8 Combinatorial species

Combinatorial species were invented by the Canadian mathematician André Joyal ($\sim$ 1980). They provide a conceptual interpretation to relations between generating functions. In a way, they restore the combinatorial meaning that is often lost when we pass to generating functions. As motivation we give first two problems which lend themselves naturally to a species analysis.

8.1 Applications – An intuitive account

Derangements

As we have seen in Chapter 2 §1, derangements are permutations without fixed points. Their exponential generating function is,

$$D(t) = D_0 + D_1 t + D_2 \frac{t^2}{2!} + D_3 \frac{t^3}{3!} + \cdots\ , \tag{1}$$

where D_n is the number of derangements of $[n] = \{1, 2, 3, \ldots, n\}$.

On the other hand, an arbitrary permutation of $[n]$ will have a cycle structure like,

$$(* * \cdots *)(* * \cdots *) \cdots (**) \cdots (**)(*)(*) \cdots (*),$$

where the cycles $(*)$ correspond to points fixed by the permutation. In other words, an arbitrary permutation is a *product* of a permutation which fixes no points (a derangement), and a permutation which fixes every point (the identity permutation). For example, (19573)(46)(2)(8)(10) is the product of (19573)(46) (a derangement of the set $\{1, 3, 4, 5, 6, 7, 9\}$) and (2)(8)(10) (the identity permutation on the set $\{2, 8, 10\}$).

Hence, intuitively,

$$\text{Permutations} \ \equiv\ \text{Derangements} \times \text{Identity Permutations}. \tag{2}$$

Let us write the exponential generating function for permutations as

$$\begin{aligned} P(t) &= 1 + t + 2!\frac{t^2}{2!} + 3!\frac{t^3}{3!} + \cdots \\ &= 1 + t + t^2 + t^3 + \cdots \\ &= (1-t)^{-1}, \end{aligned} \tag{3}$$

and for the identity permutations

$$\begin{aligned} U(t) &= 1 + t + \frac{t^2}{2!} + \frac{t^3}{3!} + \cdots \\ &= e^t. \end{aligned} \tag{4}$$

Question. Is it possible that the intuitive insight of (2) will translate over to generating functions? That is, can we say,

$$P(t) = D(t)U(t)? \tag{5}$$

We can indeed, as we can see by multiplying (5), or equivalently,

$$D(t) = P(T)U(t)^{-1} = (1 + t + t^2 + t^3 + \cdots)(1 - t + \frac{t^2}{2!} - \frac{t^3}{3!} + \cdots),$$

See the classic paper by Joyal, "*Une théorie combinatoire des series formelles*", Advances in Math. 42 (1981), 1-82.

which yields

$$D_n = n!(1 - \frac{1}{1!} + \frac{1}{2!} - \frac{1}{3!} + \cdots + (-1)^n \frac{1}{n!}),$$

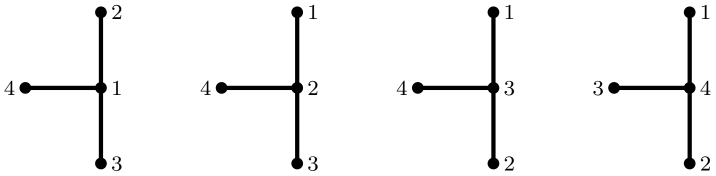
the formula which we found for the number of derangements (Chapter 2, (5)).

The insight of Joyal consists in the observation that relations between generating functions and combinations of generating functions (like (5)), already exist as relations between *structures* and ways of combining structures, like (2).

Cayley's formula

Cayley's formula gives the number of *labeled trees* on a finite set. If the set has n elements (labelled), the formula says that there are n^{n-2} distinct trees that have elements of the set as vertices.

For example, if the set is [4], the trees are,



and the twelve linear trees of the form

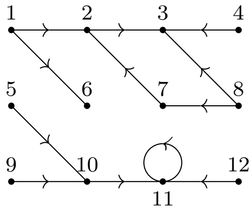
for a total of $4^2 = 16$.

A *rooted labelled tree* is a labelled tree with a distinguished element (called the root). Cayley's formula says that there are n^{n-1} rooted labelled trees, and n^n doubly rooted labelled trees.

Knowing Cayley's formula, we suspect a connection between the number of doubly rooted labelled trees, and the number of functions $f\colon [n] \to [n]$ (endofunctions of $[n]$), which, by Chapter 1 (6), is also n^n. Many proofs of Cayley's formula exploit this connection.

Let us see what the relationship between endofunctions and doubly labelled trees.

Endofunctions.. Such a function can be represented by an *arrowgraph*, consisting of the elements of $[n]$, with an arrow from x to $f(x)$, for each $x \in [n]$. The figure shows the arrowgraph for an endofunction on the set [12]. (For this example $f(6) = 1, \quad f(2) = 3, \quad f(11) = 11, \quad$ etc.)



Notice that if we repeatedly apply f to a point x we eventually obtain a cycle (since $[n]$ is finite). The points which lie on these cycles form what we may call the *final image* of f. Thus, the final image of f consists of the points $2, 3, 7, 8$ and 11. The final image gives us a natural way of partition the set $[n]$ into subsets. If x is an element of the final image, define the set U_x to consist of all elements $y\imath n[n]$ which enter the final image first at x, under repeated applications of f.

In the above figure,

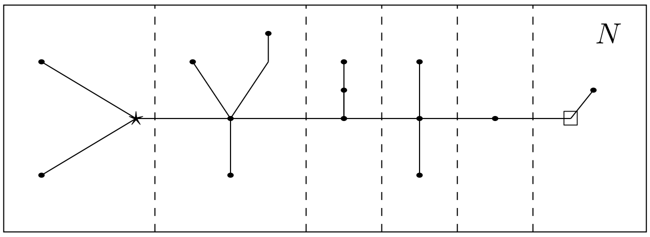
$$U_2 = \{1, 2, 6\},\ U_3 = \{3, 4\},\ U_7 = \{7\},\ U_8 = \{8\},\ U_{11} = \{5, 9, 1, 11, 12\}.$$

Then sets $U_2, U_3, U_7, U_8, U_{11}$ give a parititon of $[n]$ into 5 parts. Moreover each set U_i has a rooted tree structure on it, the root being the unique element of U_i from the final image. To complete the desription of f int remains only to know how the final image is arranged, i.e. how the parts of the partition are permuted.

