Counting fixed-point-free Cayley permutations

Giulio Cerbai and Anders Claesson

Department of Mathematics
University of Iceland

12 July 2025

Abstract

Cayley permutations generalize permutations by allowing repeated values. We use two-sort species to derive differential equations that capture the recursive structure of their functional digraphs. These equations enable us to enumerate fixed-point-free Cayley permutations. Our approach also yields combinatorial identities and counting formulas for Cayley permutations whose functional digraphs have specific graph-theoretical properties, such as being a tree, being a forest, or being connected.

1 Introduction

The *hat-check problem* lies at the intersection of combinatorics and probability theory. Pierre Rémond de Montmort [13] first described it in 1708 in his "Essai d'analyse sur les jeux de hasard". The problem asks for the probability that a uniformly random permutation is a *derangement*, that is, fixed-point-free. Montmort and Bernoulli later provided solutions featuring a striking occurrence of Napier's constant in combinatorics: as the size grows to infinity, the probability converges to $1/e$.

A lesser-known yet easier-to-prove result is that *endofunctions* have the same asymptotic behavior. Here endofunctions are maps from $[n] = \{1, 2, \dots, n\}$ to $[n]$. Among all n^n endofunctions, $(n-1)^n$ have no fixed points. As n approaches infinity, the quotient converges again to $1/e$. Indeed, $(n-1)^n/n^n = (1-1/n)^n \to 1/e$.

Cayley permutations lie between permutations and endofunctions and present a more challenging version of this problem. We define the combinatorial species of Cayley permutations in Section 4. Cayley permutations are endofunctions whose image set contains every number between one and its maximum value. They can also be seen as the lexicographically smallest representatives for equivalence classes of positive integer words modulo order isomorphism. Cayley permutations play a key role in recent research extending the theory of permutations and permutation patterns to broader classes of objects [1, 6, 7, 8, 9].

The inherent difficulty of *Cayley-derangements*—Cayley permutations with no fixed points—lies in their biased nature. Unlike permutations and endofunctions, Cayley permutations exhibit a bias toward smaller values. In permutations and endofunctions, fixed points are uniformly distributed, so the probability of having a fixed point at a certain position is independent of the position. In Cayley permutations, smaller values are more likely to appear and thus more likely to be fixed points. Furthermore, the combinatorial description of derangements is less transparent for Cayley permutations than for permutations and endofunctions. For instance, a permutation can be seen as a pair consisting of the set of its fixed points together with a derangement of the non-fixed points, which leads directly to the well-known enumeration of derangements by the subfactorial numbers.

This paper provides a counting formula for Cayley-derangements (Theorem 6.6). Namely, the number of Cayley-derangements on $[n]$ is

$$\sum_{r=0}^{n} \frac{|\mathrm{Der}[r]|}{r!} \sum_{i=0}^{n} i! \left\{ \frac{n}{i} \right\}_r,$$

in which $|\mathrm{Der}[r]|$ is the rth subfactorial and $\left\{\frac{n}{i}\right\}_r = \left\{ {n \atop i} \right\}_r - \left\{ {n \atop i} \right\}_{r+1}$ is a difference of r-Stirling numbers.

We use two-sort species to capture the combinatorial properties of functional digraphs of Cayley-derangements. In fact, we derive differential equations encoding the recursive structure of the broader class of *R-recurrent* Cayley permutations. These include Cayley permutations whose functional digraph has certain graph-theoretical properties such as being a tree, being a forest, or being connected. The appearance of species and functional digraphs in this context is unsurprising given Joyal's [12] brilliant proof of Cayley's formula for the number of trees. We outline the paper's content below.

We begin with foundational material on combinatorial species (Section 2) and endofunctions (Section 3), establishing the connection between endofunctions and their functional digraphs. We then introduce Cayley permutations as a natural bridge between permutations and general endofunctions (Section 4).

Our main contribution lies in constructing R-recurrent functional digraphs and deriving their enumerative properties (Section 5). By exploiting the recursive structure inherent in these objects, we obtain differential equations for two-sort species that yield an explicit counting formula for Cayley-derangements (Section 6). Our approach extends Joyal's [12] classical bijection between doubly-rooted trees and functional digraphs to the two-sort setting (Section 7). We also complement our two-sort analysis with unisort equations (Section 8).

Statistical properties of R-recurrent functional digraphs emerge naturally from our framework, including distributions of internal nodes and connected components (Section 9). The paper concludes by exploring generalizations that suggest promising directions for future research (Section 10), including modifications to the recursive structure.

2 Combinatorial species

A $\mathbb{B}$-*species* (or simply *species*) F is a rule that produces

- for each finite set U, a finite set $F[U]$;

- for each bijection $\sigma : U \to V$, a bijection $F[\sigma] : F[U] \to F[V]$ such that $F[\sigma \circ \tau] = F[\sigma] \circ F[\tau]$ for all bijections $\sigma : U \to V$, $\tau : V \to W$, and $F[\mathrm{id}_U] = \mathrm{id}_{F[U]}$ for the identity map $\mathrm{id}_U : U \to U$.

An element $s \in F[U]$ is called an *F-structure*, or *structure of F*, on U, and the map $F[\sigma]$ is called the *transport of F-structures along σ*. A species defines both a class of labeled combinatorial structures—$F[U]$—and how relabeling of the underlying set affects these structures—the transport of structure $F[\sigma]$. Since the transport of structure is a bijection, the cardinality of $F[U]$ depends solely on the cardinality of U, not on the nature of the elements of U.

For $n \geq 0$, let $[n] = \{1, 2, \ldots, n\}$; in particular, $[0] = \emptyset$. The *(exponential) generating series* of a species F is the formal power series

$$F(x) = \sum_{n \geq 0} |F[n]| \frac{x^n}{n!},$$

where $F[n]$ is short for $F[[n]]$.

The species G is a *subspecies* of F, written $G \subseteq F$, if for each pair of finite sets U, V and bijection $\sigma : U \to V$ we have $G[U] \subseteq F[U]$ and $G[\sigma] = F[\sigma]|_{G[U]}$. We denote by F_+ and F_n the subspecies of F consisting of F-structures on nonempty sets and on sets with cardinality n, respectively.

Throughout this paper we use the following species, whose formal definitions can be found in Joyal's seminal paper [12] or in the book by Bergeron, Labelle and Leroux [2]. These are the main references for combinatorial species theory. Shorter introductions can be found in our recent papers [9, 10, 11].

$1 = E_0$ is the species characteristic of the empty set.

$X = E_1$ is the species characteristic of singletons.

E is the species of sets.

L is the species of linear orders.

$\mathcal{S}$ is the species of permutations.

$\mathcal{C}$ is the species of cyclic permutations (cycles).

Their generating series are easy to compute. We have $1(x) = 1$, $X(x) = x$,

$$E(x) = e^x, \; L(x) = \mathcal{S}(x) = 1/(1-x) \; \text{and} \; \mathcal{C}(x) = -\log\bigl(1-x\bigr).$$

The reader familiar with category theory will see that a $\mathbb{B}$-species is nothing but a functor $F : \mathbb{B} \to \mathbb{B}$, where $\mathbb{B}$ is the category of finite sets with bijections as morphisms. Two species F and G are *(combinatorially) equal*, and we write $F = G$, if there is a natural isomorphism of functors between F and G. It is easy to see that $F(x) = G(x)$ if $F = G$. For $\mathbb{B}$-species, the converse is false. A well-known example is given by linear orders L and permutations $\mathcal{S}$, defined as follows. Let U be a finite set with $n = |U|$. A linear order of U is a bijection $f : [n] \to U$, while a permutation is a bijection $g : U \to U$. The transport of structure along a bijection $\sigma : U \to V$ is defined as

$$L[\sigma](f) = \sigma \circ f \quad \text{and} \quad \mathcal{S}[\sigma](g) = \sigma \circ g \circ \sigma^{-1}.$$

Although the number of structures over $[n]$ is $n!$ in both cases, and hence $L(x) = \mathcal{S}(x)$, the species L and $\mathcal{S}$ are not combinatorially equal [2, Ex. 10, Chapter 1.1].

Let us now introduce several operations that allow us to build new species from previously known ones. To do so, we define only the set of structures. The transport of structure, here omitted, is induced naturally, and the reader interested in a formal definition is referred to the usual book [2].

Let us start with sum and product. Given two species F and G, an $(F+G)$-*structure* is either an F-structure or a G-structure:

$$(F+G)[U] = F[U] \sqcup G[U],$$

where $\sqcup$ denotes disjoint union. An $(F \cdot G)$-*structure* on U is a pair (s,t) such that s is an F-structure on a subset $U_1 \subseteq U$ and t is a G-structure on $U_2 = U \setminus U_1$:

$$(F \cdot G)[U] = \bigsqcup_{(U_1,U_2)} F[U_1] \times G[U_2],$$

where $U = U_1 \sqcup U_2$ and "$\times$" denotes cartesian product.

Next, we define composition of species. An $(F \circ G)$-*structure*, where $G[\emptyset] = \emptyset$, is a generalized partition where each block carries a G-structure and the set of blocks is structured by F:

$$(F \circ G)[U] = \bigsqcup_{\beta \in \mathrm{Par}[U]} F[\beta] \times \prod_{B \in \beta} G[B],$$

where $\mathrm{Par}[U]$ denotes the set of partitions of U.

