

1

Introduction to Species of Structures

1.0. Introduction

This chapter contains the basic concepts of the combinatorial theory of species of structures. It is an indispensable starting point for the developments and applications presented in the subsequent chapters. We begin with some general considerations on the notion of structure, everywhere present in mathematics and theoretical computer science. These preliminary considerations lead us in a natural manner to the fundamental concept of species of structures.

The definition puts the emphasis on the transport of structures along bijections and is due to C. Ehresmann [87], but it is A. Joyal [158] who showed its effectiveness in the combinatorial treatment of formal power series and for the enumeration of labeled structures as well as unlabeled (isomorphism types of) structures.

We introduce in Section 1.2 some of the first power series that can be associated to species: generating series, types generating series, cycle index series. They serve to encode all the information concerning labeled and unlabeled enumeration.

Sections 1.3 and 1.4 form an introduction to *the algebra* of species of structures. Various combinatorial operations on species of structures are used to produce new ones, in general more complex. The operations introduced here are addition, multiplication, substitution, and differentiation of species of structures. They constitute a combinatorial lifting of the corresponding operations on formal power series. The problems of specification, classification, and enumeration of structures are then greatly simplified, using this algebra of species. Also, this approach reveals a remarkable link between the composition of functions and the plethystic substitution of symmetric functions, in the context of Pólya theory.

1.1. Notion of Species of Structures

The concept of structure is fundamental, recurring in all branches of mathematics, as well as in computer science. From an informal point of view, a *structure* s

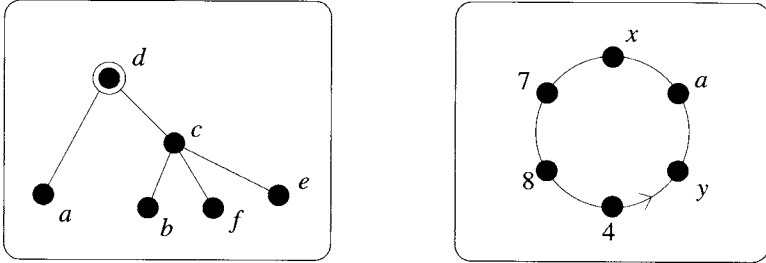


Fig. 1.

is a *construction* γ which one performs on a set U (of data). It consists of a pair

$$s = (\gamma, U). \tag{1}$$

It is customary to say that U is the underlying set of the structure s or even that s is a structure constructed from (or labeled by) the set U . Figure 1 depicts two examples of structures: a rooted tree and an oriented cycle. In a set theoretical fashion, the tree in question can be described as $s = (\gamma, U)$, where

$$\begin{aligned} U &= \{a, b, c, d, e, f\}, \\ \gamma &= (\{d\}, \{\{d, a\}, \{d, c\}, \{c, b\}, \{c, f\}, \{c, e\}\}). \end{aligned} \tag{2}$$

The singleton $\{d\}$ which appears as the first component of γ indicates the *root* of this rooted tree. As for the oriented cycle, it can be put in the form $s = (\gamma, U)$, where

$$\begin{aligned} U &= \{x, 4, y, a, 7, 8\}, \\ \gamma &= (\{4, y\}, (y, a), (a, x), (x, 7), (7, 8), (8, 4)). \end{aligned} \tag{3}$$

The abuse of notation $s = \gamma$, which consists of identifying a structure $s = (\gamma, U)$ with the construction γ , will be used if it does not cause any ambiguity with regard to the nature of the underlying set U . Here is an example which could give rise to such an ambiguity:

$$s = (\gamma, U) \quad \text{where} \quad U = \{c, x, g, h, m, p, q\}, \quad \gamma = \{x, m, p\}. \tag{4}$$

Here the structure is that of a-subset-of-a-set U . The knowledge of γ alone does not enable one to recover the set U .

A traditional approach to the concept of structure consists of generalizing the preceding examples within axiomatic set theory. However, in the present work we adopt a *functorial* approach which puts an emphasis on the *transport of structures* along bijections. Here is an example which illustrates the concept of transport of structures.