To summarize: an endofunction on $[n]$ is completely determined by the following information.

(i) a partition of $[n]$,
(ii) a rooted tree structure on each part of the partition,
(iii) a permutation of the parts,
(in the above figure, this permutation is $(U_2U_3U_8U_7)(U_{11})$.) (6)

Doubly rooted labelled trees. Atypical such tree is shown below (we have left off the labels).



The roots are denoted by $\star$ and $\square$ (in general the two roots could be the same vertex and if we interchange $\star$ and $\square$ we obtain a different tree).

Notice that in a doubly rooted tree there is a unique path, determined by the roots. This path allows us to partition the set of vertices into subsets of the form U_α consisting of all vertices connected to the point x on the path. In the figure the partition is indicated by the dotted lines. Moreover each U_x posses the structure of a rooted tree. In fac thte doubly rooted tree on the set $[n]$ is completely determined by the data

(i) a partition of $[n]$,
(ii) a rooted tree structure on each part of the partition, (7)
(iii) a linear ordering of the parts.

Note. The number of linear orderings of $[n]$ is $n!$ and hence is the same as the number of permutations of the set.This suggests, from the above two descriptions, that Cayley's formula holds.

More precisely, Let $E(t)$, $R(t)$, $R_2(t)$ be the generating functions for endofunctions, root labelled trees and doubly root labelled trees, respectively. Let $L(t)$ be the generating function for linear orderings (note that $L(t) = p(t)$ given in (3)). We will see that a notion of *substitution of structures* exists which forces $E(t) = P(R(t))$ and $R_2(t) = L(R(t))$, completing the proof of Cayley's formula.

8.2 The species of Graphs

To define a combinatorial species, we need to consider two aspects. The *structures* of the species, and the way the structures are *transported* by bijections between they underlying sets. Before making a formal definition, we look at both of these aspects for the species of graphs, denoted by Graph.

The structures of Graph

Recall that a graph Γ on a set N consists of a set of *edges* joining distinct pairs of elements of N. The elements of N are called the *vertices* of Γ. We shall consider only graphs whose edges are 2-*element* sets of vertices. These are also called *labellled* graphs, because the vertices, being elements of the set N, are *distinguishable.* For example, If $N = \{1, 2, 3\}$, the graphs

$$\Gamma_1 = \{\{1, 2\}, \{2, 3\}\}$$

and

$$\Gamma_2 = \{\{1, 3\}, \{2, 3\}\}$$

depicted in Figure 1 are distinct, even though without their labels they look the same.

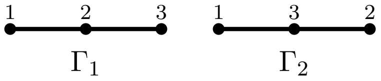


Figure 1. A pair of labelled graphs

Graph(N) is called the set of structures of the species Graph on the set N, and it will simply be the collection of all labelled graphs that can be constructed on the set N. If $|N| = n$, then N has $\binom{n}{2}$ 2-element subsets and since a labelled graph is obtained by choosing some collection of these 2-element sets and calling them edges, the total number of labelled graphs on N is $2^{\binom{n}{2}}$. That is,

$$|\text{Graph}(N)| = 2^{\binom{n}{2}}.$$

Notation. For the set $[n] = \{1, 2, 3, \ldots, n\}$, we write Graph$[n]$ for Graph$([n])$.

For example, Graph([3]) consists of the 8 labelled graphs of Figure 2.

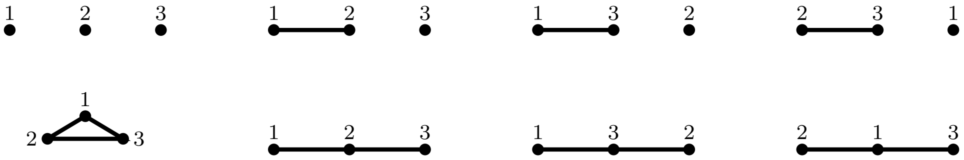


Figure 2. The labelled graphs on three vertices

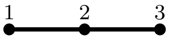
As we see from the diagrams, there are really only four *types* of graphs with 3 vertices, represented by the graphs without labels given in Figure 3.

Figure 3. The types of graphs on three vertices

The bijections of Graph

Next, let us consider the transport of structures which is a crucial part of the definition of a species. Consider two sets N and N', both of cardinality n. We know that Graph(N) and Graph(N') both consist of $2^{\binom{n}{2}}$ labelled graphs and we can see that these collections are 'essentially the same'. Two sets are 'essentially the same' if, for each bijection $u\colon N \to N'$ between the underlying sets, there exists a special bijection (i.e. one-to-one correspondence) determined by u and denoted $\mathrm{Graph}(u)$. between the sets $\mathrm{Graph}(N)$ and $\mathrm{Graph}(N')$.

We already know that there must be one-to-one correspondences between $\mathrm{Graph}(N)$ and $\mathrm{Graph}(N')$ since these sets both have $2^{\binom{n}{2}}$ elements. To define $\mathrm{Graph}(u)$, we think of u as relabelling the elements of N with the elements of N'. Then u turns each graph on N into a graph on N'. For example, if $N = \{1, 2, 3\}$, $N' = \{\star, \Box, \diamond\}$, $u(1) = \star$, $u(2) = \Box$, $u(3) = \diamond$, then u sends



to

That is, relabelling the elements of N with elements of N', u defines a natural *relabelling map* between $\mathrm{Graph}(N)$ to $\mathrm{Graph}(N')$. It is this map which we call $\mathrm{Graph}(u)$.

More formally, if $u\colon N \to N'$ is a bijection and Γ is a graph on N, we define Γ' to be the graph on N' with edges $\{u(i), u(j)\}$, whenever $\{i, j\}$ is an edge of Γ. Then $\mathrm{Graph}(u)\colon \mathrm{Graph}(N) \to \mathrm{Graph}(N')$ is the map which sends Γ to Γ'.

We say that Γ is *transported to* Γ' *via* u.

