n \in \mathbb{N}}\) we may restrict attention to those of \((\tilde{R}_n)_{n \in \mathbb{N}}\).

It follows from the general theory of Doob–Martin compactifications that the family of distributions of infinite bridges is a compact convex set, every element of which is unique mixture of extremal distributions. Moreover, every extremal infinite bridge is a limit of finite bridges (but the converse is not true in general). Furthermore, we see from Proposition 3.7 that each Markov process of the form \((\nu R_n)_{n \in \mathbb{N}}\) (resp. \((\nu \tilde{R}_n)_{n \in \mathbb{N}}\)) is an infinite bridge for \((R_n)_{n \in \mathbb{N}}\) (resp. \((\tilde{R}_n)_{n \in \mathbb{N}}\)) and, as we shall see below, these infinite bridges are extremal. The interesting questions for the infinite bridges of \((R_n)_{n \in \mathbb{N}}\) and \((\tilde{R}_n)_{n \in \mathbb{N}}\) are therefore:

- What are all the extremal infinite bridges?
- Is every limit of finite bridges extremal?
- Is every extremal infinite bridge of the form \((\nu R_n)_{n \in \mathbb{N}}\) (resp. \((\nu \tilde{R}_n)_{n \in \mathbb{N}}\)) for some \(\nu\)?

In the case of \((R_n)_{n \in \mathbb{N}}\), these questions were answered in [EW17] as follows.

**Theorem 4.1.**  
(i) An infinite bridge for \((R_n)_{n \in \mathbb{N}}\) is extremal if and only if it has an almost surely trivial tail \(\sigma\)-field.

(ii) An infinite bridge for \((R_n)_{n \in \mathbb{N}}\) is extremal if and only if it is a weak limit of finite bridges.

(iii) The extremal infinite bridges for \((R_n)_{n \in \mathbb{N}}\) coincide with collection of Markov processes of the form \((\nu R_n)_{n \in \mathbb{N}}\) for some diffuse probability measure \(\nu\).

Because of the close affinity between \((R_n)_{n \in \mathbb{N}}\) and \((\tilde{R}_n)_{n \in \mathbb{N}}\), it is natural to conjecture that the obvious analogue of Theorem 4.1 holds for \((\tilde{R}_n)_{n \in \mathbb{N}}\). The next theorem shows that analogues of parts (i) and (ii) do hold.

**Theorem 4.2.**  
(i) An infinite bridge for \((\tilde{R}_n)_{n \in \mathbb{N}}\) is extremal if and only if it has an almost surely trivial tail \(\sigma\)-field.

(ii) An infinite bridge for \((\tilde{R}_n)_{n \in \mathbb{N}}\) is extremal if and only if it is a weak limit of finite bridges.

**Proof.** See Proposition 5.19 and the proof of Corollary 5.21 in [EGW17].  \(\square\)

**Remark 4.3.** Remark 5.20 in [EGW17] says that Proposition 5.19 in that paper can also be proved along the lines of Lemma 5.3 therein. As pointed out to us by Julian Gerstenberg and Rudolf Grübel, the argument in the proof of Lemma 5.3 and consequently Remark 5.20 is incorrect. However, this does not invalidate the proof of Proposition 5.19 given directly after the statement, as this proof is along completely different lines. Also, this does not invalidate any of the remainder of [EGW17] since Lemma 5.3 and Remark 5.20 of that paper are not used further therein.

**Remark 4.4.** Each chain \((\nu \tilde{R}_n)_{n \in \mathbb{N}}\) is an infinite bridge for \((\tilde{R}_n)_{n \in \mathbb{N}}\). By Remark 2.1 and the Hewitt-Savage zero-one law, each such chain has an almost surely trivial tail \(\sigma\)-field and hence by part (i) of Theorem 4.2 is extremal. However, as Example 4.5 shows, these are not the only extremal infinite PATRICIA bridges. Hence the analogue of Theorem 4.1 (iii) does not hold for \((\tilde{R}_n)_{n \in \mathbb{N}}\). A concrete description of the extremal infinite PATRICIA bridges comprises the remainder of the paper.

**Example 4.5.** Let \(\epsilon_2, \epsilon_3, \ldots\) be a sequence of independent identically distributed \(\{0, 1\}\)-valued random variables with common distribution \(\mathbb{P}\{\epsilon_k = 0\} = \mathbb{P}\{\epsilon_k = 1\}\)
4.6, 2 \leq k < \infty$. For $N \in \mathbb{N}$, define $(\bar{R}_n^N)_{n \in \mathbb{N}}$ with $\bar{R}_n^N \in \mathbb{S}_n$ for $n \in \mathbb{N}$ by requiring that $(\bar{R}_n^N)_{n \in \mathbb{N}}$ is Markov, $\bar{R}_n^N$ has the same distribution as $\{\emptyset\} \cup \bigcup_{k=2}^N \{\epsilon_2, \ldots, \epsilon_k\}$, and the backward transition probabilities of $(\bar{R}_n^N)_{n \in \mathbb{N}}$ are the same as those of $(\bar{R}_n)_{n \in \mathbb{N}}$. We may suppose that $\bar{R}_n^N = \{\emptyset\} \cup \bigcup_{k=2}^N \{\epsilon_2, \ldots, \epsilon_k\}$. It follows from the form of the backward transition probabilities that for $M \in \mathbb{N}$, $\bar{R}_M^N = \{\emptyset\} \cup \bigcup_{k=2}^M \{\epsilon_2, \ldots, \epsilon_k\}$, where $2 \leq I_2 < I_3 < \ldots < I_M \leq N$ is a certain uniform random subset of $\{2, \ldots, N\}$ of cardinality $M$ that is independent of $\epsilon_2, \ldots, \epsilon_N$. Thus $\bar{R}_M^N$ has the same distribution as $\bar{R}_M$ and hence, by Kolmogorov’s extension theorem, there exists a Markov process $(\bar{R}_n^\infty)_{n \in \mathbb{N}}$ such that $(\bar{R}_n^\infty)_{n \in \mathbb{N}}$ has the same distribution as $(\bar{R}_n^N)_{n \in \mathbb{N}}$ for any $N \in \mathbb{N}$. Therefore $(\bar{R}_n^\infty)_{n \in \mathbb{N}}$ is an infinite bridge for $(\bar{R}_n)_{n \in \mathbb{N}}$.

We can give a pathwise construction of $(\bar{R}_n^\infty)_{n \in \mathbb{N}}$ as follows. Let $((Y_n, \eta_n))_{n = 0}^\infty$ be an infinite sequence of independent identically distributed $[0, 1] \times \{0, 1\}$-valued random variables such that $Y_n$ has the uniform distribution on $[0, 1]$, $P\{\eta_n = 0\} = P\{\eta_n = 1\} = \frac{1}{2}$, and $Y_n$ and $\eta_n$ are independent. For $2 \leq n < \infty$, let $\sigma_n$ be the permutation of $\{2, \ldots, n\}$ such that $Y_{\sigma_n(2)} < Y_{\sigma_n(3)} < \ldots < Y_{\sigma_n(n)}$. For $2 \leq k \leq n$, put $\epsilon_n,k = \eta_{\sigma_n(k)}$. Setting $\bar{R}_1^\infty = \emptyset$ and $\bar{R}_n^\infty := \{\emptyset\} \cup \bigcup_{k=2}^n \{\epsilon_n,k, \epsilon_n,2, \ldots, \epsilon_n, (k-1)\} \epsilon_n,k\}$ for $n \geq 2$ produces a process with the desired distribution. Note that if $\pi$ is a permutation of $\{2, 3, \ldots\}$ that leaves every element of $\{N + 1, N + 2, \ldots\}$ fixed for some $N \geq 2$ and we replace $((Y_n, \eta_n))_{n = 0}^\infty$ by $((Y_{\pi(n)}, \eta_{\pi(n)}))_{n = 0}^\infty$ in this construction, then the values of $\bar{R}_n^\infty$, $n \geq N$, are left unchanged. It therefore follows from the Hewitt-Savage zero-one law that the tail $\sigma$-field of $(\bar{R}_n^\infty)_{n \in \mathbb{N}}$ is $P$-a.s. trivial and thus, by part (i) of Theorem 4.2, this process is an extremal infinite bridge for $(\bar{R}_n)_{n \in \mathbb{N}}$.

