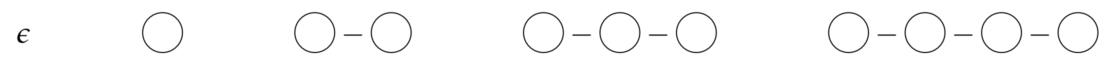
LECTURE 10 – CONSTRUCTIONS WITH LABELED OBJECTS

JAY PANTONE

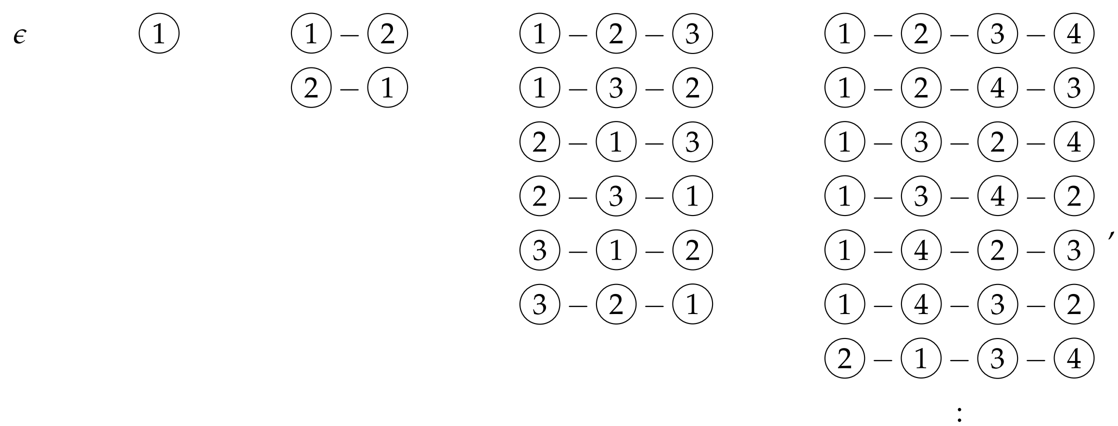
The combinatorial classes we have been working with can all be categorized as *unlabeled* objects. A *labeled object* is formed by taking an unlabeled object of size n and assigning bijectively the labels in $\{1, \dots, n\}$ to the each component of the object that contributes to size.

Before continuing, let's explore some examples of classes of labeled objects and how they differ inherently from unlabeled objects.

Example: (Labeled sequences of one object.) Let $\mathcal{Z}$ be a class that consists of a single object of size 1. The unlabeled construction $\textsc{Seq}(\mathcal{Z})$ has one element of every size:

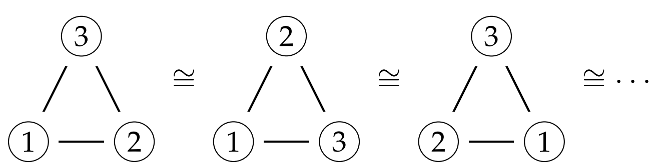


where as the class of labeled sequences has more:

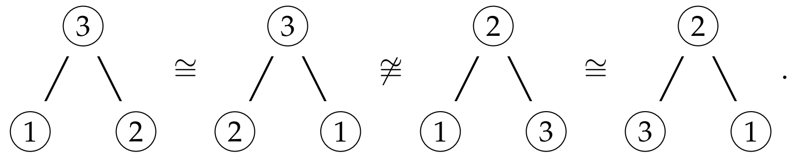


in fact there are $n!$ labeled sequences of length n.

Example: (Labeled graphs.) There are not always $n!$ times as many labeled objects as unlabeled. Consider labeled graphs on 3 vertices. There are four unlabeled graphs on three vertices, while there are eight labeled graphs. This is due to several isomorphisms between the graphs; for example

