Massachusetts Institute of Technology
6.042J/18.062J: Mathematics for Computer Science
Professor Albert R. Meyer

Handout 43
April 10, 1998

Notes for Recitation 9 (4/10/98)

1 Multinomial Coefficients

The following theorem was proved in lecture using the Division Theorem.

Theorem 1.1 *Let A be the set $\{a_1, a_2, \ldots, a_n\}$, and let $r_1, r_2, \ldots, r_n$ be non-negative integers. The number of permutations of the set A where each element a_i is repeated exactly r_i times is:*

$$\frac{(r_1 + r_2 + \ldots + r_n)!}{r_1!\ r_2!\ \ldots\ r_n!}$$

For example, the theorem says that the number of permutations of $\{P, E, R\}$ where P is repeated 3 times, E is repeated 2 times, and R is repeated 1 time is $\frac{(3+2+1)!}{3!\ 2!\ 1!} = 60$. This is the answer we found before.

What is the number of permutations of the set $A = \{a_1, a_2\}$ where a_1 is repeated r_1 times and a_2 is repeated r_2 times? According to Theorem 1.1, the answer is:

$$\frac{(r_1 + r_1)!}{r_1!\ r_2!} = \binom{r_1 + r_2}{r_1}$$

We can restate the question as follows. How many strings contain r_1 copies of the symbol a_1 and r_2 copies of the symbol a_2? This is equal to the number of ways to choose r_1 distinct positions for the a_1's from the set of all $r_1 + r_2$ positions, which is $\binom{r_1+r_2}{r_1}$. This is the same as the previous answer.

By shifting from "permutations" to "strings" in this way, we can give an alternate proof of Theorem 1.1.

Proof. *(Alternate proof of Theorem 1.1.)* There is a bijection between permutations of the set $A = \{a_1, a_2, \ldots, a_n\}$ where each element a_i appears exactly r_i times and strings where each symbol a_i appears exactly r_i times. Therefore, we can count permutations by counting strings as follows.

The number of ways to choose r_1 distinct positions for the a_1's from the set of all $r_1 + r_2 + \ldots + r_n$ positions is $\binom{r_1+r_2+\ldots+r_n}{r_1}$. Then the number of ways to choose r_2 distinct positions for the a_2's from the set of all $r_2 + \ldots + r_n$ remaining positions is $\binom{r_2+\ldots+r_n}{r_2}$, and so forth. The total number of strings is therefore:

$$\begin{aligned}
&\binom{r_1+r_2+\ldots+r_n}{r_1}\cdot\binom{r_2+\ldots+r_n}{r_2}\cdot\binom{r_3+\ldots+r_n}{r_3}\cdots\binom{r_n}{r_n} \\
=\ &\frac{(r_1+r_2+\ldots+r_n)!}{r_1!\ (r_2+\ldots+r_n)!}\cdot\frac{(r_2+\ldots+r_n)!}{r_2!\ (r_3+\ldots+r_n)!}\cdot\frac{(r_3+\ldots+r_n)!}{r_3!\ (r_4+\ldots+r_n)!}\cdots\frac{r_n!}{r_n!\ 0!} \\
=\ &\frac{(r_1+r_2+\ldots+r_n)!}{r_1!\ r_2!\ \ldots\ r_n!}
\end{aligned}$$

The first equality uses the definition of binomial coefficients, and the second follows by cancelling terms. ■

The quantity

$$\frac{(r_1+r_2+\ldots+r_n)!}{r_1!\ r_2!\ \ldots\ r_n!}$$

is called a *multinomial coefficient* and is denoted

$$\binom{r_1+r_2+\ldots+r_n}{r_1, r_2, \ldots, r_n}.$$

Given a set with $r_1 + r_2 + \ldots + r_n$ elements, the multinomial coefficient $\binom{r_1+r_2+\ldots+r_n}{r_1,r_2,\ldots,r_n}$ represents the number of ways to choose r_1 elements, then r_2 of the remaining elements, and so forth.

Multinomial coefficients also arise in the following generalization of the Binomial Theorem.