Now, for any $n \in \mathbb{N}$ the random tree $\bar{R}_n^\infty$ is a zig-zag path of length $n - 1$ with leaves attached to it, and hence it is clear from Lemma 4.6 that the extremal infinite bridge $(\bar{R}_n^\infty)_{n \in \mathbb{N}}$ is not of the form $(\bar{R}_n^\infty)_{n \in \mathbb{N}}$ for any diffuse probability measure $\nu$: if $(\bar{t}_n)_{n \in \mathbb{N}}$ is a sequence such that, for all $n \in \mathbb{N}$, $\bar{t}_n \in \mathbb{S}_n$ is a path of length $n - 1$ with leaves attached then $\bigcup_{m=1}^\infty \bigcap_{m=m}^\infty \bar{t}_m$ is contained in tree consisting of a possibly infinite path with leaves attached to it. Indeed, it is not hard to see that $P\{\epsilon_n,2 = 0 \text{ i.o.}\} = P\{\epsilon_n,2 = 1 \text{ i.o.}\} = 1$, which implies that $\bigcup_{m=1}^\infty \bigcap_{m=m}^\infty \bar{R}_n^\infty = \{0, 1\}$, $P$-a.s. The fact that $(\bar{R}_n^\infty)$ is not of the form $(\bar{R}_n^\infty)_{n \in \mathbb{N}}$ for some $\nu$ is a consequence of the following lemma:

**Lemma 4.6.** For any diffuse probability measure $\nu$,  
$$
\bigcup_{m=1}^\infty \bigcap_{m=m}^\infty \nu \bar{R}_n = \{0, 1\}^*, \quad P - \text{a.s.}
$$

**Proof.** Note that, by construction, $\nu$ has full support, where we equip $\{0, 1\}^\infty$ with the product topology (that is, the topology generated by the sets of the form $\tau(y)$, $y \in \{0, 1\}^\infty$). Thus, for any $k \in \mathbb{N}$ there exists an $N \in \mathbb{N}$ such that for all $n \geq N$ and for all $y \in \{0, 1\}^k$ there is an $m \in \mathbb{N}$ with $Z_m \in \tau(y)$. Therefore, $\nu \bar{R}_n \supseteq \{0\} \cup \bigcup_{k=1}^N \{0, 1\}^k$ for $n \geq N$, which implies that $\nu \bar{R}_n \supseteq \{0\} \cup \bigcup_{k=1}^N \{0, 1\}^k$ for $n \geq N$. Since $(\nu \bar{R}_n)_{n \in \mathbb{N}}$ and $(\nu \bar{R}_n)_{n \in \mathbb{N}}$ have the same distribution, this establishes the result. 

Our next goal is to introduce a class of extremal infinite bridges for $(\bar{R}_n)_{n \in \mathbb{N}}$ that subsumes both the class of Markov chains of the form $(\nu \bar{R}_n)_{n \in \mathbb{N}}$ and the
Markov chain in Example 4.5. In fact, using Proposition 3.7 and Remark 3.6 (showing that the common backward transitions of the PATRICIA chains coincide with those of the Rémy chain) we will be able to obtain a representation for all the extremal infinite PATRICIA bridges. This will be achieved through the concept of a didendritic system introduced in [EGW17, Definition 5.8] and revisited in the next section.

5. Didendritic systems

**Notation 5.1.** Recall the partial order $<$ defined on $\{0,1\}^* \sqcup \{0,1\}^\infty$. Given $u, v \in \{0,1\}^* \sqcup \{0,1\}^\infty$, there is a unique $w \in \{0,1\}^* \sqcup \{0,1\}^\infty$ with the properties $w \leq u$, $w \leq v$, and if $x$ is any element of $\{0,1\}^* \sqcup \{0,1\}^\infty$ with $x \leq u$ and $x \leq v$, then $x \leq w$. Denote this $w$ by $u \wedge v$. For example, if $u = u_1 \ldots u_m$ and $v = v_1 \ldots v_n$ are elements of $\{0,1\}^*$, then $u \wedge v = u_1 \ldots u_p = v_1 \ldots v_q$, where $p := \max\{q : 1 \leq q \leq m \wedge n, u_1 \ldots u_q = v_1 \ldots v_q\}$ and the maximum of the empty set is defined to be zero.

**Definition 5.2.** Define two further partial orders $<_L$ and $<_R$ on $\{0,1\}^*$ by declaring that $u_1 \ldots u_m <_L v_1 \ldots v_n$ (resp. $u_1 \ldots u_m <_R v_1 \ldots v_n$) if $m < n$ and $v_1 \ldots v_{m+1} = u_1 \ldots u_m 0$ (resp. $m < n$ and $v_1 \ldots v_{m+1} = u_1 \ldots u_m 1$).

**Definition 5.3.** Given $\bar{\ell} \in \mathfrak{S}_n$, consider a bijective labeling of the leaves of $\bar{\ell}$ by a set $\mathcal{N}$ with $\#\mathcal{N} = n$ (that is, each leaf receives a distinct label). Use the notation $\bar{\ell}$ to denote the labeled object. Define an equivalence relation $\equiv$ on $\mathcal{N} \times \mathcal{N}$ by declaring that $(i,j)$ and $(i',j')$ are equivalent if $g, h, i, j$ label respectively leaves $u, v, w, x$ such that $u \wedge v = w \wedge x$. Denote by $\langle i, j \rangle$ the equivalence class containing $(i,j) \in \mathcal{N} \times \mathcal{N}$. Note that $(i,i)$ is the only pair in the equivalence class $(i,i)$ so we will usually denote this equivalence class more simply by $i$. If $k, \ell$ label respectively the leaves $y, z$, then label the vertex $y \wedge z$ with $\langle k, \ell \rangle$.

With a slight abuse of notation, define a partial order $<_L$ (resp. $<_R$) on $\{\langle i,j \rangle : i, j \in \mathcal{N}\}$ by declaring that $(g, h) <_L \langle i,j \rangle$ if $(g, h)$ labels a vertex $u$ and $(i,j)$ labels a vertex $v$ such that $u <_L v$ (resp. $u <_R v$). Similarly, define a third partial order $<$ by declaring that $(g, h) < \langle i,j \rangle$ if $(g, h)$ labels a vertex $u$ and $(i,j)$ labels a vertex $v$ with $u < v$.

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

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

(B) For distinct $i, j \in \mathcal{N}$, 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 $i, j, k, \ell \in \mathcal{N}$, 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 $f, g, h, i, j, k \in \mathcal{N}$. 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 $\langle f,g \rangle <_R \langle j,k \rangle$. 
Definition 5.4. A didendritic system $\mathcal{D} = (\mathcal{N}, \equiv, \langle \cdot, \cdot \rangle, <_L, <_R, <)$ with the non-empty (possibly infinite) label set $\mathcal{N}$ is the set $\mathcal{N} \times \mathcal{N}$ equipped with an equivalence relation $\equiv$, equivalence classes $\langle \cdot, \cdot \rangle$, and partial orders $<_L, <_R$ and $<$ on the set of equivalence classes such that the above stated properties (A)-(E) hold.

Remark 5.5. We show in Proposition 5.8 that any didendritic system with a finite label set may be thought of as a leaf-labeled full binary tree. This claim was made in [EGW17, Remark 5.10]. The axioms in Definition 5.4 differ from those in [EGW17]: the former are equivalent to the latter plus the triplet property (C).

Example 5.6. Consider the equivalence relation on $[3] \times [3]$ defined by

$$(h, i) \equiv (j, k) \text{ if an only if } h = j \text{ and } i = k, \text{ or } h = k \text{ and } i = j.$$ 

For $(i, j) \in \{(1, 2), (1, 3), (2, 3)\}$ say that $(i, j) <_L \langle i, i \rangle$ and $(i, j) <_R \langle j, j \rangle$. Moreover, say that $(h, i) < \langle j, k \rangle$ if $(h, i) <_L \langle j, k \rangle$ or $(h, i) <_R \langle j, k \rangle$. Then $([3], \equiv, <_L, <_R, <)$ meets the axioms of [EGW17, Definition 5.8]. However, this system does not correspond to a binary tree with 3 leaves, since this would require exactly 5 equivalence classes (each one corresponding to a vertex of the tree), whereas this system has $3 + 3 = 6$ equivalence classes.