The *derivative* F' of F is obtained by putting an F-structure on a set to which we add an additional element, denoted by $\star$:

$$F'[U] = F[U \sqcup \{\star\}].$$

It is easy to see that the generating series behave well with respect to all the operations introduced thus far:

$$(F+G)(x) = F(x) + G(x); \quad (F \cdot G)(x) = F(x) \cdot G(x);$$
$$(F \circ G)(x) = F(G(x)); \quad F'(x) = \frac{d}{dx}[F(x)].$$

Composition allows us to define the species of *set partitions* and *ballots* (ordered set partitions) as $\mathrm{Par} = E \circ E_+$ and $\mathrm{Bal} = L \circ E_+$. It immediately follows that $\mathrm{Par}(x) = \exp(e^x - 1)$ and $\mathrm{Bal}(x) = 1/(2 - e^x)$, where $|\mathrm{Bal}[n]| = |\mathrm{Cay}[n]|$ equals the nth Fubini number (A000670 [14]). Finally, we introduce an operation of *pointing* that allows us to distinguish a specific element of the underlying set. The species F-*pointed*, denoted as $F^\bullet$, is defined by

$$F^\bullet[U] = F[U] \times U.$$

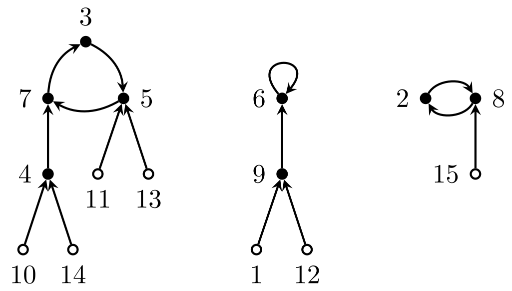


Figure 1: The functional digraph of $f = 985776326459548 \in \mathrm{End}[15]$, where the ith letter of f is $f(i)$. Internal nodes are black while leaves are white.

An $F^\bullet$*-structure* on U is thus a pair (s, u) where s is an F-structure on U and $u \in U$ is a distinguished element that we can think of as being "pointed at". Note that $|F^\bullet[n]| = n|F[n]|$. Furthermore, derivative and pointing are related by

$$F^\bullet = X \cdot F' \quad \text{and} \quad F^\bullet(x) = xF'(x).$$

Virtual species are an extension of $\mathbb{B}$-species that allow for subtraction of species. Details about virtual species are not relevant in this context, so we refer the interested reader to Section 2.5 of the usual book [2]. Two examples that will be useful in this paper are the multiplicative inverses of sets and linear orders:

$$E^{-1} = \sum_{k \geq 0} (-1)^k (E_+)^k; \qquad L^{-1} = 1 - X.$$

3 Endofunctions and functional digraphs

The species of *endofunctions* is defined by

$$\mathrm{End}[U] = \{f : U \to U\} \quad \text{and} \quad \mathrm{End}[\sigma](f) = \sigma \circ f \circ \sigma^{-1}$$

for any finite sets U and V and any bijection $\sigma : U \to V$. Throughout this paper we identify an endofunction $f \in \mathrm{End}[U]$ with its *functional digraph*, defined as the directed graph with vertex set U and directed edges $(u, f(u))$, or $u \mapsto f(u)$, $u \in U$. Alternatively, the species of functional digraphs is the subspecies of directed graphs where each node has outdegree one. The transport of structures is inherited from the transport on directed graphs: a directed graph on U with edges $\{(x, y)\}$ is mapped by the transport of structure along σ to the directed graph on V with edges $\{(\sigma(x), \sigma(y)\}$. Identifying an endofunction with its functional digraph is a natural isomorphism, and thus the species of endofunctions and functional digraphs are combinatorially equal. An instance of a functional digraph can be found in Figure 1.

Let $u \in U$ be a node in the functional digraph of f. We say that u is *internal* if it has positive indegree, that is, if $u \in \mathrm{Im}(f)$. Otherwise, u is a *leaf*. A *fixed point* of f is an element $u \in U$ such that $f(u) = u$, which corresponds to a *loop* in the functional digraph.

A *permutation* is a bijective endofunction. It is well known that any permutation can be written as a product of disjoint cycles. In the language of species, a permutation

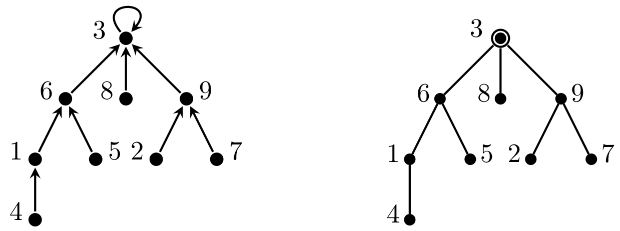


Figure 2: The functional digraph of $f = 693163933 \in \mathrm{End}[9]$, on the left, and the corresponding rooted tree, on the right.

is a set of cycles: $\mathcal{S} = E(\mathcal{C})$, where $\mathcal{C}$ is the species of *cyclic permutations*, that is, permutations consisting of a single cycle. In terms of generating series, we have

$$\mathcal{S}(x) = \exp(\mathcal{C}(x)) \quad \text{and} \quad \mathcal{C}(x) = \log\big(\mathcal{S}(x)\big) = \log\left(\frac{1}{1-x}\right).$$

Alternatively, a permutation is a pair consisting of the set of its fixed points together with a *derangement* of the non-fixed points. A permutation with no fixed point is called a *derangement* and we denote the species of derangements (of permutations) by Der. This decomposition yields the species equation $\mathcal{S} = E \cdot \mathrm{Der}$. Now, solving $\mathcal{S}(x) = E(x) \cdot \mathrm{Der}(x)$ for $\mathrm{Der}(x)$ one obtains

$$\mathrm{Der}(x) = \frac{e^{-x}}{1-x} \quad \text{and} \quad |\mathrm{Der}[n]| = n! \sum_{i=0}^{n} \frac{(-1)^i}{i!},$$

a familiar expression for the number of derangements (A000166 [14]).

Next, we introduce the species $\mathcal{A}$ of rooted trees. It plays a key role in this paper. A *tree* is a connected acyclic simple graph and we let $\mathfrak{a}$ denote the species of trees. A *rooted tree* is a tree with a distinguished node, called its *root*. Clearly, a rooted tree is a synonym for a pointed tree, $\mathcal{A} = \mathfrak{a}^\bullet$. A rooted tree consists of a root (an X-structure) to which is appended a set of rooted trees (an $E(\mathcal{A})$-structure). That is,

$$\mathcal{A} \,=\, X \cdot E(\mathcal{A}).$$

A slight tweak of the above definition allows us to see rooted trees as a subspecies of functional digraphs. Note that there is exactly one cycle in every connected component of a functional digraph. We identify rooted trees with connected functional digraphs whose only cycle is a loop (a fixed point); this loop plays the role of root, and the edges lose their orientation. Conversely, by putting a loop around the root of a tree and orienting the edges from the leaves towards the root we obtain a connected functional digraph whose only cycle is a loop. A rooted tree and its digraph counterpart are depicted in Figure 2.

A formula first discovered by Borchardt [3] and later attributed to Cayley [5] shows that there are n^{n-2} (labeled) trees on $[n]$. Joyal [12] exploited the interplay between rooted trees and endofunctions to provide a brilliant proof of Cayley's formula (see also Section 2.1 of the book by Bergeron, Labelle and Leroux [2]). We recall the core idea of Joyal's proof, which we will extend to two-sort species in Section 7. First, observe that an endofunction is a permutation of rooted trees:

$$\mathrm{End} = \mathcal{S}(\mathcal{A}). \tag{1}$$

To see this, let us distinguish two types of points in its functional digraph:

- *Recurrent points*: u such that $f^k(u) = u$ for some $k \geq 1$; they are precisely the points located on a cycle. Taken as whole they form a set of cycles, that is, a permutation.
- *Nonrecurrent points*: u such that $f^k(u) \neq u$ for every $k \geq 1$; they are arranged in (rooted) trees each of which hangs from a recurrent point, which can be seen as its root.

The underlying set U is thus partitioned into a set of rooted trees, structured by a permutation of recurrent points. For example, the permutation of the recurrent points of the endofunction in Figure 1 is

$$2 \mapsto 8; \quad 3 \mapsto 5; \quad 5 \mapsto 7; \quad 6 \mapsto 6; \quad 7 \mapsto 3; \quad 8 \mapsto 2.$$

Joyal proved that doubly-rooted trees and nonempty linear orders of rooted trees are combinatorially equal. Indeed, let t be a structure of $\mathcal{A}^\bullet = \mathfrak{a}^{\bullet\bullet}$; that is, a tree where two nodes (not necessarily distinct) are distinguished. Define the *spine* of t as the directed path from the first distinguished node, the *tail*, to the second distinguished node, the *head*. Joyal used the suggestive name "vertebrates" for the species of doubly-rooted trees. The spine determines a nonempty linear order, with a (rooted) tree hanging from each of its nodes. The resulting structure is a nonempty linear order of rooted trees:

$$\mathcal{A}^\bullet = L_+(\mathcal{A}).$$

Two species F and G are said to be *equipotent*, written $F \equiv G$, if $F(x) = G(x)$. Since $L \equiv \mathcal{S}$ we have

$$\mathcal{A}^\bullet = L_+(\mathcal{A}) \equiv \mathcal{S}_+(\mathcal{A}) = \mathrm{End}_+$$

and, consequently, $|\mathcal{A}^\bullet[n]| = |\mathrm{End}[n]| = n^n$ for every $n \geq 1$. Finally,

$$n^n = |\mathcal{A}^\bullet[n]| = |\mathfrak{a}^{\bullet\bullet}[n]| = n^2 \cdot |\mathfrak{a}[n]|$$

from which Cayley's formula immediately follows.

If we assume that $U = \{u_1, \ldots, u_n\}$ is totally ordered, with $u_1 < u_2 < \cdots < u_n$, then Joyal's proof induces a bijection between doubly-rooted trees on U and functional digraphs on U. In this case, we map the spine $f \in L[U]$ of a doubly-rooted tree to the permutation $g \in \mathcal{S}[U]$ such that $g(u_i) = u_j \Leftrightarrow f(i) = u_j$, leaving the rooted trees attached to the spine untouched. This way we obtain a permutation of rooted trees, that is, an endofunction. For instance, the doubly-rooted tree associated with the endofunction of Figure 1 is depicted in Figure 3.