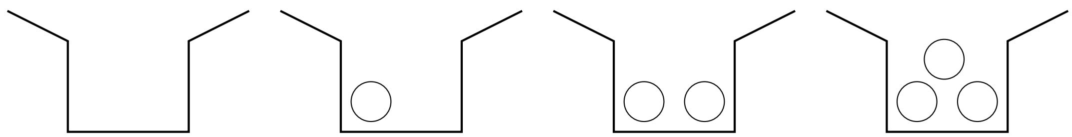


and

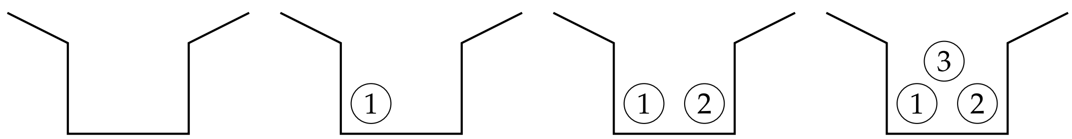


We'll prove later that the number of labeled graphs on n nodes is exactly $2^{\binom{n}{2}}$.

Example: (Balls in an urn.) Consider the class of indistinguishable objects placed in a single urn. In the unlabeled case, there is one of each size.



The labeled case also has exactly one object of each size.



While this is a fairly trivial construction, it is an important building block for labeled objects.

Among the three examples, we have seen a situation in which the number of labeled objects of size n is exactly $n!$ times the number of unlabeled objects of size n, one in which they are equal, and one that falls in between (this is the most common).

Exponential Generating Functions

Definition: We have previously defined the generating function of a sequence $\{a_n\}_{n\geq 0}$ to be the formal power series

$$f(z) = \sum_{n=0}^{\infty} a_n z^n.$$

We now refer to this as the *ordinary generating function* or *OGF*. Define the *exponential generating function* or *EGF* of a sequence $\{a_n\}_{n\geq 0}$ to be the formal power series

$$f(z) = \sum_{n=0}^{\infty} \frac{a_n}{n!} z^n.$$

As we'll soon see, the link between admissible constructions and generating functions for labeled objects is much easier to describe using EGFs instead of OGFs.

Example: (Labeled sequences of one object.) In the first example above, we described permutations as labeled sequences of a single object of size 1. As we know there are $n!$ permutations of length n, we conclude that the EGF for permutations is

$$\sum_{n=0}^{\infty} \frac{n!}{n1} z^n = \sum_{n=0}^{\infty} z^n = \frac{1}{1-z}$$

whereas the OGF for the same class is

$$\sum_{n=0}^{\infty} n! z^n,$$

which does not have a closed form and has a radius of convergence of zero.

Example: (Balls in an urn.) The class of indistinguishable balls in a single urn has counting sequence $a_n = 1$ for all n. Therefore, the OGF is,

$$\sum_{n=0}^{\infty} z^n = \frac{1}{1-z},$$

while the EGF is

$$\sum_{n=0}^{\infty} \frac{1}{n!} z^n = e^z.$$

Admissible Constructions: Disjoint Unions, Labeled Products, and Sequences

Our goal now is to replicate the theory of unlabeled constructions. We seek to describe a number of constructions that can be performed on a set of classes of labeled objects with the property that the EGF for the resulting class can be deduced from only knowledge of the EGFs of the input classes.

Disjoint Union. Our most basic construction, disjoint union, works as before. There are no labeling issues (of the kind we will soon see), because each object in $\mathcal{A} + \mathcal{B}$ appears identically as it does in either $\mathcal{A}$ or $\mathcal{B}$. Letting $\mathcal{C} = \mathcal{A} + \mathcal{B}$ and noting that $c_n = a_n + b_n$, we see that the construction is admissible, because

$$\begin{aligned} C(z) &= \sum_{n=0}^{\infty} \frac{c_n}{n!} z^n \\ &= \sum_{n=0}^{\infty} \frac{a_n + b_n}{n!} z^n \\ &= \sum_{n=0}^{\infty} \left(\frac{a_n}{n!} + \frac{b_n}{n!} \right) z^n \\ &= \sum_{n=0}^{\infty} \frac{a_n}{n!} z^n + \sum_{n=0}^{\infty} \frac{b_n}{n!} z^n \\ &= A(z) + B(z). \end{aligned}$$

Labeled Product. In order to make sense in the realm of labeled objects, we must modify the Cartesian product. With unlabeled objects, each object in the cartesian product $\mathcal{A} \times \mathcal{B}$ is a pair (α, β) with $\alpha \in \mathcal{A}$ and $\beta \in \mathcal{B}$. This doesn't quite make sense for labeled objects. Suppose $\mathcal{A}$ is the class of labeled sequences of a single object (i.e., permutations) discussed in the first example above. Then, one element in the Cartesian product $\mathcal{A} \times \mathcal{A}$ is

$$\left(② - ① - ③, \quad ① - ② \right).$$

This is *not* a labeled combinatorial objects, as it has size 5 but the labels are not assigned bijectively from the set $\{1, 2, 3, 4, 5\}$.