Lemma 5.7. Any didendritic system with $\#\mathcal{N} \geq 3$ has the following property

\[(F) \quad \text{For distinct } i, j, k \in \mathcal{N}, \text{ there exists a bijective mapping from } \{(g, h) : g, h \in \{i, j, k\}\} \text{ to the set of vertices of a full binary tree } t \text{ with three leaves labeled by } i, j, k \text{ which preserves the three partial orders and, for } g \in \{i, j, k\}, \text{ maps } \langle g, g \rangle \text{ to the leaf of } t \text{ labeled } g.\]

Proof. Assume that the first of the three possible set of relations in (C) is valid, i.e. $\langle i, j \rangle = \langle i, k \rangle < \langle j, k \rangle$. Then we must have by (B) that $\langle i, j \rangle = \langle i, k \rangle < \langle j, k \rangle < j,$ and $(i, j) < i$.

Applying (B) to $i$ and $j$, we have either $(i, j) <_L i$ and $(i, j) <_R j$ or $(i, j) <_L j$ and $(i, j) <_R i$. Assume the former. Then (E) (together with (B)) enforces $(i, j) <_R \langle j, k \rangle$.

Now applying (B) to $j$ and $k$, we have either $(j, k) <_L j$ and $(j, k) <_R k$ or $(j, k) <_L k$ and $(j, k) <_R j$. Assume the former. This then results in the full binary tree $t$ with vertex set $\{\emptyset, 0, 1, 10, 11\}$, where $\emptyset$ is labeled by $\langle i, j \rangle = \langle i, k \rangle$, 1 is labeled by $\langle j, k \rangle$, and the three leaves 0, 10 and 11 are labeled by $i$, $j$ and $k$, respectively.

Combining all the ways allowed by (C) and (B), we arrive at the $3 \times 2 \times 2 = 12$ possible leaf-labeled full binary trees with three leaves and label set $\{i, j, k\}$. 

Let $\hat{t}$ be a leaf-labeled full binary tree, $\mathcal{D}$ be a didendritic system with finite label set $\mathcal{N}$, and $\psi$ be a mapping from the set of equivalence classes of $\mathcal{D}$ to the set of vertices of $\hat{t}$. We call $\psi$ an isomorphism from $\mathcal{D}$ to $\hat{t}$ if it is bijective and preserves the three partial orders. Necessarily, $\psi$ then maps the label set $\mathcal{N}$ bijectively to the set of leaves of $\hat{t}$. 

**Proposition 5.8.** Let $\mathcal{D} = (\mathcal{N}, \equiv, \langle \cdot , \cdot \rangle, <_L, <_R, <)$ be a finite didendritic system. Then $\mathcal{D}$ is isomorphic to some leaf-labeled full binary tree $\hat{t}$. Moreover, both $\hat{t}$ and the isomorphism are unique.

**Proof.** Denote the cardinality of $\mathcal{N}$ by $n$.

For $n = 2$ the assertion follows directly from properties (A)-(C).

For $n \geq 3$ we will use that $\mathcal{D}$ also has property (F) because of Lemma 5.7.

Indeed, for $n = 3$ the assertion is immediate from that property.

We now proceed by induction and assume that the assertion is true for some $n \geq 3$.

For the induction step, assume without loss of generality that $\mathcal{N}_{n+1} = [n+1]$, and let $\mathcal{D}_{n+1}$ obey the assumptions of the proposition with label set $[n+1]$. We are going to construct a unique leaf-labeled full binary tree $\hat{t}_{n+1}$ and isomorphism $\psi$ from $\mathcal{D}_{n+1}$ to $\hat{t}_{n+1}$.

To this end we denote the restriction of $\mathcal{D}_{n+1}$ to $[n]$ by $\mathcal{D}_n$. Since $\mathcal{D}_n$ meets the conditions of the induction hypothesis, there exists a unique leaf-labeled full binary tree $\hat{t}_n$ and isomorphism $\phi$ from $\mathcal{D}_n$ to $\hat{t}_n$. By a slight abuse of notation we will use the symbol $(i,j)$ both for the equivalence classes in $\mathcal{D}_n$ and for the equivalence classes in $\mathcal{D}_{n+1}$.

Now choose $a, b \in [n]$ such that $\phi((a, b)) = \emptyset \in \hat{t}_n$. Focusing on $a, b, n+1$, the triplet property (C) allows to distinguish between two cases. The first case is when $(n+1, a) = (n+1, b) < (a, b)$ and the second case is when $(a, b) = (n+1, b) < (n+1, a)$ or $(a, b) = (n+1, a) < (n+1, b)$

**Case 1:** $(n+1, a) = (n+1, b) < (a, b)$.

Arguing by the isomorphy of $\mathcal{D}_n$ and $\hat{t}_n$, and recalling that $\phi((a, b)) = \emptyset$, we see that for all $h \in [n]$ we have that either $(a, b) = (a, h) < (a, h)$ or $(a, b) = (a, h) < (b, h)$. In the former case, we obtain from $(n+1, a) = (n+1, b) < (a, b)$, $(a, b) = (a, h) < (a, h)$, and the transitivity of $<$ that $(n+1, a) < (a, h)$ and then, from the triplet property (C) applied to the triplet $a, h, n+1$, that $(n+1, h) = (n+1, a) < (a, h)$. In the latter case, we obtain by a similar argument that $(n+1, h) = (n+1, b) < (b, h)$. Thus, in both cases we have for all $h \in [n]$ that $(n+1, h) = (n+1, a) = (n+1, b)$. Since for all $i, j \in [n]$ with $(i, j) \neq (a, b)$ we have $(a, b) < (i, j)$ by the induction assumption, the transitivity of $<$ implies that $(n+1, h) = (n+1, a) = (n+1, b) < (i, j)$ for all $i, j \in [n]$. Thus, we are led to define

$$\psi((n+1, h)) := \emptyset \text{ for } h \in [n].$$

To define the other values of $\psi$, we consider two subcases:

**Case 1a:** $(n+1, a) = (n+1, b) <_L (a, b)$.

Recall from above that $(a, b) \leq (i, j)$ for all $i, j \in [n]$. Property (E) then ensures that $(n+1, a) = (n+1, b) <_L (i, j)$ for all $i, j \in [n]$. Moreover, the triplet property (C) applied to the triplet $n+1, a, b$ combined with property (D) guarantees that $(n+1, a) = (n+1, b) <_R (n+1, n+1)$. Consequently, we are led to set

$$\psi((n+1, n+1)) := 0,$$

$$\psi((i, j)) := 1\phi((i, j)) \text{ for } i, j \in [n].$$
Case 1b: \( \langle n + 1, a \rangle = \langle n + 1, b \rangle \prec_R \langle a, b \rangle \).

By an argument similar to that in Case 1b, we are led to set

\[
\psi((n + 1, n + 1)) := 1,
\psi((i, j)) := 0 \phi((i, j)) \text{ for } i, j \in [n].
\]

Case 2: \( \langle a, b \rangle = \langle n + 1, b \rangle < \langle n + 1, a \rangle \) or \( \langle a, b \rangle = \langle n + 1, a \rangle < \langle n + 1, b \rangle \).

From property (F) for \( a, b, n + 1 \), we see that either of the two above subcases implies that \( \langle a, b \rangle < n + 1 \). Also, we may assume without loss of generality from property (B) (by exchanging the roles of \( a \) and \( b \) if necessary) that \( \langle a, b \rangle < L a \) and \( \langle a, b \rangle < R b \).