By adjusting equation (1), one obtains species identities for several classes of endofunctions defined in terms of properties of their functional digraph. For instance, the recurrent points of a connected endofunction form a cyclic permutation, and hence the species of *connected endofunctions* is $\mathcal{C}(\mathcal{A})$. Similarly, if the recurrent points are structured in a set, we obtain *forests* (acyclic functional digraphs), $E(\mathcal{A})$. If the permutation of recurrent points is a derangement, we obtain functional digraphs with no fixed points, $\mathrm{Der}(\mathcal{A})$. In some cases, pairing these equations with Cayley's formula leads to explicit enumeration of the corresponding digraphs. All the examples mentioned above have been widely discussed in the literature. See also Table 1 for the corresponding counting sequences, in which A-numbers refer to the OEIS [14].

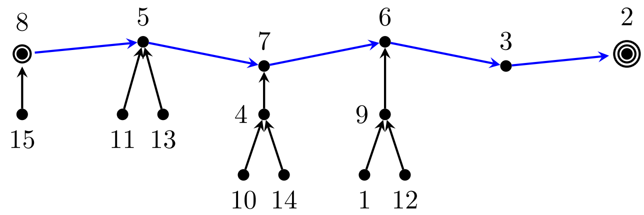


Figure 3: The doubly-rooted tree associated with the endofunction of Figure 1 under Joyal's bijection. The spine is colored in blue; tail and head are distinguished with one and two circles, respectively.

	Permutations		Endofunctions	
All structures	$\mathcal{S}$	$n!$	End	n^n
Trees	1	$\delta_{n,0}$	$\mathcal{A}$	n^{n-1}
Forests	E	1	$E(\mathcal{A})$	$(n+1)^{n-1}$
Connected	$\mathcal{C}$	$(n-1)!(1-\delta_{n,0})$	$\mathcal{C}(\mathcal{A})$	A001865
Derangements	Der	A000166	$\mathrm{Der}(\mathcal{A})$	$(n-1)^n$

Table 1: Enumeration of functional digraphs of permutations and endofunctions, where δ is the Kronecker delta.

4 Cayley permutations

We now define the species of Cayley permutations [10]. A *Cayley permutation* on U is a function $f : U \to [n]$ where $\mathrm{Im}(f) = [k]$ for some $k \le n = |U|$. The transport of structure along a bijection $\sigma : U \to V$ is given by

$$\mathrm{Cay}[\sigma](f) = f \circ \sigma^{-1}.$$

Cayley permutations and ballots are combinatorially equal [10]. In the special case where $U = [n]$, we have the inclusions

$$\mathcal{S}[n] \subseteq \mathrm{Cay}[n] \subseteq \mathrm{End}[n]$$

and all these structures can be identified with their functional digraphs. There is, however, no obvious way to define the functional digraph of Cayley permutations on an arbitrary finite set U. Furthermore, the transport of structure on $\mathcal{S}$ and End is conjugation, which cannot be extended to Cayley permutations due to the requirement $\mathrm{Im}(f) = [k]$. For these reasons, we now assume $\mathbb{L}$-species, defined below.

Let $\mathbb{L}$ denote the category of finite totally ordered sets with order-preserving bijections as morphisms. An $\mathbb{L}$-*species* is a functor $F : \mathbb{L} \to \mathbb{B}$. Since there is a unique order-preserving bijection between any pair of finite totally ordered sets of the same cardinality, transport of structure is immaterial in the context of $\mathbb{L}$-species. We often assume that the underlying total order is $[n]$ with the standard order and regard permutations and Cayley permutations as subspecies of endofunctions. This allows us to identify all these structures with their functional digraph. Another consequence of having a unique order-preserving bijection between totally ordered sets with the same cardinality is that for $\mathbb{L}$-species $F = G$ if and only if $F(x) = G(x)$. For example,

$L = \mathcal{S}$ as $\mathbb{L}$-species even though $L \neq \mathcal{S}$ as $\mathbb{B}$-species. In this case, a permutation f in $\mathcal{S}[n]$ is naturally identified with the linear order $f(1)f(2)\dots f(n)$ in $L[n]$.

Operations on $\mathbb{B}$-species extend naturally to $\mathbb{L}$-species. New operations also become possible. Given two totally ordered sets $\ell_1 = (U_1, \preceq_1)$ and $\ell_2 = (U_2, \preceq_2)$, their *ordinal sum* $\ell = \ell_1 \oplus \ell_2$ is the totally ordered set $\ell = (U, \preceq)$ where $U = U_1 \sqcup U_2$, $\preceq$ respects $\preceq_1$ and $\preceq_2$, and all elements of U_1 are smaller than the elements of U_2. The totally ordered sets obtained by adding a new minimum or maximum element to ℓ are denoted by $1 \oplus \ell$ and $\ell \oplus 1$, respectively. Given a totally ordered set $\ell = (U, \preceq)$, the derivative F' of an $\mathbb{L}$-species F is defined by $F'[\ell] = F[1 \oplus \ell]$. Additionally, the *integral* of F, denoted $\int_0^X F(T)dT$ or $\int F$, is defined by

$$\left(\int F\right)[\ell] = \begin{cases} \emptyset & \text{if } \ell = \emptyset, \\ F[\ell \setminus \{\min(\ell)\}] & \text{if } \ell \neq \emptyset. \end{cases}$$

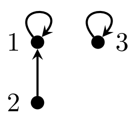
In other words, an F'-structure is an F-structure on a total order endowed with a new minimum, and an $\int F$-structure is an F-structure on a total order deprived of its minimum. Derivative and integral can be alternatively defined in terms of maximum elements: $F'[\ell] = F[\ell \oplus 1]$ and $(\int F)[\ell] = F[\ell \setminus \{\max(\ell)\}]$ for $\ell \neq \emptyset$. Similarly to $\mathbb{B}$-species, operations on $\mathbb{L}$-species behave well with respect to generating series, including a species version of the fundamental theorem of calculus [2, equation 35]:

$$\int F' = F_+; \qquad \left(\int F\right)' = F; \qquad \left(\int_0^X F(T)dT\right)(x) = \int_0^x F(t)dt.$$

Let us now return to Cayley permutations. We have observed in Section 3 that functional digraphs of endofunctions are precisely those directed graphs where each node has outdegree one. It is easy to see that a functional digraph is the functional digraph of a permutation if and only if each node has both indegree and outdegree equal to one. The following lemma holds because the internal nodes of $f \in \mathrm{End}[n]$ are labeled by $\mathrm{Im}(f)$, and we omit its proof. It characterizes the functional digraphs of Cayley permutations as those whose internal nodes are labeled with the smallest elements of the underlying total order.

Lemma 4.1. *Let $f \in \mathrm{End}[n]$. Then f is a Cayley permutation with $\mathrm{Im}(f) = [k]$ if and only if the internal nodes in the functional digraph of f have labels $[k]$.*

Although $\mathbb{L}$-species address some of the cosmetic issues of Cayley permutations, there are structural problems that require a more sophisticated approach. As an example, consider the following functional digraph:



It is the functional digraph of the endofunction f in End[3] where $f(1) = f(2) = 1$ and $f(3) = 3$. Each connected component is a rooted tree that satisfies the requirement of Lemma 4.1: there is only one leaf in the tree on the left, which is labeled with the largest element of the underlying set $\{1, 2\}$. However, the resulting functional digraph does *not* satisfy this property, and indeed f is not a Cayley permutation. More generally, an endofunction is a permutation of rooted trees, but a Cayley permutation is *not* a permutation of rooted trees of Cayley permutations. Similarly, we cannot use composition to obtain species equations for Cayley permutations whose functional digraph is connected, a forest, or a derangement. Generalizing the equation $\mathcal{S} = E \cdot \mathrm{Der}$ is also not possible: a Cayley permutation is *not* a pair consisting of the set of its fixed points together with a Cayley permutation without fixed points.

To circumvent these issues, we assume that internal nodes and leaves are nodes of different sorts. This can be formalized using two-sort species.

A *two-sort* $\mathbb{B}$-*species* (or $\mathbb{B}$-*species of two sorts*) is a functor $F : \mathbb{B} \times \mathbb{B} \to \mathbb{B}$, where $\mathbb{B} \times \mathbb{B}$ denotes the product category with pairs of finite sets as objects and pairs of bijections as arrows. Composition of arrows is componentwise and the identity arrow on the pair (U, V) is $\mathrm{id}_{(U,V)} = (\mathrm{id}_U, \mathrm{id}_V)$. Two-sort species allow us to build combinatorial structures whose underlying set contains elements of two different sorts: those belonging to U and those belonging to V. For a two-sort species F, we write $F = F(X, Y)$ to indicate that the two sorts are called X and Y. The letters X and Y also serve as singleton species of the corresponding sort:

$$X[U, V] = \begin{cases} \{u\}, & \text{if } U = \{u\} \text{ and } V = \emptyset; \\ \emptyset, & \text{otherwise.} \end{cases}$$

$$Y[U, V] = \begin{cases} \{v\}, & \text{if } U = \emptyset \text{ and } V = \{v\}; \\ \emptyset, & \text{otherwise.} \end{cases}$$

The *(exponential) generating series* of $F(X, Y)$ is

$$F(x, y) = \sum_{i,j \geq 0} |F[i, j]| \frac{x^i}{i!} \frac{y^j}{j!},$$

where we let $F[i, j] = F[[i], [j]]$. Most of the operations introduced for unisort species can be naturally extended to the two-sort context, and definitions can be found in the usual references [2, 11]. An example relevant in the next sections is partial differentiation:

$$\left(\frac{\partial F}{\partial X}\right)[U, V] = F[U \cup \{*\}, V]; \qquad \left(\frac{\partial F}{\partial Y}\right)[U, V] = F[U, V \cup \{*\}].$$

Similarly, we have two pointing operations:

$$F^{\bullet_X} = X \cdot \left(\frac{\partial F}{\partial X}\right); \qquad F^{\bullet_Y} = Y \cdot \left(\frac{\partial F}{\partial Y}\right). \tag{2}$$

Composition of two-sort species is more delicate and requires additional care. We define it below in the only case needed in this paper. Let F be a unisort species and let $G = G(X, Y)$ be a two-sort species. Then $F \circ G = (F \circ G)(X, Y)$ is the two-sort species with structures

$$(F \circ G)[U, V] = \bigcup_{\beta \in \mathrm{Par}[U \sqcup V]} F[\beta] \times \prod_{B \in \beta} G[B \cap U, B \cap V].$$

An $\big(F \circ G\big)$-structure is thus a generalized partition on elements of two sorts, where each block carries a G-structure (of two-sorts), and the blocks are structured by the (unisort) species F.

