LECTURE 12 – LABELED TREES, MAPPINGS, AND GRAPHS

JAY PANTONE

Labeled Trees

Like unlabeled trees, labeled trees can be either planar or non-planar. Also like unlabeled trees, the planar case proves easier than the non-planar case. All trees are rooted and the size of a tree is the total number of nodes unless otherwise stated.

Labeled planar trees. Let $\mathcal{A}$ be the class of labeled planar trees. Every tree in $\mathcal{A}$ can be decomposed as a root with a possibly-empty sequence of subtrees, giving

$$\mathcal{A} = \mathcal{Z} \star \textsc{Seq}(\mathcal{A}).$$

This implies

$$A(z) = \frac{z}{1 - A(z)},$$

giving the familiar generating function

$$A(z) = \frac{1 - \sqrt{1-4z}}{2}.$$

This implies that

$$a_n = n! \cdot \frac{1}{n}\binom{2n-2}{n-1} = \frac{(2n-2)!}{(n-1)!}.$$

In fact, in the more general case in which the out-degree of any node is constrained to lie in $\Omega \subseteq \mathbb{Z}_{\geq 0}$, the symbolic construction

$$\mathcal{A} = \mathcal{Z} \star \textsc{Seq}_{\Omega}(\mathcal{A})$$

yields the functional equation

$$A(z) = z\phi(A(z))$$

where $\phi(u) = \sum_{\omega \in \Omega} u^{\omega}$. By Lagrange inversion.

$$a_n = n![z^n]A(z) = n! \cdot \frac{1}{n}[u^{n-1}]\phi(u)^n.$$

This proves that number of labeled planar trees of size n under a very wide set of possible restrictions is exactly $n!$ times the number of unlabeled planar trees of the same type. This is, of course, combinatorially evident: every two ways of labeling a given unlabeled planar tree results in distinct labeled planar trees. For this reason, we skip right to the analysis of labeled non-planar trees, for which this property no longer holds.

Labeled non-planar trees. The symbolic construction of labeled non-planar trees is as easy as that for labeled planar trees: all SEQ operators become SET operators. Yet, this leads to great differences in the EGFs produced and the subsequence asymptotic analysis.

The class $\mathcal{T}$ of all labeled non-planar trees is

$$\mathcal{T} = \mathcal{Z} \star \textsc{Set}(\mathcal{T}),$$

producing the functional equation

$$T(z) = ze^{T(z)}.$$

Using Lagrange inversion with $\phi(u) = e^u$, we can extract

$$\begin{aligned}
t_n &= n![z^n]T(z) \\
&= n! \cdot \frac{1}{n}[u^{n-1}](e^u)^n \\
&= n! \cdot \frac{1}{n}[u^{n-1}]e^{un} \\
&= n! \cdot \frac{1}{n}[u^{n-1}]\left(1 + un + \frac{u^2n^2}{2!} + \frac{u^3n^3}{3!} + \cdots\right) \\
&= n! \cdot \frac{1}{n}\frac{n^{n-1}}{(n-1)!} \\
&= n^{n-1}.
\end{aligned}$$

This was first proved by Cayley, and as a result such trees are often known as *Cayley trees*. A similar question asks for the number of *non-rooted* labeled non-planar trees. One need only observe that from each non-rooted labeled non-planar tree on n vertices, one can form n distinct rooted labeled non-planar trees (each choice of root gives a different tree). Therefore, the number of non-rooted Cayley trees on n vertices is n^{n-2}.

Versions of labeled non-planar trees in which the out-degree of nodes are restricted are easily constructed by using the restricted SET construction.