Note. If Γ is transported to Γ' via u, then

Properties of $\mathrm{Graph}(u)$

If N'' is another set of n elements and $v\colon N' \to N''$ is a bijections, then $vu\colon N \to N''$ is a bijection also. Moreover, $\mathrm{Graph}(v)$ takes the graph Γ' on N' to the graph Γ'' with edges $\{v(p), v(q)\}$ whenever $\{p, q\}$ is an edge of Γ', and so $\mathrm{Graph}(v)\mathrm{Graph}(u)$ takes the graph Γ on N to the graph with edges $\{v(u(i)), v(u(j))\}$, whenever $\{i, j\}$ is an edge of Γ. But $v(u(i)) = (vu)(i)$ and therefore,

$$\mathrm{Graph}(v)\mathrm{Graph}(u)\Gamma = \mathrm{Graph}(vu)\Gamma. \tag{8}$$

Also, if $1_n\colon N \to N$ denotes the identity mapping, i.e. $i_N(i) = i$ for all $i \in N$, then $\mathrm{Graph}(1_N)$ is the identity mapping on $\mathrm{Graph}(N)$. That is,

$$\mathrm{Graph}(1_N) = 1_{\mathrm{Graph}(N)}. \tag{9}$$

Note. Equations (8) and (9) show that $\mathrm{Graph}(u)$ is a bijection. To see this, notice,

$$\begin{aligned}\mathrm{Graph}(u)\mathrm{Graph}(u^{-1}) &= \mathrm{Graph}(uu^{-1}) \\ &= \mathrm{Graph}(1_N) = 1_{\mathrm{Graph}(N)}.\end{aligned}$$

So, $\mathrm{Graph}(u^{-1})$ is the inverse of $\mathrm{Graph}(u)$.

8.3 Definition of Species

We now give the main definition of the subject.

Definition. A *combinatorial species* S (or just *species*), consists of the following data.

1. For each finite set N there is a finite set $S(N)$.
2. For each bijection $u\colon N \to N'$ there is a function $S(u)\colon S(N) \to S(n')$ satisfying the conditions
   - (i) $S(1_N) = 1_{S(N)}$,

(ii) If $u\colon N \to N'$ and $v\colon N' \to N''$ are bijections, then $S(vu) = S(v)S(u)$.

Notation. The set $S(N)$ is called the set of *structures* of *species* S on N or the set of S-structures on N. If $\sigma \in S(N)$ is an S-structure and $u\colon N \to N'$ is a bijection if is often possible, as in the case of graphs, to think of $S(u)\sigma$ as obtained from σ by 'relabeling via u'. We say that $S(u)\sigma$ is the *S-structure* on N' obtained by *transporting σ via u.*

Remark. Thinking of $S(u)$ as relabelling works particularly well when the S-structures on N are constructed from the elements of N in some systematic way. Fortunately our examples will have this property. The generic S-structure can be represented by a *Labelle diagram* as in Figure 4, where the box contains the set N and the arc represents the S-structure σ supported by the set.

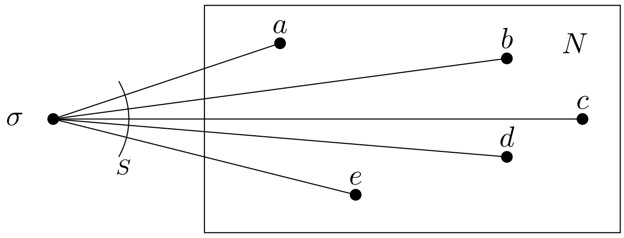


Figure 4. The Labelle diagram for an S-structure

Property (2)(ii) is illustrated in Figure 5.

$$\begin{tikzcd}
N \arrow[dr, "u"] \arrow[dd, "vu"] & & & S(N) \arrow[dr, "S(u)"] \arrow[dd, "S(vu)"] & \\
 & N' \arrow[dl, "v"] & \rightsquigarrow & & S(N') \arrow[dl, "S(v)"] \\
N'' & & & S(N'') &
\end{tikzcd}$$

Figure 4. Transport of structures.

Remark. Intuitively, two labeled graphs are 'essentially the same' if they look alike when their labels are removed. In the language of species, two S-structures $\sigma \in S(N)$ and $\tau \in S(N')$ are of the same *type* if there is a bijecton $u\colon N \to N'$ such that $\tau = S(u)\sigma$. In this case we also say that u is an *isomorphism* from σ to τ.

Remark. The concept of a species is a very special case of another mathematical concept of much wider application: namely the example of a *functor* between to *categories.* The collection of finite sets and the bijections between them is an example of a category. A specites is just a functor, between the categories $\mathrm{Sets}_{fin,bij}$ of finite sets and bijections, and Sets_{fin}, of finite sets and functions.

To complete this section, we give one more definitions, which will establish the connection between species and generating functions.

Definition. Suppose S is a species and let $|S(n)|$ denote the number of S-structures supported by $[n]$. Define the *exponential generating function* (or just generating function) for S to be the formal power series

$$\#S(t) = \sum_{n\geq 0} |S(n)| \frac{t^n}{n!} \tag{10}$$

This particular way of represntating a species first appeared in the paper by Gilbert Labelle, "*The Computation of the cycle index series of some combinatorial species*", Combinatorial Year, M.I.T. (1984).

Definition. We say that two species are *equipotent* if they have the same generating functions.

Example 1. We have seen that there are $2^{\binom{n}{2}}$ graphs supported by $[n]$ and so the generating function for Graph is

$$\#\mathrm{Graph}(t) = \sum_{n\geq 0} 2^{\binom{n}{2}} \frac{t^n}{n!}. \tag{11}$$

8.4 Examples of species

We have already seen one important example of a species, the species Graph of §2. We now increase our repertoire by introducing several other important examples.

Example 2. The species Perm is defined by taking $\mathrm{Perm}(N)$ to be the set of all permutations of the set N.

If $u\colon N \to N'$ is a bijection and $\sigma \in \mathrm{Perm}(N)$, then $\mathrm{Perm}(u)$ is defined by

$$\mathrm{Perm}(u)\sigma = u\sigma u^{-1}.$$

[Perm(u) is just the relabeling bijection. For example, if $u\colon [5] \to [5]$ is a bijection then

$$\mathrm{Perm}(u)((124)(35)) = (u1u2u4)(u3u5),$$

where we have written ui for $u(i)$ for simplicity.]