5 R-recurrent functional digraphs

Recall from Section 3 that rooted trees of one sort are defined by the species equation

$$\mathcal{A}(X) \,=\, X \cdot E\big(\mathcal{A}(X)\big).$$

We use the same letter here to denote the two-sort species $\mathcal{A} = \mathcal{A}(X, Y)$ of rooted trees with internal nodes of sort X and leaves of sort Y. Then

$$\mathcal{A}(X, Y) \,=\, X \cdot E\big(\mathcal{A}(X, Y) - X + Y\big), \tag{3}$$

where $\mathcal{A}(X, 0) = X$ is the tree consisting of a single root (of sort X). Indeed—as in the unisort case—a rooted tree of two sorts is obtained by appending a set of rooted trees of two sorts to the root X. There is one caveat: when we append a single node X, it becomes a leaf in the resulting tree and thus its sort changes from X to Y. This explains the additional term $-X + Y$ in equation (3).

In the unisort case, an endofunction is a permutation of rooted trees, $\mathrm{End} = \mathcal{S} \circ \mathcal{A}$. More precisely, functional digraphs of endofunctions are obtained by structuring the roots of a set of rooted trees in an $\mathcal{S}$-structure. Since this construction preserves internal nodes and leaves, the composition $\mathcal{S} \circ \mathcal{A}(X, Y)$ yields functional digraphs of two sorts where internal nodes have sort X and leaves have sort Y. The same holds for the species of Table 1 if we replace $\mathcal{S}$ with E (forests, that is, acyclic functional digraphs), $\mathcal{C}$ (connected functional digraphs), and Der (derangements, that is, functional digraphs with no fixed points). This brings us to the following definition of R-recurrent functional digraphs of two sorts (see also Figure 4).

Definition 5.1. Let R be a unisort species. The species $\Psi_R = \Psi_R(X, Y)$ of *R-recurrent functional digraphs of two sorts* is defined by

$$\Psi_R(X, Y) = R \circ \mathcal{A}(X, Y). \tag{4}$$

In other words, a Ψ_R-structure is obtained by putting an R-structure on the roots of a set of rooted trees of two sorts; in this context, the R-structure is called the *recurrent part* and its points are said to be *recurrent*. Some significant choices of R are listed in Table 2. Note also that $\Psi_X(X, Y) = \mathcal{A}(X, Y)$, which gives

$$\Psi_R = R \circ \Psi_X. \tag{5}$$

The two-sort species Ψ_R simultaneously encompasses (subspecies of) Cayley permutations and endofunctions, allowing us to enumerate functional digraphs of both types at once. From an enumerative standpoint, using different sorts for internal nodes and leaves is equivalent to labeling the internal nodes with the smallest elements of the underlying total order (and the remaining elements with the largest), which characterizes functional digraphs of Cayley permutations by Lemma 4.1. On the other hand, using the same sort $X = Y$ for internal nodes and leaves yields all functional digraphs, that is, all endofunctions. We shall make this observation more formal, but first a couple of definitions.

Let $F = F(X, Y)$ be a two-sort species. The unisort species $F(X, X)$ is obtained by identifying the sorts X and Y. Note that

$$F(X, X)[U] = \bigcup_{U_1 \sqcup U_2 = U} F[U_1, U_2] \tag{6}$$

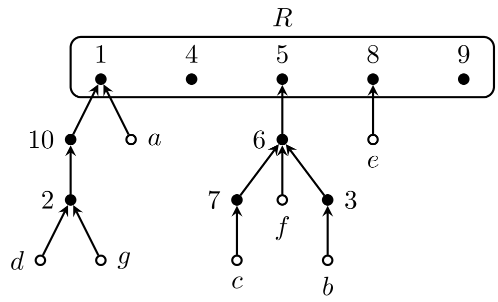


Figure 4: An R-recurrent functional digraph of two sorts. Nodes of sort X (internal nodes) are black and labeled by numbers; nodes of sort Y (leaves) are white and labeled by letters.

R	$\Psi_R(X,Y)$-structures	Endofunctions	Cayley permutations
$\mathcal{S}$	All structures	End	Cay
X	Trees	End_X	Cay_X
E	Forests	End_E	Cay_E
$\mathcal{C}$	Connected	$\mathrm{End}_{\mathcal{C}}$	$\mathrm{Cay}_{\mathcal{C}}$
Der	Derangements	$\mathrm{End}_{\mathrm{Der}}$	$\mathrm{Cay}_{\mathrm{Der}}$

Table 2: Unisort species of endofunctions and Cayley permutations for different choices of the recurrent structure R in $\Psi_R(X,Y)$.

and

$$|F(X,X)[n]| = \sum_{i+j=n} \binom{n}{i} |F[i,j]|. \tag{7}$$

Assuming $\mathbb{L}$-species, we also let $\widehat{F}(Z)$ be the unisort species whose structures are

$$\widehat{F}[\ell] = \bigcup_{\ell_1 \oplus \ell_2 = \ell} F[\ell_1, \ell_2]. \tag{8}$$

That is, an $\widehat{F}$-structure on ℓ is simply an $F(X,Y)$-structure on $[\ell_1, \ell_2]$, where $\ell = \ell_1 \oplus \ell_2$. As a result,

$$|\widehat{F}[n]| = \sum_{i+j=n} |F[i,j]|. \tag{9}$$

Now, a $\Psi_{\mathcal{S}}$-structure where internal nodes and leaves have the same sort is simply a functional digraph of one sort, and we have the combinatorial equality

$$\Psi_{\mathcal{S}}(X,X) = \mathrm{End}(X).$$

On the other hand, by Lemma 4.1 there is a bijection between $\widehat{\Psi}_{\mathcal{S}}[n]$ and $\mathrm{Cay}[n]$. More precisely, structures in $\Psi_{\mathcal{S}}[i,j]$ correspond to functional digraphs of Cayley permutations on $[n] = [i+j]$ that have i internal nodes and j leaves. Therefore,

$$\widehat{\Psi}_{\mathcal{S}}(X) = \mathrm{Cay}(X).$$

This motivates the following definition.

Definition 5.2. Let R be a unisort species. The species of *R-recurrent endofunctions* and the species of *R-recurrent Cayley permutations* are defined as follows:

$$\begin{aligned}\mathrm{End}_R(X) &= \Psi_R(X,X);\\ \mathrm{Cay}_R(X) &= \widehat{\Psi}_R(X).\end{aligned}$$

Our main goal is to find a formula for $|\Psi_R[i,j]|$. Then, by choosing R as one of $\mathcal{S}$, X, E, $\mathcal{C}$, or Der, and letting $F = \Psi_R$ in equations (7) and (9), we obtain counting formulas for the endofunctions and Cayley permutations of Table 2. Comparing (7) and (9) also shows that the number of endofunctions with i internal nodes and j leaves is simply obtained by multiplying the corresponding number of Cayley permutations by the binomial coefficient $\binom{i+j}{i}$; this additional term accounts for the possibility of labeling internal nodes and leaves of endofunctions without the constraint of using the smallest i labels for the internal nodes.

In Section 4, we observed that neither the equation $\mathrm{End} = \mathcal{S} \circ \mathcal{A}$ nor the equation $\mathcal{S} = E \cdot \mathrm{Der}$ readily extends to Cayley permutations. Both constructions are, however, valid for the two-sort species $\Psi_{\mathcal{S}}$. Indeed, $\Psi_{\mathcal{S}} = \mathcal{S} \circ \Psi_X$ by definition, and using $\mathcal{S} = E \cdot \mathrm{Der}$ it follows that

$$\begin{aligned}\Psi_{\mathcal{S}} &= (E \cdot \mathrm{Der}) \circ \Psi_X\\ &= E(\Psi_X) \cdot \mathrm{Der}(\Psi_X) = \Psi_E \cdot \Psi_{\mathrm{Der}}.\end{aligned}$$

Therefore, in the two-sort sense, functional digraphs are permutations of rooted trees. They are also products of sets of rooted trees (forests) and functional digraphs with no fixed points.

6 A differential equation

Equation (4) defines the species $\Psi_R(X,Y)$ of R-recurrent functional digraphs as the composition of R with rooted trees of two sorts. As anticipated, we wish to derive a formula for the coefficients $|\Psi_R[i,j]|$. Although the combinatorial construction manifested in (4) is transparent, it does not provide a straightforward path to derive the desired coefficients. Below, we give an alternative recursive construction of Ψ_R-structures that leads to a more convenient differential equation.