Forests. A *forest* is an (unordered) set of rooted trees, and a *k-forest* is a forest with precisely k trees. The symbolic construction for k-forests is

$$\mathcal{F}^{(k)} = \textsc{Set}_k(\mathcal{T}),$$

giving

$$F^{(k)}(z) = \frac{T(z)^k}{k!}.$$

Coefficient extraction proceeds as follows

$$\begin{aligned} f_n^{(k)} &= n![z^n]\frac{T(z)^k}{k!} \\ &= \frac{n!}{k!}[z^n]T(z)^k \\ &= \frac{n!}{k!}\cdot\frac{k}{n}[z^{n-k}]\phi(u)^n \\ &= \frac{(n-1)!}{(k-1)!}[z^{n-k}]e^{un} \\ &= \frac{(n-1)!}{(k-1)!}\frac{n^{n-k}}{(n-k)!} \\ &= \binom{n-1}{k-1}n^{n-k}. \end{aligned}$$

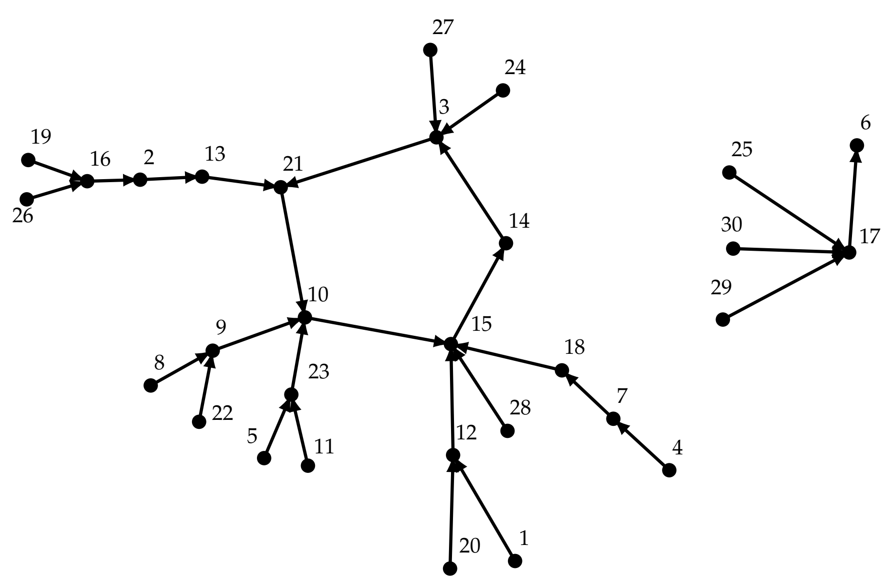
It follows that the number of all forests is

$$f_n = \sum_{k=1}^{n}\binom{n-1}{k-1}n^{n-k} = (n+1)^{n-1}.$$

Exercise: Prove directly that $f_n = (n+1)^{n-1}$ using the formula n^{n-2} for unrooted labeled non-planar trees. It should basically be a one-sentence proof.

Functional Graphs

Let $\mathcal{F}$ be the class of mappings from $[1\,..\,n]$ to $[1\,..\,n]$. The size of such a mapping is n. A mapping can be represented as a directed graph on n vertices with edges $x \to y$ if and only if $y = f(x)$. Such a graph is called a functional graph; an example is shown below.



Perhaps somewhat surprisingly, we can express the class $\mathcal{F}$ with a symbolic construction. Upon inspection of the picture above, we see that:

(1) A functional graph is a *set* of connected functional graphs.

(2) A connected functional graph is a *cycle* of rooted non-planar labeled trees.

Therefore, letting $\mathcal{K}$ be the class of connected functional graphs:

$$\begin{cases} \mathcal{F} &= \textsc{Set}(\mathcal{K}) \\ \mathcal{K} &= \textsc{Cyc}(\mathcal{T}) \\ \mathcal{T} &= \mathcal{Z} \star \textsc{Set}(\mathcal{T}) \end{cases}.$$

Translating to generating functions:

$$\begin{cases} F(z) &= e^{K(z)} \\ K(z) &= \log\left(\dfrac{1}{1-T(z)}\right) \\ T(z) &= ze^{T(z)} \end{cases}.$$

