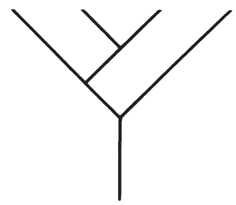
C H A P T E R

TWENTY



Catalan Numbers

If an infinite sequence of positive integers is simple enough, such as the doubling series (1, 2, 4, 8, 16, . . .) or the squares (1, 4, 9, 16, 25, . . .), it is easily recognized. And few mathematicians would fail to recognize the Fibonacci numbers (1, 1, 2, 3, 5, 8, . . .) or the triangular numbers (1, 3, 6, 10, 15, 21, . . .). If the sequence is unfamiliar, however, an enormous amount of time can be wasted searching for a recursive or nonrecursive procedure that generates the sequence. (A procedure is recursive if calculating a next term calls for knowledge of the preceding terms; a nonrecursive formula gives the nth term without such knowledge.)

It is hard to believe, but it was not until 1973 that *A Handbook of Integer Sequences* (Academic Press, 1973) was published. This invaluable tool, compiled by N. J. A. Sloane of Bell Laboratories, lists more than 2300 integer sequences in numerical order. A mathematician who encounters a puzzling sequence no longer needs to spend hours trying to find its generating formula. He simply looks for the sequence in Sloane's book. The chances are excellent that it is there, followed by a list of references where the reader can check on the nature of the beast.

Our topic here is the *Handbook's* sequence 577: 1, 2, 5, 14, 42, 132, 429, 1430, 4862, 16796, ... The components of this sequence are called Catalan numbers. They are not as well known as Fibonacci numbers, but they have the same delightful propensity for popping up unexpectedly, particularly in combinatorial problems. In 1971 Henry W. Gould, a mathematician at West Virginia University, privately issued a bibliography of 243 references on Catalan numbers; in many cases the authors were not even aware they were dealing with a sequence known for more than two centuries. In 1976 Gould increased the number of references to 450. Indeed, the Catalan sequence is probably the most frequently encountered sequence that is still obscure enough to cause mathematicians, lacking access to Sloane's *Handbook,* to expend inordinate amounts of energy rediscovering formulas that were worked out long ago.

It was Leonhard Euler who first discovered the Catalan numbers after asking himself: In how many ways can a fixed convex polygon be divided into triangles by drawing diagonals that do not intersect? An example can be provided with triangles, quadrilaterals, pentagons, and hexagons (*see* Figure 125). Note that 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 $n - 2$. It is easy to prove that this relation holds in general. The number of possible triangulations for each of these four polygons are the first four terms of the Catalan sequence.

Applying an induction process that he described as "quite laborious," Euler obtained the following recursive formula:

$$\frac{2 \times 6 \times 10 \times \cdots (4n - 10)}{(n - 1)!}$$

Numbers above the line have the form $(4n - 10)$, where n is a positive integer greater than 2. The exclamation mark is of course the factorial sign. It stands for the product of all positive integers from 1 through the preceding expression. For example, if $n = 6$ (the sides of a hexagon), the formula becomes

$$\frac{2 \times 6 \times 10 \times 14}{5!} = 14$$

Unusually simple recursive formulas are obtained by putting another 1 in front of the series: 1, 1, 2, 5, 14, ... Let k be the last number of a partial sequence and n the position of the next number. The next number is then