Notice that

$$\begin{aligned}
&\mathrm{Perm}(0) = \{e\},\\
&\mathrm{Perm}(1) = \{e\},\\
&\mathrm{Perm}(2) = \{e, (12)\},\\
&\mathrm{Perm}(3) = \{e, (12), (13), (23), (123), (132)\},
\end{aligned}$$

where e is the identity permutation and two permutations σ and τ on a set N are of the same 'type' (or isomorphic)precisely if they are *conjugate* i.e.

$$\tau = u\sigma u^{-1}$$

for some bijection $u\colon N \to N$.

The generating function for Perm is

$$\#\mathrm{Perm}(t) = \sum_{n\geq 0} t^n. \tag{12}$$

[#Perm(t) was denoted $P(t)$ in §1.]

Example 3. The species Lin is he species of all *linear orderings.*

Thus $\mathrm{Lin}(N)$ is the set of all orderings of elements of N. For example,

$$\mathrm{Lin}[3] = \{123, 132, 213, 231, 312, 321\}$$

If $u\colon N \to N'$ is a bijection and $x_1x_2\ldots x_n$ is a linear ordering of the elements of N then

$$\mathrm{Lin}(u)(x_1x_2\ldots x_n) = ux_1ux_2\ldots ux_n.$$

[Again, Lin(u) is the relabelling bijection.]

The generating function for Lin is the same as that for Perm.

$$\#\mathrm{Lin}(t) = \sum_{n\geq 0} t^n. \tag{13}$$

[#Lin(t) was denoted by $L(t)$ in §1.]

Note. If $x_1x_2\dots x_n$ and $y_1y_2\dots y_n$ are linear orderings of the set N, the bijection

$$u\colon x_i \to y_i$$

is an isomorphism between them. Thus there is precisely *one* type of linear ordering on a set N. Since there are several types of permutation on N, this shows that Perm and Lin are clearly different species in an essential sense. We make this more precise in the next section.

8.5 Isomorphism of species

If S and T are equipotent species then sets of S -structures and T-structures supported by a finite set N have the same cardinality. However, as the example of Perm and Lin shows, the number of types of structures may be different, and in such a case we do not want to regard the two species as being 'the same' or 'equivalent'. On the other hand, it will often happen that we can find alternative descriptions of the same species. Thus we need a precise criterion to establish when two species are the same.

To recapitulate, S and T are equipotent if for each finite set N there exists a bijection

$$\epsilon_n\colon S(N) \to T(N).$$

This does not force S and T to have the same number of types. Even if S and T are equipotent and have the same number of types on each finite set we may still wish to regard them as different species (see Problems 1 and 2 at the end of this chapter).

To be equivalent species:

- we will want the bijections ϵ_N to 'preserve' transport of structures. That is, if $\sigma \in S(N)$ and $u\colon N \to N'$ is a bijection, then σ is transported to $\tau = S(u)\sigma$ via u;
- we will want $\epsilon_N(\sigma)$ to be transported to $\epsilon_{N'}(\tau)$ also via u. In other words, we want

  $$\epsilon_{N'}(\tau) = T(u)\epsilon_N(\sigma)$$

  or

  $$\epsilon_{N'}S(u)(\sigma) = T(u)\epsilon_N(\sigma),$$

  for each $\sigma \in S(N)$.

Definition. We say that two species S nd T are *isomorphic* (or *equivalent*) if for each finite set N, there is a bijection

$$\epsilon_N\colon S(N) \to T(N)$$

such that

$$\epsilon_{N'}S(u) = T(u)\epsilon_N \tag{14}$$

for each bijection $u\colon N \to N'$. This property is illustrated in Figure 6.

$$\begin{tikzcd}
S(N) \arrow[r, "\epsilon_N"] \arrow[d, "S(u)"] & T(N) \arrow[d, "T(u)"] \\
S(N') \arrow[r, "\epsilon_{N'}"] & T(N')
\end{tikzcd}$$

Figure 6. Isomorphism of Species.

If (14) holds, the diagram in Figure 6 is said to commute.

8.6 Sums and products of species

We will see in the following section that the operations of addition, multiplication, substitution and differentiation on formal power series all suggest natural operations on species. Thus if S and T are species, $S+T$, $S \cdot T$, $S(T)$ and ∂S are species whose generating functions are obtained by performing the corresponding operations on the individual generating functions.

Let us begin with the simplest ways to combine two species.

Definition. (Sum of Species.) Let S and T be two species. Then we define $S+T$ to be the species whose structures are the disjoint union of all S-structures and all T-structures, i.e.

$$(S+T)(N) = S(N) \sqcup T(N)$$

for each finite set N. The diagram in Figure 7 represents the generic structure of $S+T$.

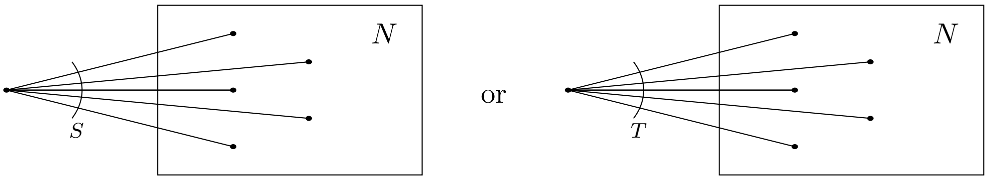


Figure 7. The Labelle diagram for the sum of species.