Example 1. Consider the rooted tree $s = (\gamma, U)$ of Figure 1 whose underlying set is $U = \{a, b, c, d, e, f\}$. Replace each element of U by those of

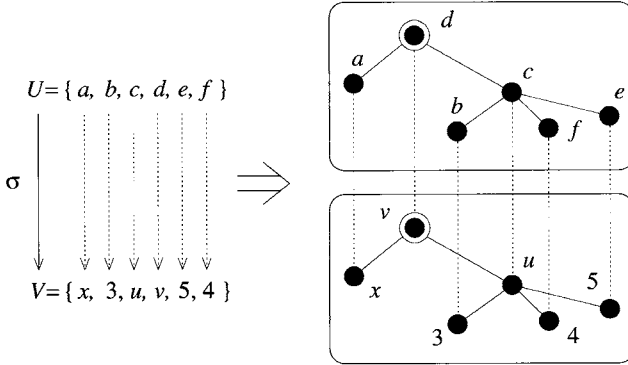


Fig. 2.

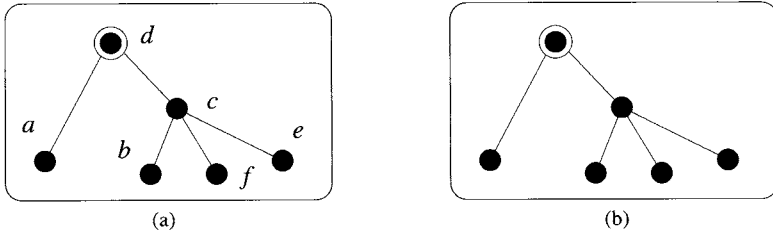


Fig. 3.

$V = \{x, 3, u, v, 5, 4\}$ via the bijection $\sigma : U \rightarrow V$ described by Figure 2. This figure clearly shows how the bijection σ allows the transport of the rooted tree s onto a corresponding rooted tree $t = (\tau, V)$ on the set V , simply by replacing each vertex $u \in U$ by the corresponding vertex $\sigma(u) \in V$. We say that the rooted tree t has been obtained by transporting the rooted tree s along the bijection σ , and we write

$$t = \sigma \cdot s. \tag{5}$$

From a purely set theoretical point of view, this amounts to replacing simultaneously each element u of U appearing in γ by the corresponding element $\sigma(u)$ of V in the expression of γ . The rooted trees s and t are said to be *isomorphic*, and σ is called an *isomorphism* of s to t .

Intuitively two isomorphic structures can be considered as identical if the nature of the elements of their underlying sets is ignored. This “general form” that isomorphic structures have in common is their *isomorphism type*. It often can be represented by a diagram (see, for example, Figure 3(b)) in which the elements of the underlying set are represented by “indistinguishable” points. The structure is then said to be *unlabeled*. Figure 4 illustrates a rooted tree *automorphism*. In this case, the sets U and V coincide, the bijection $\sigma : U \rightarrow U$ is a permutation

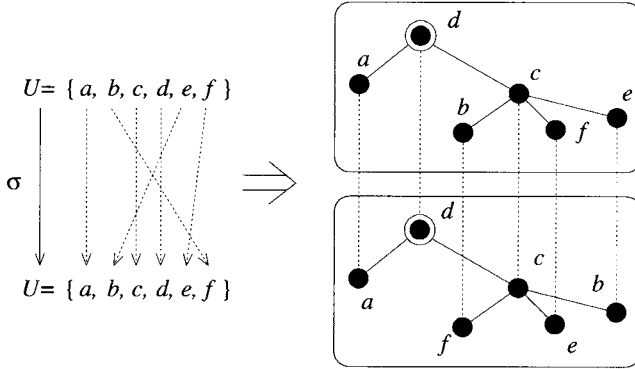


Fig. 4.

of U , and the transported rooted tree $\sigma \cdot s$ is identical to the tree s , that is to say, $s = \sigma \cdot s$.

The preceding examples show that the concept of transport of structures is of prime importance since it enables one to define the notions of isomorphism, isomorphism type, and automorphism. In fact, the transport of structures is at the very base of the general concept of species of structures.