To build any Ψ_R-structure, we first build its recurrent part, an R-structure of recurrent points of sort X. The nonrecurrent points are located on rooted trees of two sorts, each of which hangs from a recurrent point. Such a rooted tree can be obtained recursively. Starting from its root, one repeatedly appends to a specified internal node a path of internal nodes that ends with a leaf; that is, a structure of $L(X) \cdot Y$, which we call a *branch*. This recursive construction of Ψ_R-structures is described by the differential equation

$$\frac{\partial}{\partial Y}\Psi_R(X,Y) = L(X) \cdot \Psi_R^{\bullet_X}(X,Y), \qquad \Psi_R(X,0) = R(X). \tag{10}$$

The initial condition determines the recurrent structure. On the left-hand side of the main equation, we have a structure of $(\partial/\partial Y)\Psi_R$; that is, a Ψ_R-structure with an additional leaf. Equivalently, consider the maximal path starting from the additional

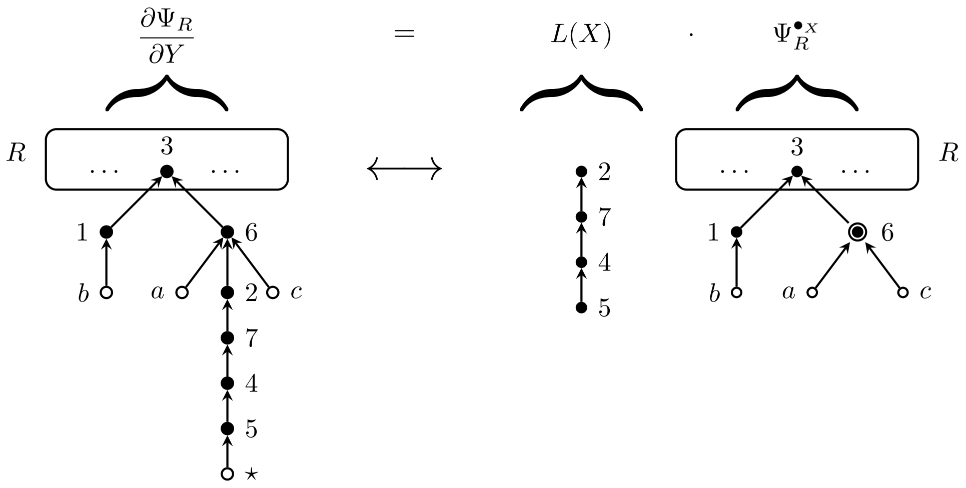


Figure 5: The construction manifested in equation (10).

leaf (excluded) that consists of internal nodes with indegree at most one. It is an $L(X)$-structure, which is appended to an internal node of a Ψ_R-structure. By pointing to that internal node we obtain a $\Psi_R^{\bullet_X}$-structure.

Assuming $\mathbb{L}$-species, we can rewrite equation (10) using the integral operator as

$$\Psi_R(X,Y) = R(X) + \int L(X) \cdot \Psi_R^{\bullet_X}(X,Y) dY.$$

Figure 5 illustrates this construction. We give a formal proof of equation (10) next.

Theorem 6.1. *The two-sort species $\Psi_R(X,Y)$ of R-recurrent functional digraphs satisfies equation* (10).

Proof. Recall from equations (3) and (5) that $\Psi_R = R \circ \Psi_X$ and $\Psi_X = X \cdot E\big(\Psi_X - X + Y\big)$, where $\Psi_X = \mathcal{A}(X,Y)$ and $\Psi_X(X,0) = X$. That Ψ_R satisfies the initial condition of (10) is easy to see:

$$\Psi_R(X,0) = R \circ \Psi_X(X,0) = R(X).$$

Next, we calculate the partial derivatives of Ψ_X. On differentiating the equation $\Psi_X = X \cdot E\big(\Psi_X - X + Y\big)$ we get

$$\frac{\partial \Psi_X}{\partial X} = \Big(\frac{\partial \Psi_X}{\partial X} + X - 1\Big) \cdot E(\Psi_X - X + Y)$$

or, equivalently,

$$X \cdot (1 - \Psi_X) \cdot \frac{\partial \Psi_X}{\partial X} = (1 - X) \cdot \Psi_X.$$

Since $L \cdot (1 - X) = 1$ and $L_+ = XL$ we can rewrite this as

$$L_+(X) \cdot \frac{\partial \Psi_X}{\partial X} = L_+\big(\Psi_X\big).$$

By a similar calculation we find that

$$\frac{\partial \Psi_X}{\partial Y} = L_+\big(\Psi_X\big)$$

and putting these last two equations together we arrive at

$$\frac{\partial \Psi_X}{\partial Y} = L_+(X) \cdot \frac{\partial \Psi_X}{\partial X}.$$

Finally, we differentiate $\Psi_R = R \circ \Psi_X$ to obtain

$$\begin{aligned}\frac{\partial \Psi_R}{\partial Y} &= \frac{\partial \Psi_X}{\partial Y} \cdot R'(\Psi_X) \\ &= L_+(X) \cdot \frac{\partial \Psi_X}{\partial X} \cdot R'(\Psi_X) \\ &= L_+(X) \cdot \frac{\partial \Psi_R}{\partial X} = L(X) \cdot \Psi_R^{\bullet_X},\end{aligned}$$

which concludes the proof. $\square$

Equation (10) immediately leads to a recursive formula for $|\Psi_R[i,j]|$.

Lemma 6.2. *We have*

$$|\Psi_R[i,j]| = \begin{cases} |R[i]| & j = 0; \\ 0 & j \geq 1, i = 0; \\ i \cdot \big(|\Psi_R[i,j-1]| + |\Psi_R[i-1,j]|\big) & i,j \geq 1. \end{cases}$$

Proof. It is easy to see that the first two identities hold. Next, we rewrite equation (10) of Theorem 6.1 as follows:

$$\begin{aligned}\frac{\partial \Psi_R}{\partial Y} = L \cdot X \cdot \frac{\partial \Psi_R}{\partial X} \quad &\Longleftrightarrow \quad \frac{\partial \Psi_R}{\partial Y} \cdot (1 - X) = X \frac{\partial \Psi_R}{\partial X} \\ &\Longleftrightarrow \quad \frac{\partial \Psi_R}{\partial Y} = X \cdot \left(\frac{\partial \Psi_R}{\partial X} + \frac{\partial \Psi_R}{\partial Y} \right).\end{aligned}$$

Identifying coefficients of the last equation gives the third identity. $\square$

Next, we use Lemma 6.2 to obtain the anticipated formula for $|\Psi_R[i,j]|$. It involves the r-Stirling numbers of the second kind, defined as follows. Let $n, m, r \geq 0$. The *r-Stirling number of the second kind* $\left\{ {n \atop m} \right\}_r$ is the number of set partitions of $[n]$ with m blocks such that the numbers $1, 2, \ldots, r$ are contained in distinct blocks. Broder [4, Theorem 2] showed that these numbers obey a recurrence analogous to the one satisfied by the classic Stirling numbers:

$$\left\{ {n \atop m} \right\}_r = m \left\{ {n-1 \atop m} \right\}_r + \left\{ {n-1 \atop m-1} \right\}_r, \quad r < n, \tag{11}$$

with initial conditions $\left\{ {n \atop m} \right\}_n = \delta_{m,n}$, and $\left\{ {n \atop m} \right\}_r = 0$ if $r > n$. We define[1]

$$\left\{ \frac{n}{m} \right\}_r := \left\{ {n \atop m} \right\}_r - \left\{ {n \atop m} \right\}_{r+1}$$

[1]We choose this notation mainly to remind us of the r-Stirling numbers. A fraction also makes some logical sense in that a partition is a *division* of an n-element set into m blocks. The fraction bar may alternatively symbolize the minus sign on the right-hand side. Finally, overloading the fraction notation has historical precedents in the Legendre and Jacobi symbols.

This difference of r-Stirling numbers counts set partitions of $[n]$ with m blocks such that the numbers $1, 2, \ldots, r$ are contained in distinct blocks, but $1, 2, \ldots, r, r+1$ are not. In other words, r is maximal such that $1, 2, \ldots, r$ appear in distinct blocks. From (11) it follows that

$$\left\{ \frac{n}{m} \right\}_r = m \left\{ \frac{n-1}{m} \right\}_r + \left\{ \frac{n-1}{m-1} \right\}_r, \qquad r < n \tag{12}$$

except $\left\{ \frac{n}{n} \right\}_{n-1} = 0$, with initial conditions $\left\{ \frac{n}{m} \right\}_n = \delta_{m,n}$, and $\left\{ \frac{n}{m} \right\}_r = 0$ if $r > n$. In Table 3 we give the numbers $\left\{ \frac{n}{m} \right\}_r$ for $r = 1, 2, 3, 4$, $n = 1, 2, \ldots, 8$ and $m = 1, 2, \ldots, 8$.

$r = 1$							
1							
1	0						
1	1	0					
1	3	1	0				
1	7	6	1	0			
1	15	25	10	1	0		
1	31	90	65	15	1	0	
1	63	301	350	140	21	1	0

$r = 2$							
0							
0	1						
0	2	0					
0	4	2	0				
0	8	10	2	0			
0	16	38	18	2	0		
0	32	130	110	28	2	0	
0	64	422	570	250	40	2	0

$r = 3$							
0							
0	0						
0	0	1					
0	0	3	0				
0	0	9	3	0			
0	0	27	21	3	0		
0	0	81	111	36	3	0	
0	0	243	525	291	54	3	0

$r = 4$							
0							
0	0						
0	0	0					
0	0	0	1				
0	0	0	4	0			
0	0	0	16	4	0		
0	0	0	64	36	4	0	
0	0	0	256	244	60	4	0

Table 3: Triangles of the numbers $\left\{ \frac{n}{k} \right\}_r$, for $r = 1, 2, 3, 4$.

Theorem 6.3. *For $i, j, r \geq 0$, we have*

$$|\Psi_{R_r}[i,j]| = \frac{i!|R[r]|}{r!} \left\{ \frac{i+j}{i} \right\}_r.$$

Proof. We shall prove that the term on the right-hand side satisfies the recursion of Lemma 6.2 with $R = R_r$. The initial conditions are easy to verify, so let us assume that $i, j \geq 1$. We need to prove that

$$\frac{i!|R[r]|}{r!} \left\{ \frac{i+j}{i} \right\}_r = i \cdot \left(\frac{(i-1)!|R[r]|}{r!} \left\{ \frac{i+j-1}{i-1} \right\}_r + \frac{i!|R[r]|}{r!} \left\{ \frac{i+j-1}{i} \right\}_r \right).$$

But this is equivalent to $\left\{ \frac{i+j}{i} \right\}_r = \left\{ \frac{i+j-1}{i-1} \right\}_r + i \left\{ \frac{i+j-1}{i} \right\}_r$, which follows from (12). $\square$

Using Theorem 6.3 and summing over r we arrive at the following corollary.

