Catalan Numbers in Process Synthesis

Mahnoosh Shoaei, J. T. Sommerfeld
School of Chemical Engineering
Georgia Institute of Technology
Atlanta, GA 30332

In one of the early papers on process synthesis, Thompson and King (1972) present a closed-form equation for determining the number of possible sequences for separating N components from each other into single-component products. This note shows that this determination is a simple application of Catalan numbers, which date back to 1751 in the mathematics field.

Separation Sequences

Thompson and King showed that the number of possible sequences S for separating N components from each other into single-component products, using only one simple method of separation (e.g., simple binary distillation), is given by the following closed-form equation:

$$S = \sum_{j=1}^{N-1} S_j S_{N-j} = \frac{[2(N-1)]!}{N!(N-1)!} \tag{1}$$

Thus, for merely a two-component mixture (AB), only one separation sequence is possible namely, A/B. For a three-component mixture (ABC), there are two distinct separation sequences: AB/C, followed by A/B; and A/BC followed by B/C. The five possibilities for a four-component mixture ($ABCD$) are listed below:

1. A/BCD, B/CD, C/D
2. A/BCD, BC/D, B/C
3. AB/CD, A/B, C/D
4. ABC/D, A/BC, B/C
5. ABC/D, AB/C, A/B

The number of possible sequences increases rapidly as the number of components increases, as can be seen from the data of Table 1, from Henley and Seader (1981).

Equation 1 can also be generalized to accommodate multiple separation methods. Thus, Nishida and coauthors (1981) present the following formula:

$$S = \frac{[2(N-1)]!}{N!(N-1)!} M^{N-1} \tag{2}$$

for computing the number of possible separation sequences S for a mixture of N components to be separated into N pure-component products, using M separation methods.

Catalan Numbers

The sequence of numbers appearing in the second column of Table 1—1, 2, 5, 14, 42, 132, 429, . . .—occurs frequently in the field of mathematics, particularly in combinatorial problems; they are known as Catalan numbers (a_n). One of the formulas for the nth Catalan number is (Alter, 1971):

$$a_n = \frac{(2n-2)!}{n!(n-1)!} \tag{3}$$

Other equivalent expressions are:

$$a_n = \frac{1}{n}\binom{2n-2}{n-1} \tag{4}$$

and:

$$a_n = \frac{1}{2n-1}\binom{2n-1}{n-1} \tag{5}$$

The above three expressions are valid for integer values of $n \geq 1$. Successive Catalan numbers can also be generated by recursion formulas, invoking previous Catalan numbers. Thus, with $a_1 = 1$ and for $n \geq 2$:

$$a_n = \sum_{i=1}^{n-1} a_i a_{n-i} \tag{6}$$

and:

$$a_n = \left(\frac{4n-6}{n}\right) a_{n-1} \tag{7}$$

Some authors choose to begin the indexing of the Catalan sequence with $n = 0$. In this case, expressions for the Catalan

Table 1. Number of Separation Sequences Using Only One Simple Method of Separation

N, No. of Components	S, No. of Sequences
2	1
3	2
4	5
5	14
6	42
7	132
8	429
9	1,430
10	4,862
11	16,796

numbers are:

$$a_n = \frac{1}{n+1}\left[\frac{(2n)!}{n!n!}\right] \qquad (8)$$

and:

$$a_n = \frac{1}{n+1}\binom{2n}{n} \qquad (9)$$

and the sequence is 1, 1, 2, 5, 14, 42, . . .

In any event, it is clear that Eqs. 1 and 3 are identical. The conclusion is that the number of possible sequences for separating N components ($N \geq 2$) from each other into single-component products, using only one simple method of separation, is equal to the ($N - 1$)th Catalan number.

It was the German mathematician Leonhard Euler who first discovered the Catalan number sequence in 1751. He posed the following question: In how many ways can a convex n-sided polygon (n-gon) be divided into triangles by drawing diagonals that do not intersect? In every case, regardless of how the n-gon is triangulated, the number of diagonals is always $n - 3$ and the number of triangles is $n - 2$. For $n \geq 3$, the answer to this question is given by the ($n - 2$)th Catalan number. Thus, the solutions are 1 for a triangle, 2 for a quadrilateral, 5 for a pentagon, 14 for a hexagon, and so on. The name of this sequence of numbers, however, actually derives from the Belgian mathematician Eugene Charles Catalan. In 1838, he found this sequence to be the solution to the problem of how many ways a chain of n alphabetical letters can be parenthesized.

Forder (1961) presents an interesting way of visualizing and constructing Catalan numbers, which is shown in the triangular diagram of Figure 1. The construction of this triangular matrix begins with the insertion of values of unity at all locations in the first or top row. Each of the remaining entries is then formed as the sum of all elements in the row immediately above, from the leftmost or diagonal element to the element immediately above, inclusively. More simply, a given interior element can be formed as the sum of the element immediately to its left (west) and the element immediately above (north), as indicated in Figure 1. The sequence of Catalan numbers appears along the main diagonal of this triangular matrix. Again, as shown in Figure 1, these diagonal elements are merely the sums of the first two elements in the row immediately above.