Theorem 1.2 (Multinomial Theorem)

$$(x_1 + x_2 + \ldots + x_n)^r = \sum^{r_1+r_2+\ldots+r_n=r} \binom{r}{r_1, r_2, \ldots, r_n} x^{r_1} x^{r_2} \ldots x^{r_n}$$

2 An Apparently Complicated Counting Problem

Here is an absurd counting problem. In how many ways can we fill a bag with n fruits subject to the following constraints?

- The number of apples must be even.
- The number of bananas must be a multiple of 5.
- There can be at most four oranges.
- There can be at most one pear.

For example, there are 7 ways to form a bag with 6 fruits:

Apples	6	4	4	2	2	0	0
Bananas	0	0	0	0	0	5	5
Oranges	0	2	1	4	3	1	0
Pears	0	0	1	0	1	0	1

This problem appears complicated and without much hope for a closed form solution for selecting n fruits. But we can actually use generating functions to find an answer.

First, we need a generating function for the number of fruit bags. Initially, we can consider each fruit separately. That is, we can count the number of fruit bags with just apples, then just bananas, pears, *etc.* For example, the number of fruit bags with 0 apples is one (the empty bag), the number of bags with 1 apple is zero (the number of apples must be even), the number of bags with 2 apples is one (the bag with two apples), etc. The generating functions for the four types of fruit are worked out below.

$$\begin{aligned}
A(x) &= 1 + x^2 + x^4 + x^6 + \ldots &&= 1/(1-x^2),\\
B(x) &= 1 + x^5 + x^{10} + x^{15} + \ldots &&= 1/(1-x^5),\\
O(x) &= 1 + x + x^2 + x^3 + x^4 &&= (1-x^5)/(1-x),\\
P(x) &= 1 + x.
\end{aligned}$$

Now from lecture, we know

Theorem 2.1 *Let $A(x)$ and $B(x)$ be ordinary generating functions for the number of ways to choose items, subject to given constraints, from disjoint sets $\mathcal{A}$ and $\mathcal{B}$, respectively. The generating function for the number of ways to choose objects from the union $\mathcal{A} \cup \mathcal{B}$ subject to the same constraints is the product $A(x) \cdot B(x)$.*

So the generating function for the number of ways to pick n fruits subject to the constraints above is the product of A, B, O and P:

$$\begin{aligned}
\text{OGF for fruit bags} &= A(x) \cdot B(x) \cdot O(x) \cdot P(x) \\
&= \frac{1}{1-x^2} \frac{1}{1-x^5} \frac{1-x^5}{1-x}(1+x) \\
&= \frac{1}{(1-x)^2} \\
&= 1 + 2x + 3x^2 + 4x^3 + \ldots
\end{aligned}$$

The coefficient of x^n is simply $n+1$; that is, the number of ways to form a bag of n fruits is just $n+1$. This is consistent with the example we worked out, since there were 7 fruit bags with 6 fruits. Amazing!

3 Paths and Matching Parentheses

In Recitation 2 we examined the well-matched strings of parentheses. For example, `(()((())()))` is well-matched. On the other hand, no string that begins with a left parenthesis or ends with a right parenthesis can be well-matched; neither can any proper prefix of a well-matched string be well-matched.

We want to compute the number, S_n, of such well-matched parenthesis strings of length $2n$. It's easy to check that $S_1 = 1$, $S_2 = 1$, $S_3 = 2$. With a little case analysis, the only well-matched strings of 8 parentheses are

`(()(())), (()()()), ((())()), (((()))), ((()())),`

so $S_4 = 5$.

The team problem develops an ingenious combinatorial approach to arrive at a simple formula for S_n. Here we show an alternative derivation of the formula by manipulating generating functions.

It is not hard to write up a recurrence for S_n. For $n > 0$, a well matched string of $2n$ parentheses must start with a left parenthesis and end with a right parenthesis. Also, for $n > 1$, there must be a $k \in \mathbb{N}$ and a well-matched string of $2k$ parentheses immediately after the initial left parenthesis where $0 < k < n$. This substring must be followed by a *sequence* of well-matched strings of total length $2n - 2k - 2$:

$$\texttt{(} \quad \texttt{(}_1 \ldots \texttt{)}_{2k} \quad \texttt{(}_1\texttt{)}_2\texttt{(}_3 \ldots \texttt{)}_{2n-2k-2} \quad \texttt{)}$$

Since we can turn a sequence of well-matched strings into a single string by enclosing it in parentheses, the number of such sequences of well-matched strings is the same as the number of well-matched strings of length $2n - 2k$. So there are S_{n-k} such sequences of well-matched strings. Since there are, by definition, S_k well-matched initial strings of $2k$ parentheses, we have altogether

$$S_n = \sum_{k=1}^{n-1} S_k S_{n-k}$$

for $n > 1$.