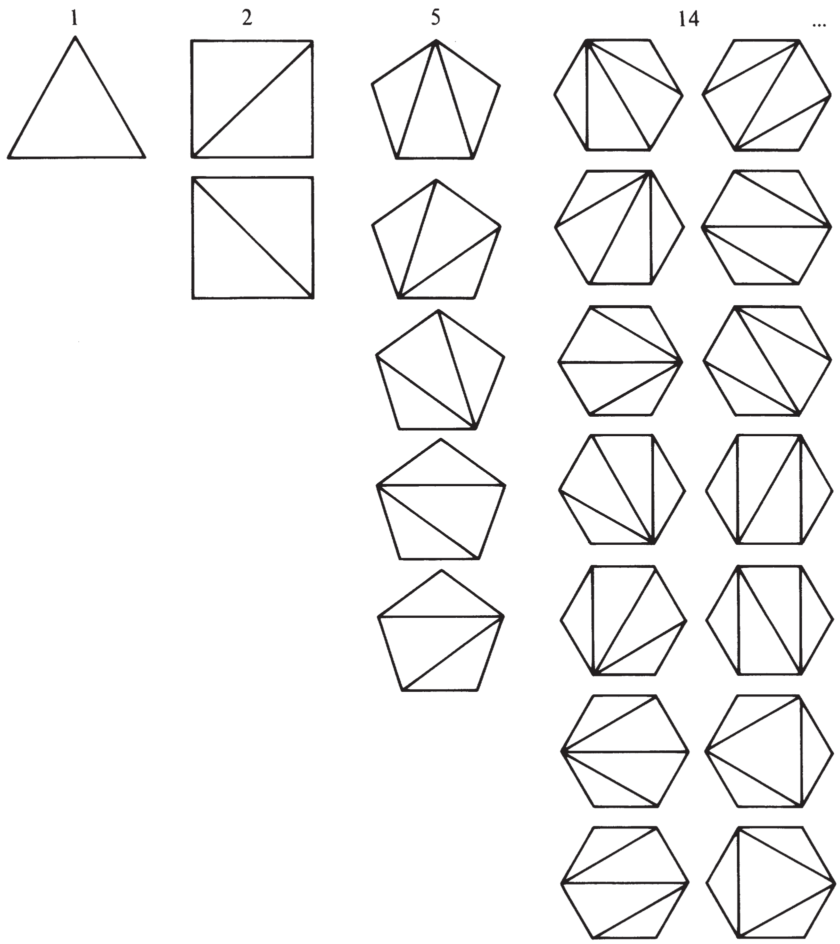


Figure 125 Leonhard Euler's polygon-triangulation problem

$$\frac{k(4n-6)}{n}$$

Johann Andreas von Segner, Euler's eighteenth-century contemporary, found a whimsical recursive procedure for the same form of the Catalan sequence. Write the partial sequence forward, then put below it the same numbers in backward order. Multiply each top number by the one below it and add all the products; the result is the next number of the sequence. For example,

$$\begin{array}{rrrrr} 1 & 1 & 2 & 5 & 14 \\ \times\underline{14} & \underline{5} & \underline{2} & \underline{1} & \underline{1} \end{array}$$
$$14 + 5 + 4 + 5 + 14 = 4\dot{2}$$

Euler's polygon triangulation is isomorphic with many seemingly unrelated problems. It was Eugene Charles Catalan, the Belgian mathematician for whom the sequence is named, who in 1838 solved the following problem. We have a chain of n letters in a fixed order. We want to add $n - 1$ pairs of parentheses so that inside each pair of left and right parentheses there are two "terms." These paired terms can be any two adjacent letters, or a letter and an adjacent parenthetical grouping, or adjacent groupings. In how many ways can the chain be parenthesized?

For two letters, *ab,* there is only one way: (*ab*). For three letters there are two ways: ((*ab*)*c*) and (*a*(*bc*)). For four letters there are five ways: ((*ab*)(*cd*)), (((*ab*)*c*)*d*), (*a*(*b*(*cd*))), (*a*((*bc*)*d*)), and ((*a*(*bc*))*d*). The numbers of these ways, 1, 2, and 5, are the first three Catalans, and the Catalan sequence enumerates the ways of parenthesizing all longer chains.

H. G. Forder, writing on Catalan numbers in 1961, showed a simple way to establish one-to-one correspondence between the triangulated polygons and the parenthesized expressions. An example is a triangulated heptagon (*see* Figure 126). Label its sides (excluding the base) *a* through *f*. Every diagonal spanning adjacent sides is labeled with the letters of those sides in parentheses. Each remaining diagonal is then lettered in similar fashion by combining the labels on the other two sides of the triangle. The base is lettered last. The

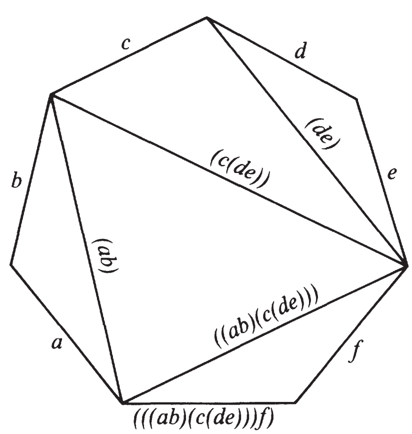


Figure 126 Parenthesized triangulation of a heptagon

expression for the base is uniquely determined by the dissection. If you apply this technique to the polygons portrayed in Figure 125, you will obtain the parenthesized expressions shown at the right in Figure 127.

