Labelled combinatorial structures

Many objects of classical combinatorics present themselves a labelled structure, where the atoms of an object carry different labels (that we may assume that is the set $\{1,\dots,n\}$). Our objective is to develop a "Symbolic Method" for these combinatorial classes.

Labelled classes and exponential generating functions

Def / A combinatorial class A is labelled if each element is built by atoms, such that each atom has a different label. The size of an element is the number of atoms of the element.

So let $\alpha \in A$ be an element in the labelled class A of size n. In this case the labels used to label α are in $\{1,\dots,n\}$. Under this assumption we say that α is well-labelled. We will usually consider an element ε with size 0 (without any label).

As in the unlabelled case, write $A_n = \{\alpha \in A : |\alpha| = n\}$, and $a_n = |A_n|$. Then, we have the following definition:

Def / If $a_n < \infty$ for all n, then the exponential generating function (EGF) associated to A is:

$$A(x) = \sum_{\alpha \in A} \frac{1}{|\alpha|!} x^{|\alpha|} = \sum_{n \geq 0} a_n \frac{x^n}{n!}$$

The introduction of the term $n!$ is due to the existence of labels, and in particular in order to be able to deal with the cartesian product.

Operations between labelled class

As we saw in the unlabelled setting, we will be able to introduce a dictionary between combinatorial families and the corresponding EGF.

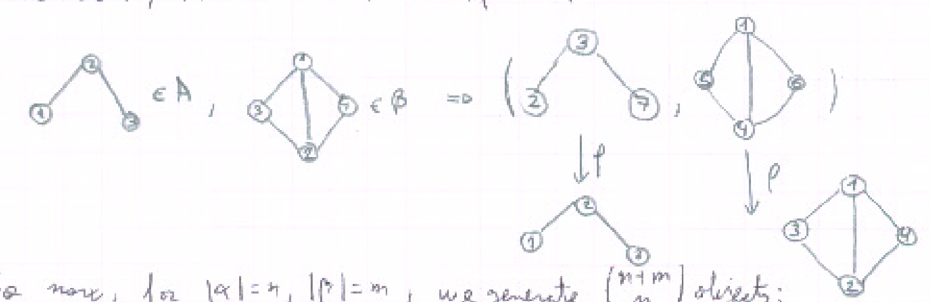
Ⓐ Disjoint union: the construction is exactly the SAME as in the unlabelled case.

Ⓑ (Labelled) product: here we need to work a little more due to the existence of labels. Let α be an element of A, of size n, which is labelled with integers not necessarily in $\{1,\dots,n\}$. Then we write $\rho(\alpha)$ for the well-labelled object which is obtained from α by relabelling it.

In this case, given two labelled structures A and B, and $\alpha \in A, \beta \in B$, $|\alpha| = n$, $|\beta| = m$, the labelled product of α and β is:

$$\alpha * \beta = \{(\alpha', \beta') : \rho(\alpha') = \alpha,\ \rho(\beta') = \beta,\ (\alpha', \beta') \text{ is well-labelled}\}$$

And then, $A * B = \{\alpha * \beta : \alpha \in A, \beta \in B\}$.



So now, for $|\alpha| = n$, $|\beta| = m$, we generate $\binom{n+m}{n}$ objects:

$$\frac{x^n}{n!} \cdot \frac{x^m}{m!} = \binom{n+m}{n} \frac{x^{n+m}}{(n+m)!} \Rightarrow A(x)\, B(x)$$

Ⓒ Sequence Construction: as in the unlabelled case: $Sq(A) = \mathcal{E} \cup A \cup A*A \cup \ldots \Rightarrow \frac{1}{1-A(x)}$. Similarly, if we restrict ourselves to indices in a certain subset Δ:

$$Seq_{\Delta}(A) = \bigcup_{\delta \in \Delta} A * \overset{\delta)}{\ldots} * A \Rightarrow \sum_{\delta \in \Delta} A(x)^{\delta}$$

Ⓓ Set Construction: is defined as in the Powerset construction (essentially, the sequence construction by forgetting the order). Here is easier than in the case of unlabelled families because we do not have internal symmetries (we have the labels):

$$Set(A) = \mathcal{E} \cup A \cup A*A/\sim \cup A*A*A/\sim \cup \ldots \Rightarrow 1 + A(x) + \frac{A(x)^2}{2} + \ldots = \exp(A(x))$$

Ⓔ Pointing and Substitution: same situation as in the unlabelled setting.

