Enumerative Combinatorics 8: Species

Peter J. Cameron

Autumn 2013

In this lecture I will discuss a very nice unifying principle for a number of topics in enumerative combinatorics, the theory of species, introduced by André Joyal in 1981. Species have been used in areas ranging from infinite permutation groups to statistical mechanics, and I can't do more here than barely scratch the surface.

Joyal gave a category-theoretic definition of species; I will take a more informal approach.

There is a book on species, by Bergeron, Labelle and Leroux, entitled *Combinatorial Species and Tree-Like Structures*; but I think that Joyal's original paper in *Advances in Mathematics* is hard to beat.

8.1 What is a species?

As I said earlier, a typical combinatorial structure of the type we wish to count is often built on a finite set; we are interested in counting labelled structures (the different structures built on a fixed set) and also the unlabelled structures (essentially the isomorphism types of structures).

A *species* is a functor $\mathbf{F}$ (this word is used by Joyal in its technical sense from category theory; I will be less formal but will explain what is going on) which takes an n-element set and produces the set of objects in which we are interested; it should also have the property that the functor transforms any bijection between n-element sets A and B to a bijection between the sets $\mathbf{F}(A)$ and $\mathbf{F}(B)$ of objects built on these sets. Because of this condition, we can use the standard n-element set $\{1, 2, \ldots, n\}$, but don't have to worry if during the argument we have a non-standard set (such as a proper subset of the standard set).

Joyal's intuition is that we think of a formal power series where the coefficients are not numbers, but sets of combinatorial objects:

$$\mathbf{F} = \sum_{n\geq 0} F(\{1,2,\ldots,n\})x^n.$$

Suitable specialisations will give us the generating functions for unlabelled and unlabelled objects.

The first specialisation is to replace the set $\mathbf{F}(A)$ by the sum of the cycle indices of the automorphism groups of the unlabelled structures in $\mathbf{F}(A)$: let us call this $Z(\mathbf{F})$. This will be a formal power series in infinitely many variables $s_1, s_2, \ldots$. Now it turns out that the specialisations

$$\begin{aligned} f(x) &= Z(\mathbf{F}; s_n \leftarrow x^n \text{ for all } n), \\ F(x) &= Z(\mathbf{F}; s_1 \leftarrow x, x_n \leftarrow 0 \text{ for } n > 1), \end{aligned}$$

give us, respectively, the ordinary generating function for the unlabelled structures in the species $\mathbf{F}$, and the exponential generating function for the labelled structures.

8.2 Examples

If this is a bit abstract, hopefully some examples will bring it back to earth.

Sets Let **Set** denote the "identity" species, where the structure on the finite set A is simply a labelling of A. Thus, for each n, there is one unlabelled srtucture, and one labelled structure. So the generating functions are

$$\begin{aligned} \mathrm{set}(x) &= \sum_{n\geq 0} x^n = \frac{1}{1-x}, \\ \mathrm{Set}(x) &= \sum_{n\geq 0} \frac{x^n}{n!} = \exp(x) \end{aligned}$$

respectively.

The cycle index of the species **Set** can be computed as follows. First,

$$Z(S_n) = \frac{1}{n!} \sum \frac{n!}{1^{a_1}\cdots n^{a_n} a_1! \cdots a_n!} s_1^{a_1} \cdots s_n^{a_n},$$

where the sum is over all partitions of n having a_i parts of size i for $i = 1, 2, \ldots, n$ (the coefficient is the number of permutations with this cycle structure). Summing this over all n seems a formidable task, but a remarkable simplification occurs: since $n!$ cancels we can sum over the variables $a_1, \ldots, a_n$ independently. We obtain

$$Z(\mathbf{Set}) = \exp\left(\sum_{i\geq 1}\left(\frac{s_i}{i}\right)\right).$$

Now substituting X^i for s_i for all i gives

$$\begin{aligned}
\mathrm{set}(x) &= \exp\left(\sum_{i\geq 1}\left(\frac{x^i}{i}\right)\right) \\
&= \exp(-\log(1-x)) \\
&= \frac{1}{1-x}, \\
\mathrm{Set}(x) &= \exp(x),
\end{aligned}$$

as expected.

Note that the formula for the sum of the cycle indices of the symmetric groups was known in the combinatorial enumeration community before Joyal provided it with this nice interpretation.

Linear orders A much easier case is the species **Lin** of linear (or total) orders. There are $n!$ labelled linear orders on n points; all are isomorphic, and there are no non-trivial automorphisms, so we have

$$Z(\mathbf{Lin}) = \sum_{n\geq 0} s_1^n = \frac{1}{1-s_1},$$

from which the generating functions are $\mathrm{lin}(x) = \mathrm{Lin}(x) = 1/(1-x)$.

8.3 Operations on species

There are three important ways that we can add two species $\mathbf{F}$ and $\mathbf{G}$.

Sum $\mathbf{F}+\mathbf{G}$ is the species which constructs on the set A all the $\mathbf{F}$-objects and all the $\mathbf{G}$-objects (we assume these two classes to be disjoint). Clearly the cycle index and the generating functions for unlabelled and labelled objects are simply obtained by adding those for $\mathbf{F}$ and $\mathbf{G}$.