Moreover, if

$$u\colon N \to N'$$

is a bijection, we define the transport of structures by

$$(S+T)(u)\sigma = \begin{cases} S(u)\sigma, & \text{if } \sigma \in S(N), \\ T(u)\sigma, & \text{if } \sigma \in T(N). \end{cases}$$

More generally, an arbitrary family of species, $(S_i)_{i\in I}$ is *summable* if for all finite sets N the set of indices $i \in I$ for which $S_i(N) \neq \emptyset$ is finite. In this case we have

$$\Big(\sum_{i\in I} S_i\Big)(N) = \sum_{i\in I} S_i(N).$$

It follows immediately that

$$\#(S+T)(t) = \#S(t) + \#T(t)$$

and if $(S_i)_{i\in I}$ is summable then

$$\#\Big(\sum_{i=0}^{\infty} S_i\Big)(t) = \sum_{i=0}^{\infty} \#S_i(t).$$

The generating function of a sum of species is the sum of their generating functions

Definition. The *ordered partition* of a set N into two pieces is a pair (N_1, N_2) such that $N_1 \cup N_2 = N$ and $N_1 \cap N_2 = \emptyset$. Similarly, and ordered partition of N into n parts is an n-tuple $(N_1, \ldots, N_n)$ with

$$N = N_1 \cup N_2 \cup \cdots \cup N_n$$

and

$$N_i \cap N_j = \emptyset \quad \text{for } i \neq \text{j}.$$

Definition. We define the *product* $S \cdot T$ of two species S and T to be the species whose structures on a set N are all quadruples (N_1, N_2, σ, τ), where (N_1, N_2) is an ordered partition of N, σ is an S-structure on N_1 and τ is a T-structure on N_2.

The generic $S \cdot T$ structure is shown in Figure 8.

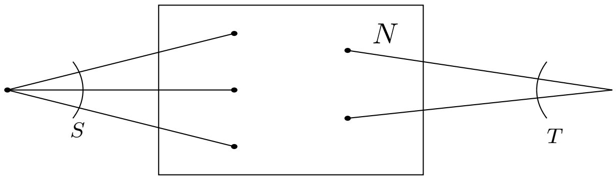


Figure 8. The Labelle diagram for $S \cdot T$.

If $u\colon N \to N'$ is a bijection then

$$(S \cdot T)(u)(N_1, N_2, \sigma, \tau) = (uN_1, uN_2, S(u)\sigma, T(u)\tau).$$

Alternatively,

$$(S \cdot T)(N) = \bigsqcup_{K \subseteq N} S(K) \times T(N - K).$$

Hence if $|N| = n$ we have

$$|(S \cdot T)(N)| = \sum_{k=0}^{n} \binom{n}{k} |S([k]| \cdot |T[n-k]|$$

and so

$$\begin{aligned}
\#(S \cdot T)(t) &= \sum_{n \geq 0} |(S \cdot T)[n]| \frac{t^n}{n!} \\
&= \sum_{n \geq 0} \frac{1}{n!} \Big(\sum_{k=0}^{n} \binom{n}{k} |S[k]| \cdot |T[n-k]| \Big) t^n \qquad (15) \\
&= \sum_{n \geq 0} \Big(\sum_{k=0}^{n} \frac{|S[k]|}{k!} \frac{|T[n-k]|}{(n-k)!} \Big) t^n \\
&= \#S(t) \cdot \#T(t). \qquad (8.1)
\end{aligned}$$

The generating function of a product of species is the product of their generating functions

In particular, If S is a species and N a finite set, the kth power of S, is the species S^k, whose structures on N are all *ordered partitions* of N into k parts, say $(N_1, \ldots, N_k)$, together with an S-structure on each of the N_i's.

8.7 Some simple species

To exploit the sum and product properly, we need some simple 'building block' species.

The Singleton Species X

We define X by

$$X(N) = \begin{cases} \{\star\}, & \text{if } |N| = 1, \\ \emptyset, & \text{if } |N| \neq 1. \end{cases}$$

Clearly,

$$\#X(t) = t.$$

The X^k-structures on a finite set N are ordered partitions of N into k parts, together with an X-structure on each part. Thus there are no X^k-structures on N unless $|N| = k$, in which case the structures are ordered partitions of N into k singleton parts with a unique $\star$ on each part. THus for $|N| = k$, a typical X^k structure on N could be written

$$((\{x_1\}, \star), (\{x_2\}, \star), \ldots, (\{x_n\}, \star)), \tag{16}$$

where $N = \{x_1, \ldots, x_n\}$. Notice that (16) is just a complicated way of writing the linear ordering $x_1 x_2 \cdots x_n$ of N.

The Constant Species A

Given any finite set A we define the species A by

$$A(N) = \begin{cases} A, & \text{if } N = \emptyset, \\ \emptyset, & \text{if } N \neq \emptyset. \end{cases}$$

If $u\colon N \to N'$ is a bijection, then

$$A(u) = \begin{cases} 1_A, & \text{if } N = \emptyset, \\ 1_\emptyset, & \text{if } N \neq \emptyset. \end{cases}$$

The Empty Species I

This is the species which has a single structure on the emptyset and no structure on any other set. Thus

$$I = A, \quad \text{where } A = \{\star\}.$$

Notice that

$$\#A(t) = |A|$$

and in particular

$$\#I(t) = 1.$$

The Uniform Species U

This species has a single structure on each finite set. One way of describing this unique structure is to regard it as the identity permutation on the set. That is

$$U(N) = \{1_N\}.$$

Notice that the structure on a set can be transported to any set with the same cardinality (along any bijection between the sets). The generating function for U is clearly

$$\#U(t) = \exp(t).$$

Note. Because of its generating function, U si often called the *exponential species*, Exp.

Example 4. Consider the first example of §1. Lt D be the *species of derangements.* The structures of D are permutations without fixed points. For example,

$$D[3] = \{(123), (132)\}.$$

The intuitive formula (1) now becomes

$$D \cdot U = \bigsqcup_{K \subseteq N} D(K) \times U(N - K).$$

In other words, we can interpret $D \cdot U$ as the species Perm of permutations. Then we have immedaitely, from (15),

$$\#D(t) \cdot \#U(t) = \#\mathrm{Perm}(t),$$

and we find

$$\begin{aligned}\#D(t) &= \exp(-t)(1-t)^{-1}\\ &= \Big(1 - t + \frac{t^2}{2!} - \frac{t^3}{3!} + \cdots\Big)\Big(1 + t + t^2 + \cdots\Big).\end{aligned}$$

Thus

$$|D[n]| = \Big(1 - 1 + \frac{1}{2!} - \frac{1}{3!} + \frac{1}{4!} - \cdots (-1)^n \frac{1}{n!}\Big) n!,$$

which is the usual formula for the number of derangemnets of n elements.