The British mathematician Arthur Cayley proved that Catalan numbers count the number of trees that are planar, trivalent, and planted. A "tree" is a connected graph (points joined by edges) that has no circuits. "Planar" means that it can be drawn on the plane without intersections. "Planted" means that it has one "trunk," the end of which is called the "root." The graph can thus be drawn to simulate a tree growing up from the ground. "Trivalent" means that at each point (except at the root and at the ends of branches), the tree forks to create a spot where three edges meet.

The illustration of this structure (Figure 127) is almost self-explanatory. The gray lines show how each triangulation corresponds to a planted trivalent tree. Next to the polygons, corresponding trees are drawn in conventional form. It is easy to see how the grouping of a tree's branches corresponds to its parenthesized expression. Below each expression, we convert it to a binary number by replacing every left-hand parenthesis with 1 and every letter with zero, ignoring all right-hand parentheses. These binary numbers are convenient shorthand ways of designating the polygon dissections and their trees. Right-hand parentheses are not needed because, given the left-hand ones and the method of grouping letters, the right-hand parentheses can always be added in a unique manner.

The Polish mathematician Jan Lukasiewicz found a pleasant way to obtain each tree's binary number (*see* Figure 128). Picture a tree with four top ends. They are labeled 0 and the trivalent points are labeled 1. Imagine a worm crawling up the trunk and around the entire tree along the broken path in the illustration. At each point, the worm calls out the label. Once a point is called, it is not called again. In this example the worm calls out 1101000, which proves to be the very binary number we obtained from the tree's parenthesized expression.

In 1964 it was discovered that normal planted trees are also counted by Catalan numbers. They are planted trees of n points, including the ends but not the root. They can also be described as planted trees of n edges. A point in such a tree can have any valence.

Many ways have been found for showing a one-to-one correspondence between trees of this kind and the planted trivalent trees. The simplest one was pointed out by Frank Bernhart (*see* Figure 129). The trivalent trees are drawn so that at each point of valence 3 the edges go up, down, and to the right. Imagine that each horizontal edge shrinks to a point and disappears. If there is a trivalent point at the right end of the edge, it is carried to the left to merge

