7

Catalan Numbers

Author: Thomas A. Dowling, Department of Mathematics, Ohio State University.

Prerequisites: The prerequisites for this chapter are recursive definitions, basic counting principles, recurrence relations, rooted trees, and generating functions. (See Sections 3.4, 4.1, 6.1, 6.2, 6.4, and 8.1 of *Discrete Mathematics and Its Applications*, Fifth Edition, by Kenneth H. Rosen.)

Introduction

Sequences and arrays whose terms enumerate combinatorial structures have many applications in computer science. Knowledge (or estimation) of such integer-valued functions is, for example, needed in analyzing the complexity of an algorithm. Familiar examples are the polynomials in n, exponential functions (with an integer base) with exponent a polynomial in n, factorials, and the binomial coefficients. Less familiar are the Stirling numbers considered elsewhere in this book. The sequence of positive integers to be met here, called the Catalan* numbers, enumerate combinatorial structures of many different types. Those include nonnegative paths in the plane, well-formed sequences of parentheses, full binary trees, well-parenthesized products of variables, stack permutations, and triangulations of a convex polygon.

After defining the Catalan numbers explicitly by formula, we will show by a *combinatorial argument* that they count nonnegative paths in the plane. The size of each set of structures subsequently considered is shown to be a Catalan number by establishing a one-to-one correspondence between that set and a set of structures shown earlier to be of that size.

* Eugène Charles Catalan (1814–1894) was a prominent Belgian mathematician who had numerous publications on multiple integrals, general theory of surfaces, mathematical analysis, probability, geometry, and superior arithmetic.

Using *recurrence relations* and *generating functions*, we will show by a different approach that the size of one of these sets (well-parenthesized products of variables) is a Catalan number. All the sets considered are then enumerated by Catalan numbers in view of the existence of the one-to-one correspondences previously established. The asymptotic behavior of the sequence is investigated, and we obtain the order of magnitude of the nth Catalan number.

Paths and Catalan Numbers

Suppose $m = a + b$ votes were cast in an election, with candidate A receiving a votes and candidate B receiving b votes. The ballots are counted individually in some random order, giving rise to a sequence of a As and b Bs. The number of possible ballot sequences is the number $C(m, a)$ of a-element subsets of the set $\{1, 2, \ldots, m\}$, since each such subset indicates which of the m ballots were cast for candidate A. Assuming all such sequences are equally likely, what is the probability that candidate A led throughout the counting of the ballots?

Every sequence of m ballots can be represented by an ordered m-tuple $(x_1, x_2, \ldots, x_m)$ with $x_i = 1$ if the ith ballot was a vote for A and $x_i = -1$ if it was for B. Then after i ballots are counted, the ith *partial sum* $s_i = x_1 + x_2 + \ldots + x_i$ (with $s_0 = 0$) represents A's "lead" over B. If we denote by $P(a, b)$ the number of sequences $(x_1, x_2, \ldots, x_m)$ with the partial sums $s_i > 0$ for $i = 1, 2, \ldots, m$, then the probability that A led throughout is $P(a, b)/C(m, a)$.

The Catalan numbers arise in the case where $a = b$, which we shall now assume. Denote the common value of a and b by n, so that $m = 2n$. Suppose we seek the probability that A never trailed throughout the counting. There are $C(2n, n)$ possible sequences of n 1s and n -1s. We seek the number for which $s_i \geq 0$ for $i = 1, 2, \ldots, 2n - 1$.

We can represent the sequence $(x_1, \ldots, x_{2n})$ by a *path* in the plane from the origin to the point $(2n, 0)$ whose *steps* are the line segments between $(i-1, s_{i-1})$ and (i, s_i), for $i = 1, 2, \ldots, 2n$. The ith step in this path then has slope $s_i - s_{i-1} = x_i \in \{1, -1\}$.

Example 1 Let $n = 5$, so the paths run from the origin to the point $(10, 0)$. We display three such sequences and their partial sum sequences. For clarity, the 1s in the sequence are represented by plus signs and the -1s by minus signs. Draw the corresponding paths.