Example 5. By convention we write X^0 for I. Then the sequence

$$\{X^k\}_{k=0,1,2,\ldots}$$

is summable since for each finite set N, $X^k(N)$ is empty unless $|N| = k$. Then structures of species $I + X + X^2 + \cdots$ are elements of the form (16) and the bijection

$$\epsilon_N \colon ((\{x_1\}, \star), \ldots, (\{x_N\}, \star)) \mapsto x_1 \cdots x_n, \tag{17}$$

'preserves' the transport of structures in the sense of §3 (see Problem 3).
Thus we write

$$\mathrm{Lin} = I + X + X^2 + \cdots .$$

Example 6. The species U^2 has structures on a finite set N which are quadruples $(K, N-K, 1_K, 1_{N-k})$ and these can be identified with subsets of N. The transport-preserving bijection which allows us to make this statement is

$$\epsilon_N \colon (K, N-K, 1_K, 1_{N-k}) \mapsto K.$$

Thus we call U^2 the *subset species*. Its generating function is

$$\#U^2(t) = \exp(t)^2 = \sum_{n\geq 0} \frac{2t)^n}{n!},$$

which confirms the fact that and n-set has 2^n subsets.

Example 7. To describe simply the *species of m-subsets* we introduce first the species U_m defined by

$$U_m(N) = \begin{cases} \{1_N\}, & \text{if } |N| = m, \\ \emptyset, & \text{if } |N| \neq m. \end{cases}$$

The sequence $\{U_m\}_{m=0,1,\ldots}$ is clearly summable and

$$U = \sum_{m\geq 0} U_m.$$

Moreover the structures of the product species $U_m \cdot U$ on the set N are quadruples $(K, N-K, \sigma, \tau)$, where $\sigma \in U_m(K)$ and $\tau \in U(N-K)$. Thus the structures correspond precisely to the m-subsets of N.

Since

$$\#U_m(t) = \frac{t^m}{m!},$$

it follows that

$$\begin{aligned} \#(U_m \cdot U)(t) &= \sum_{k\geq 0} \frac{t^{k+m}}{k!m!} \\ &= \sum_{n\geq 0} \binom{n}{m} \frac{t^n}{n!}. \end{aligned}$$

where $\binom{n}{m} = 0$ unless $n \geq m$.

Comment. The above examples are simple but important illustrations of the interplay between species and their generating functions. They already show how combinatorial information at the level of species (in this case, the fact that species of subsets decompose into a product of very simple species) an be efficiently exploited using generating fucntions and the fact that species and their generating functions combine the same way.

8.8 Substitution of species

We now come to the most important way of combining species, *substitution*.

Definition. First, suppose that S is a species with no structure on the empty set. An *S-partition* of a finite set N into k parts is a set

$$\{(N_1, \sigma_1), (N_2, \sigma_2), \ldots, (N_k, \sigma_k)\}$$

where $\{N_1, N_2, \dots, N_k\}$ is a partition of N into k *non-empty* parts and $\sigma_1, \sigma_2, \dots, \sigma_k$ are S-structures on $N_1, N_2, \dots, N_k$, respectively. Notice that k can be $0, 1, 2, \dots$ or n and that k must be 0 if $N = \emptyset$.

Definition. (Substitution of Species.) Given two species, S and T, er define the *substitution of S into T* to be the species $T(S)$ whose structures on a set N are pairs (a, ρ) where a s an S-partition of N and ρ is a T-structure on the classes of a. Figure 9 gives a representation of the generic structures of $T(S)$.

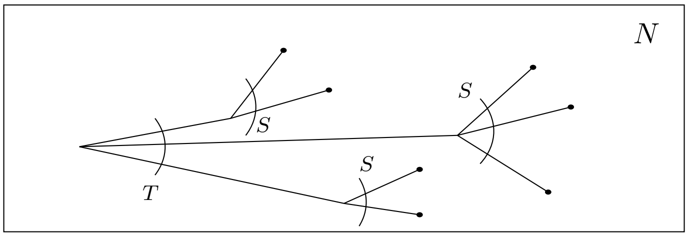


Figure 9. The Labelle diagram for $T(S)$.

If $u\colon N \to N'$ is a bijection and $T(u)$, $S(u)$ are relabelling bijections then so is $T(S)(u)$.
In general,

$$T(S)(u)\colon (a, \rho) \mapsto (a', \rho'),$$

where

$$a' = \{(uN_1, S(u)\sigma_1), \dots, (uN_k, S(u)\sigma_k)\}$$

if

$$a = \{(N_1, \sigma_1), \dots, (N_k, \sigma_k)\}$$

and $\rho' = T(u')\rho$ where $u'\colon a \to a'$ is defined by

$$u'\colon (N_i, \sigma_i) \mapsto (uN_i, S(u)\sigma_i).$$

We will se that $T(S)$ is precisely the species whose generating function is obtained by substituting $\#S(t)$ into $\#T(t)$. That is

$$\#T(S)(t) = \#T(\#S(t)). \tag{18}$$

This also explains why there cannot be any S-structures on the empty set.

Definition. Frist define the species T_k by

$$T_k(N) = \begin{cases} T(N), & \text{if } k = |N|, \\ \emptyset, & \text{if } k \neq |N|. \end{cases}$$

Then

$$T = \sum_{k \geq 0} T_k$$

and

$$T(S) = \sum_{k \geq 0} T_k(S).$$

But for a set N, $T_k(S)(N)$ consists of those pairs (a, ρ), where a is an S-partition of N into exactly k parts and ρ is a T-structure on the elements of a. Thus

$$|T_k(S)(N)| = |U_k(S)(N)|\,|T[k]| \tag{19}$$

and we find

$$\begin{aligned}\#T_k(S)(t) &= |T[k]| \sum_{n\geq 0} |U_k(S)[n]| \frac{t^n}{n!} \\ &= |T[k]| \frac{\#S^k(t)|}{k!}.\end{aligned}$$

Hence ti follows that

$$\begin{aligned}\#T(S)(t) &= \# \sum_{k\geq 0} T_k(S)(t) \\ &= \sum_{k\geq 0} \#T_k(S)(t) \\ &= \sum_{k\geq 0} |T[k]| \frac{\#S^k}{k!} \\ &= \#T(\#S(t)).\end{aligned}$$

This completes the proof of (18).