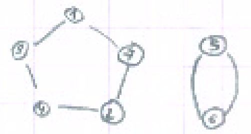
Ⓕ Cycle construction: again, by the existence of labels we do not have internal symmetries:

$$Cyc(A) = \mathcal{E} \cup A \cup A*A/_C \cup A*A*A/_C \cup \ldots \Rightarrow 1 + A(x) + \frac{A(x)^2}{2} + \ldots + \frac{A(x)^k}{k} + \ldots = -\log(1-A(x))$$

Examples

Let us see some examples

① Permutations: let S and $\mathcal{C}$ denote the class of permutations and cyclic permutations, and write S_n and $\mathcal{C}_n$ for the ones of size n:

$$\begin{matrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 \\ \downarrow & \downarrow & \downarrow & \downarrow & \downarrow & \downarrow & \downarrow \\ 3 & 7 & 4 & 2 & 6 & 5 & 1 \end{matrix} = (1,3,4,2,7)(5,6) =$$  $\Rightarrow$ A permutation is a set of cyclic permutations

Hence, $S = Set(\mathcal{C})$. As we know that $|S_n| = n!$ and $\mathcal{C}_n = (n-1)!$, we have that

$$S(x) = \sum_{n \geq 0} n! \frac{x^n}{n!} = \frac{1}{1-x}, \quad C(x) = \sum_{n \geq 0} (n-1)! \frac{x^n}{n!} = -\log(1-x),$$

and $(1-x)^{-1} = \exp(-\log(1-x))$.

We can also consider restricted type of permutations:

a) Derangement $\Rightarrow$ No fixed points $\Rightarrow$ $D = Set(\mathcal{C}_{\geq 1}) = \exp\left(\sum_{n \geq 2} (n-1)! \frac{x^n}{n!}\right) = \frac{1}{1-x} e^{-x}$

b) Involution $\Rightarrow$ Only cycles of length 1 and 2: $\mathcal{I} = Set(\bullet \cup 0) = \exp(x + x^2/2)$

② Labelled trees: let $\mathcal{T}$ be the class of rooted labelled trees. We consider trees with an embedding and without and embedding.

a) With an embedding $\Rightarrow$ $\mathcal{T} = \mathcal{Z} * Seq_{\geq 0}(\mathcal{T}) \Rightarrow T(x) = \frac{x}{1-T(x)} \Rightarrow$ Catalan function.

b) Without an embedding $\Rightarrow$ $\mathcal{T} = \mathcal{Z} * Set(\mathcal{T}) \Rightarrow T(x) = x e^{T(x)}$

This second case is new, and we can study it by means of Lagrange inversion formula:

$$[x^n]\, T(x) = \frac{1}{n} [t^{n-1}]\, e^{tn} = \frac{n^{n-1}}{n!} \Rightarrow t_n = n^{n-1} \text{ (Cayley)}.$$

From this counting we can get the GF of unrooted trees. Write it as $U(x)$. Then $xU'(x) = T(x)$. Hence,

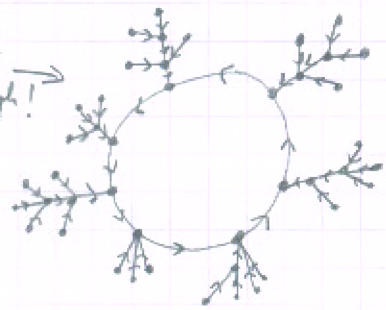
$$U(x) = \int_0^x \frac{T(y)}{y} dy = \left\{ \begin{matrix} T(y) = u \\ T'(y)dy = du \end{matrix} \right\} \Rightarrow \int_0^x e^{T(y)} dy = \int_{T(0)}^{T(x)} 1 - u \, du = T(x) - \frac{1}{2} T(x)^2$$

$$T'(y) = \frac{e^{u}}{\ldots}$$

$\Rightarrow$ Combinatorial

③ Mappings and functional graphs: Consider the set of functions $f:[n]\to[n]$. For each one of them we can construct an oriented graph in the following way:

it can be a multigraph!



$\Rightarrow$ Take x_0, and consider the sequence $x_0, f(x_0), f^2(x_0), \dots$

$\Rightarrow$ As the sequence is finite, it must loop at some moment.