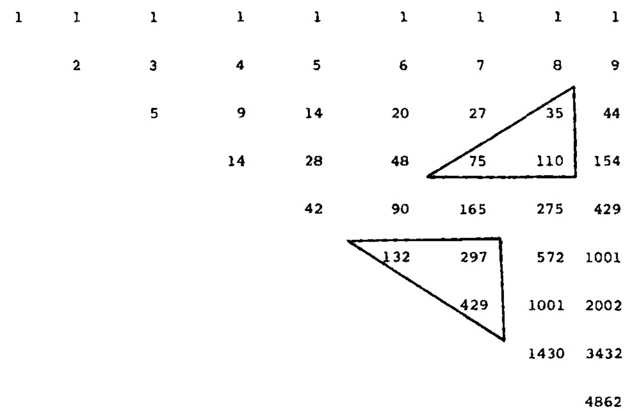

Figure 1. Triangular representation of Catalan numbers.

Gardner (1976) points out an interesting property in the Catalan number sequence. Namely, odd Catalan numbers appear at all positions, and only at those positions, that are powers of two. Thus, the 1st, 2nd, 4th, 8th, 16th, and so on Catalan numbers are odd.

Applications

Alter (1971) states that more than 200 papers have been written on the Catalan numbers, due to their frequent appearance in various combinatorial problems. These applications range from election possibilities and postage stamps to geometry (e.g., Euler's original problem) to continued fractions and Legendre polynomials. The Catalan number sequence is really a special case of the class of problems known as Ballot problems, which originated in the year 1887. Other important applications of these numbers are in the theory of random walks, in order statistics, in enumerating decision patterns, and in queuing theory.

Many of the more recent applications of Catalan numbers derive from graph theory, which has enjoyed considerable rejuvenation in many fields, including chemical engineering. Gardner (1976) observers that Catalan numbers really count the number of trees that are planar, trivalent, and planted. A tree is a connected graph (vertices joined by edges) that has no circuits. Planar means that it is drawn on the oriented plane without intersections. Planted means that it has one trunk, the end of which is called the root (the degree of the root is one). Thus, the graph can be drawn to simulate a tree growing up from the ground. Trivalent means that at each vertex (except at the root and at the ends of the branches) the tree forks to create a point where three edges meet; alternatively, a tree is trivalent if the degree of every vertex is either one or three.

Fibonomial Catalan Numbers

Gould (1972) has developed an analogy between the Catalan and Fibonacci numbers. Applications of the latter to the solution of optimization problems in chemical engineering have been discussed by Wilde (1964), for example. The Fibonacci sequence of numbers (0, 1, 1, 2, 3, 5, 8, . . .) is simply defined by the following formula:

$$F_n = F_{n-1} + F_{n-2} \qquad (10)$$

and $F_0 = 0$, $F_1 = 1$. Fibonomial coefficients are then defined by:

$$\begin{Bmatrix} n \\ k \end{Bmatrix} = \frac{F_n F_{n-1} \cdots F_{n-k+1}}{F_k F_{k-1} \cdots F_1} \tag{11}$$

where:

$$\begin{Bmatrix} n \\ o \end{Bmatrix} = 1 \tag{12}$$

Specifically then, Gould demonstrates that the expression:

$$\frac{1}{F_{n+1}} \begin{Bmatrix} 2n \\ n \end{Bmatrix} = \frac{1}{F_{n+1}} \left[\frac{F_{n+1} F_{n+2} \cdots F_{2n}}{F_1 F_2 \cdots F_n} \right] \tag{13}$$

generates a sequence of integer numbers, namely, 1, 1, 3, 20, 364, 17017, . . . , known as the Fibonomial Catalan sequence. This sequence is an exact analog of the Catalan sequence, as comparison of Eqs 9 and 13 readily indicates.

Acknowledgment

The authors wish to acknowledge the role of Lawrence Riddle, Department of Mathematics and Computer Science, Emory University, in first calling our attention to the Catalan number sequence.

Literature Cited

Alter, R., "Some Remarks and Results on Catalan Numbers," *Proc. 2nd Louisiana Conf. Combinatorics, Graph Theory, and Computing,* Baton Rouge, 109 (1971).

Forder, H. G., "Some Problems in Combinatorics," *Math. Gazette,* **45,** 199 (1961).

Gardner, M., "Mathematical Games," *Sci. Amer.,* **244**(6), 120 (June, 1976).

Gould, H. W., "A New Primality Criterion of Mann and Shanks and Its Relation to a Theorem of Hermite with Extension to Fibonomials," *Fibonacci Q.,* **10,** 355 (1972).

Henley, E. J., and J. D. Seader, *Equilibrium-Stage Separation Operations in Chemical Engineering,* Wiley, New York (1981).

Nishida, N., G. Stephanopoulos, and A. W. Westerberg, "A Review of Process Synthesis," *AIChE J.,* **27,** 321 (1981).

Thompson, R. W., and C. J. King, "Systematic Synthesis of Separation Schemes," *AIChE J.,* **18,** 941 (1972).

Wilde, D. J., *Optimum Seeking Methods,* Prentice-Hall, Englewood Cliffs, NJ (1964).

Manuscript received Nov. 25, 1985.