Put \( L := \{ g \in [n + 1] : \langle a, b \rangle < L g \} \), \( R := \{ h \in [n + 1] : \langle a, b \rangle < R h \} \), and note from property (D) that \( n + 1 \in L \cup R \). Moreover, \( a \in L \) and \( b \in R \), hence \( \#L \leq n \) and \( \#R \leq n \). We may thus apply the induction assumption to \( D_L \) and \( D_R \), defined to be the restrictions of \( D_{n+1} \) to \( L \) and \( R \), respectively. Let \( t_L \) and \( t_R \) be the corresponding leaf-labeled full binary trees, and let \( \chi_L \) and \( \chi_R \) be the corresponding isomorphisms. The desired leaf-labeled full binary tree \( t_{n+1} \) has the vertex set \( \{ \emptyset \} \cup \{ 0v : v \in t_L \} \cup \{ 1v : v \in t_R \} \), and the isomorphism \( \psi \) from \( D_{n+1} \) to \( t_{n+1} \) is given by

- \( \psi((g, h)) := \emptyset \) for \( g \in L \), \( h \in R \),
- \( \psi((g, h)) := 0v \) for \( g, h \in L \), with \( v := \chi_L((g, h)) \)
- \( \psi((g, h)) := 1v \) for \( g, h \in R \), with \( v := \chi_R((g, h)) \).

\( \square \)

The next result shows that a finite didendritic system can be constructed in two stages. The first stage determines the partial order \( \prec \) while the second stage resolves each instance of \( \prec \) as either \( \prec_L \) or \( \prec_R \) in a consistent manner. A randomized version of this construction appears in the statement of the main representation result, [EGW17, Theorem 8.2]. In the next proposition we elaborate on the “deterministic heart” of that construction.

**Proposition 5.9.** Let \( \mathcal{N} \) be a finite set, \( \equiv \) be an equivalence relation on \( \mathcal{N} \times \mathcal{N} \) and \( \prec \) be a partial order on the set of equivalence classes. Suppose also that for distinct \( i, j \in \mathcal{N} \) there are elements \( w(i, j) \) of the set \( \{ \ominus, \cap \} \). Assume properties (A) and (C) from Definition 5.4, as well as the following properties

- (B') For distinct \( i, j \in \mathcal{N} \), \( \langle i, j \rangle < i \) and \( \langle i, j \rangle < j \).
- (B'') For \( i \neq j \), \( w(i, j) = \cap \) if and only if \( w(j, i) = \cap \).
- (E') For distinct \( i, j, k \in \mathcal{N} \), if \( \langle i, j \rangle = \langle i, k \rangle < \langle j, k \rangle \), then \( w(i, j) = w(i, k) \).

Then there is a unique pair of partial orders \( \prec_L \) and \( \prec_R \) on \( \{ (i, j) : i, j \in \mathcal{N} \} \) such that

\( \langle i, j \rangle \prec_L i \) and \( \langle i, j \rangle < R j \iff w(i, j) = \cap \)

and the ensemble \( \langle \mathcal{N} \rangle, \equiv, \langle, \cap \rangle, \prec_L, \prec_R, \prec \rangle \) is a didendritic system.

**Proof.** We may assume w.l.o.g. that \( \mathcal{N} = [n] \) for some \( n \in \mathbb{N} \).

Step 1. For \( n = 1 \) there is nothing to prove.

Step 2. For \( n = 2 \) we have two cases according to property (B'):
If \( w(1, 2) = \cap \) then \( \langle 1, 2 \rangle \prec_L 1 \) and \( \langle 1, 2 \rangle < R 2 \). If \( w(1, 2) = \ominus \) then \( \langle 1, 2 \rangle < L 2 \) and \( \langle 1, 2 \rangle < R 1 \). In both cases, \( \langle \{1, 2\}, \equiv, \langle, \cap \rangle, \prec_L, \prec_R, \prec \rangle \) is a didendritic system.
Step 3. Now assume $n \geq 3$.

To show property (B) in the definition of a didendritic system we can argue as in the case $n = 2$.

Next we show for distinct $i, j, k \in [n]$ that the assumptions of the proposition define partial orders $<_L,<_R$ on $\{ (g, h) : g, h \in \{ i, j, k \} \}$ which meet condition (F) formulated in Lemma 5.7.

Indeed, because of condition (C) we have one of the following cases:

Case 1: $\langle i, j \rangle = \langle i, k \rangle < \langle j, k \rangle$,
Case 2: $\langle j, k \rangle = \langle j, i \rangle < \langle k, i \rangle$,
Case 3: $\langle k, i \rangle = \langle k, j \rangle < \langle i, j \rangle$.

Because of property (E'), in Case 1 we have $w(i, j) = w(i, k)$, in Case 2 we have $w(j, k) = w(j, i)$, and in Case 3 we have $w(k, i) = w(k, j)$. Thus, in the three cases, we have the four different choices $\langle \cup, \cap \rangle, \langle \cap, \cap \rangle, \langle \cap, \cap \rangle, \langle \cap, \cap \rangle$ for the pair $w(i, j), w(j, k)$ in Case 1, for the pair $w(j, i), w(k, i)$ in Case 2, and for the pair $w(k, i), w(i, j)$ in Case 3. Any of these $3 \times 4 = 12$ sub-cases 1a), . . . , 3d) leads to partial orders $<_L,<_R$ on $\{ (g, h) : g, h \in \{ i, j, k \} \}$ and to a (distinct) leaf-labeled full binary tree $\tilde{t}$ whose five vertices are bijectively labeled by the elements of $\{ (g, h) : g, h \in \{ i, j, k \} \}$ in an order preserving way. For example, in Case 1a) $\tilde{t}$ consists of the vertex set $\{ \emptyset, 0, 1, 10, 11 \}$, with 0 labeled $i$, 10 labeled $j$ and 11 labeled $k$. In Case 3c), $\tilde{t}$ consists of the vertex set $\{ \emptyset, 0, 00, 01, 1 \}$, with 00 labeled $i$, 01 labeled $j$ and 1 labeled $k$.

Step 4. We now proceed inductively to show also the remaining properties (D) and (E) in the definition of a didendritic system.

Let $\mathcal{N}_{n+1} = [n+1]$ and consider the restrictions of $\equiv$ and $<$ to $[n] \times [n]$. By the induction hypothesis there is a unique didendritic system $\mathcal{D}_n$ with label set $[n]$ that has the properties stated in the proposition. By Proposition 5.8 there is a unique leaf-labeled full binary tree $t_n$ whose $n$ vertices are bijectively labeled by the elements of $[n]$ such that the partial orders $<_L,<_R$ and $<$ are preserved. Again by Proposition 5.8, for the induction step it suffices to construct out of $t_n$ a leaf-labeled full binary tree $\tilde{t}_{n+1}$ whose $2n+1$ vertices are bijectively labeled by the equivalence classes $\{i,j\}$, $i,j \in [n+1]$, such that the partial orders $<_L,<_R$ and $<$ are preserved, and to show that the construction of $\tilde{t}_{n+1}$ is unique.

Let $a, b \in [n]$ be such that the root $\emptyset$ of $t_n$ is labeled by $\langle a, b \rangle$. Because of condition (C) applied to the triplet $n+1, a, b$, one of the following three cases applies:

Case 1: $\langle n+1, a \rangle = \langle n+1, b \rangle < \langle a, b \rangle$
Case 2a: $\langle a, b \rangle = \langle a, n+1 \rangle < \langle b, n+1 \rangle$
Case 2b: $\langle b, n+1 \rangle = \langle b, a \rangle < \langle n+1, a \rangle$

In Case 1, because of the transitivity of $<$ and the isomorphy between $\mathcal{D}_n$ and $\tilde{t}_n$, we have $\langle n+1, a \rangle = \langle n+1, b \rangle < \langle i, j \rangle$ for all $i,j \in [n]$. Therefore, within $\tilde{t}_{n+1}$, the root $\emptyset$ must be labeled by $\langle n+1, a \rangle = \langle n+1, b \rangle$, and either the vertex 1 or the vertex 0 must be a leaf labeled by $n+1$. The former is the case if $w(n+1, a) = \cup$, the latter if $w(n+1, a) = \cap$. In the former case we have $\langle n+1, a \rangle = \langle n+1, b \rangle <_L \langle a, b \rangle$, and the addresses in $\tilde{t}_{n+1}$ of the equivalence classes $\{i,j\}$, $i,j \in [n+1]$, are then given just as in Case 1a in the proof of Proposition 5.8. In the latter case we have
\[ \langle n + 1, a \rangle = \langle n + 1, b \rangle <_R \langle a, b \rangle, \] and the addresses in \( \hat{t}_{n+1} \) of the equivalence classes \( \langle i, j \rangle, i, j \in [n+1], \) are then given just as in Case 1b in the proof of Proposition 5.8.