Corollary 6.4. *For $i, j \geq 0$, we have*

$$|\Psi_R[i,j]| = i! \sum_{r=0}^{i} \frac{|R[r]|}{r!} \left\{ \frac{i+j}{i} \right\}_r. \tag{13}$$

Recall from Definition 5.2 that $\mathrm{End}_R(X) = \Psi_R(X, X)$ and $\mathrm{Cay}_R(X) = \widehat{\Psi}_R(X)$. The number of R-recurrent endofunctions and Cayley permutations on $[n] = [i + j]$ are now obtained by combining equation (13) with identities (7) and (9).

Corollary 6.5. *For $n \geq 0$, we have*

$$|\mathrm{End}_R[n]| = \sum_{r=0}^{n} \frac{|R[r]|}{r!} \sum_{i=0}^{n} (n)_i \left\{ {n \atop i} \right\}_r$$

and

$$|\mathrm{Cay}_R[n]| = \sum_{r=0}^{n} \frac{|R[r]|}{r!} \sum_{i=0}^{n} i! \left\{ {n \atop i} \right\}_r$$

The expression $(n)_i$, above, is the *falling factorial*; it is defined by

$$(n)_i = n(n-1)(n-2)\cdots(n-i+1) = \binom{n}{i} i!.$$

Letting $R = E_r$ in Corollary 6.5, we obtain two identities for the number of forests of r rooted trees of endofunctions and Cayley permutations:

$$|\mathrm{End}_{E_r}[n]| = \sum_{i=0}^{n} \frac{(n)_i}{r!} \left\{ {n \atop i} \right\}_r;$$

$$|\mathrm{Cay}_{E_r}[n]| = \sum_{i=0}^{n} \frac{i!}{r!} \left\{ {n \atop i} \right\}_r.$$

These allow us to rewrite the formulas of Corollary 6.5 as

$$|\mathrm{End}_R[n]| = \sum_{r=0}^{n} |R[r]| \cdot |\mathrm{End}_{E_r}[n]|; \tag{14}$$

$$|\mathrm{Cay}_R[n]| = \sum_{r=0}^{n} |R[r]| \cdot |\mathrm{Cay}_{E_r}[n]|. \tag{15}$$

The combinatorics underlying equations (14) and (15) is clear: An End_R-structure is obtained by putting an R-structure over a set of rooted trees. The same holds for Cay_R-structures. Furthermore, a forest of endofunctions is a set of rooted trees: $\mathrm{End}_E = E \circ \mathrm{End}_X$ and $\mathrm{End}_{E_r} = E_r \circ \mathrm{End}_X = \frac{1}{r!}(\mathrm{End}_X)^r$. Thus, hidden underneath equation (14) we find the species composition $\mathrm{End}_R = R \circ \mathrm{End}_X$, which is is equation (4) when $Y = X$. On the other hand, $\mathrm{Cay}_E \neq E \circ \mathrm{Cay}_X$.

Alternatively, letting $R = \mathcal{S}_r$ we obtain

$$|\mathrm{End}_R[n]| = \sum_{r=0}^{n} \frac{|R[r]|}{r!} |\mathrm{End}_{\mathcal{S}_r}[n]|;$$

$$|\mathrm{Cay}_R[n]| = \sum_{r=0}^{n} \frac{|R[r]|}{r!} |\mathrm{Cay}_{\mathcal{S}_r}[n]|,$$

where $|\mathrm{End}_{\mathcal{S}_r}[n]|$ and $|\mathrm{Cay}_{\mathcal{S}_r}[n]|$ are simply the number of endofunctions and Cayley permutations with r recurrent points.

By substituting the number $|R[r]|$ of R-structures of size r into Theorem 6.3, we obtain counting formulas for the two-sort version of the species of Table 2. Using Corollary 6.5 we also obtain identities for the corresponding unisort structures listed in Table 4. A noteworthy example is a formula for the number of Cayley-derangements, the main result of this paper:

Theorem 6.6. *The number of fixed-point-free Cayley permutations of* $[n]$ *is*

$$|\mathrm{Cay}_{\mathrm{Der}}[n]| \;=\; \sum_{r=0}^{n} \frac{|\mathrm{Der}[r]|}{r!} \sum_{i=0}^{n} i! \left\{ \frac{n}{i} \right\}_r .$$

The first few numbers of this sequence are

$$1, 0, 1, 4, 25, 184, 1617, 16492, 191721, 2503040, 36267393, 577560596.$$

At the time of writing, this sequence is not in the OEIS [14].

We highlight three additional noteworthy instances of Corollary 6.5 below.

Proposition 6.7. *The number of Cayley permutations of* $[n]$ *whose functional digraph is a tree equals the number of all Cayley permutations of* $[n-1]$*. That is,*

$$|\mathrm{Cay}_X[n]| = |\mathrm{Cay}[n-1]|.$$

The number of Cayley permutations of $[n]$ *whose functional digraph is a forest is*

$$|\mathrm{Cay}_E[n]| = \sum_{r=0}^{n} \frac{1}{r!} \sum_{i=0}^{n} i! \left\{ \frac{n}{i} \right\}_r .$$

The number of Cayley permutations of $[n]$ *whose functional digraph is connected is*

$$|\mathrm{Cay}_C[n]| = \sum_{r=0}^{n} \frac{1}{r} \sum_{i=0}^{n} i! \left\{ \frac{n}{i} \right\}_r .$$

Proof. Setting $R = X$ in Corollary 6.5 we get the first identity:

$$|\mathrm{Cay}_X[n]| = \sum_{r=0}^{n} \frac{|X[r]|}{r!} \sum_{i=0}^{n} i! \left\{ \frac{n}{i} \right\}_r = \sum_{i=0}^{n} i! \left\{ \frac{n}{i} \right\}_1 = |\mathrm{Cay}[n-1]|,$$

where in the last step $\left\{ \frac{n}{i} \right\}_1$ equals the $(n-1, i)$th Stirling number of the second kind. The second and third identities are immediate consequences of Corollary 6.5. $\square$

In the next section we generalize Joyal's bijection between doubly-rooted trees and endofunctions to the two-sort setting.

7 A two-sort species extension of Joyal's bijection

Recall from Section 3 Joyal's bijection between doubly-rooted trees and functional digraphs of endofunctions. It identifies the spine of a doubly-rooted tree with a

	R	$\lvert\Psi_{R_r}[i,j]\rvert$	$\lvert\mathrm{End}_R[n]\rvert$	$\lvert\mathrm{Cay}_R[n]\rvert$
All structures	$\mathcal{S}$	$i!\left\{{i+j \atop i}\right\}_r$	n^n	$\lvert\mathrm{Cay}[n]\rvert$
Trees	X	$i!\delta_{r,1}\left\{{i+j \atop i}\right\}_r$	n^{n-1}	$\lvert\mathrm{Cay}[n-1]\rvert$
Forests	E	$\frac{i!}{r!}\left\{{i+j \atop i}\right\}_r$	n^{n-2}	Proposition 6.7
Connected	$\mathcal{C}$	$\frac{i!}{r}\left\{{i+j \atop i}\right\}_r$	A001865	Proposition 6.7
Derangements	Der	$\lvert\mathrm{Der}[r]\rvert\frac{i!}{r!}\left\{{i+j \atop i}\right\}_r$	$(n-1)^n$	Theorem 6.6

Table 4: Number of R-recurrent functional digraphs with r recurrent points, i internal nodes and j leaves.

nonempty permutation, leaving the rooted trees that hang from the spine untouched. This bijection implies the unisort $\mathbb{L}$-species identity

$$\mathcal{A}^\bullet = \mathrm{End}_+. \tag{16}$$

We shall extend Joyal's construction to the two-sort case by revealing a link between functional digraphs $\Psi_{\mathcal{S}}$ and rooted trees $\Psi_X = \mathcal{A}(X,Y)$. In the proof of Proposition 6.1, we showed that

$$\frac{\partial \Psi_X}{\partial Y} = L_+\big(\Psi_X\big).$$

Assuming $\mathbb{L}$-species, we have $L = \mathcal{S}$ and hence $L_+(\mathcal{A}) = \mathcal{S}_+(\mathcal{A}) = \Psi_{\mathcal{S}_+}$. That is,

$$\frac{\partial \Psi_X}{\partial Y} = \Psi_{\mathcal{S}_+}. \tag{17}$$

We have found a combinatorial proof of (17) in the spirit of Joyal's proof of (16). On the left-hand side, we have a rooted tree of two sorts with an additional leaf. Define its *spine* as the path from this leaf to the root. Ignoring the initial leaf, this path is a nonempty linear order of (internal) nodes of sort X, where a rooted tree of two-sorts hangs from each of its nodes. Identifying the spine with a permutation in the usual way yields a nonempty permutation of rooted trees, that is, a structure of $\mathcal{S}_+ \circ \Psi_X = \Psi_{\mathcal{S}_+}$, the species on the right-hand side. In other words, the additional leaf of a structure of $(\partial/\partial Y)\Psi_X$ plays the role of *tail*, while the root plays the role of *head*. A two-sort vertebrate is born in the process, and by reading a permutation off its spine we obtain a permutation of rooted trees of two sorts; that is, a structure of $\Psi_{\mathcal{S}_+}$. This construction—an instance of which can be found in Figure 6—is easily invertible, and thus it yields a combinatorial proof of equation (17).

8 Unisort species of Cayley functional digraphs

In this section, we build on the enumerative results obtained for Ψ_R in the two-sort case to obtain species equations for its unisort counterpart Cay_R. These rely on the *ordinal product* of $\mathbb{L}$-species, defined as

$$(F \odot G)[\ell] = \sum_{\ell=\ell_1\oplus\ell_2} F[\ell_1] \times G[\ell_2].$$

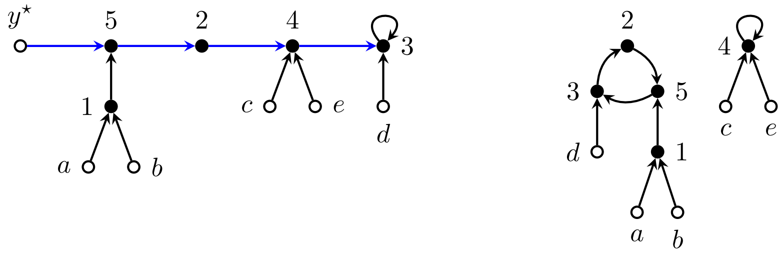


Figure 6: A rooted tree of two sorts with an additional leaf, denoted $y^\star$, and the corresponding functional digraph under the extension of Joyal's construction. The spine of the tree, in blue, is the linear order 5243; it is mapped to the recurrent part of the functional digraph on the right.