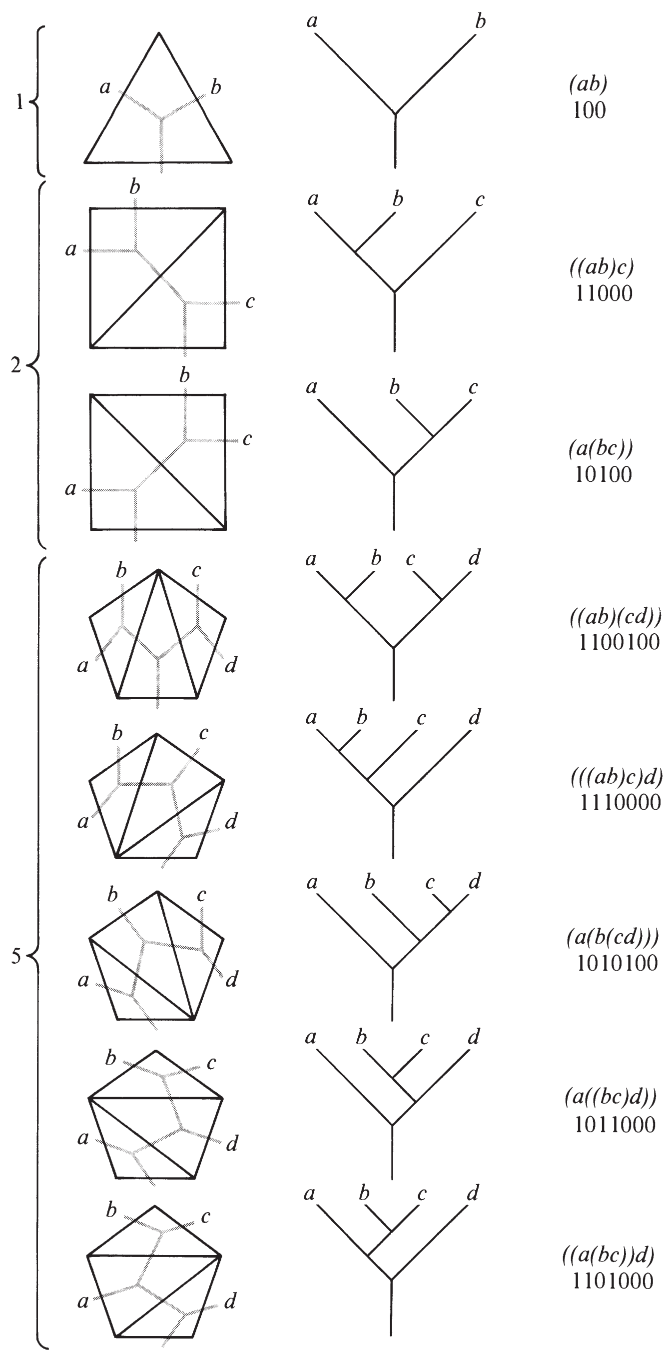


Figure 127 Triangulation and planted trivalent trees

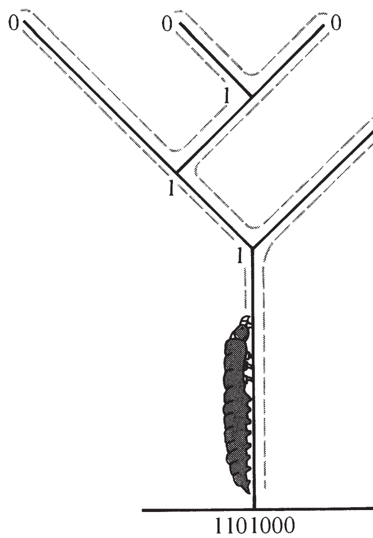


Figure 128 A worm generating a tree's binary number

with the point at the left. All the vertical edges remain distinct. This simple transformation changes all planted trivalent trees of n ends into all planted trees of n edges.

A worm crawling up and around any tree in this new set (the trees at the right in the illustration) will call out the same binary number as the tree's partner if it alters its procedure as follows. The worm starts at the bottom point rather than at the root. Each time it crawls up an edge, it calls 1, and each time it crawls down an edge, it calls 0.

Consider chessboards of sides 2, 3, 4, ... All squares north and west of a main diagonal are shaded (*see* Figure 130). We are to move the rook from the lower left corner to the upper right corner. It cannot enter a shaded cell, and its only allowed movements are north or east. For a board of side n, how many different paths can the rook take?

Once more the Catalans give the answer. Below each board of side n write the binary number for the planted trivalent tree of n ends. Taking the binary digits from left to right, move the rook one square to the right for each 1 and one square up for each 0. (The final digit is ignored.) This pattern generates a path, and in this way all the rook paths are obtained.

Here are seven more recreational problems solved by the Catalans. For the first five, I shall indicate how the corresponding binary numbers (ignoring final digits) solve the problem.