Example 2. As an introduction to the formal definition of species of structures, here is a detailed description of the *species* \mathcal{G} of all *simple graphs* (i.e., undirected graphs without loops or multiple edges). For each finite set U , we denote by $\mathcal{G}[U]$ the set of all structures of simple graph on U . Thus

$$\mathcal{G}[U] = \{g \mid g = (\gamma, U), \gamma \subseteq \wp^{[2]}[U]\}, \tag{6}$$

where $\wp^{[2]}[U]$ stands for the collection of (unordered) pairs of elements of U . In the simple graph $g = (\gamma, U)$, the elements of U are the *vertices* and γ is the set of *edges*. Clearly $\mathcal{G}[U]$ is a finite set. The following three expressions are considered to be equivalent:

- g is a simple graph on U ;
- $g \in \mathcal{G}[U]$;
- g is a \mathcal{G} -structure on U .

Moreover, each bijection $\sigma : U \rightarrow V$ induces, by transport of structure (see Figure 5), a function

$$\mathcal{G}[\sigma] : \mathcal{G}[U] \rightarrow \mathcal{G}[V], \quad g \mapsto \sigma \cdot g, \tag{7}$$

describing the transport of graphs along σ . Formally, if $g = (\gamma, U) \in \mathcal{G}[U]$, then $\mathcal{G}[\sigma](g) = \sigma \cdot g = (\sigma \cdot \gamma, V)$, where $\sigma \cdot \gamma$ is the set of pairs $\{\sigma(x), \sigma(y)\}$ of elements of V obtained from pairs $\{x, y\} \in \gamma$. Thus each edge $\{x, y\}$ of g finds itself relabeled $\{\sigma(x), \sigma(y)\}$ in $\sigma \cdot g$.

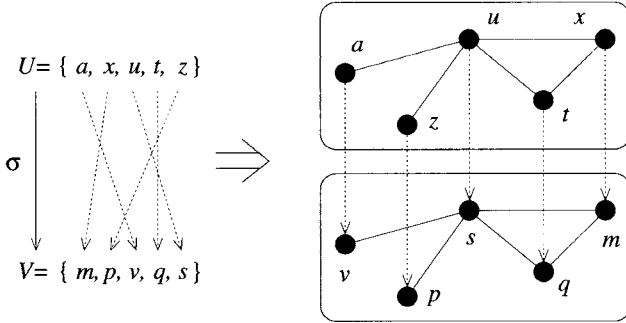


Fig. 5.

Since this transport of graphs along σ is only a relabeling of the vertices and edges by σ , it is clear that for bijections $\sigma : U \rightarrow V$ and $\tau : V \rightarrow W$, one has

$$\mathcal{G}[\tau \circ \sigma] = \mathcal{G}[\tau] \circ \mathcal{G}[\sigma], \tag{8}$$

and that, for the identity map $\text{Id}_U : U \rightarrow U$, one has

$$\mathcal{G}[\text{Id}_U] = \text{Id}_{\mathcal{G}[U]}. \tag{9}$$

These two equalities express the *functoriality* of the transports of structures $\mathcal{G}[\sigma]$. It is this property which is abstracted in the definition of species of structures.

Definition of Species of Structures

Definition 3. A *species of structures* is a rule F which

- i) *produces*, for each finite set U , a finite set $F[U]$,
- ii) *produces*, for each bijection $\sigma : U \rightarrow V$, a function

$$F[\sigma] : F[U] \rightarrow F[V]. \tag{10}$$

The functions $F[\sigma]$ should further satisfy the following functorial properties:

- a) for all bijections $\sigma : U \rightarrow V$ and $\tau : V \rightarrow W$,

$$F[\tau \circ \sigma] = F[\tau] \circ F[\sigma], \tag{11}$$

- b) for the identity map $\text{Id}_U : U \rightarrow U$,

$$F[\text{Id}_U] = \text{Id}_{F[U]}. \tag{12}$$

An element $s \in F[U]$ is called an F -*structure on* U (or even a *structure of species* F on U). The function $F[\sigma]$ is called the *transport of* F -structures along σ .