Now we assume Case 2a and we will parallel the reasoning in Case 2 in the proof of Proposition 5.8. From Step 3 we have the property (F) in hand; this applied to the triplet \( \{i, j, k\} = \{a, b, n + 1\} \) ensures that the set of equivalence classes \( \{\langle g, h \rangle : g, h \in \{a, b, n + 1\}\} \) is isomorphic to a labeled binary tree \( \hat{t} \) with 3 leaves, for which either 0 or 1 is a leaf. Assuming without loss of generality that \( w(a, b) = \infty \) we see that the vertex 0 in \( \hat{t} \) must be labeled with \( a \) and the vertex 1 in \( \hat{t} \) must be labeled with \( b, n + 1 \). Thus, we arrive that \( \langle a, b \rangle = \langle a, n + 1 \rangle <_R \langle b, n + 1 \rangle \) and \( \langle a, b \rangle <_L \langle a, n + 1 \rangle <_R \langle b, n + 1 \rangle \).

Put \( L := \{ h \in [n] : \langle a, b \rangle <_L h \} \) and \( R := \{ h \in [n] : \langle a, b \rangle <_R h \} \cup \{ n + 1 \} \). Note that \( \#L \leq n \) and \( \#R \leq n \) because \( a \in L \) and \( b \in R \). We may thus apply the induction assumption to the restrictions of \( \equiv, < \) and \( \rightarrow \) to the two sets \( L \) and \( R \), respectively. Let \( L_L \) and \( R_R \) be the corresponding leaf-labeled full binary trees, and let \( \chi_L \) and \( \chi_R \) be the corresponding isomorphisms. The desired leaf-labeled full binary tree \( \hat{t}_{n+1} \) has the vertex set \( \{ \emptyset \} \cup \{ 0v : v \in \hat{t}_L \} \cup \{ 1v : v \in \hat{t}_R \} \), and the addresses of the equivalence classes \( \langle g, h \rangle, g, h \in [n + 1] \), in the tree \( \hat{t}_{n+1} \) are given by

- \( \psi(\langle g, h \rangle) := \emptyset \) for \( g \in L, h \in R \),
- \( \psi(\langle g, h \rangle) := 0v \) for \( g, h \in L \), with \( v := \chi_L(\langle g, h \rangle) \),
- \( \psi(\langle g, h \rangle) := 1v \) for \( g, h \in R \), with \( v := \chi_R(\langle g, h \rangle) \).

The partial orders \( <_L \) and \( <_R \) are then simply read off from the corresponding orders in the leaf-labeled full binary tree \( \hat{t}_{n+1} \), and clearly also the properties (D) and (E) are inherited.

It remains to deal with the case 2b. This is, however, completely analogous to case 2a, and we refrain from giving the parallel arguments here. \( \square \)

We have seen that finite didendritic systems are essentially finite leaf-labeled full binary trees. One way to produce didendritic systems with infinite label sets is via a projective limit construction as detailed in the following two lemmas. We omit the (simple) proofs. The first lemma says that the natural “projection” of a didendritic system to a subset of its label set is again a didendritic system.

**Lemma 5.10.** Consider a didendritic system \( (N, \equiv, \langle \cdot, \cdot \rangle, <_L, <_R, <) \). Let \( N' \) be a nonempty subset of \( N \). Define an equivalence relation \( \equiv' \) on \( N' \times N' \) by declaring that \( (i, j) \equiv' (k, \ell) \) if and only if \( (i, j) \equiv (k, \ell) \). Write \( \langle i, j \rangle' \) for the equivalence class \( (i, j) \in N' \times N' \). Define a partial order \( \langle \cdot, \cdot \rangle' \) on the equivalence classes of \( \equiv' \) by declaring that \( (i, j) \equiv' (k, \ell) \) if and only if \( \langle i, j \rangle \equiv (k, \ell) \). Define partial orders \( \langle L, \rangle', \langle R, \rangle' \) analogously. Then \( (N', \equiv', \langle \cdot, \cdot \rangle', <_L', <_R', \langle \cdot, \cdot \rangle') \) is a didendritic system.

**Definition 5.11.** Suppose that the didendritic systems \( (N, \equiv, \langle \cdot, \cdot \rangle, <_L, <_R, <) \) and \( (N', \equiv', \langle \cdot, \cdot \rangle', <_L', <_R', \langle \cdot, \cdot \rangle') \) are as in Lemma 5.10. We then say that the latter didendritic system is the restriction of the former to \( N' \).

The following lemma asserts the existence of a projective limit for a projective family of didendritic systems.

**Lemma 5.12.** Suppose that \( (N^n, \equiv n, \langle \cdot, \cdot \rangle n, <_L, <_R, <) \), \( n \in \mathbb{N} \), are didendritic systems such that for \( m < n \) we have \( N^m \subseteq N^n \) and \( (N^m, \equiv m, \langle \cdot, \cdot \rangle m, <_L, <_R, <) \) is the restriction of \( (N^n, \equiv n, \langle \cdot, \cdot \rangle n, <_L, <_R, <) \) to \( N^m \). Put \( N^\infty := \bigcup_{n \in \mathbb{N}} N^n \).
Then there is a unique didendritic system \((N^\infty, \equiv^\infty, \langle \cdot, \cdot \rangle^\infty, \langle \cdot \rangle^\infty, <_L^\infty, <^\infty, <^\infty)\) such that for each \(n \in \mathbb{N}\) the didendritic system \((N^n, \equiv^n, \langle \cdot, \cdot \rangle^n, \langle \cdot \rangle^n, <_L^n, <^n, <^n)\) is the restriction of \((N^\infty, \equiv^\infty, \langle \cdot, \cdot \rangle^\infty, \langle \cdot \rangle^\infty, <_L^\infty, <^\infty, <^\infty)\) to \(N^n\).

6. Infinite PATRICIA bridges and exchangeable random didendritic systems

It is shown in [EGW17, Section 5] that if \((\bar{R}_n^\infty)_{n \in \mathbb{N}}\) is an infinite PATRICIA bridge, then there is a Markov chain \((\bar{R}_n^\infty)_{n \in \mathbb{N}}\) such that for each \(n \in \mathbb{N}\) the random element \(\bar{R}_n^\infty\) is a leaf-labeled full binary tree with \(n\) leaves labeled by \([n]\) and the following hold.

- The full binary tree obtained by removing the labels of \(\bar{R}_n^\infty\) is \(\bar{R}_n^\infty\).
- For every \(n \in \mathbb{N}\), the conditional distribution of \(\bar{R}_n^\infty\) given \(\bar{R}_n^\infty\) is uniform over the \(n!\) possible leaf-labelings of \(\bar{R}_n^\infty\) by \([n]\).
- In going backward from time \(n + 1\) to time \(n\), \(\bar{R}_{n+1}^\infty\) is transformed into \(\bar{R}_n^\infty\) according to the following deterministic procedure:
  - Delete the leaf labeled \(n + 1\), along with its sibling (which may or may not be a leaf).
  - If the sibling of the leaf labeled \(n + 1\) is a leaf, then assign the sibling’s label to the common parent (which is now a leaf).
  - If the sibling of the leaf labeled \(n + 1\) is not a leaf, then attach the subtree below the sibling to the common parent with its leaf labels unchanged and leave all other leaf labels unchanged.