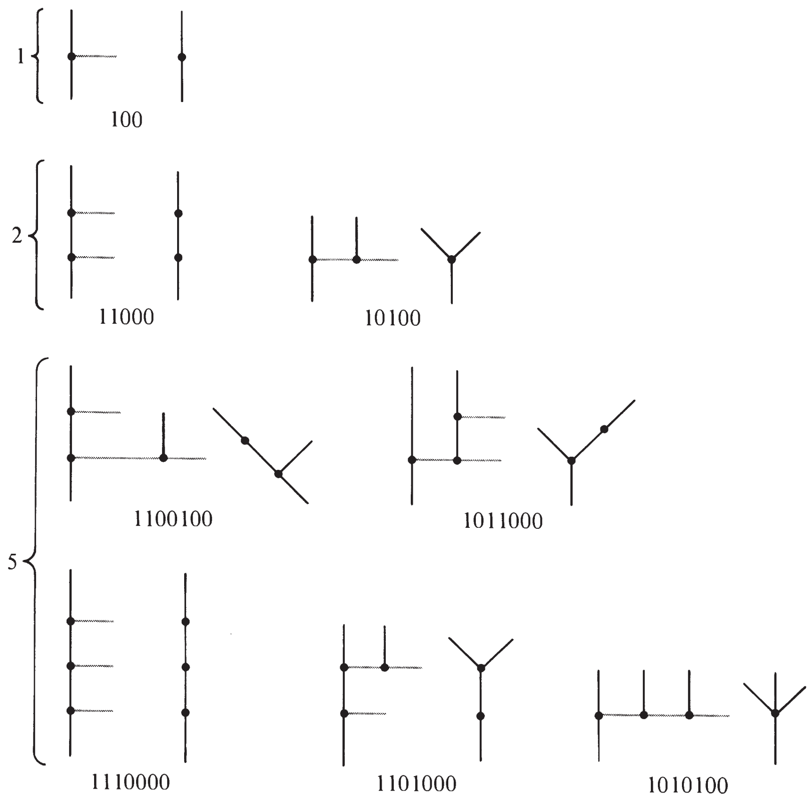


Figure 129 Transformation of planted trivalent trees to planted normal trees

1. Two men, A and B, are running for office. Each man gets n votes. In how many ways can the $2n$ votes be counted so that at no time is A behind B? ($1 =$ vote for A, $0 =$ vote for B.)

2. Place a penny, a nickel, and a dime in a row. On the penny put a stack of n face-up playing cards with values in consecutive order from bottom to top. The cards are moved one at a time from the penny to the nickel or from the nickel to the dime. (No other moves are allowed.) By mixing these two types of moves, you will end, after $2n$ moves, with all the cards on the dime. Given n cards, how many different permutations can you achieve on the dime? ($1 =$ move from penny to nickel, $0 =$ move from nickel to dime.)

3. An inebriated man leaves the door of a bar and staggers straight ahead. His steps are equal, but before each step, he has a random choice of going forward or backward. How many ways can he take $2n$ steps that will return him to the door? ($1 =$ step forward, $0 =$ step back.)

This random walk can be given other forms. A king starts on the first row of a chessboard and moves one square forward or back along the file to end on the starting square after $2n$ moves. Draw a space-time diagram of the moves, with time measured along the horizontal base line. The zigzag path can be viewed as the profile of a mountain with peaks an integral number of miles high and a base length of $2n$ miles. The paths depict all mountain ranges of this type.

4. An even number ($2n$) of soldiers, no two the same height, line up in two equal rows, A and B. How many ways can they do it so that from left to right in each row the heights are in ascending order and each soldier in row B is taller than his counterpart in row A? (Number the soldiers 1, 2, 3, . . . according to increasing height, and number the digits of the binary numbers from left to right. The 1 digits give the numbers for row A, and the 0 digits give the numbers for row B. The problem is easily modeled with playing cards.)

5. Tickets are 50 cents, and $2n$ customers stand in a queue at the ticket window. Half of them have \$1 each and the others have 50 cents each. The cashier starts with no money. How many arrangements of the queue are possible with the proviso that the cashier always be able to make change? (1 = 50 cents, 0 = \$1.)