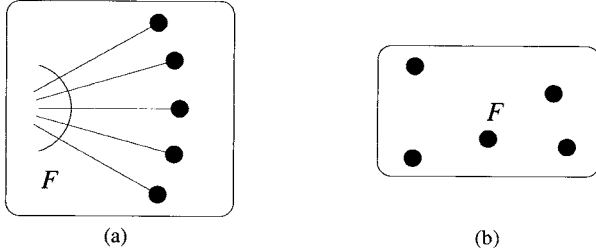


Fig. 6.

- As for graphs, the three following expressions are considered to be equivalent:
- s is a structure of species F on U ;
 - $s \in F[U]$;
 - s is an F -structure on U .

It immediately follows from the functorial properties that each transport function $F[\sigma]$ is necessarily a bijection (see Exercise 2). We use the notation $\sigma \cdot s$, or sometimes $\sigma \cdot_F s$ to avoid ambiguity, to designate $F[\sigma](s)$.

In order to represent a generic F -structure, we often utilize drawings like those of Figures 6(a) and 6(b). The black dots in these figures represent the (distinct) elements of the underlying set. The F -structure itself is represented in 6(a) by a circular arc labeled F , and in 6(b) by the superposition of the symbol F .

Observe that the notions of isomorphism, isomorphism type, and automorphism of F -structures are implicitly contained in the definition of the species F .

Definition 4. Consider two F -structures $s_1 \in F[U]$ and $s_2 \in F[V]$. A bijection $\sigma : U \rightarrow V$ is called an *isomorphism* of s_1 to s_2 if $s_2 = \sigma \cdot s_1 = F[\sigma](s_1)$. One says that these structures have the same *isomorphism type*. Moreover, an isomorphism from s to s is said to be an *automorphism* of s .

The advantage of this definition of species is that the rule F which produces the structures $F[U]$ and the transport functions $F[\sigma]$ can be described in any fashion, provided that the functoriality conditions (11) and (12) are satisfied. For example, one can either use axiomatic systems, explicit constructions, algorithms, combinatorial operations, functional equations, or even simple geometric figures to specify a species. We will illustrate each of these approaches with some examples.

Examples of Species Defined Using Set Theoretic Axioms

A species F can be defined by means of a system \mathfrak{A} of well-chosen axioms by requiring

$$s = (\gamma, U) \in F[U] \quad \text{if and only if} \quad s = (\gamma, U) \quad \text{is a model of} \quad \mathfrak{A}.$$

The transport $F[\sigma]$ is carried out in the natural fashion illustrated earlier. Clearly one can introduce in this manner a multitude of species, including the following (see Exercise 4):

- the species \mathcal{A} , of *rooted trees*;¹
- the species \mathcal{G} , of *simple graphs*;
- the species \mathcal{G}^c , of *connected simple graphs*;
- the species \mathfrak{a} , of *trees* (connected simple graphs without cycles);¹
- the species \mathcal{D} , of *directed graphs*;
- the species Par , of *set partitions*;
- the species \emptyset , of *subsets*, i.e.,

$$\emptyset[U] = \{S \mid S \subseteq U\}; \tag{13}$$

- the species End , of *endofunctions*, i.e.,

$$\text{End}[U] = \{\psi \mid \psi : U \longrightarrow U\}; \tag{14}$$

- the species Inv , of *involutions*, i.e., those endofunctions ψ such that $\psi \circ \psi = \text{Id}$;
- the species \mathcal{S} , of *permutations* (i.e., bijective endofunctions);
- the species \mathcal{C} , of *cyclic permutations* (or oriented cycles);
- the species \mathcal{L} , of *linear* (or *total*) *orders*, etc.

For example, identifying each endofunction ψ with its (γ, U) , we can describe the species End of all endofunctions as follows: $\psi = (\gamma, U) \in \text{End}[U]$ if and only if

$$\gamma \subseteq U \times U \quad \text{and} \quad (\forall x)[(x \in U) \implies (\exists! y)[(y \in U) \quad \text{and} \quad ((x, y) \in \gamma)]]. \tag{15}$$

Directed graphs γ satisfying (15) are called *functional digraphs*. We also say that γ is the *sagittal graph* of the endofunction ψ .