As we saw in Definition 5.3, the leaf-labeled full binary tree \(\bar{R}_n^\infty\) defines a random didendritic system with label set \([n]\) for any \(n \in \mathbb{N}\). Moreover, for any \(n \in \mathbb{N}\) the didendritic system defined by \(\bar{R}_n^\infty\) is the restriction to \([n]\) of the didendritic system defined by \(\bar{R}_{n+1}^\infty\). It follows from Lemma 5.12 that there is a random didendritic system \((\bar{N}, \equiv, \langle \cdot, \cdot \rangle, <_L, <_R, <)\) such that the restriction of this random didendritic system to \([n]\) is the random didendritic system defined by \(\bar{R}_n^\infty\) for all \(n \in \mathbb{N}\). Because of Proposition 5.8 we can recover \(\bar{R}_n^\infty\) and hence \(\bar{R}_n^\infty\) from this restriction and therefore we can recover \((\bar{R}_n^\infty)_{n \in \mathbb{N}}\) and \((\bar{R}_n^\infty)_{n \in \mathbb{N}}\) from \((\bar{N}, \equiv, \langle \cdot, \cdot \rangle, <_L, <_R, <)\).

The random didendritic system defined by \((\bar{R}_n^\infty)_{n \in \mathbb{N}}\) is not arbitrary: it inherits distributional symmetries from the uniform labeling in the construction of \((\bar{R}_n^\infty)_{n \in \mathbb{N}}\) from \((\bar{R}_n^\infty)_{n \in \mathbb{N}}\). We now develop some terminology to describe these symmetries.

**Definition 6.1.** Given a didendritic system \(D = (\bar{N}, \equiv; \langle \cdot, \cdot \rangle; <_L, <_R, <)\) with label set \(\bar{N}\) and a permutation \(\sigma\) of \(\bar{N}\) such that \(\sigma(i) = i\) for all but finitely many \(i \in \bar{N}\), the didendritic system \(D^\sigma = (\bar{N}, \equiv^\sigma; \langle \cdot, \cdot \rangle^\sigma; <_L^\sigma, <_R^\sigma, <^\sigma)\) is defined by

- \((i', j') \equiv^\sigma (i'', j'')\) if and only if \((\sigma(i'), \sigma(j')) \equiv (\sigma(i''), \sigma(j''))\).
- \((i, j)^\sigma\) is the equivalence class of the pair \((i, j)\) for the equivalence relation \(\equiv^\sigma\).
- \((h, i)^\sigma <_L^\sigma (j, k)^\sigma\) if and only if \((\sigma(h), \sigma(i)) < (\sigma(j), \sigma(k))\).
- \((h, i)^\sigma <_R^\sigma (j, k)^\sigma\) if and only if \((\sigma(h), \sigma(i)) < (\sigma(j), \sigma(k))\).
- \((h, i)^\sigma <^\sigma (j, k)^\sigma\) if and only if \(\sigma(h), \sigma(i) < \sigma(j), \sigma(k)\).

A random didendritic system \(D = (\bar{N}, \equiv; \langle \cdot, \cdot \rangle; <_L, <_R, <)\) with label set \(\bar{N}\) is exchangeable if for each permutation \(\sigma\) of \(\bar{N}\) such that \(\sigma(i) = i\) for all but finitely many \(i \in \bar{N}\) the random didendritic system \(D^\sigma\) has the same distribution as \(D\).
The following result is [EGW17, Lemma 5.12].

**Lemma 6.2.** The random dendritic system on $\mathbb{N}$ corresponding to the labeled version of an infinite PATRICIA bridge is exchangeable. Conversely, if we have an exchangeable random dendritic system (on $\mathbb{N}$), apply Lemma 5.10 to restrict it to a sequence of random dendritic systems on $[n]$ for $n \in \mathbb{N}$, and apply Proposition 5.8 to construct a full binary tree that is leaf-labeled by $[n]$, then the resulting sequence of leaf-labeled full binary trees is a PATRICIA bridge.

Our aim is to find concrete representations of the extremal in finite Rémy bridges (recall that an infinite Rémy bridge is extremal if it has a trivial tail $\sigma$-field). To this end, it will be useful to relate the extremality of an infinite Rémy bridge to properties of the associated exchangeable random dendritic system. We say that an exchangeable random dendritic system $D$ is ergodic if

$$P\{D \in A \triangle \{D^\sigma \in A\}\} = 0$$

for all permutations $\sigma$ of $\mathbb{N}$ such that $\sigma(i) = i$ for all but finitely many $i \in \mathbb{N}$ implies that

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

The following result is [EGW17, Proposition 5.19].

**Proposition 6.3.** An infinite Rémy bridge is extremal if and only if the associated exchangeable random dendritic system is ergodic.

**Example 6.4.** We return to Example 4.5 and give a concrete representation of the corresponding random dendritic system. Let $U_1, U_2, \ldots$ be independent and uniformly distributed on $[0, 1]$, and let $\varepsilon_1, \varepsilon_2, \ldots$ be i.i.d. with values in $\{\lt, \gt\}$ and with $P\{\varepsilon_i = \lt\} = P\{\varepsilon_i = \gt\} = \frac{1}{2}$. We define the equivalence relation $\equiv$ on $\mathbb{N} \times \mathbb{N}$ by declaring that

$$(i, j) \equiv (k, k) \iff i = j = k$$

and if $i \neq j$ and $k \neq \ell$ then

$$(i, j) \equiv (k, \ell) \iff U_i \wedge U_j = U_k \wedge U_\ell.$$

The partial order $<$ is defined by declaring that

- for $i, j, k, \ell \in \mathbb{N}$, $(i, j) < (k, k) \iff i \neq j$ and $U_i \wedge U_j \leq U_k$,
- for $i \neq j$, $k \neq \ell$, $(i, j) < (k, \ell) \iff i \neq j$ and $U_i \wedge U_j < U_k \wedge U_\ell$.

We say that $(i, j) \lt_L (k, \ell)$ (respectively, $(i, j) \lt_R (k, \ell)$) if and only if $(i, j) < (k, \ell)$ and $\varepsilon_h = \lt$ (respectively, $\varepsilon_h = \gt$), where

$$h = \begin{cases} i, & \text{if } U_i < U_j, \\ j, & \text{if } U_j < U_i. \end{cases}$$

One can check that the restriction of this random dendritic system to the label set $[n]$ defines a random full binary tree which has the same distribution as the random full binary tree $R^n_{\infty}$ in 4.5. The labeling of this tree by $[n]$ is clearly uniform, and the result of passing from the restriction to $[n+1]$ to the restriction to $[n]$ is given by the deterministic transformation described at the beginning of this section. Together this shows that the corresponding full binary tree-valued process is distributed as the labeled version of the PATRICIA bridge from Example 4.5.

It is clear by the symmetry inherent in the construction that the above random dendritic system is exchangeable. We already saw in Example 4.5 that the above
PATRICIA bridge is extremal, and hence by Proposition 6.3 the exchangeable random didendritic system is ergodic. The ergodicity can also be seen directly from the observation that the restrictions of the random didendritic system to disjoint finite subsets of \( \mathbb{N} \) are independent (see [EGW17, Remark 5.18]).

7. INFINITE PATRICIA BRIDGES AND REAL TREES

The first part of [EGW17, Theorem 8.2] constructs an ergodic exchangeable random didendritic system \( \mathcal{D} \) (and hence an extremal infinite PATRICIA bridge) from an \( \mathbb{R} \)-tree \( \mathbf{S} \) with root \( \rho \), a probability measure \( \mu \) on \( \mathbf{S} \), and a function \( W : \mathbf{S} \times [0,1] \times \mathbf{S} \times [0,1] \to \{\nearrow, \searrow\} \). For the purposes of this construction, we first fix some notation. For \( x, y \in \mathbf{S} \) let \( x \prec y \) denote that element in the segment \([x, y]\) that is closest to the root \( \rho \) of \( \mathbf{S} \) (equivalently, \( |\rho, x \cap y| = |\rho, x| \cap |\rho, y| \)). We say that \( x \prec y \) if \( x \in [\rho, y] \).

The following properties of \( \mathbf{S}, \rho, \mu, \) and \( W \) are essential:

(T) Let \( (\xi_n)_{n \in \mathbb{N}} \) be i.i.d. with common distribution \( \mu \). Then almost surely for distinct \( i, j, k \in \mathbb{N} \), one of

\[
\xi_i \prec \xi_j = \xi_i \prec \xi_k \prec \xi_j \prec \xi_k,
\]

\[
\xi_j \prec \xi_k = \xi_j \prec \xi_i \prec \xi_k \prec \xi_i,
\]

or

\[
\xi_k \prec \xi_i = \xi_k \prec \xi_j \prec \xi_i \prec \xi_j
\]

holds.

