NYT Digits Puzzle

Abstraction of problem at the root of NYT digits puzzle

[combinatorics - numbers]

We have N integers $d_1 < d_2 < \ldots < d_N$. We have four possible operations $\times$, $\div$, $+$, and $-$. What is the number of different real numbers that are greater than 1 that can be formed from combining all N integers using the available operations.

(a) Say we have two different integers. How many different real numbers (that are greater than 1) can we form from these integers? (*Note: Since we are looking for numbers that are greater than* 1*, we don't consider negative numbers or fractions less than 1*).

(b) Same question as above, but we have three different integers.

(c) What is the general number of real-numbers we can create with N integers?

(d) Solve the recursive relation in (c) using an exponential generating function $Q(x) = \sum_{\ell=1}^{\infty} x^{\ell} Q_{\ell}/\ell!$.

Solution:

1. If we have two integers, there are four different real numbers that we can create from them using the operations $\times$, $\div$, $+$, and $-$. Namely,

$$d_2 + d_1,\ d_2 - d_1,\ d_2 \times d_1,\ d_2/d_1. \tag{1}$$

Note that we are excluding the two operations $(d1/d2)$ and $d_1 - d_2$ because we require the output of subtraction to be greater than zero and the output of division to be greater than 1.

Let Q_k be the number of unique integer combinations we can make with k numbers (according to the specification that subtraction only yields non-zero integers and division only yields numbers greater than 1). We thus have that

$$Q_2 = 4. \tag{2}$$

■

2. Say we have three numbers. There are $\binom{3}{2}$ ways we can select two numbers from that set of three. When we select those two numbers, there are Q_2 ways we can generate operation combinations from them in order to yield a single number. We would then be left with the single number generated from the operation combinations and the number that we did not choose from the "3 choose 2" operations. For these two numbers, there are Q_2 ways to combine them. Therefore, Q_3 is

$$Q_3 = Q_2 \binom{3}{2} Q_2 = 48. \tag{3}$$

■

3. We now want to generalize this result to the case where we have an arbitrary N numbers to combine. To get to closer to this generalization, we consider the answer for Q_4. Akin to the result for Q_3, we note that there are $\binom{4}{3}$ ways to select three numbers from a set of four. After we select those numbers, there are Q_3 ways we can generate operation combinations that set of three. After these combinations, we would have a single number generated from the operation combination of the three numbers and the

number we did not choose from the "4 choose 3" operations. There are Q_2 ways to combine these final two numbers and thus we find that one term of Q_4 is

$$Q_4 = Q_2 \binom{4}{3} Q_3 + \cdots . \tag{4}$$

To find the other term, we do this operation again with a larger set of selected numbers. There are $\binom{4}{2}$ ways to select two numbers from a set of four. For the selected numbers there are Q_2 ways to combine them to generate a single number. For the unselected two numbers, there is also Q_2 ways to combine them into a single number. After these two separate combinations we would be left with single numbers drawn from each set, which we can then combine in Q_2 ways to yield a final result. Thus, there are $Q_2\binom{4}{2}Q_2Q_2$ ways to do this combination and the full expression for Q_4 is

$$Q_4 \stackrel{?}{=} Q_2 \binom{4}{3} Q_3 + Q_2 \binom{4}{2} Q_2 Q_2. \tag{5}$$

However, we have done a bit of double counting. Say we have the number $(1, 2, 3, 4)$. If we choose $(1, 2)$ we have $(3, 4)$ left over. Alternatively, if we choose $(3, 4)$ we have $(1, 2)$ left over. Combining $(1, 2)$ in all possible ways and then combining the result with all possible combinations of $(3, 4)$ is the same as combining $(3, 4)$ in all possible ways and then combining the result with all possible combinations of $(1, 2)$. However, the factor $\binom{4}{2}$ counts both sets of combinations as distinct. To correct for this we need to multiply it by $1/2$. We then have the result

$$Q_4 = Q_2 \binom{4}{3} Q_3 + \frac{1}{2} Q_2 \binom{4}{2} Q_2 Q_2. \tag{6}$$

We can write this result more cleanly(and pave the way towards generalization) by introducing $Q_1 = 1$. We then have

$$Q_4 = \frac{1}{2} Q_2 \left[\binom{4}{3} Q_3 Q_1 + \binom{4}{2} Q_2 Q_2 + \binom{4}{1} Q_1 Q_3 \right] = 960. \tag{7}$$