	Sequence	**Partial sum sequence**
(i)	$(-,-,+,+,+,+,-,-,-,+)$	$(0,-1,-2,-1,0,1,2,1,0,-1,0)$
(ii)	$(+,+,-,+,-,-,+,+,-,-)$	$(0,1,2,1,2,1,0,1,2,1,0)$
(iii)	$(+,+,-,+,+,-,+,-,-,-)$	$(0,1,2,1,2,3,2,3,2,1,0)$

Solution: The corresponding paths are shown in Figure 1. □

Figure 1. Some paths from (0,0) to (10,0).

Note that in sequence (i) the partial sum sequence has both positive and negative entries, so the path lies both above and below the x-axis. In sequence (ii) all the partial sums are nonnegative, so the path does not go below the x-axis. But in sequence (iii), we have $s_i > 0$ for $i = 1, 2, \ldots, 2n-1$. Hence, except for its two endpoints, the path lies entirely above the x-axis.

We call a path from $(0,0)$ to $(2n,0)$ **nonnegative** if all $s_i \geq 0$ and **positive** if $s_i > 0$ for $i = 1, 2, \ldots, 2n-1$. Thus, a nonnegative path corresponds to a ballot sequence in which candidate A and candidate B both received n votes, but candidate A never trailed. Similarly, a positive path represents a ballot sequence in which candidate A was leading all the way until the final ballot brought them even.

We shall see that the number of nonnegative paths and positive paths from the origin to the point $(2n,0)$ are both Catalan numbers. Before proceeding with this ballot problem, let us define the Catalan numbers.

Definition 1 The *Catalan number* c_n, for $n \geq 0$, is given by

$$c_n = \frac{1}{n+1}\,C(2n,n). \qquad \square$$

The values of c_n for $n \leq 10$ are given in Table 1.

n	0	1	2	3	4	5	6	7	8	9	10
c_n	1	1	2	5	14	42	132	429	1,430	4,862	16,796

Table 1. The first eleven Catalan numbers.

Example 2 Find all the nonnegative paths from the origin to $(6,0)$. Which of these are positive paths?

Solution: There are $C(4,2) = 6$ slope sequences of length 6 that start with 1 and end with -1. To check for nonnegativity, we compute their partial sums, and we find the nonnegative paths have slope sequences

$$(+,-,+,-,+,-),\quad (+,-,+,+,-,-),\quad (+,+,-,-,+,-),$$
$$(+,+,-,+,-,-),\quad (+,+,+,-,-,-),$$

with respective partial sum sequences

$$(0,1,0,1,0,1,0),\quad (0,1,0,1,2,1,0),\quad (0,1,2,1,0,1,0),$$
$$(0,1,2,1,2,1,0),\quad (0,1,2,3,2,1,0).$$

These paths are shown in Figure 2. □

Figure 2. The nonnegative paths from the origin to (6,0).

Let us now show that the numbers of nonnegative and positive paths are Catalan numbers.

Theorem 1 The number of paths from the origin to $(2n,0)$ that are
(i) positive is the Catalan number c_{n-1},
(ii) nonnegative is the Catalan number c_n.

Proof: We shall first establish a one-to-one correspondence between positive paths of length $2n$ and nonnegative paths of length $2n-2$. Let $(s_0, s_1, \ldots, s_{2n})$ be the partial sum sequence of a positive path P. Then $s_0 = s_{2n} = 0$ and $s_i \geq 1$ for $i = 1, 2, \ldots, 2n-1$. Let $\mathbf{x} = (x_1, x_2, \ldots, x_{2n})$ be the corresponding slope sequence of the steps, so $x_i = s_i - s_{i-1} \in \{1, -1\}$. Since $s_1 \geq 1$ and $s_0 = 0$, we must have $s_1 = 1$. Hence the path P passes through the point $(1,1)$. Similarly, since $s_{2n-1} \geq 1$ and $s_{2n} = 0$, we have $s_{2n-1} = 1$. Therefore P passes through the point $(2n-1, 1)$. If we omit the first and last terms from the slope sequence, we have a sequence $\mathbf{x}'$ that has $n-1$ 1s and $n-1$ -1s. Further, the partial sums for $\mathbf{x}'$ satisfy

$$\begin{aligned} s'_i &= x'_1 + x'_2 + \ldots + x'_i \\ &= x_2 + x_3 + \ldots + x_{i+1} \\ &= s_{i+1} - 1 \\ &\geq 0 \end{aligned}$$

for $0 \leq i \leq 2n-2$. Thus the positive path P from the origin to $(2n,0)$ corresponds to a nonnegative path P' from the origin to $(2n-2,0)$. Geometrically, the path P' is produced from P by taking the segment of P from $(1,1)$ to $(2n-1,1)$ relative to the new coordinate system obtained by translating the origin to the point $(1,1)$.