An $F \odot G$- structure is thus obtained by putting an F-structure on the initial segment ℓ_1 and a G-structure on the terminal segment ℓ_2.

We shall use $\mathbb{L}$-species to show that

$$\mathrm{Cay}_R = \frac{1}{2}\sum_{r\geq 0} R_r \odot (1 + E^r \mathrm{Bal}^r).$$

This specializes to give expressions for the unisort species of Cayley-derangements and the other species in Table 2:

$$\mathrm{Cay}_{\mathrm{Der}} = \frac{1}{2}\sum_{r\geq 0} \mathrm{Der}_r \odot (1 + E^r \mathrm{Bal}^r)\,. \tag{18}$$

Let us start with a simple lemma.

Lemma 8.1. *For $r \geq 1$, we have*

$$\sum_{i\geq 0} i! \cdot \left\{\frac{n+r+1}{i}\right\}_r = r! \cdot r \cdot \left|\big(E^r \mathrm{Bal}^{r+1}\big)[n]\right|.$$

Proof. The left-hand side counts ballots of $[n{+}r{+}1]$ with i blocks, where r is maximal such that $1, \ldots, r$ are contained in distinct blocks. Any such ballot is structured

$$\beta_1 \; \{u_1, \ldots\} \; \beta_2 \; \{u_2, \ldots\} \; \ldots \; \beta_r \; \{u_r, \ldots\} \; \beta_{r+1},$$

where $u_1 u_2 \ldots u_r$ is a permutation of $1, \ldots, r$ and the β_i's are ballots filling the space between the blocks containing the u_i's. The permutation of $1, \ldots, r$ gives the $r!$ term on the right-hand side. Since $r{+}1$ belongs to one of the blocks containing the u_i's, we have r choices to arrange it. Finally, the remaining elements, $r{+}2, r{+}3, \ldots, n{+}r{+}1$, are arranged in a structure of $E^r \cdot \mathrm{Bal}^{r+1}$ (of size n): elements in the same block as some u_j determine an E^r-structure, while the other elements determine the ballots $\beta_1, \ldots, \beta_{r+1}$, that is, a structure of Bal^{r+1}. $\square$

Theorem 8.2. *We have the $\mathbb{L}$-species identity*

$$\mathrm{Cay}_R = R + \sum_{r\geq 1} (R_r)^\bullet \odot \left(\int E^r \mathrm{Bal}^{r+1}\right).$$

Proof. Let F denote the species on the right-hand side. To prove the desired $\mathbb{L}$-species identity, it suffices to show that $|F[n]| = |\mathrm{Cay}_R[n]|$ for $n \geq 0$. Expanding the definition of F yields

$$\begin{aligned}
|F[n]| &= |R[n]| + \sum_{r=1}^{n} |R^\bullet[r]| \cdot \left| \left(\int E^r \mathrm{Bal}^{r+1} \right) [n-r] \right| \\
&= |R[n]| + \sum_{r=1}^{n-1} r|R[r]| \cdot \left| \left(E^r \mathrm{Bal}^{r+1} \right) [n-r-1] \right| \cdot \frac{r!}{r!} \\
&= |R[n]| + \sum_{r=1}^{n-1} \frac{|R[r]|}{r!} \sum_{i \geq 0} i! \left\{ {n \atop i} \right\}_r && \text{(by Lemma 8.1)} \\
&= \sum_{r=1}^{n} \frac{|R[r]|}{r!} \sum_{i \geq 0} i! \left\{ {n \atop i} \right\}_r \\
&= |\mathrm{Cay}_R[n]| && \text{(by Corollary 6.5)}
\end{aligned}$$

where in the penultimate equality we used that the term corresponding to $r = n$ in the summation equals

$$\frac{|R[n]|}{n!} \sum_{i \geq 0} i! \left\{ {n \atop i} \right\}_n = \frac{n!|R[n]|}{n!} \left\{ {n \atop n} \right\}_n = |R[n]|.$$

This concludes the proof. □

An arguably simpler species equation for Cay_R is provided in the next result.

Corollary 8.3. *We have*

$$\mathrm{Cay}_R = \frac{1}{2} \sum_{r \geq 0} R_r \odot (1 + E^r \mathrm{Bal}^r)\,.$$

Proof. The $\mathbb{L}$-species identity

$$\int E^r \mathrm{Bal}^{r+1} = \frac{1}{2r} (E^r \mathrm{Bal}^r - 1). \tag{19}$$

is easily verified using differentiation and the species version of the fundamental theorem of calculus. Next,

$$\begin{aligned}
\mathrm{Cay}_R &= R + \sum_{r \geq 1} (R_r)^\bullet \odot \left(\int E^r \mathrm{Bal}^{r+1} \right) && \text{(by Theorem 8.2)} \\
&= \sum_{r \geq 0} R_r + (R_r)^\bullet \odot \left(\int E^r \mathrm{Bal}^{r+1} \right) && \text{(since } R = \textstyle\sum_{r \geq 0} R_r \text{ \& } R_0^\bullet = \emptyset) \\
&= \sum_{r \geq 0} R_r + r R_r \odot \frac{1}{2r} (E^r \mathrm{Bal}^r - 1) && \text{(by equation (19))} \\
&= \frac{1}{2} \sum_{r \geq 0} R_r \odot (1 + E^r \mathrm{Bal}^r)\,,
\end{aligned}$$

where the last identity follows by $R_r = R_r \odot 1$ and linearity of $\odot$. □

Finding a combinatorial proof of this corollary remains an open problem.

9 Statistics and asymptotics

Theorem 6.3 provides a closed formula for the number of Ψ_R-structures with i internal nodes, among which r are recurrent, and j leaves. By suitably summing over r, i and j, we obtain the distribution of each of these statistics over R-recurrent functional digraphs. For instance, the number of R-recurrent functional digraphs with n nodes and r recurrent points is

$$\sum_{i+j=n} |\Psi_{R_r}[i,j]| = \frac{|R[r]|}{r!} \sum_{i=0}^{n} i! \left\{ \frac{n}{i} \right\}_r .$$

The distributions of internal nodes and leaves are obtained similarly. The total number of recurrent points, internal nodes and leaves can be computed in the same fashion by multiplying the term inside the summation by r, i or j. An example is the total number of recurrent points over Ψ_R-structures on $[i,j]$:

$$\sum_{r=0}^{i} |\Psi_{R_r}[i,j]| \cdot r = \sum_{r=0}^{i} \frac{|R[r]|}{r!} i! \left\{ \frac{i+j}{i} \right\}_r \cdot r.$$

Multiplying by r has the same effect as pointing to the R-structure. This gives us an elegant species expression for the total number of recurrent points:

$$\Psi_{R^\bullet}(X,Y), \quad \text{where } |\Psi_{R^\bullet}[i,j]| = \sum_{r=0}^{i} |\Psi_{R_r}[i,j]| \cdot r. \tag{20}$$

If two species F and G are related by $F = E(G)$, with $G(0) = 0$, then G is said to be the species of *connected* F-structures. In this sense, an F-structure is a set of G-structures each of which is called a *connected component*. We are by now familiar with the species of connected permutations, that is, cycles $\mathcal{C}$ where $\mathcal{S} = E(\mathcal{C})$. Not every species admits a notion of connected structure in the realm of $\mathbb{B}$-species, but it is always possible [2, Proposition 18], under the assumption $F(0) = 1$, to construct a virtual species G such that $F = E(G)$. The species G is uniquely determined and is called the *combinatorial logarithm* of F, denoted by $G = \log F$.

For any unisort species R such that $R(0) = 1$, we define the two-sort species of *connected R-recurrent functional digraphs* by letting the recurrent part be equal to the logarithm of R:

$$\Psi_{\log R}(X,Y).$$

Pointing to the E-structure in $E(\log R) = R$ has the effect of multiplying by the number of connected components. Therefore, the total number of connected components over Ψ_R-structures is obtained as $\Psi_{E^\bullet(\log R)}$. Note that $E^\bullet(X) = X \cdot E(X)$, from which

$$E^\bullet(\log R) = R \cdot \log R \quad \text{and} \quad \Psi_{E^\bullet(\log R)} = \Psi_{R \log R} \tag{21}$$

follows. Finally, Theorem 6.3 gives the total number of connected components over Ψ_R-structures on $[i,j]$:

$$\sum_{r=0}^{i} |\Psi_{R \log R}[i,j]| = \sum_{r=0}^{i} \frac{|(R \log R)[r]|}{r!} i! \left\{ \frac{i+j}{i} \right\}_r . \tag{22}$$

To illustrate the techniques introduced in this section, we will show that permutations, Cayley permutations and endofunctions share the following property:

The total number of recurrent points over connected structures of a given size equals the total number of structures of that size.

First, we consider permutations. There are $(n-1)!$ connected permutations (cycles) of size $n \geq 1$. Since every point in a permutation is recurrent, the total number of recurrent points over connected structures is $(n-1)! \cdot n = n! = |\mathcal{S}[n]|$. To deal with Cayley permutations and endofunctions we use the species of connected functional digraphs, $\Psi_{\mathcal{C}}$. Since $\mathcal{C}' = L$, we find that $\mathcal{C}^{\bullet} = X \cdot \mathcal{C}' = X \cdot L = L_{+}$. Thus, the total number of recurrent points over $\Psi_{\mathcal{C}}$ is given by

$$|\Psi_{\mathcal{C}^{\bullet}}[i,j]| = |\Psi_{L_+}[i,j]| = |\Psi_{\mathcal{S}_+}[i,j]|.$$

Finally, we translate the previous identity to functional digraphs of Cayley permutations and endofunctions to establish the desired property:

$$\begin{aligned} |\mathrm{End}_{\mathcal{C}^{\bullet}}[n]| &= |\mathrm{End}_{\mathcal{S}}[n]| = |\mathrm{End}[n]|; \\ |\mathrm{Cay}_{\mathcal{C}^{\bullet}}[n]| &= |\mathrm{Cay}_{\mathcal{S}}[n]| = |\mathrm{Cay}[n]|. \end{aligned}$$

The crucial property that makes the above derivation work is $\mathcal{C}^{\bullet} \equiv \mathcal{S}_{+}$. In other words, the species $F = \mathcal{S}_+$ satisfies $(\log F)^{\bullet} \equiv F_{+}$. The solutions to this differential equation are precisely the *k-colored permutations*. That is, $F \equiv \mathcal{S}(kX)$ for some positive integer k.