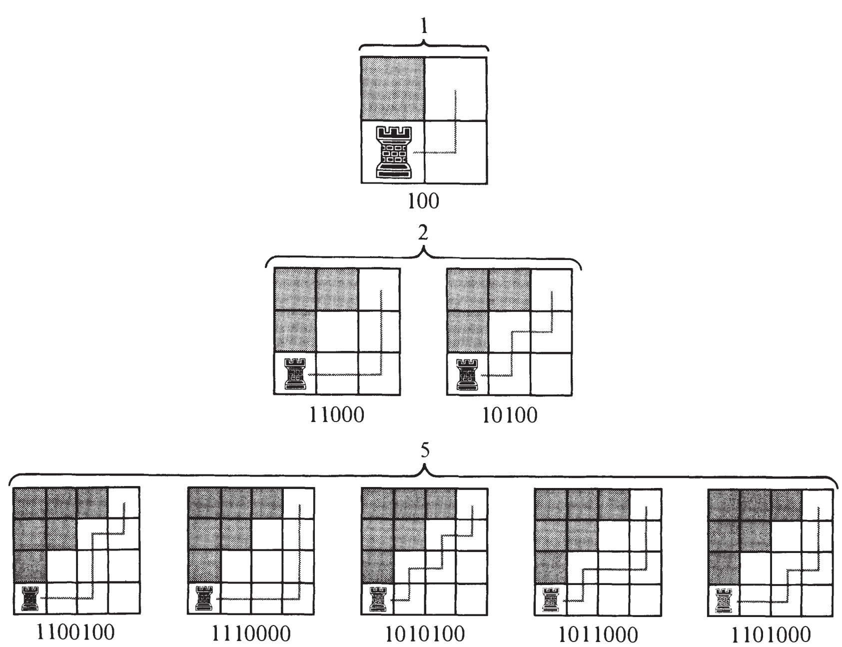


Figure 130 How Catalan numbers count a rook's paths

6. Hexaflexagons are curious toys made by folding straight or crooked strips of paper into a hexagonal structure that alters its "faces" when it is flexed. (They are described in the first chapter of my *Scientific American Book of Mathematical Puzzles & Diversions.* Simon & Schuster, 1959.) A regular hexaflexagon of a specified type passes through different states as it is flexed. The total number of states, for all varieties of regular hexaflexagons of n faces, is a Catalan number. For example, a hexahexaflexagon (six faces) can be made in three ways. The total number of states is the Catalan number 42.

If we ignore the states and ask in how many essentially different ways a regular hexaflexagon of n faces can be made, the answer is provided by a sequence that counts the triangulations of convex polygons when rotations and reflections are excluded. This remarkable sequence (No. 942 in Sloane's *Handbook*) is 1, 3, 4, 12, 27, 82, 228, 733, 2282, 7528, ...

In unpublished papers, Bernhart and other flexagation addicts describe ways of mapping the changes of states, as a flexagon of n faces is flexed, by tracing paths around the lines of a triangulated polygon of $n + 1$ sides.

7. An even number of people are seated around a circular table. Each extends one arm and they clasp hands in pairs, but in such a way that no pair of joined arms crosses another. Given the number of pairs, in how many ways can this be done? More precisely, place $2n$ spots in fixed positions on the circumference of a circle and then find all the ways they can be paired by drawing nonintersecting chords.

Can you find a simple geometric way to establish a one-to-one correspondence between this problem and any of the above problems?

A nonrecursive formula for the nth Catalan has different forms depending on how the positions of the Catalans are numbered. The formula is simplest if the sequence begins 1, 2, 5,. . . . In this numbering, the nth Catalan is

$$\frac{(2n)!}{n!(n+1)!}$$

If the series begins 1, 1, 2, 5, . . . , it turns out that odd Catalan numbers greater than 1 appear at all positions, and only at those positions, that are powers of two. Thus the fourth, eighth, sixteenth, and so on Catalans are odd. This is only one of many unusual properties of the sequence that have been discovered.

A word of caution: When one works on combinatorial problems, it is easy to confuse the Catalan sequence with a closely related one: 1, 2, 5, 15, 52, 203, 877, ... As Gould points out in notes on his bibliography (which also includes a separate listing of references on the above series), when structures are complicated, it is easy to miss a fifteenth structure (when $n = 4$) and to be tricked into supposing you have encountered a Catalan sequence. The numbers are called Bell numbers after Eric Temple Bell, who published a lot about them. They count the partitions of n elements. For example, the number of rhyme schemes for a stanza of n lines is a Bell number. A quatrain has fifteen possible rhyme schemes. A 14-line sonnet, if convention is thrown to the winds, can have 190,899,322 (the fourteenth Bell) distinct rhyme schemes. But, you may object, who would write a sonnet with a rhyme scheme such as *aaaaaaaa aaaaaa?* Allowing a word to rhyme with itself, James Branch Cabell conceals just such a sonnet (each line ending with "love") in Chapter 14 of *Jurgen* (Grosset and Dunlap, 1919). I would guess it no accident that Cabell's 14-line poem starts as the fourteenth paragraph of Chapter 14.