(LR) For an independent sequence of i.i.d. \([0,1]\)-valued random variables \( (\vartheta_n)_{n \in \mathbb{N}} \) with common uniform distribution one has almost surely

- for \( i \neq j \), \( W(\xi_i, \vartheta_i, \xi_j, \vartheta_j) = \nearrow \) if and only if \( W(\xi_j, \vartheta_j, \xi_i, \vartheta_i) = \searrow \);
- for distinct \( i, j, k \), if \( \xi_i \prec \xi_j = \xi_i \prec \xi_k \prec \xi_j \prec \xi_k \), then \( W(\xi_i, \vartheta_i, \xi_j, \vartheta_j) = W(\xi_i, \vartheta_i, \xi_k, \vartheta_k) \).

Assume (T) and (LR) hold. Let \( (\xi_i, \vartheta_i) \) be i.i.d. copies of a random variable with distribution \( \mu \otimes \lambda \), where \( \lambda \) is the uniform distribution on \([0,1]\). Using the random input \( (\xi_i, \vartheta_i)_{i \in \mathbb{N}} \), we define

- the equivalence relation \( \equiv^S \) on \( \mathbb{N} \times \mathbb{N} \) by declaring
  for \( i, k, \ell \in \mathbb{N} \) that \( (i, i) \equiv^S (k, \ell) \) if and only if \( i = k = \ell \),
  and for \( i \neq j \) and \( k \neq \ell \), that \( (i, j) \equiv^S (k, \ell) \) if and only if \( \xi_i \prec \xi_j = \xi_k \prec \xi_\ell \),
- the partial order \( <^S \) on the equivalence classes \( (\cdot, \cdot)^S \) of \( \equiv^S \) by declaring
  for \( i, k, \ell \in \mathbb{N} \) that \( (i, j)^S \prec^S (k, \ell)^S \) if and only if \( i \neq j \) and \( \xi_i \prec \xi_j \prec \xi_k \prec \xi_\ell \),
  and for \( i \neq j, k \neq \ell \), that \( (i, j) \prec^S (k, \ell) \) if and only if \( \xi_i \prec \xi_j \prec \xi_k \prec \xi_\ell \),
- the mappings \( w : \{(i, j) : i, j \in \mathbb{N}, i \neq j\} \to \{\nearrow, \searrow\} \) by putting
  \( w(i, j) := W(\xi_i, \vartheta_i, \xi_j, \vartheta_j) \).

Proposition 5.9 then extends \((\mathbb{N}, \equiv^S, <^S, w)\) into a random didendritic system \( \mathcal{D}^S = (\mathbb{N}, \equiv^S, \{\cdot, \cdot\}^S, <^S, <^S_R, <^S) \). The exchangeability of \( \mathcal{D}^S \) is clear because the random variables \( (\xi_i, \vartheta_i), i \in \mathbb{N} \), are independent and identically distributed. The ergodicity of \( \mathcal{D} \) holds because the restrictions of the random didendritic system to disjoint finite subsets of \( \mathbb{N} \) are independent (see [EGW17, Remark 5.18]).

We can now formulate the assertion of [EGW17, Theorem 8.2] in a still more explicit and “constructive” manner.
Theorem 7.1. Let $S$ be a complete separable $\mathbb{R}$-tree with root $\rho$, $\mu$ be a probability measure on $S$, and $W$ be a Borel measurable function from $S \times [0,1] \times S \times [0,1]$ to $\{\prec, \preceq\}$. Suppose that the properties (T) and (LR) hold. Then $D^S = (N, \equiv^S, \langle \cdot, \cdot \rangle^S, <^L_S, <^R_S, <^S)$ is an ergodic exchangeable random didendritic system. Conversely, for any ergodic exchangeable random didendritic system $D$ there exists a 4-tuple $(S, \rho, \mu, W)$ with the abovementioned properties such that $D^S$ has the same distribution as $D$.

In short: The construction described at the beginning of the section builds an ergodic exchangeable random didendritic system (and hence an extremal infinite PATRICIA bridge) from a rooted $\mathbb{R}$-tree endowed with a sampling measure $\mu$ and a “left-right prescription” $W$ that obey the consistency properties (T) and (LR). Conversely, any ergodic exchangeable random didendritic system (and hence any extremal infinite PATRICIA bridge) arises from such a construction.

The first part of Theorem 7.1 and its proof has already been explained. We now briefly review the proof of the second part. The arguments in [EGW17, Section 6] construct from a given ergodic exchangeable random didendritic system $D = (N, \equiv, \langle \cdot, \cdot \rangle, <^L, <^R, <)$ a complete separable (ultrametric) $\mathbb{R}$-tree $T$ with a distinguished point $\rho \in T$, and with an injective mapping from the set of equivalence classes $\{i,j \in N\}$ into $T$ such that the partial order on $T$ defined by the root $\rho$ extends the partial order $<$ on $\{i,j \in N\}$. The core of $T$, denoted by $\Gamma(T)$, is the closure of the set of points of attachment the leaves of $T$. Here, $\Pi(x)$, the point of attachment of a leaf $x$, equals $x$ if the leaf is not isolated, whereas for an isolated leaf it is such that the line segment $[x, \Pi(x))$ is the maximal one among all line segments $[x,y)$ that arise as intersections of $T$ with open balls centered at $x$. Now [EGW17, Proposition 7.4] constructs on the complete separable $\mathbb{R}$-tree $S := \Gamma(T)$ with root $\rho$ a diffuse probability measure $\mu$ having property (T), and such that $(N, \equiv, \langle \cdot, \cdot \rangle, <)$ has the same distribution as $(N, \equiv^S, \langle \cdot, \cdot \rangle^S, <^S)$.

In [EGW17, Section 8], using the Aldous-Hoover-Kallenberg theory on exchangeable random arrays, a Borel measurable function $W : S \times [0,1] \times S \times [0,1] \to \{\prec, \preceq\}$ is constructed which has the above property (LR), and which is such that the resulting ergodic exchangeable random didendritic system $D$ has the same distribution as the resulting ergodic exchangeable random didendritic system $D^S$.

Example 7.2. We continue Examples 4.5 and 6.4. Here, the ultrametric $\mathbb{R}$-tree $T$ may be taken to be the interval $[0, \frac{1}{2}]$, along with disjoint segments of length $\frac{1}{2}(1 - U_i)$ attached to the points $\xi_i = \frac{1}{2}U_i$, and with the root $\rho = 0$. This is the same description of $T$ as in [EGW17, Example 6.7], except that the roles of 0 and $\frac{1}{2}$ have been interchanged, in order to tie in with the way the order $<$ is constructed in Example 6.4. The core $S = \Gamma(T)$ is then the interval $[0, \frac{1}{2}]$, and the sampling measure $\mu$ is the uniform distribution on this interval. The prescription of “left versus right” is then determined by the function