The generating function of the substitution of one species into another is the formal power series obtained by substituting the generating function of one into that of the other.

Example 8. Cayley's formula. We need to define three new species. First the species Endo of endofunctions. Its structures on a set N are (of course) the endofunctions of N, and if

$$u\colon N \to N'$$

is a bijection, then $\mathrm{Endo}(u)$ is defined by

$$\mathrm{Endo}(u)\colon f \mapsto ufu^{-1} \tag{20}$$

(20) is illustrated in Figure 10.

$$\begin{tikzcd} N \arrow[d, "u"] & f \in \mathrm{End}_O(N) \arrow[d, "\mathrm{End}_O(u)"] \\ N' & ufu^{-1} \in \mathrm{End}_O(N') \end{tikzcd}$$

Figure 10. Transport of endofunctions.

Next we define Root to be the species whose structures are rooted trees (tress with a distinguished vertex, called the *root*). For $u\colon N \to N'$ a bijection, $\mathrm{Root}(u)$ will be the relabelling bijection.

We see from the data of (6), that Endo and Perm (Root) structures are the same. Hence the species Endo and Perm(Root) could only differ in the way their structures are transported. But tranport of structures in both cases is just relabelling via a bijection so we conclude that

$$\mathrm{Endo} = \mathrm{Perm}(\mathrm{Root}) \tag{21}$$

Finally, let Root_2 be the species of doubly rooted trees.

From the data of (7), we see that

$$I + \mathrm{Root}_2 = \mathrm{Lin}(\mathrm{Root}). \tag{22}$$

From (21) and (22) we obtain the identity

$$1 + \#\mathrm{Root}_2(t) = \#\mathrm{Endo}(t)$$

which completes the proof of Cayley's formula.

Example 9. Suppose S is a species with $S(\emptyset) = \emptyset$ and let U_k be the species defined in Example 7. Then $U_k(S)$ is the species whose structures on a set N are the S-partitions of N into exactly k parts. For example, if Tree is the species of labelled trees, then $U_k(\mathrm{Tree})(N)$ is the set of all *forests* of k trees on the set N. (Here we do not tolerate trees on the empty set!)

If $S(\emptyset) = \emptyset$ then the elements of $S^k(N)$ may be regarded as *ordered* S-partitions and thus we have immediately,

$$\begin{aligned}\#U_k(S) &= \frac{1}{k!}\#S^k(t)\\ &= \frac{1}{k!}(\#S(t))^k.\end{aligned} \tag{23}$$

Also the family of species

$$(U_k(S))_{k=0,1,\ldots}$$

is summable and (23) suggests that we define the *exponential species* $\exp(S)$ to be

$$\exp(S) = \sum_{k\ge 0} U_k(S). \tag{24}$$

The structures of $\exp(S)$ are all S-partitions. In particular,

$$U = \exp(X). \tag{25}$$

Example 10. Let U be the uniform species and S an arbitrary species. Then $U(S)$ is th species of all S-partitions and so

$$U(S) = \exp(S). \tag{26}$$

Let W be the species defined by

$$W = \sum_{k>0} U_k. \tag{27}$$

Then $W(\emptyset) = \emptyset$, (in fact, $U = I + W$) and we can substitute W into U_k and U. The structures of $U_k(W)$ on a set N are W-partitions of N into k nonempty parts. But there is only one W-structure on each nonempty set, so W-partitions are really 'enriched' partitions, and we are justified in calling $U_k(W)$ the *species of partitions into exactly k nonempty parts* (see Problem 4). That is, the number of $U_k(W)$-structures on $[n]$ is just $S(n,k)$, the corresponding Stirling number of the second kind. We can immediately write down the generating function for $S(n,k)$ for fixed k. Indeed,

$$\begin{aligned}\#U_k(W)(t) &= \frac{1}{k!}(\#W(t))^k \quad \text{(by (23))}\\ &= \frac{1}{k!}(\#U - \#I)^k\\ &= \frac{1}{k!}(\exp(t) - 1)^k.\end{aligned} \tag{28}$$

Similarly we call $U(W)$ the *species of all partitions* (into nonempty parts). The number of $U(W)$-structures on the set $[n]$ is the Bell number B_n and so the generating function for the Bell numbers is

$$\begin{align*}
\#U(W)(t) &= \#\exp(W)(t) \\
&= \#\big(\sum_{k\geq 0} U_k(W)\big)(t) \\
&= \sum_{k\geq 0} \#U_k(W)(t) \tag{29} \\
&= \sum_{k\geq 0} \frac{1}{k!}(\exp(t)-1)^k \\
&= \exp(\exp(t)-1). \tag{8.2}
\end{align*}$$

Example 11. The singleton species X satisfies the condition $X(\emptyset) = \emptyset$ and for each nonempty set N there is just one X-partition of N, namely the partition of N into $|N|$ parts, in other words the set N itself. Thus

$$S(X) = S \tag{30}$$

for any species S.

8.9 Derivative of a species

In this section we consider one final operation on species.

Definition. (Derivative of a Species.) Given a species S, its *dervative* ∂S is the species whose structures on the set N are the S-structures of the disjoint union $N \sqcup \{\star\}$ with a fixed singleton set $\{\star\}$.

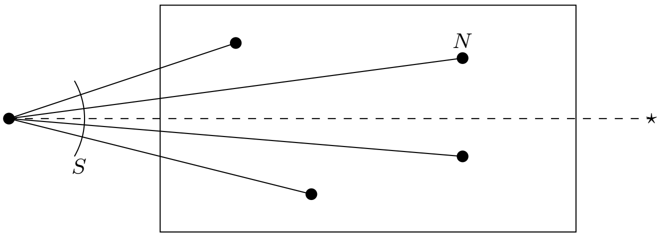


Figure 11. The Labelle diagram for ∂S.

If $u\colon N \to N'$ is a bijection, then

$$\partial S(u) = S(u^*)$$

where

$$u^*\colon N \sqcup \{\star\} \to N' \sqcup \{\star\}$$

is the bijection which fixes $\star$ and such that

$$u^*(x) = u(x)$$

for $x \in N$.