Given the form Eq.(7), we can now generalize this result to the case where we have N integers. Matching the form of the summations, we have

$$Q_N = \frac{1}{2} Q_2 \sum_{k=1}^{N-1} \binom{N}{k} Q_{N-k} Q_k, \tag{8}$$

which is a pretty clean result. If one plugs the sequence $1, 48, 960, \ldots$ into the OEIS, one finds that Q_N represents

$$Q_N = 2^{N-1} N! C_{N-1}, \tag{9}$$

where $C_N = \frac{1}{N+1}\binom{2N}{N}$ are the Catalan numbers.

The Catalan numbers are a sequence of numbers that often appear in combinatorial problems with some element of recursion. Why do they appear in the context of the NYT digits puzzle?

An explanation of their appearance comes from one of the applications of Catalan numbers described in the associated Wikipedia entry (which is itself taken from *Enumerative Combinatorics* by Richard Stanley).

> C_n is the number of different ways $n + 1$ factors can be completely parenthesized (or the number of ways of associating associating n applications of a binary operator, as in the matrix chain multiplication problem). For $n = 3$, for example, we have the following five different parenthesizations of four factors:
>
> $$((ab)c)d \quad (a(bc))d \quad (ab)(cd) \quad a((bc)d) \quad a(b(cd)) \tag{10}$$

When we are doing computations for the NYT digits game, we are essentially doing "parenthesizations" for a given ordering of numbers, where the operation that defines the parenthesization (ab) is chosen from one of the $Q_2 = 4$ operations that allow us to combine a pair of numbers. From here, the connection to Catalan numbers is more clear. We can construct Q_N the number of ways to computationally combine N numbers by multiplying the factors

- C_{N-1}: Number of ways to "parenthesize" N numbers for a given ordering
- $N!$: Number of possible orderings of N numbers
- $(Q_2)^{N-1}$: Number of operation combinations for a given "parenthesized" combination
- $1/2^{N-1}$: Combinatorial correction for double counting re-orderings of elements in a pair

which yields

$$Q_N = N! \left(\frac{Q_2}{2}\right)^{N-1} C_{N-1},$$

as we previously found.

4. We want to solve Eq.(8) using exponential generating functions. Our result should reproduce Eq.(9). First, we define the exponential generating function $Q(x)$ as

$$Q(x) \equiv \sum_{\ell=1}^{\infty} \frac{x^\ell}{\ell!} Q_\ell. \tag{11}$$

We note that $Q(x)$ satisfies $Q(x=0) = 0$. We will use this result to constrain the possible solutions to the equation. Multiplying both sides of Eq.(8) by $x^N/N!$ and summing the result from $N = 2$ to $N = \infty$, we have

$$\begin{aligned}\sum_{N=2}^{\infty} \frac{x^N}{N!} Q_N &= \frac{1}{2} Q_2 \sum_{N=2}^{\infty} \sum_{k=1}^{N-1} \frac{1}{k!(N-k)!} Q_{N-k} Q_k \, x^{N-k} x^k \\ &= \frac{1}{2} Q_2 \sum_{N=2}^{\infty} \sum_{k=1}^{\infty} \frac{1}{k!(N-k)!} Q_{N-k} Q_k \, x^{N-k} x^k,\end{aligned} \tag{12}$$

where in the second line we took the summation over k to infinity since we assume $Q_j = 0$ for $j < 1$. Introducing the identity $\sum_{\ell=1}^{\infty} \delta(N, \ell + k)$, we then obtain

