LECTURE 13 – THE BOXED PRODUCT

JAY PANTONE

In this lecture we'll explore a variant of the labeled product called the *boxed product*. It is a very useful construction that can in fact be used to define sets and cycles in a more natural way than we have done previous.

Definition: Given combinatorial classes $\mathcal{B}$ (with $b_0 = 0$) and $\mathcal{C}$, define

$$\mathcal{A} = \mathcal{B}^{\square} \star \mathcal{C}$$

to be the subset of objects in the labeled product $\mathcal{B} \star \mathcal{C}$ with the property that the smallest label is given to the $\mathcal{B}$ object.

Theorem 13.1. *The boxed product construction is admissible.*

Proof. Let $\mathcal{A} = \mathcal{B}^{\square} \star \mathcal{C}$. We have to show that $A(z)$ can be computed from $B(z)$ and $C(z)$.

For the regular labeled product $\mathcal{D} = \mathcal{B} \star \mathcal{C}$, we noted that

$$d_n = \sum_{k=0}^{n} \binom{n}{k} b_k c_{n-k}.$$

The boxed product is similar, with only the restriction that the smallest labeled must occur on the object from $\mathcal{B}$. (Note that once you agree to apply the smallest label to the object from $\mathcal{B}$, its location on that object is determined. It must go where the smallest label already was on the $\mathcal{B}$ object as it existed in $\mathcal{B}$.) Therefore, our binomial coefficient is altered:

$$a_n = \sum_{k=1}^{\infty} \binom{n-1}{k-1} b_k c_{n-k}.$$

To convert this to the level of generating functions we have to be a bit tricky. Rewrite as

$$\begin{aligned} a_n &= \sum_{k=1}^{\infty} \binom{n-1}{k-1} b_k c_{n-k} \\ &= \sum_{k=1}^{\infty} \frac{(n-1)!}{(k-1)!(n-k)!} b_k c_{n-k} \\ &= \sum_{k=1}^{\infty} \frac{k}{n} \frac{n!}{k!(n-k)!} b_k c_{n-k} \\ &= \frac{1}{n} \sum_{k=1}^{\infty} \binom{n}{k} (k b_k) c_{n-k} \\ &= \frac{1}{n} \sum_{k=0}^{\infty} \binom{n}{k} (k b_k) c_{n-k}. \end{aligned}$$

Clearly a_n is n^{-1} times the coefficient of z^n in product of $C(z)$ with the generating function for the terms kb_k, which is $zB'(z)$. Therefore

$$zA'(z) = zB'(z)C(z).$$

This provides the desired formula:

$$A(z) = \int_0^z B'(t)C(t)\,dt.$$

□

Exercise: How does the box product prove the familiar integration by parts formula

$$\int_0^z A(t)B'(t)dt = A(z)B(z) - \int_0^z A'(t)B(t)dt?$$

A variant of the boxed product is the *max-boxed product* $\mathcal{B}^{\blacksquare} \star \mathcal{C}$ in which we require that the *largest* label is applied to the object from $\mathcal{B}$. The generating function transformation is the same.

Records in Permutations. Let π be a permutation. A *record* or a *left-to-right maximum* in π is an entry $\pi(j)$ such that $\pi(i) < \pi(j)$ for all $i < j$. In other words, when reading the permutation left-to-right the records are those entries that are larger than any entry seen before.

Let $\mathcal{Q}$ be the class of permutations that start their largest elements. Then,

$$\mathcal{Q} = \mathcal{Z}^{\blacksquare} \star \mathcal{P},$$

where $\mathcal{P}$ is the class of all permutations. From this we recover

$$Q(z) = \int_0^z \left(\frac{d}{dt}t\right)\frac{1}{1-t}\,dt = \log\left(\frac{1}{1-z}\right).$$

Let $\mathcal{P}^{(k)}$ be the class of permutations with k records. We may be tempted to say

$$\mathcal{P}^{(k)} \stackrel{?}{=} \textsc{Seq}_k(\mathcal{Q}),$$

but that is not quite right. While $(512, 7346)$ is a sequence of length 2 of things from $\mathcal{Q}$ and 5127346 has two records, another sequence of length 2 of things from $\mathcal{Q}$ is $(7346, 512)$ yet 7346512 has only one record. Really, we want a sequence in which the maximum labels increase as you get further down the sequence. In this sense, a sequence of length 2 of elements from $\mathcal{Q}$ is

$$(\mathcal{Z}^{\blacksquare} \star \mathcal{P}) \star (\mathcal{Z}^{\blacksquare} \star \mathcal{P})^{\blacksquare}$$

and a sequence of length 3 is

$$((\mathcal{Z}^{\blacksquare} \star \mathcal{P}) \star (\mathcal{Z}^{\blacksquare} \star \mathcal{P})^{\blacksquare}) \star (\mathcal{Z}^{\blacksquare} \star \mathcal{P})^{\blacksquare}.$$