This is a messy, nonlinear recurrence and might seem hard to solve. But the sum above looks exactly like the *convolution* of $(S_0, \ldots, S_n)$ with itself — except that the sum goes from 1 to $n-1$ instead of 0 to n. Let's fix it so the sum goes from 0 to n:

$$S_n = \left(\sum_{k=0}^{n} S_k S_{n-k}\right) - S_0 S_n - S_n S_0.$$

We'll agree that by convention there is exactly one *empty* string, and it is well-matched, so $S_0 = 1$ by convention. So we conclude that

$$S_n = \left(\sum_{k=0}^{n} S_k S_{n-k}\right) / 3$$

for $n > 1$. Now we really do have the convolution, and this suggests trying to fiddle with the *square* of the generating function for the S_n's.

That is, let

$$S(x) ::= \sum_{n=0}^{\infty} S_n x^n.$$

So

$$\begin{aligned}
S(x) &= S_0 + xS_1 + \sum_{n=2}^{\infty}\Big(\sum_{k=0}^{n} S_k S_{n-k}/3\Big)x^n \\
&= 1 + x + \sum_{n=2}^{\infty}\left(\sum_{k=0}^{n} S_k S_{n-k}\right) x^n/3 \\
&= 1 + x + \left(-S_0^2 - (S_0 S_1 + S_1 S_0)x + \sum_{n=0}^{\infty}\sum_{k=0}^{n} S_k S_{n-k} x^n\right)/3 \\
&= 1 + x + (-1 - 2x + S(x)^2)/3.
\end{aligned}$$

Therefore

$$S^2/3 - S + x/3 + 2/3 = 0,$$

and finally

$$S^2 - 3S + x + 2 = 0.$$

At this point we can solve for a closed form for S:

$$S(x) = \frac{3 \pm \sqrt{9 - 4(x+2)}}{2} = \frac{3}{2} \pm \frac{\sqrt{1-4x}}{2}.$$

Since $S(0) = S_0 = 1$, we need the negative root, namely:

$$S(x) = \frac{3}{2} - \frac{(1-4x)^{1/2}}{2}.$$

To get the coefficients S_n, we need to expand this quantity in a Taylor series. That is, we know that $S_n = S^{(n)}(0)/n!$ where $S^{(n)}$ is the nth derivative of S. For example,

$$\begin{aligned}
S^{(1)}(x) &= \frac{-(1/2)(1-4x)^{(1/2)-1}(-4)}{2}, \\
S^{(2)}(x) &= \frac{-(1/2)((1/2)-1)(1-4x)^{(1/2)-2}(-4)^2}{2}, \\
S^{(n)}(x) &= \frac{-(1/2)((1/2)-1)((1/2)-2)\cdots(1/2-(n-1))(1-4x)^{(1/2)-n}(-4)^n}{2}.
\end{aligned}$$

So for $n > 0$,

$$S_n = \frac{S^{(n)}(0)}{n!} = \frac{((1/2)-1)((1/2)-2)\cdots(1/2-(n-1))(-4)^{n-1}}{n!}.$$