Hence,

$$\begin{aligned} f_n &= n![z^n]\frac{1}{1-T(z)} \\ &= n![z^n]\sum_{k=0}^{\infty} T(z)^k \\ &= n!\sum_{k=0}^{\infty}[z^n]T(z)^k \\ &= n!\sum_{k=0}^{\infty}\frac{k}{n}\frac{n^{n-k}}{(n-k)!} \\ &= n^n, \end{aligned}$$

as expected from the definition of the class. The last equality above results from general results on binomial summation. Moreover, one can show in a similar way that

$$k_n = (n-1)!\sum_{k=1}^{n}\frac{n^{n-k}}{(n-k)!}.$$

Next we consider various restricted versions of mappings.

Maps with no fixed points. A *fixed point* is an element $x \in [1\,..\,n]$ such that $f(x) = x$. The class $\mathcal{N}$ of maps with no fixed points is given by

$$\begin{cases} \mathcal{N} &= \textsc{Set}(\mathcal{K}) \\ \mathcal{K} &= \textsc{Cyc}_{>1}(\mathcal{T}) \\ \mathcal{T} &= \mathcal{Z} \star \textsc{Set}(\mathcal{T}) \end{cases}.$$

which gives the identity

$$N(z) = \frac{e^{-T(z)}}{1-T(z)}.$$

Lagrange inversion confirms that $N_n = (n-1)^n$, as expected. For comparison, the asymptotic behavior is

$$N_n \sim e^{-1} n^n.$$

Similarly, let $\mathcal{M}$ be mappings with no fixed points and no 2-cycles, so that $f(x) \neq x$ and $f(f(x)) \neq x$ for all x. One similarly finds

$$M(z) = \frac{e^{-T-T^2/2}}{1-T}$$

with

$$m_n \sim e^{-3/2} n^n.$$

Idempotent maps. A map is *idempotent* if $f(f(x)) = f(x)$ for all x. In other words, for every x, either x or $f(x)$ is a fixed point. Such mappings are described by functional graphs in which each connected component consists of a cycle of length 1 with a set of single vertices directed at the cycle of length one. Hence the class $\mathcal{I}$ of idempotent mappings is

$$\mathcal{I} = \textsc{Set}(\mathcal{Z} \star \textsc{Set}(\mathcal{Z}))$$

giving

$$I(z) = e^{ze^z}.$$

Saddle-point asymptotics give

$$I_n \sim \frac{n!}{\sqrt{2\pi n \zeta}} \zeta^{-n} e^{(n+1)/(\zeta+1)},$$

where ζ is the positive solution of $\zeta(\zeta+1)e^{\zeta} = n+1$.

One can prove in fact that in a random mapping of size n, one expects to reach a cycle in $O(\sqrt{n})$ steps from a randomly chosen point. As random number generators are typically large mappings, this gives rise to the catchy warning: *A random random number generator is almost surely bad.*

LABELED GRAPHS

Acyclic graphs. An *acyclic graph* is one with no cycles. Since a connected graph is acyclic if and only if it is a tree, it follows that a labeled acyclic graph is a set of labeled unrooted non-planar trees.

Let $T(z)$ be the EGF of labeled rooted non-planar trees, and recall that

$$T(z) = ze^{T(z)}.$$

Let $U(z)$ be the EGF of labeled unrooted non-planar trees, and let $A(z)$ be the EGF of acyclic graphs. The argument above shows that

$$A(z) = e^{U(z)},$$

but we haven't yet found $U(z)$. Given that $T(z)$ is best expressed implicitly, and that $t_n = nu_n$, it's unlikely that $U(z)$ has a nice closed-form expression. Instead, we will express $U(z)$ in terms of $T(z)$. First, note that

$$U(z) = \int_0^z \frac{T(w)}{w} dw.$$

(Hint: just look at the formal power series.) Now, using the identity $T(w) = we^{T(w)}$ and integration by parts,

$$\begin{aligned} U(z) &= \int_0^z \frac{T(w)}{w} dw \\ &= \int_0^z e^{T(w)} dw \\ &= we^{T(w)} - \int_0^z wT'(w)e^{T(w)} dw \\ &= T(w) - \int_0^z T'(w)T(w) dw \\ &= T(w) - \frac{1}{2}T(w)^2. \end{aligned}$$