Note that the transport $\text{End}[\sigma] : \text{End}[U] \longrightarrow \text{End}[V]$ along the bijection $\sigma : U \longrightarrow V$ is given by the formula

$$\text{End}[\sigma](\psi) = \sigma \circ \psi \circ \sigma^{-1}, \tag{16}$$

for each $\psi \in \text{End}[U]$. Indeed, upon setting $\theta = \text{End}[\sigma](\psi) \in \text{End}[V]$, the pairs $(u, \psi(u))$ run over the functional digraph determined by ψ if and only if the pairs $(\sigma(u), \sigma(\psi(u)))$ run over the sagittal graph determined by θ . Moreover, the relation $v = \sigma(u)$ is equivalent to the relation $u = \sigma^{-1}(v)$. We then deduce that the functional digraph of θ is given by pairs of the form $(v, \sigma \circ \psi \circ \sigma^{-1}(v))$ with $v \in V$.

Examples of Explicit Constructions

When the structures of a species F are particularly simple or not numerous, it can be advantageous to define the species by an explicit description of the sets $F[U]$

¹ Translator's note: *arbre* and *arborescence* are the French words for tree and rooted tree, respectively; hence the notation \mathfrak{a} and \mathcal{A} .

and transport functions $F[\sigma]$. The following species fall under this category. In each case the transport of structures $F[\sigma]$ is obvious.

- The species E , of *sets*, defined by $E[U] = \{U\}$. For each finite set U , there is a unique E -structure, namely the set U itself.
- The species ε , of *elements*, defined by $\varepsilon[U] = U$, where the structures on U are the elements of U .
- The species X , characteristic of *singletons*, defined by

$$X[U] = \begin{cases} \{U\}, & \text{if } |U| = 1, \\ \emptyset, & \text{otherwise.} \end{cases} \quad (17)$$

Here \emptyset denotes the empty set. As a consequence, there are no X -structures on a set U when $|U| \neq 1$.

- The species 1 , characteristic of the *empty set*, defined by

$$1[U] = \begin{cases} \{U\}, & \text{if } U = \emptyset, \\ \emptyset, & \text{otherwise.} \end{cases} \quad (18)$$

- The *empty* species, denoted by 0 , defined by $0[U] = \emptyset$ for all U .
- The species E_2 , characteristic of sets of cardinality 2, defined by

$$E_2[U] = \begin{cases} \{U\}, & \text{if } |U| = 2, \\ \emptyset, & \text{otherwise.} \end{cases} \quad (19)$$

Example of Algorithmic Construction

One can specify structures in an algorithmic fashion. For instance, an algorithm can be given which generates all the binary rooted trees on a given set of vertices. If we designate this algorithm by \mathcal{B} , and by $\mathcal{B}[U]$ the set of structures produced by \mathcal{B} for a given U , then we have

Algorithm 5. INPUT: A FINITE SET U ; OUTPUT: THE SET $\mathcal{B}[U]$ OF BINARY ROOTED TREES ON U .

- 1) $\mathcal{B}[U] := \{\emptyset\}$, if the set U is empty;
- 2) $\mathcal{B}[U] :=$ the set of triples (g, x, d) ,
 obtained by choosing in all possible ways:
 - a) x an element of U ;
 - b) S a subset of $U \setminus \{x\}$;
 then let $T :=$ the complement of S in $U \setminus \{x\}$;
 - c) g an element of $\mathcal{B}[S]$;
 - d) d an element of $\mathcal{B}[T]$.

End.

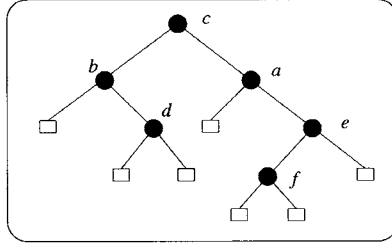


Fig. 7.