To remedy this, we define the labeled product $\mathcal{A} \star \mathcal{B}$ to be set of all pairs (α, β) in the (unlabeled) Cartesian product, with all possible *consistent relabelings*. A *consistent relabeling* is one in which the labels of each of α and β, while maybe changed, are in the same relative order. Continuing the example above, in the labeled product $\mathcal{A} \star \mathcal{A}$ contains one object for each consistent relabeling of the pair

$$\left(\textcircled{2} - \textcircled{1} - \textcircled{3}, \quad \textcircled{1} - \textcircled{2}\right).$$

In other words, the pair above yields in $\mathcal{A} \star \mathcal{A}$ all pairs

$$\left(\textcircled{\alpha_1} - \textcircled{\alpha_2} - \textcircled{\alpha_1}, \quad \textcircled{\beta_1} - \textcircled{\beta_2}\right)$$

such that $\alpha_2 < \alpha_1 < \alpha_3$ and $\beta_1 < \beta_2$. Examples of such elements are

$$\left(\textcircled{2} - \textcircled{1} - \textcircled{3}, \quad \textcircled{4} - \textcircled{5}\right),$$
$$\left(\textcircled{4} - \textcircled{3} - \textcircled{5}, \quad \textcircled{1} - \textcircled{2}\right),$$
$$\left(\textcircled{3} - \textcircled{5} - \textcircled{2}, \quad \textcircled{1} - \textcircled{4}\right),$$

etc. The number of consistent relabelings of a pair (α, β) is

$$\binom{\alpha + \beta}{\alpha} = \binom{\alpha + \beta}{\beta}.$$

This construction may not feel natural at the moment, but we'll soon see how it plays a critical role in the establishment of further constructions.

To see that the labeled product is admissible, let $\mathcal{C} = \mathcal{A} \star \mathcal{B}$ and note that

$$\begin{aligned}
C(z) &= \sum_{n=0}^{\infty} \frac{c_n}{n!} z^n \\
&= \sum_{n=0}^{\infty} \frac{1}{n!} \left[\sum_{k=0}^{n} \binom{n}{k} a_k b_{n-k} \right] z^n \\
&= \sum_{n=0}^{\infty} \frac{1}{n!} \left[\sum_{k=0}^{n} \frac{n!}{k!(n-k)!} a_k b_{n-k} \right] z^n \\
&= \sum_{n=0}^{\infty} \left[\sum_{k=0}^{n} \frac{a_k}{k!} \frac{b_{n-k}}{(n-k)!} \right] z^n \\
&= A(z) \cdot B(z).
\end{aligned}$$

This is *exactly* why we defined exponential generating functions to have the form they do: to make this formula work nicely.

Sequence. Sequences are defined analogously to the unlabeled case, with a similar requirement as above that every sequence is consistently relabeled. Hence,

$$\text{SEQ}(\mathcal{A}) = \epsilon + \mathcal{A} + (\mathcal{A} \star \mathcal{A}) + (\mathcal{A} \star \mathcal{A} \star \mathcal{A}) + \cdots,$$

and so if $\mathcal{C} = \text{SEQ}(\mathcal{A})$ then

$$C(z) = 1 + A(z) + A(z)^2 + A(z)^3 + \cdots = \frac{1}{1 - A(z)}.$$

We require that $a_0 = 0$. Again, we find a construction that yields the same transformation on generating functions as for unlabeled objects. Restricted sequences, too, give the same transformations as their unlabeled counterparts. We should reiterate one more time that this is not coincidence, nor does it signify that working with labeled objects is really the same as unlabeled objects; rather, the definition of the EGF of a sequence was very carefully chosen so that these transformations would be as they are.