Figure 131 shows how Bell numbers count the rhyme schemes for stanzas of one line through four lines. Lines that rhyme are joined by curves. Note that not until we get to quatrains does a pattern (No. 8) require an intersection. Joanne Growney, who worked this arrangement out in 1970 for her doctoral thesis, calls the schemes that do not force an intersection of curves "planar rhyme schemes." Bell numbers count all rhyme schemes. Catalan numbers are a sub-sequence that counts planar rhyme schemes.

The Bell sequence is No. 585 in Sloane's *Handbook*. But the Bells chime another story that we must postpone for a future book.

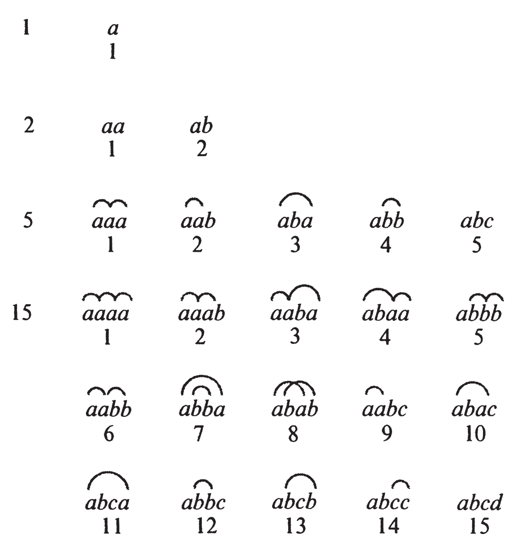


Figure 131 How Bell numbers count rhyme schemes

ANSWERS

The problem was to show that if $2n$ spots on a circle are joined in pairs by nonintersecting chords, the number of ways of doing so are counted by the Catalan numbers. Figure 132 shows how Frank Bernhart establishes a one-to-one correspondence of the chord patterns with planted normal trees of $n + 1$ edges. It also gives their binary numbers. Since these trees are counted by Catalan numbers, as I explained, the same sequence counts the chord patterns. To make the diagrams easier to interpret, the small chords are shown curved.

Imagine that the circles and chords are elastic strings. Break each circle at the top and bend it into a straight line. The chord problem then becomes: $2n$ points on the line are paired in all possible ways by joining them with curves above the line that do not intersect. This is equivalent to finding all "planar rhyme schemes" for $2n$ lines that consist of n couplets, allowing coupled lines to be separated by other lines.