$\Rightarrow$ The resulting object is a set of objects as the one described in the picture $\Rightarrow$ functional graph

We denote by $\mathcal{T}$ the class of rooted labelled trees, by K the family of connected functional graphs and by $\mathcal{F}$ the class of general functional graphs. Then,

$$\left.\begin{array}{l} \mathcal{F} = \mathrm{Set}(K) \\ K = \mathrm{Cyc}(\mathcal{T}) \\ \mathcal{T} = \bullet \times \mathrm{Set}(\mathcal{T}) \end{array}\right\} \Rightarrow \begin{array}{l} F(x) = \exp(K(x)) \\ K(x) = -\log(1-T(x)) \\ T(x) = x e^{T(x)} \end{array} \Rightarrow F(x) = (1-T(x))^{-1} = \sum_{k\geq 0} T(x)^k$$

We can now apply that $1 + xT' = (1-T)^{-1}$, getting that $[x^n]F(x) = n^n$. We can encode also the number of components and the number of cyclic points.

Probability distributions

Once we know how to encode combinatorial families, the next step is based in studying the typical behaviour of the shape of an object in the family.

More precisely, let A be a combinatorial class, and let $X: A \to \mathbb{N}$ be a parameter. Let A_n be the set of elements in A whose size is n. We consider the following question: consider A_n as a probability space where each element has the same probability of being chosen. Then X defines a random variable over A_n; so one would wonder to know if it is possible to study its probability distribution. Denote by X_n the corresponding variable.

Recall that $\mathbb{E}[X_n] = \sum_{i=0}^{\infty} i\,\mathbb{P}(X_n = i)$ and, in general, $\mathbb{E}[X_n^k] = \sum_{i=0}^{\infty} i^k\,\mathbb{P}(X_n=i)$.
So, we would like to know if:

a) Can we obtain these values?
b) Can we say something about the convergence of $\{X_n\}_{n\geq 0}$? $\dashrightarrow$ Analytic Methods.

Assume that we have the GF $A(x,u)$. Then:

$$\mathbb{E}[X_n] = \sum_{i=0}^{\infty} i \frac{a_{n,i}}{\sum_k a_{n,k}} = \frac{\sum_{i=0}^{\infty} i\, a_{n,i}}{\sum_k a_{n,k}} = \frac{[x^n]\frac{\partial}{\partial u}A(x,1)}{[x^n]A(x,1)}$$

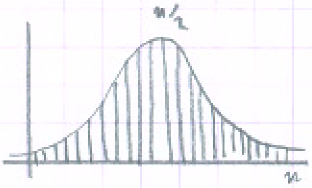
And, in general, $\mathbb{E}[X_n(X_n-1)\dots(X_n-k+1)] = \frac{1}{[x^n]A(x,1)} \cdot [x^n]\frac{\partial^k}{\partial u^k}A(x,1)$. So, the GFs encode all the information needed in order to get probability information.

Ex/ Binary words. The first example is the following: consider the set of binary words B_n and pick a word uniformly at random. Which is the distribution for the number of 0's?

$$\Rightarrow A(x,u) = \frac{1}{1-(1+u)x} \Rightarrow [x^n u^k]A(x,u) = \binom{n}{k} \Rightarrow \mathbb{E}[X_n] = \frac{\sum_{k=0}^{n} k\binom{n}{k}}{2^n} = \frac{n}{2}$$

In fact, the expression for $\frac{\partial^r}{\partial u^r} A(x,1)$ is $\frac{r!\,x^r}{(1-2x)^{r+1}}$ $\Rightarrow$ $\mathbb{E}[X_n(X_n-1)\dots(X_n-r+1)] = \frac{r!}{2^r}\binom{n}{r}$

So, the mean and the variance are $\frac{1}{2}n$ and $n/4$:



$\Rightarrow$ we will see that under general conditions we can assure normal limiting distributions, and in other cases ... we won't know!

Example / Consider now a uniformly at random permutation of size n. Denote by X_n the number of cycles. Then, we have seen that:

$$A(x,u) = \exp(-u \log(1-x)) = (1-x)^{-u}.$$

Then, we have that $[x^n]A(x,1) = \frac{n!}{n!}$, and $\frac{\partial}{\partial u} A(x,1) = \frac{-\log(1-x)}{1-x}$. Hence, $\mathbb{E}[X_n] = H_n$, where $H_n = \sum_{i=1}^{n} \frac{1}{i}$.