$$W(x, s, y, t) = \begin{cases} \prec, & \text{if } x < y \text{ and } s < \frac{1}{2}, \\ \preceq, & \text{if } x < y \text{ and } s > \frac{1}{2}, \\ \prec, & \text{if } y < x \text{ and } t < \frac{1}{2}, \\ \preceq, & \text{if } y < x \text{ and } t > \frac{1}{2}, \\ \prec, & \text{otherwise.} \end{cases}$$
Example 7.3. We know from Remark 4.4 that \( (\bar{\mathcal{R}}_n)_{n \in \mathbb{N}} \) is an extremal infinite PATRICIA bridge for each diffuse probability measure \( \nu \) on \( \{0,1\}^\infty \). Here the ultrametric \( \mathbb{R} \)-tree \( T \) may be taken as follows (cf. [EGW17, Examples 6.8 & 8.4]): Take the complete binary tree \( \{0,1\}^* \), join two elements of the form \( v_1 \ldots v_k \) and \( v_1 \ldots v_k v_{k+1} \) with a segment of length \( 1/2^{k+1} \), and let \( T \) be the completion of this \( \mathbb{R} \)-tree. The root \( \rho \in T \) is the point corresponding to the root \( \emptyset \in \{0,1\}^* \). The core \( S = \Gamma(T) \) is just \( T \) itself. There is a bijective correspondence between \( \{0,1\}^\infty \) and the points “added” in passing to the completion. The sampling measure \( \mu \) on \( S \) is identified via this correspondence with the probability measure \( \nu \) on \( \{0,1\}^\infty \). Let \( x \) and \( y \) be two points in the support of \( \mu \) that correspond to the points \( u \) and \( v \) in \( \{0,1\}^\infty \). The most recent common ancestor of \( x \) and \( y \) in \( T \) (that is, the point \( z \) such that \( [\rho,z] = [\rho,x] \cap [\rho,y] \)) is the point in \( T \) corresponding to the most recent common ancestor of \( u \) and \( v \) in \( \{0,1\}^* \). The “left versus right” rule is then given by

\[
W(x, y, t) = \begin{cases} \wedge, & \text{if } u \wedge v <_L u \text{ and } u \wedge v <_R v, \\
\vee, & \text{if } u \wedge v <_L v \text{ and } u \wedge v <_R u. 
\end{cases}
\]

Acknowledgement. We thank the referee of [EW17] for suggesting that we extend our analysis of the radix sort chain to the PATRICIA chain. We acknowledge the hospitality of the Institute for Mathematical Sciences, National University of Singapore, where part of this paper was written.

References

[And92] Arne Andersson, Comments on: “On the balance property of Patricia tries: external path length viewpoint” [Theoret. Comput. Sci. 68 (1989), no. 1, 1–17; MR1022654 (91a:68036)] by P. Kirschenhofer, H. Prodinger and W. Szpankowski, Theoret. Comput. Sci. 106 (1992), no. 2, 391–400, With a reply by Kirschenhofer, Prodinger and Szpankowski. MR 1192779

[AS95] Laurent Alonso and René Schott, Random generation of trees, Kluwer Academic Publishers, Boston, MA, 1995, Random generators in computer science. MR 1331596 (96k:68154)

[Bou01] Jérémie Bourdon, Size and path length of Patricia tries: dynamical sources context, Random Structures Algorithms 19 (2001), no. 3-4, 289–315, Analysis of algorithms (Krynica Morska, 2000). MR 1871557

[Dev92] Luc Devroye, A note on the probabilistic analysis of Patricia trees, Random Structures Algorithms 3 (1992), no. 2, 203–214. MR 1151362

[Dev02] ______, Laws of large numbers and tail inequalities for random tries and PATRICIA trees, J. Comput. Appl. Math. 142 (2002), no. 1, 27–37, Probabilistic methods in combinatorics and combinatorial optimization. MR 1910516

[Dev05] ______, Universal asymptotics for random tries and PATRICIA trees, Algorithmica 42 (2005), no. 1, 11–29. MR 2131826

[EGW17] Steven N. Evans, Rudolf Grübel, and Anton Wakolbinger, Doob-Martin boundary of Rémy’s tree growth chain, Ann. Probab. 45 (2017), no. 1, 225–277. MR 3601650

[EW17] Steven N. Evans and Anton Wakolbinger, Radix sort trees in the large, Electron. Commun. Probab. 22 (2017), Paper No. 68, 13. MR 3734107

[FL14] Michael Fuchs and Chung-Kuei Lee, A general central limit theorem for shape parameters of m-ary tries and PATRICIA trees, Electron. J. Combin. 21 (2014), no. 1, Paper 1.68, 26. MR 3192399

[Ger18] Julian Gerstenberg, Exchangeable interval hypergraphs and limits of ordered discrete structures, 2018, eprint arXiv:1802.09015 [math.PR].

[Gwe68] G. Gwehenberger, Anwendung einer binären Verweiskettenmethode beim Aufbau von Listen, Elektronische Rechenanlagen 10 (1968), 223–226.
[Knu73] Donald E. Knuth, *The art of computer programming. Volume 3*, Addison-Wesley Publishing Co., Reading, Mass.-London-Don Mills, Ont., 1973, Sorting and searching, Addison-Wesley Series in Computer Science and Information Processing. MR 0445948 (56 #4281)

[KPS89] Peter Kirschenhofer, Helmut Prodinger, and Wojciech Szpankowski, *On the balance property of Patricia tries: external path length viewpoint*, Theoret. Comput. Sci. 68 (1989), no. 1, 1–17. MR 1022654

[KS02] Charles Knessl and Wojciech Szpankowski, *Limit laws for the height in PATRICIA tries*, J. Algorithms 44 (2002), no. 1, 63–97, Analysis of algorithms. MR 1932678

[Lec15] Kevin Leckey, *Probabilistic Analysis of Radix Algorithms on Markov Sources*, Ph.D. thesis, Johann Wolfgang Goethe-Universität in Frankfurt am Main, 2015.

[LG99] Jean-François Le Gall, *Spatial branching processes, random snakes and partial differential equations*, Lectures in Mathematics ETH Zürich, Birkhäuser Verlag, Basel, 1999. MR 1714707 (2001g:60211)

[Mag15] Abram N. Magner, *Profiles of Patricia Tries*, ProQuest LLC, Ann Arbor, MI, 2015, Thesis (Ph.D.)–Purdue University. MR 3517923

[Mar03] Philippe Marchal, *Constructing a sequence of random walks strongly converging to Brownian motion*, Discrete random walks (Paris, 2003), Discrete Math. Theor. Comput. Sci. Proc., AC, Assoc. Discrete Math. Theor. Comput. Sci., Nancy, 2003, pp. 181–190 (electronic). MR 2042386 (2004m:60099)

[MKS14] Abram Magner, Charles Knessl, and Wojciech Szpankowski, *Expected external profile of PATRICIA tries*, ANALCO14—Meeting on Analytic Algorithmics and Combinatorics, SIAM, Philadelphia, PA, 2014, pp. 16–24. MR 3248350

[Mor68] Donald R. Morrison, *Patricia—Practical Algorithm To Retrieve Information Coded in Alphanumeric*, Journal of the ACM 15 (1968), 514–534.

[MS18] Abram Magner and Wojciech Szpankowski, *Profiles of PATRICIA tries*, Algorithmica 80 (2018), no. 1, 331–397. MR 3741547

[Rém85] Jean-Luc Rémy, *Un procédé itératif de dénombrement d’arbres binaires et son application à leur génération aléatoire*, RAIRO Inform. Théor. 19 (1985), no. 2, 179–195. MR 803997 (87h:68132)

[RJS93] Bonita Rais, Philippe Jacquet, and Wojciech Szpankowski, *Limiting distribution for the depth in PATRICIA tries*, SIAM J. Discrete Math. 6 (1993), no. 2, 197–213. MR 1215228

[RP81] L. C. G. Rogers and J. W. Pitman, *Markov functions*, Ann. Probab. 9 (1981), no. 4, 573–582. MR 624684

[Sta15] Richard P. Stanley, *Catalan numbers*, Cambridge University Press, New York, 2015. MR 3467982

[Szp90] Wojciech Szpankowski, *Patricia tries again revisited*, J. Assoc. Comput. Mach. 37 (1990), no. 4, 691–711. MR 1083648

[TGL14] Weitian Tong, Randy Goebel, and Guohui Lin, *On the smoothed heights of Trie and Patricia index trees*, Computing and combinatorics, Lecture Notes in Comput. Sci., vol. 8591, Springer, Cham, 2014, pp. 94–103. MR 3247579

[TGL16] ______, *Smoothed heights of tries and patricia tries*, Theoret. Comput. Sci. 609 (2016), no. part 3, 620–626. MR 3429094

Department of Statistics, University of California, 367 Evans Hall #3860, Berkeley, CA 94720-3860, U.S.A.

E-mail address: evans@stat.berkeley.edu

Institut für Mathematik, Goethe-Universität, 60054 Frankfurt am Main, Germany

E-mail address: wakolbinger@math.uni-frankfurt.de