In fact there is a slightly easier way. Every permutation with k records can be thought of not as a sequence of k permutations that start with their largest elements, but as a set. From this set, the permutation is recovered by ordering the permutations in the set in increasing order of their largest element. For example

$$\{7346, 512\} \longrightarrow 5127346$$

and

$$\{52, 31, 746\} \longrightarrow 3152746.$$

Therefore,

$$\mathcal{P}^{(k)} = \text{SET}_k(\mathcal{Q}),$$

and so

$$P^{(k)}(z) = \frac{1}{k!}\left(\log\left(\frac{1}{1-z}\right)\right)^k.$$

The coefficients of this are the Stirling cycle numbers that we've seen before:

$$p_n^{(k)} = \begin{bmatrix} n \\ k \end{bmatrix}.$$

Redefinitions with the boxed product. Earlier, we didn't have a nice symbolic construction for SET and CYC. With the boxed product, we can fix this. If $\mathcal{F} = \text{SET}(\mathcal{G})$, then

$$\mathcal{F} = \{\epsilon\} + (\mathcal{G}^{\blacksquare} \star \mathcal{F}),$$

because in every set of things from $\mathcal{G}$ you can identify the one with the largest label, and the rest form another set. From this we see

$$F(z) = 1 + \int_0^z G'(t)F(t)dt,$$

which has solution

$$F(z) = \exp(G(z)).$$

Similarly, every cycle of things from $\mathcal{G}$ can be decomposed as a thing from $\mathcal{G}$ with the largest label followed by a sequence of more things from $\mathcal{G}$, giving for $\mathcal{F} = \text{CYC}(\mathcal{G})$

$$\mathcal{F} = (\mathcal{G}^{\blacksquare} \star \text{SEQ}(\mathcal{G}))$$

and thus

$$F(z) = \int_0^z G'(t) \cdot \frac{1}{1-G(t)}dt = \log\left(\frac{1}{1-G(z)}\right).$$

Parking functions. A street has n empty parking spots. One-by-one, n cars drive down the street. Each has a first choice of place to park. If that spot is full, then they proceed to the next open spot if there is one. This is effectively a map $\rho : [1\,..\,n] \to [1\,..\,n]$, where $\rho(i) = j$ if the ith car that enters prefers spot j.

Let $\mathcal{F}$ be the class of *parking functions*, those maps ρ with the property that everyone gets to park. In order to build a specification for $\mathcal{F}$, we need to introduce a construction that we have heretofore omitted.

Let $\mathcal{A}$ be a class. The *pointing construction* $\Theta\mathcal{A}$ consists of objects from $\mathcal{A}$, each with a particular component marked, or "pointed at". For example, if $\mathcal{P}$ is the class of permutations, then $\Theta\mathcal{P}$ is the class of permutations with a single entry highlighted. Letting $\mathcal{C} = \Theta\mathcal{A}$, it's clear that

$$c_n = na_n,$$

and so the generating functions translation is

$$C(z) = zA'(z).$$

Returning to parking functions, let's pretend that there is an imaginary $(n+1)$th parking spot in the row. If n cars enter, all with a preference between 1 and n, then sometimes they may end up in the $(n+1)$th parking spot. So, with this perspective a parking function is one in which all n cars park and the $(n+1)$th spot remains empty. (This parking function has size n, not $n+1$.) A parking function is labeled object; the labels on each parking space denote which car parked there.

Let ρ be a parking function, and consider the state of the parking spots directly prior to the nth car entering. The $(n+1)$th spot must be empty, as well as some spot in the middle, say j. Then, the line of parking spots from 1 to j and the line of parking spots from $j+1$ to $n+1$ each form smaller parking functions (i.e., the cars are parked, leaving the last spot open). If ρ is to be a parking function, then the preference of the last car must be between 1 and j.

This decomposition justifies the recursive construction

$$\mathcal{F} = (\Theta\mathcal{F} + \mathcal{F}) \star \mathcal{Z}^{\blacksquare} \star \mathcal{F}.$$

The term $\Theta\mathcal{F} + \mathcal{F}$ accounts for the fact that the incoming nth car must have a preference between 1 and j, but the size of the parking function induced on the spots 1 through j is only $j-1$. So, $\Theta\mathcal{F}$ counts when his preference is between 1 and $j-1$, while $\mathcal{F}$ counts when his preference is j. Then, $\mathcal{Z}^{\blacksquare}$ represents that the jth spot will now be filled with the largest label, and the remaining $\mathcal{F}$ term accounts for the spots between $j+1$ and $n+1$.

The construction gives the functional equation

$$F(z) = \int_0^z (tF'(t) + F(t))F(t)\,dt.$$

Hence,

$$F'(z) = (tF'(z) + F(z))F(z)$$

and so

$$\frac{F'(z)}{F(z)} = zF'(z) + F(z) = (zF(z))'.$$

Integrating both sides gives

$$\log(F(z)) = zF(z),$$

or

$$F(z) = e^{zF(z)}.$$

This is not the same functional equation as Cayley trees, but it's close. Similar coefficient extraction gives

$$f_n = (n+1)^{n-1}.$$