Example 12. What is the derivative of the species Tree? For example, the structures of ∂Tree on the set $[2]$ are displayed in Figure 12.

Figure 12. The three structures of ∂Tree$[2]$

Another way of looking at these structures is to eliminate the vertex labeled $\star$ but *distinguishing* the vertices to which $\star$ is attached, as we have done in Figure 13.

Figure 13. Another way of looking at the ∂Tree-structures on $[2]$.

Observe that the structures in Figure 13 are the Root-partitions of $[2]$.

In general, eliminating $\star$ in this way from a ∂Tree-structure on the set N will yield a Root-partition on N, and viec versa. Hence we have the result

$$\partial \text{Tree} = \exp(\text{Root}). \tag{31}$$

Example 13. Let Circ be the species of *cyclic orderings* whose structures on a set N are circles with the elements of N positioned on the circle, where only the relative position is important. Transport of structure is relabelling. For example, Figure 14 shows the two structures of Circ$\{1, 2, \star\}$ or ∂Circ$[2]$.

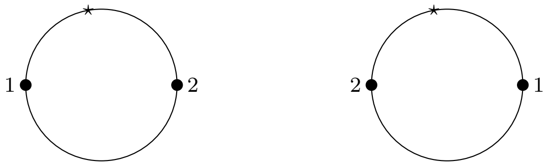


Figure 14. The structures of ∂Circ$[2]$

Eliminating the vertex$\star$ yield the linear orderings of $[2]$. And in general we find,

$$\partial \text{Circ} = \text{Lin} \tag{32}$$

Next we verify that the generating functions of the derivative behave as they should.

The generating function of the derivative of a species is the derivative of its generating function.

The proof goes as follows. Let S be a species. Then

$$\begin{aligned}
\#\partial S &= \sum_{k\geq 0} |\partial S[k]| \frac{t^k}{k!} \\
&= \sum_{k\geq 0} |S[k+1]| \frac{t^k}{k!} \\
&= \frac{d}{dt} \sum_{k\geq 0} |S[k+1]| \frac{t^{k+1}}{(k+1)!} \\
&= \frac{d}{dt}(\#\partial S).
\end{aligned} \tag{33}$$

As we would expect the derivative satisfies the usual properties:

$$\begin{aligned} &\text{(a)} \quad \partial(S+T) = \partial S + \partial T, \\ &\text{(b)} \quad \partial(ST) = \partial S \cdot T + S \cdot \partial T, \\ &\text{(c)} \quad \partial(S(T)) = \partial S(T) \cdot \partial T. \end{aligned} \tag{34}$$

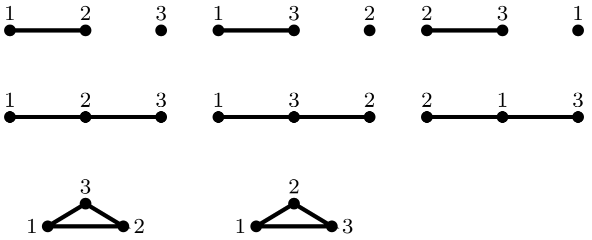
The proofs of (a) and (c) have been left to the exercises. To prove (b), let N be any finite set. The $\partial(S \cdot T)$-structures on N are quadruples of the form $(K, N \sqcup \{\star\} - K, \sigma, \tau)$ which we can classify into two families, depending on whether or not $\star$ is contained in K. If $\star$ is contained in K we have the family of $(\partial S) \cdot T$-structures on N and if $\star$ is not contained in K then the family of structures is that of $S \cdot (\partial T)$. Hence

$$\partial(S \cdot T)(N) = (\partial S \cdot T + S \cdot \partial T)(N).$$

It is clear that transport of structure is identical for the two species $\partial(S \cdot T)$ and $\partial S \cdot T + S \cdot \partial T$ and hence (b) is satisfied.

8.10 Problems 7

1. (D.E. Taylor) Let S and T be species with structures only on sets of three elements. On $[3]$ the structures of S are



The structures of T on a three element set are

and the six structures of the form

So S and T transport of structures is by relabeling. Show that S and T each have the same number of types of structures, but are not isomorphic.

2. (D.E. Taylor) Let S and T be species with structures on sets of four elements and no structures on any other sets. The structures of S are all those of the form

and those of T have the form

Show that S and T have the same number of structures, the same number of types of structures and the same number of structures of each type, but are still not isomorphic.

3. Show that the bijection (17) in Example 5 defines an isomorphism between the species Lin and $1 + X + X^2 + \cdots$.

4. Show that $U_k(W)$ is the species of partitions into k nonempty parts by displaying the isomorphism.

5. Let Edge be the species of trees with a distinguished edge. Show that

$$\mathrm{Edge} = U_2(\mathrm{Root}).$$

Deduce that

$$\#\mathrm{Edge}(t) = \frac{1}{2}\#(\mathrm{Root}(t))^2$$

and hence calculated the number of edged tress with $2, 3$, and 5 vertices.

6. Show that

$$\mathrm{Root} = X \cdot \exp(\mathrm{Root}).$$

7. Let Permf be the species of permutations which have a specified fixed point. Show that $I + \mathrm{Permf}$, Perm and Lin have the same generating functions but are not isomorphic.

8. Let S be any species and λ and μ any positive integers. Suppose

$$\#S^\lambda(t) = \sum_{n\ge 0} s_n(\lambda)\frac{t^n}{n!}$$

and

$$\#S^\mu(t) = \sum_{n\ge 0} s_n(\mu)\frac{t^n}{n!}.$$

Show that

$$s_n(\lambda+\mu) = \sum_{i=0}^{n} \binom{n}{i} s_i(\lambda)s_{n-i}(\mu). \qquad (*)$$

(The sequence of functions $\{s_n(\lambda)\}_{n\ge 0}$ satisfying (*) is calle dthe *sequence of binomial type* associated with the species S.)

9. Find the binomial sequences associated with the species U, X, $I + X$, and Lin.

10. Show that

$$\exp(\mathrm{Circ}) = \mathrm{Perm},$$

and hence deduce that

$$\#\mathrm{Circ}(t) = \log(1-t)^{-1}.$$

11. Prove the identities (a) and (c) of (34).