Set. In the realm of labeled objects, the distinction between PSET and MSET is meaningless. This is because all components of a labeled object have distinct labels, and so one does not have to worry about duplicates appearing in a set: even if there are two occurrences of objects with the same underlying unlabeled form, they must have different labels.

A set can be thought of as a sequence in which the ordering is ignore. It's not hard to see that any set of k labeled objects can be arranged in $k!$ different ways, and in particular the labelings force each of these arrangements to be different. As usual, we define the size of a set to be the sum of the sizes of the components of the set.

To find the appropriate generating function, first note that

$$\text{SET}(\mathcal{A}) = \mathcal{E} + \mathcal{A} + \text{SET}_2(\mathcal{A}) + \text{SET}_3(\mathcal{A}) + \cdots,$$

where $\text{SET}_k(\mathcal{A})$ is the class of sets of size k of objects from $\mathcal{A}$. Moreover, if $\mathcal{B} = \text{SET}_k(\mathcal{A})$ then

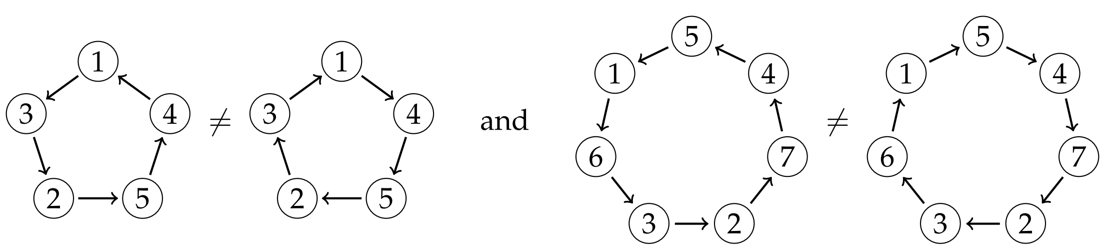
$$B(z) = \frac{1}{k!} A(z)^k.$$

Thus, if $\mathcal{C} = \text{SET}(\mathcal{A})$, we have

$$C(z) = 1 + A(z) + \frac{A(z)^2}{2!} + \frac{A(z)^3}{3!} + \cdots = \exp(A(z)).$$

This requires $a_0 = 0$.

Cyc. Given a class $\mathcal{A}$, the class $\text{CYC}(\mathcal{A})$ consists of all sequences of labeled objects from $\mathcal{A}$ where the beginning and end of the sequence are identified, so that the objects can be thought of as arranged in a circle and given an orientation. For example, letting $\mathcal{Z}$ be the class containing a single object of size 1, some objects in $\text{CYC}(\mathcal{Z})$ are:



A cycle can be formed by starting with a sequence and identifying the beginning and the end. Conversely, a sequence can be formed from a cycle of length k by picking any of the k objects to be the first in line. So, if $\mathcal{B} = \textsc{Cyc}_k(\mathcal{A})$, then

$$B(z) = \frac{1}{k} A(z)^k,$$

and if $\mathcal{C} = \textsc{Cyc}(\mathcal{A})$,

$$C(z) = \sum_{k=0}^{\infty} \frac{1}{k} A(z)^k = \log\left(\frac{1}{1 - A(z)}\right).$$

This requires $a_0 = 0$.

Additional Notes.

(1) The formulas for SET and CYC are far easier than their unlabeled counterparts (PSET and MSET, which we saw, and CYC, which we didn't). This is again due to the labels that force everything to be distinct. To make this more explicit, try counting by hand the number of ways to pick a subset of size 3 of elements from the set $\{\square, \diamond, \circ\}$. Unlabeled, there are three subsets of size 2. Labeled, how many are there?

(2) The deductions above of the form

$$B(z) = \frac{1}{k!} A(z)^k \qquad \text{and} \qquad B(z) = \frac{1}{k} A(z)^k$$

are operating on the level of generating functions. This is why we do not see the $\frac{1}{n!}$ of the exponential generating function appearing. It is already in the $A(z)^k$ term.