$$\begin{aligned}\sum_{N=2}^{\infty} \frac{x^N}{N!} Q_N &= \frac{1}{2} Q_2 \sum_{N=2}^{\infty} \sum_{k=1}^{\infty} \sum_{\ell=1}^{\infty} \frac{1}{k!\ell!} Q_\ell Q_k \, x^\ell x^k \delta(N, k+\ell), \\ &= \frac{1}{2} Q_2 \sum_{k=1}^{\infty} \frac{1}{k!} Q_k \, x^k \sum_{\ell=1}^{\infty} \frac{1}{\ell!} Q_\ell,\end{aligned} \tag{13}$$

which (using $Q_1 = 1$) gives us the equation

$$Q(x) - x = \frac{Q_2}{2} Q(x)^2. \tag{14}$$

Solving this equation for $Q(x)$ in turn gives us

$$Q(x) = \frac{1}{Q_2} \left(1 - \sqrt{1 - 2Q_2 x}\right), \tag{15}$$

where we discarded the extraneous solution that does not yield $Q(x = 0) = 0$. Next, we expand the square root as a Taylor series. We write out this calculation in full

$$\begin{aligned}
\sqrt{1-u} &= \sum_{k=0}^{\infty} \binom{1/2}{k} (-u)^k \\
&= \sum_{k=0}^{\infty} \frac{1/2(1/2-1)(1/2-2)\cdots(1/2-k+1)}{k!}(-u)^k \\
&= \sum_{k=0}^{\infty} \frac{1}{2^k k!}(1-2)(1-4)\cdots(1-(2k-2))(-u)^k \\
&= \sum_{k=0}^{\infty} \frac{(-1)^{k-1}}{2^k k!} 1\cdot 3\cdots(2k-3)(-u)^k \\
&= -\sum_{k=0}^{\infty} \frac{1}{2^k k!}\frac{(2k-1)!!}{(2k-1)} u^k \\
&= -\sum_{k=0}^{\infty} \frac{1}{2^k k!}\frac{(2k)!}{2^k k!(2k-1)} u^k, \qquad (16)
\end{aligned}$$

which yields

$$\sqrt{1-u} = -\sum_{k=0}^{\infty} \frac{1}{2k-1}\binom{2k}{k}\left(\frac{u}{4}\right)^k. \qquad (17)$$

This result makes Eq.(15) become

$$\begin{aligned}
Q(x) &= \frac{1}{Q_2}\left(1 + \sum_{k=0}^{\infty} \frac{1}{2k-1}\binom{2k}{k}\left(\frac{Q_2 x}{2}\right)^k\right) \\
&= \frac{1}{Q_2}\sum_{k=1}^{\infty} \frac{1}{2k-1}\binom{2k}{k}\left(\frac{Q_2 x}{2}\right)^k. \qquad (18)
\end{aligned}$$

Some combinatorial manipulation gives us

$$\frac{1}{2k-1}\binom{2k}{k} = \frac{2}{k}\binom{2k-2}{k-1} = 2C_{k-1}, \qquad (19)$$

where C_k are the Catalan numbers. With this new combinatorial coefficient, Eq.(18) becomes

$$Q(x) = \sum_{k=1}^{\infty} C_{k-1}\left(\frac{Q_2}{2}\right)^{k-1} x^k. \qquad (20)$$

And from equating this result to Eq.(15), we can thus infer

$$Q_N = N!\left(\frac{Q_2}{2}\right)^{N-1} C_{N-1}, \qquad (21)$$

where C_{N-1} are the Catalan Numbers. The closed form expression for Q_N is in turn

$$Q_N = \left(\frac{Q_2}{2}\right)^{N-1}\frac{(2N-2)!}{(N-1)!}, \qquad [\#\text{ of } N \text{ number calculation combinations}] \qquad (22)$$

which has the asymptotic behavior $Q_N \sim O\left((2Q_2)^N\right)$. In the case of the NYT digits game (for which $Q_2 = 4$), we have $Q_N \sim O(2^{3N})$.

■