Next, we consider the total number of cycles/connected components over structures in $\Psi_{\mathcal{S}}[i,j]$. It is simply obtained by letting $R = \mathcal{S}$ in equation (21), where $E^{\bullet}(\log \mathcal{S}) = E^{\bullet}(\mathcal{C})$ gives the total number of cycles over all permutations. It is well known that the average number of cycles over permutations of $[n]$ is the nth harmonic number $H_n = \sum_{k=0}^{n} 1/k$. To see this, note that $E^{\bullet}(\mathcal{C}) = \mathcal{S} \cdot \log(\mathcal{S}) = \mathcal{S} \cdot \mathcal{C}$ and

$$E^{\bullet}(\mathcal{C})(x) = \big(\mathcal{S} \cdot \mathcal{C}\big)(x) = \frac{1}{1-x} \cdot \log\left(\frac{1}{1-x}\right).$$

Therefore,

$$\big|E^{\bullet}(\mathcal{C})[n]\big| = \big|(\mathcal{S} \cdot \mathcal{C})[n]\big| = \sum_{k=1}^{n} \binom{n}{k} (n-k)!(k-1)! = n! H_n.$$

Now, by identity (22), the total number of cycles over $\Psi_{\mathcal{S}}$-structures on $[i,j]$ is

$$|\Psi_{\mathcal{S}\log\mathcal{S}}[i,j]| = i! \sum_{r=0}^{i} \frac{r! H_r}{r!} \left\{ \frac{i+j}{i} \right\}_r = i! \sum_{r=0}^{i} H_r \left\{ \frac{i+j}{i} \right\}_r.$$

Corollary 6.5 yields the total number of cycles over endofunctions (see A190314 [14]) and Cayley permutations of size n:

$$\begin{aligned} |\mathrm{End}_{E^{\bullet}(\mathcal{C})}[n]| &= \sum_{i=0}^{n} (n)_i \sum_{r=0}^{i} H_r \left\{ \frac{n}{i} \right\}_r; \\ |\mathrm{Cay}_{E^{\bullet}(\mathcal{C})}[n]| &= \sum_{i=0}^{n} i! \sum_{r=0}^{i} H_r \left\{ \frac{n}{i} \right\}_r. \end{aligned}$$

The first few numbers of the latter sequence are

$$0, 1, 4, 20, 126, 966, 8754, 91686, 1090578, 14528502, 214337874, 3469418646.$$

At the time of writing, this sequence is not in the OEIS [14]. We conclude this section by presenting two open problems of an asymptotic nature, which we will address in future work.

We opened this paper by recalling the answer to the hat-check problem: the probability that a permutation chosen uniformly at random is a derangement converges to $1/e$ as its size grows to infinity. Can we prove that Cayley-derangements have the same asymptotic behavior? That is, can we prove that

$$\lim_{n\to\infty} \frac{|\mathrm{Cay}_{\mathrm{Der}}[n]|}{|\mathrm{Cay}[n]|} = \frac{1}{e}?$$

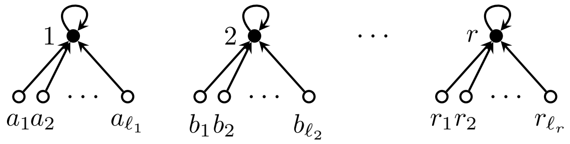
The average number of cycles over $\mathrm{End}[n]$ is asymptotic to $\frac{1}{2}(\log(2n) + \gamma)$, where Euler's constant γ is defined as the limiting difference between the harmonic series and the natural logarithm. What is the corresponding limit for the average number of cycles over Cayley permutations? Is there a generalization of Euler's constant for the average number of cycles over R-recurrent functional digraphs?

10 Generalizations

Equation (10) in Section 6 characterizes the two-sort species Ψ_R. There, linear orders determine the shape of the branches—$L(X) \cdot Y$—that are recursively appended to the recurrent part to obtain Ψ_R-structures. Replacing L with any unisort species T allows us to manipulate the shape of these branches. Define the species $\Psi_{R,T}(X, Y)$ accordingly as the solution to the system of combinatorial equations

$$\frac{\partial}{\partial Y}\Psi_{R,T}(X, Y) = T(X) \cdot \Psi^{\bullet_X}_{R,T}(X, Y), \qquad \Psi_{R,T}(X, 0) = R(X).$$

For instance, if we let $T = 1$ be the empty species, then we obtain structures where all the leaves have distance one from a recurrent point (since each branch $T(X) \cdot Y = Y$ consists of a single leaf). Letting $R = E$, we end up with a "brush-shaped" endofunction, that is, a set of roots (of sort X) to each of which is appended a set of leaves (of sort Y):



Note that

$$\Psi_{E,1}(X, Y) = E\big(X \cdot E(Y)\big).$$

The corresponding endofunctions of $\mathrm{End}_{E,1}$ and Cayley permutations of $\mathrm{Cay}_{E,1}$ are precisely those f such that $f^{(2)} = f$, where $f^{(2)} = f \circ f$ denotes the composition of f

with itself (see also A000248 and A026898 [14]). In other words, $\Psi_{E,1}$ is the species of *idempotent functional digraphs* of two-sorts.

Let us push this construction a bit further. Setting $R = \mathcal{S}$ has the effect of allowing cycles of any length, yielding maps f where $f^{(k)} = f$ for some $k > 1$; in this case,

$$\Psi_{\mathcal{S},1}(X,Y) = \mathcal{S}\big(X \cdot E(Y)\big).$$

In the same spirit, we fix k and let

$$R = E \circ \sum_{i|k} \mathcal{C}_i$$

be the species of permutations whose cycle lengths are divisible by k. Then $\Psi_{R,1}$ yields maps f such that $f^{(k)} = f$. What other choices of R and T lead to interesting examples?

In Section 5, we have defined the two-sort species $\Psi_R = R \circ \mathcal{A}$ whose structures are obtained by arranging a set of rooted trees in an R-structure. We can tweak this definition by replacing rooted trees with any other class of trees. For instance, let $\mathcal{B} = \mathcal{B}(X,Y)$ be the two-sort species of rooted trees where each node has at most two (unordered) children. They are obtained as

$$\mathcal{B}(X,Y) = X \cdot \Big(\big(1 + X + E_2\big) \circ \big(\mathcal{B}(X,Y) - X + Y\big)\Big).$$

Then, $\mathcal{S} \circ \mathcal{B}(X,Y)$ yields functional digraphs where each recurrent point has indegree at most three, and each nonrecurrent point has indegree at most two. In the corresponding endofunctions and Cayley permutations, obtained as usual via (6) and (8), recurrent points have at most three preimages, and nonrecurrent points have at most two preimages (see also A201996 [14]). More generally, rooted trees where each node has at most k children

$$\mathcal{K}(X,Y) = X \cdot \Big(\big(1 + X + E_2 + \cdots + E_k\big) \circ \big(\mathcal{K}(X,Y) - X + Y\big)\Big)$$

determine maps where recurrent points have at most $k+1$ preimages and nonrecurrent points have at most k preimages. What can we say about other classes of trees and choices of R?

References

[1] C. Bean, P. C. Bell, A. Ollson, *The insertion encoding of Cayley permutations*, arXiv: 2505.08480.

[2] F. Bergeron, G. Labelle, P. Leroux, *Combinatorial species and tree-like structures*, Volume 67 of Encyclopedia of Mathematics and its Applications, Cambridge University Press, Cambridge, 1998.

[3] C. W. Borchardt, *Über eine Interpolationsformel für eine Art symmetrischer Functionen und über deren Anwendung*, Mathematische Abhandlungen der Königlichen Akademie der Wissenschaften zu Berlin, 1860.

[4] A. Z. Broder, *The r-Stirling numbers*, Discrete Mathematics, Vol. 49, pp. 241–259, 1984.

[5] A. Cayley, *A theorem on trees*, Quarterly Journal of Pure and Applied Mathematics, Vol. 23, pp. 376–378, 1889.

[6] G. Cerbai, *Sorting Cayley permutations with pattern-avoiding machines*, The Australasian Journal of Combinatorics, Vol. 80, 2021.

[7] G. Cerbai, A. Claesson, *Transport of patterns by Burge transpose*, European Journal of Combinatorics, Vol. 108, 2023.

[8] G. Cerbai, A. Claesson, *Caylerian polynomials*, Discrete Mathematics, Vol. 347, 2024.

[9] G. Cerbai, A. Claesson, *Enumerative aspects of Caylerian polynomials*, arXiv:2411.08426.

[10] G. Cerbai, A. Claesson, D. C. Ernst, H. Golab, *Pattern-avoiding Cayley permutations via combinatorial species*, arXiv:2407.19583, 2024.

[11] A. Claesson, *A species approach to Rota's twelvefold way*, Expositiones Mathematicae, Vol. 38(4), pp. 515–536, 2019.

[12] A. Joyal, *Une théorie combinatoire des séries formelles*, Advances in Mathematics, Vol. 42(1), pp. 1–82, 1981.

[13] P. R. Montmort, *Essay d'analyse sur les jeux de hazard*, Paris: Quillau, p. 189, 1708.

[14] N. J. A. Sloane, *The on-line encyclopedia of integer sequences*, at oeis.org.