Product $\mathbf{FG}$ is the species whose objects on a set A are constructed in the following way: partition A into two (possibly empty) parts B and C; put an F-object on B, and a G-object on C. A slightly harder calculation shows that the cycle index, and hence the generating functions for unlabelled and labelled objects, are obtained by multiplying those for $\mathbb{F}$ and $\mathbb{G}$.

Here is an example. What is $\mathbf{Set}^2$? Given a set A, we partition it into a subset B and its complement $A \setminus B$. So we can regard this as the species **Subset**. The numbers of unlabelled and labelled objects in this species on n points are $n+1$ and 2^n respectively, and their generating functions are (as expected) $1/(1-x)^2$ and $\exp(2x)$.

Substitution As with power series in general, there is a formal restriction on substitution: we can only substitute $\mathbf{G}$ into $\mathbf{F}$ provided that $\mathbf{G}(\emptyset)=\emptyset$. If this condition holds, then we define $\mathbf{F}[\mathbf{G}]$-objects on A as follows: partition A (into non-empty parts); put a $\mathbf{G}$-structure on each part; and put a $\mathbf{F}$-structure on the set of parts.

The cycle index is given by substituting the cycle index of $\mathbf{G}$ into that of $\mathbf{F}$ in the following way:

$$Z(\mathbf{F}[\mathbf{G}]) = Z(\mathbf{F} : s_n \leftarrow Z(\mathbf{G}, s_m \leftarrow s_{nm})).$$

In other words, for the indeterminate s_n in $Z(\mathbf{F}$, we substitute the cycle index of $\mathbf{G}$ but in the indeterminates $s_n, s_{2n}, \ldots$ in place of $s_1, s_2, \ldots$.

The effect on the generating functions for labelled objects is simple substitution: $F[G](x) = F(G(x))$. For unlabelled objects it is a bit more complicated, we need the cycle index for $\mathbf{F}$:

$$fg(x) = Z(\mathbf{F}; s_n \leftarrow g(x^n) \text{ for all } n).$$

For example, let $\mathbf{Set}^*$ be the species of non-empty sets. Then the e.g.f. for labelled objects is $\mathrm{Set}^*(x) = \exp(x) - 1$. Now $\mathbf{Set}[\mathbf{Set}^*]$ is the species of set partitions, where the labelled objects are counted by the Bell numbers: the exponential generaing function is thus $\exp(\exp(x)-1)$, as we saw earlier. As an exercise, obtain the ordinary generating function for partitions of the integer n from this approach.

Remark The fact that substituting a species into **Set** exponentiates the generating function for labelled structures is sometimes called the *exponential principle* in enumerative combinatorics. We see that substitution of species is much more general.

Rooted structures This means structures where one point is distinguished. It can be shown that the effect of rooting a species is to apply the operator $s_1 \dfrac{\partial}{\partial s_1}$ to the cycle index, and hence to apply the operator $x\,\mathrm{d}/\mathrm{d}x$ to the generating function for labelled structures. I will denote the operation of rooting a species by R, and the operation of rooting and then removing the root (i.e., deleting a point) by D: this just corresponds to differentiation.

There are many other nice examples, some of which are described in the exercises.

8.4 Exercises

1 Define the species **Circ** of circular orders and the species **Perm** of permutations, and calculate the generating functions for unlabelled and labelled objects in these species.

Show that

$$Z(\mathbf{Circ}) = -\sum_{m\geq 1} \frac{\phi(m)}{m} \log(1 - s_m),$$

where ϕ is Euler's totient function.

Use the decomposition of permutations into disjoint cycles to show that

$$\mathbf{Set}[\mathbf{Circ}] = \mathbf{Perm},$$

and verify the appropriate identities for the generating functions.

Remark It is not so easy to calculate the cycle index of **Perm** directly, but using the above expression it is not too hard to show that

$$Z(\mathbf{Perm}) = \prod_{n\geq 1} (1 - s_n)^{-1}.$$

2 Use the fact that Catalan objects are rooted binary trees to show that the species **Cat** of Catalan objects satisfies

$$\mathbf{Cat} = \mathbf{E} + \mathbf{Cat}^2,$$

where **E** denotes the species of singleton sets (that is, it returns its input if this has cardinality 1, and the empty set otherwise).

Show similarly that the species **W** of rooted binary trees without the left-right distinction (counted by Wedderburn–Etherington numbers) satisfies

$$\mathbf{W} = \mathbf{E} + \mathbf{Set}_2[\mathbf{W}],$$

where $\mathbf{Set}_2$ is the species of 2-element sets.

3 Let **F** denote the species of "1-factors" or partitions of a set into subsets of size 2. Show that

$$\begin{aligned} D(\mathbf{F}) &= \mathbf{EF}, \\ \mathbf{F} &= \mathbf{Set}[\mathbf{Set}_2]. \end{aligned}$$

Use each of these equations to show that the exponential generating function for labelled 1-factors is $\exp(x^2/2)$.

4 Let **Graph** and **ConnGraph** be the species of graphs and connected graphs respectively. (Here, assume that a connected graph has at least one vertex.) Show that

$$\mathbf{Graph} = \mathbf{Set}[\mathbf{ConnGraph}].$$

(It follows from this that the e.g.f. for connected graphs is the logarithm of the e.g.f. for graphs.)