So we have a formula for S_n, as desired. But this formula has dots and looks messy, so let's work on simplifying it:

$$\begin{aligned}
S_n &= \frac{[(-1)(1/2)(2-1)][(-1)(1/2)(4-1)]\cdots[(-1)(1/2)(2(n-1)-1)](-4)^{n-1}}{n!} \\
&= \frac{(-1)^{n-1}(1/2)^{n-1} 1 \cdot 3 \cdot 5 \cdots (2(n-1)-1)(-4)^{n-1}}{n!} \\
&= \frac{1 \cdot 3 \cdot 5 \cdots (2(n-1)-1)}{n!} \cdot 2^{n-1} \\
&= \frac{1 \cdot 3 \cdot 5 \cdots (2(n-1)-1)}{n!} \cdot 2^{n-1} \cdot \frac{2 \cdot 4 \cdot 6 \cdots 2(n-1)}{2^{n-1}(n-1)!} \\
&= \frac{1 \cdot 2 \cdot 3 \cdots (2(n-1))}{n!(n-1)!} \\
&= \frac{(2(n-1))!}{n(n-1)!(n-1)!} \\
&= \frac{1}{n}\binom{2n-2}{n-1} \quad \text{— hey, that's not bad.}
\end{aligned}$$

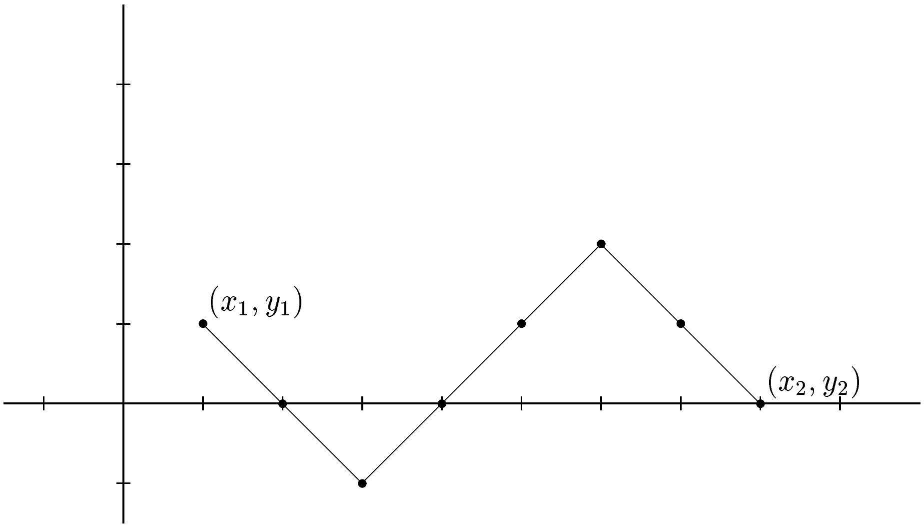

Figure 1: *A path from* $(1,1)$ *to* $(8,0)$.

The number

$$S_{n+1} = \frac{1}{n+1}\binom{2n}{n}$$

is called the n^{th} *Catalan number.*

So generating functions let us find this answer without much insight into the problem. We just did a lot of formal manipulation with derivatives and factorials.

4 Team Problem

Now we consider another ingenious approach to counting well-matched parenthesis strings. Our approach will involve considering *paths* in the plane from a point (x_1, y_1) to (x_2, y_2) for $x_i, y_i \in \mathbb{Z}$, where each step in the path is at a ± 45 degree angle, increasing the x-coordinate by 1 and the y-coordinate by ± 1. For example, a path might start at $(1,1)$ and make seven steps to the successive points $(2,0)$, $(3,-1)$, $(4,0)$, $(5,1)$, $(6,2)$, $(7,1)$, $(8,0)$ as in Figure 1.

Problem 1. **(a)** Explain why the number of well-matched parenthesized strings of length $2n$ is the same as the number of paths from $(0,0)$ to $(2n,0)$ which only touch the x-axis at their endpoints. (That is, if $0 < i < 2n$ and (i,j) is a point on the path, then $j > 0$.)

Solution: Interpret left paren as a "move up" step and right paren as a "move down" step. Then every string of k parens corresponds to a path