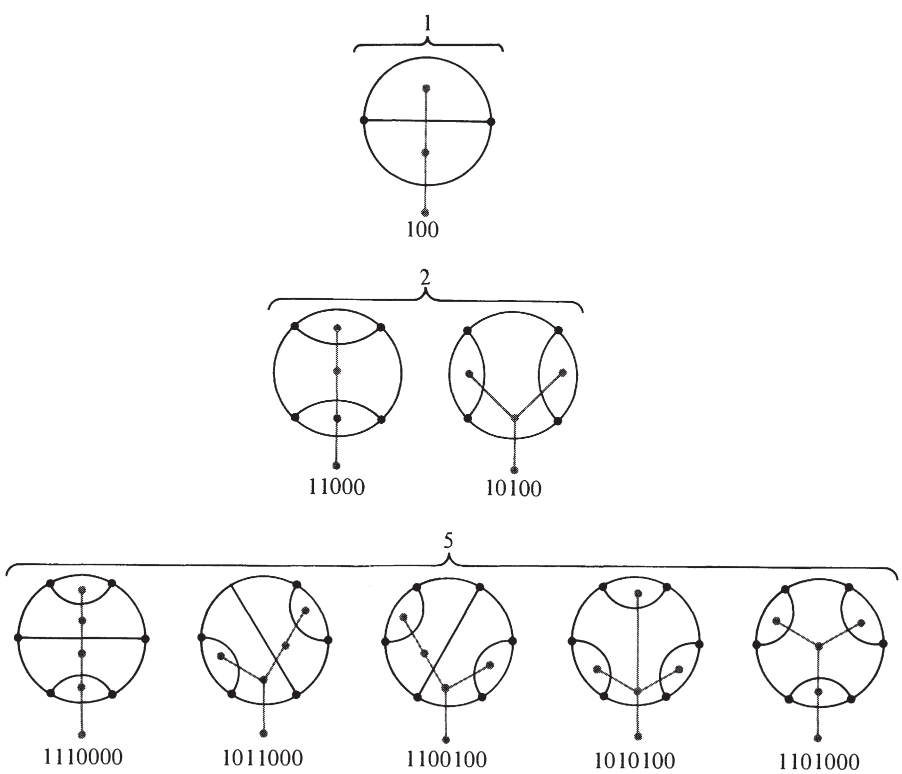


Figure 132 Correspondence between nonintersecting chords and planted trees

If the line is now closed by bringing its ends together *below,* we get an inside-out version of the original problem. It could represent a lake with $2n$ houses on the perimeter, paired in all possible ways by nonintersecting paths. (All paths joining a given pair of houses are assumed to be the same; think of each pair as being joined by an elastic string that can be lifted up, stretched or shrunk and replaced anywhere on the plane outside the lake.)

To obtain the binary numbers, imagine a worm at the bottom of each circle, inside the circle and facing west. As the worm crawls clockwise around the circle, it calls out 1 when it encounters a chord for the first time and 0 when it encounters a chord a second time. The procedure works in reverse. Given the binary number, the worm labels the spots 1 and 0. There will be only one way to join each 1 to a 0 without an intersection of chords.

ADDENDUM

My column on Catalan numbers produced so many letters telling me about other applications of the numbers, and other properties, that I can mention only a few of special interest.

Vern Hoggatt, Jr., who edited *The Fibonacci Quarterly,* explained how easily the Catalans can be found in Pascal's famous number triangle. Merely go down the center column (1, 2, 6, 20, 70, . . .) and from each number subtract the adjacent number (numbers on left and right of the central number are, of course, the same). Result: the Catalan sequence!

Paul Stockmeyer called my attention to Jack Levine's "Note on the Number of Pairs of Non-Intersecting Routes" (Scripta Mathematica 24, Winter 1959, pp. 335–338). Stockmeyer's colorful interpretation of Levine's result is to imagine two people at the same intersection of a square grid. Each simultaneously goes one step, randomly choosing to go either north or east. When their paths intersect they outline a polyomino. The Catalans count the number of distinct polyominoes that can be formed after each person has taken n unit steps. The theorem underlies many later papers, such as "A Catalan Triangle," by Louis Shapiro (*Discrete Mathematics* 14, 1976, pp. 83–90.)

Shapiro also sent me his paper, "A Short Proof of an Identity of Touchard's Concerning Catalan Numbers" (*The Journal of Combinatorial Theory,* 20, May 1976, pp. 375–376.) Here is Shapiro's interpretation of the identity. Put n points on a circle. Each point is either painted red, or green, or joined by a line to another point. The lines must not cross. Catalan numbers count the number of different patterns for each n.

My column on the related sequence of Bell numbers ran in *Scientific American* (May 1978) and will be included in a later book collection.

B I B L I O G R A P H Y

"Some Problems in Combinatorics." H. G. Forder, in *The Mathematical Gazette* 45, October 1961, pp. 199–201.

"A Note on Plane Trees." N. G. de Bruijn and B. J. M. Morselt, in *Journal of Combinatorial Theory* 2, January 1967, pp. 27–34.

"Prime and Prime Power Divisibility of Catalan Numbers." Ronald Alter and K. K. Kubota, in *Journal of Combinatorial Theory* 15, November 1973, pp. 243–256.

"Catalan Structures and Correspondences." Mike Kuchinski. Master's thesis (University of West Virginia), privately published by the author in Morgantown, West Virginia, 1977.

"The Computation of Catalan Numbers." Douglas Campbell, in *Mathematics Magazine* 57, September 1984, pp. 195–208.