Catalan Numbers
Andreas Klappenecker

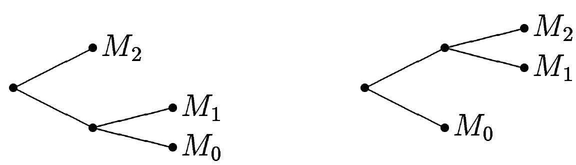
Suppose that you are given $n+1$ rectangular matrices $M_0, \ldots, M_n$, where M_i is of dimension $r_i \times r_{i+1}$. We want to find the matrix product $M_0 M_1 \ldots M_n$. In how many different ways can you build this product by multiplying two matrices at a time? Let us have a look at some small cases:

$\mathbf{n = 1}$: $(M_0 M_1)$

$\mathbf{n = 2}$: $((M_0 M_1) M_2)$, $(M_0(M_1 M_2))$

$\mathbf{n = 3}$: $(((M_0 M_1) M_2) M_3)$, $((M_0(M_1 M_2)) M_3)$, $((M_0 M_1)(M_2 M_3))$, $(M_0((M_1(M_2 M_3))$, $(M_0(M_1(M_2 M_3)))$.

Notice that the expressions above correspond to binary trees with $n+1$ leaves. For instance, in the case $n = 2$ we have the trees



It is straightforward to give a recursive formula for the number of trees. If the tree has $n+1$ leaves, $n > 0$, then the root node has two subtrees with $k+1$ leafs and $n-k$ leaves respectively. Thus the number of trees $T(n)$ with $n+1$ leaves is given by

$$T(n) = \begin{cases} 1 & \text{if } n = 0, \\ \sum_{k=0}^{n-1} T(k)T(n-k-1) & \text{otherwise.} \end{cases} \quad (1)$$

A solution in closed form is given by the **Catalan numbers** C_n, defined by

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

This can be proved with the help of **generating functions** – a powerful technique that is often helpful in the analysis of algorithms. We will study generating function in more detail later on. I will give an outline of the proof to show you the flavor of this technique.

The generating function for $T(n)$ is a formal power series

$$T(z) = \sum_{n\geq 0} T(n)z^n,$$

that comprises the information about the number of solutions for all n. The advantage of this representation is that the complicated expression (1) can be expressed simply as

$$T(z) = T(z)zT(z) + 1,$$

or

$$(T(z))^2 - \frac{T(z)}{z} + \frac{1}{z} = 0.$$

Solving this quadratic equation gives

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

We can rule out the solution corresponding to the + sign, since this would imply $T(0) = \infty$. Therefore, we have

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

It can be shown that

$$\sqrt{1-4z} = 1 - 2\sum_{n\geq 0} \frac{1}{n+1}\binom{2n}{n} z^{n+1}.$$

Therefore,

$$T(z) = \sum_{n\geq 0} T(n)z^n = \sum_{n\geq 0} \frac{1}{n+1}\binom{2n}{n} z^n.$$

And this shows that $T(n)$ is indeed equal to the Catalan number C_n.