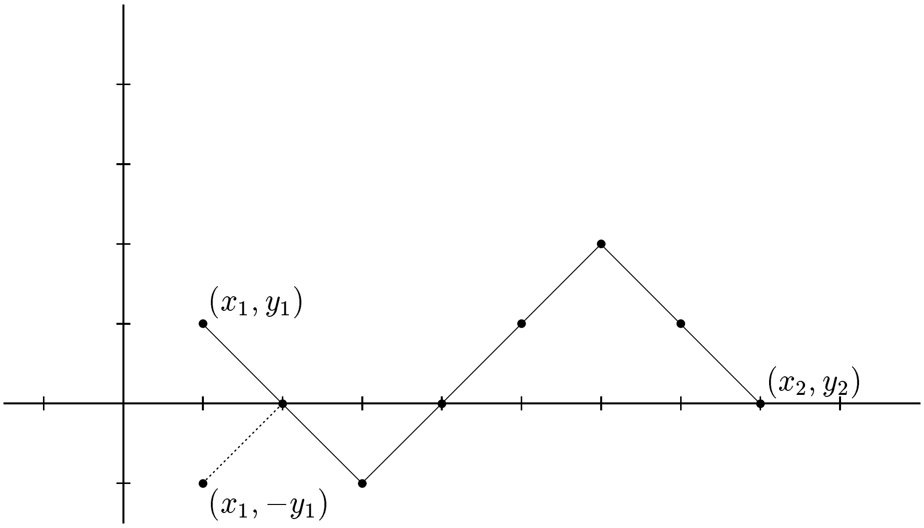

Figure 2: *Reflection of the initial part of a path.*

of length k from the origin and *vice-versa* — a bijection. Remember in Recitation 2 Team Problem 2, we showed that the well-matched strings were the same as the well-counting strings. Well-counting strings correspond exactly to paths from the origin back to the x axis which don't touch the x-axis in between endpoints.

(b) How many paths are there from (x_1, y_1) to (x_2, y_2)? (Express the answer as a binomial coefficient.)

Solution:

$$\begin{aligned}
\#\text{moves} &= x_2 - x_1 \\
\#\text{up moves} - \#\text{down moves} &= y_2 - y_1 \\
\#\text{up moves} + \#\text{down moves} &= \#\text{moves} = x_2 - x_1 \\
\#\text{up moves} &= \frac{(x_2 + y_2) - (x_1 + y_1)}{2} ::= u \\
\#\text{down moves} &= \frac{(x_2 - y_2) - (x_1 - y_1)}{2} ::= d \\
\#\text{paths} &= \binom{u+d}{u} = \binom{u+d}{d}
\end{aligned}$$

(c) For $y_1, y_2 > 0$, show that the number of paths from (x_1, y_1) to (x_2, y_2) which touch or cross the x-axis is equal to the number of paths from $(x_1, -y_1)$ to (x_2, y_2).

Solution: See Figure 2.

(d) Conclude that the number of well-matched strings of $2n$ parentheses equals the $n-1$st Catalan number

$$\frac{1}{n}\binom{2n-2}{n-1}.$$

Solution:

$$\#\text{paths}\big((1,1) \to (2n-1,1)\big) = \binom{2n-2}{n-1}$$

because

$$\begin{aligned} u &= \frac{2n-2}{2} = n-1 \\ d &= \frac{(2n-2)-0}{2} = n-1 \end{aligned}$$

$$\begin{aligned} \#\text{paths}\big((1,1) \to (2n-1,1) \text{ touching x-axis}\big) &= \#\text{paths}\big((1,-1) \to (2n-1,1)\big) \\ &= \binom{2n-2}{n} = \binom{2n-2}{n-2} \end{aligned}$$

because

$$\begin{aligned} u &= \frac{2n-0}{2} = n \\ d &= \frac{(2n-2)-2}{2} = n-2 \end{aligned}$$

$$\begin{aligned} &\#\text{paths}\big((1,1) \to (2n-1,1) \text{ above x-axis}\big) \\ &\quad = \#\text{paths}\big((1,1) \to (2n-1,1)\big) - \#\text{paths}\big((1,1) \to (2n-1,1) \text{ touching}\big) \\ &\quad = \binom{2n-2}{n-1} - \binom{2n-2}{n} \\ &\quad = \binom{2n-2}{n-1} - \frac{(2n-2)!}{n!\,(n-2)!} \\ &\quad = \binom{2n-2}{n-1} - \frac{(2n-2)!(n-1)}{n(n-1)!\,(n-1)!} \\ &\quad = \binom{2n-2}{n-1} - \frac{n-1}{n}\binom{2n-2}{n-1} \\ &\quad = \binom{2n-2}{n-1} \cdot \frac{1}{n} \end{aligned}$$