We now conclude

$$A(z) = e^{T(z)-T(z)^2/2}.$$

Asymptotic analysis shows

$$A_n \sim \sqrt{e} \cdot n^{n-2}.$$

Unicyclic graphs. A graph is *unicyclic* if it contains exactly one cycle. A unicyclic graph is very much like a function graph, except

(1) it must be connected,

(2) the cycles must have length at least 3,

(3) the cycles are undirected.

We thus invent a new construction for undirected cycles: UCYC. Now, we claim that the class $\mathcal{W}$ of unicyclic graphs is

$$\mathcal{W} = \text{UCYC}_{\geq 3}(\mathcal{T}),$$

where $\mathcal{T}$ is the class of labeled rooted non-planar trees. We now claim that

$$W(z) = \frac{1}{2} \log\left(\frac{1}{1-T(z)}\right) - \frac{1}{2}T(z) - \frac{1}{4}T(z)^2.$$

We leave this to the reader to prove[1]. Combining this with the previous result on acyclic graphs, the class of graphs made up of acyclic and unicyclic component has EGF

$$e^{A(z)+W(z)} = \frac{e^{T(z)/2-3T(z)^2/4}}{\sqrt{1-T(z)}},$$

yielding the asymptotic form

$$n![z^n]e^{A(z)+W(z)} \sim \Gamma(3/4)(2e\pi)^{-1/4}n^{n-1/4}.$$

[1]Essentially, just divide by two to go from directed cycles to undirected cycles.

All labeled graphs. The class of all labeled graphs cannot be specified with the constructions at hand, starting from atomic classes.[2] Despite this, the simple fact that a labeled graph is a set of connected graphs can shed light on asymptotic quality of the number of connected graphs.

Let $\mathcal{G}$ be the class of all labeled graphs and let $\mathcal{K}$ be the class of connected labeled graphs. Then,

$$\mathcal{G} = \textsc{Set}(\mathcal{K}),$$

so

$$G(z) = e^{K(z)}.$$

It is combinatorially evident that $g_n = 2^{\binom{n}{2}}$. Therefore,

$$K(z) = \log\left(1 + \sum_{n\geq 1} 2^{\binom{n}{2}} \frac{z^n}{n!}\right).$$

This series also has a radius of convergence of 0, but expanding gives

$$k_n = 2^{\binom{n}{2}} - \frac{1}{2} \sum_{\substack{n_1+n_2=n \\ n_i>0}} \binom{n}{n_1, n_2} 2^{\binom{n_1}{2}+\binom{n_2}{2}} + \frac{1}{3} \sum_{\substack{n_1+n_2+n_3=n \\ n_i>0}} \binom{n}{n_1, n_2, n_3} 2^{\binom{n_1}{2}+\binom{n_2}{2}+\binom{n_3}{2}} - \cdots$$

From this we can show that

$$k_n = 2^{\binom{n}{2}} \left(1 - \frac{n}{2^{n-1}} + o(2^{-n})\right).$$ [3]

This means that almost all labeled graphs are connected (perhaps, surprising). For example, among all labeled graphs on 18 vertices, only about 0.014% are not connected.

[2] This can be proved: the EGF for all labeled graphs has a radius of convergence of 0, but all classes that can be built using labeled constructions from atomic classes have a non-zero radius of convergence. This is straight-forward to prove for iterative structures, but proving it for recursive structures is quite involved.

[3] The term "$+o(2^{-n})$" loosely means "plus some more things strictly smaller than 2^{-n}".