For each bijection $\sigma : U \rightarrow V$ and each binary rooted tree $(g, x, d) \in \mathcal{B}[U]$, algorithm \mathcal{B} also produces a transport of structures, by setting

- 1) $\mathcal{B}[\sigma](\emptyset) := \emptyset$, if $U = \emptyset$;
- 2) $\mathcal{B}[\sigma](g, x, d) := (l, y, r)$ with:
 - a) $y := \sigma(x)$;
 - b) $l := \mathcal{B}[\sigma](g)$;
 - c) $r := \mathcal{B}[\sigma](d)$.

End.

For example, one of the structures produced by algorithm \mathcal{B} on the set

$$U = \{a, b, c, d, e, f\}$$

is

$$((\emptyset, b, (\emptyset, d, \emptyset)), c, (\emptyset, a, ((\emptyset, f, \emptyset), e, \emptyset)))$$

(see Figure 7 for a representation of this binary rooted tree). The transport of this structure along the bijection

$$\sigma : \{a, b, c, d, e, f\} \rightarrow \{A, B, C, D, E, F\},$$

which replaces each letter by its corresponding capital letter, clearly gives

$$\begin{aligned} &\mathcal{B}[\sigma]((\emptyset, b, (\emptyset, d, \emptyset)), c, (\emptyset, a, ((\emptyset, f, \emptyset), e, \emptyset))) \\ &= ((\emptyset, B, (\emptyset, D, \emptyset)), C, (\emptyset, A, ((\emptyset, F, \emptyset), E, \emptyset))). \end{aligned}$$

In general, species of structures satisfying a “functional equation” can be defined in an algorithmic or recursive manner. See below and in Chapter 3.

Examples of Species Defined Using Combinatorial Operations

Another way of producing species of structures is by applying operations to known species. These operations (addition, multiplication, substitution, differentiation,

etc.) will be described in detail in Sections 1.3, 1.4, 2.1, and 2.2. Here are some examples:

– The species E^3 , of *tri-colorings*, $E^3 = E \cdot E \cdot E.$ (20)

– The species E_+ , of *nonempty sets*, $1 + E_+ = E.$ (21)

– The species \mathcal{H} , of *hedges* (or lists) of *rooted trees*, $\mathcal{H} = L(\mathcal{A}).$ (22)

– The species Der , of *derangements*, $E \cdot \text{Der} = \mathcal{S}.$ (23)

– The species Bal , of *ballots* (ordered partitions), $\text{Bal} = L(E_+).$ (24)

Examples of Functional Equations

It frequently happens that a species of structures is described or characterized recursively by a functional equation. Here are some examples:

– The species \mathcal{A} , of *rooted trees*, $\mathcal{A} = X \cdot E(\mathcal{A}).$ (25)

– The species L , of *linear orders*, $L = 1 + X \cdot L.$ (26)

– The species \mathcal{A}_L , of *ordered rooted trees*, $\mathcal{A}_L = X \cdot L(\mathcal{A}_L).$ (27)

– The species \mathcal{B} , of *binary rooted trees*, $\mathcal{B} = 1 + X \cdot \mathcal{B}^2.$ (28)

– The species \mathcal{P} , of *commutative parenthesizations*, $\mathcal{P} = X + E_2(\mathcal{P}).$ (29)

The interpretation and analysis of these equations is the object of Chapter 3. Describing species of structures with the help of functional equations plays a central role in this book.

Geometric Descriptions

One can sometimes gain in simplicity or clarity by describing a species F with the help of one (or several) figure(s) which schematically represents a typical F -structure. Figure 8 represents a typical structure belonging to the species P of polygons (i.e., nonoriented cycles) on a set of cardinality 5. By definition, $P[U]$ is the set of polygons on U .

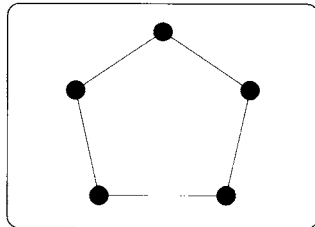


Fig. 8.