Since the first and last terms of P must be 1 and -1, respectively, the operation of deleting those terms is reversed by adding 1 before and -1 after P'. Thus the function that takes P to P' is invertible. Hence, if a_n is the number of positive paths and b_n the number of nonnegative paths from the origin to $(2n, 0)$, then $a_n = b_{n-1}$. Part (ii) will then follow from (i) once we establish that $a_n = c_{n-1}$, since we would then have $b_n = a_{n+1} = c_n$.

A positive path P from the origin to $(2n, 0)$ is determined uniquely by its segment P' from $(1, 1)$ to $(2n - 1, 1)$, which lies entirely above the x-axis. But the number of such paths P' is the difference between $C(2n - 2, n - 1)$, the total number of paths from $(1, 1)$ to $(2n - 1, 1)$, and the number of paths Q from $(1, 1)$ to $(2n - 1, 1)$ that meet the x-axis.

Thus, we can determine the number of paths P' by counting the number of such paths Q. Suppose that Q first meets the x-axis at the point $(k, 0)$. Then $k \geq 2$ since Q begins at the point $(1, 1)$. If we reflect the segment Q_1 of Q from $(1, 1)$ to $(k, 0)$ about the x-axis, we obtain a path Q_1' from $(1, -1)$ to $(k, 0)$. Then if we adjoin to Q_1' the segment Q_2 of Q from $(k, 0)$ to $(2n-1, 1)$, we obtain a path $Q^* = Q_1'Q_2$ from $(1, -1)$ to $(2n - 1, 1)$. But every such path Q^* must cross the x-axis, so we can obtain the path Q from Q^* by reflecting the segment of Q^* from $(1, -1)$ to its first point on the x-axis.

Hence we have a one-to-one correspondence between the paths Q and Q^*, so the number of paths Q from $(1, 1)$ to $(2n-1, 1)$ that meet the x-axis is equal to the number of paths Q^* from $(1, -1)$ to $(2n - 1, 1)$. Every such path Q^* must have two more increasing steps than decreasing steps, so it has n increasing steps and $n - 2$ decreasing steps. The number of such paths Q^* is therefore $C(2n - 2, n - 2)$, and this is equal to the number of such paths Q.

The number, a_n, of positive paths P from the origin to $(2n, 0)$ can now be calculated. Since it is equal to the number of positive paths P' from $(1, 1)$ to $(2n-1, 1)$, and this is the difference between the total number $C(2n-2, n-1)$ of paths from $(1, 1)$ to $(2n - 1, 1)$ and the number $C(2n - 2, n - 2)$ of paths Q from $(1, 1)$ to $(2n - 1, 1)$ that meet the x-axis, we have

$$
\begin{aligned}
a_n &= C(2n-2, n-1) - C(2n-2, n-2) \\
&= \frac{(2n-2)!}{(n-1)!(n-1)!} - \frac{(2n-2)!}{(n-2)!n!} \\
&= \frac{(2n-2)!}{(n-2)!(n-1)!}\left(\frac{1}{n-1} - \frac{1}{n}\right) \\
&= \frac{(2n-2)!}{(n-2)!(n-1)!}\,\frac{1}{n(n-1)} \\
&= \frac{1}{n}\,\frac{(2n-2)!}{(n-1)!(n-1)!} \\
&= \frac{1}{n}C(2n-2, n-1) \\
&= c_{n-1}.
\end{aligned}
$$

■

Well-formed Sequences of Parentheses

The Catalan numbers count several structures important in computer science. Let us consider first the set of *well-formed sequences of parentheses*, which can be defined recursively (see Section 3.4 in *Discrete Mathematics and Its Applications*, Fifth Edition, by Rosen).

Definition 2 A sequence of parentheses is *well-formed* if and only if it can be derived by a finite sequence of the following rules:

The empty sequence is well-formed.
If A is well-formed, then (A) is well-formed.
If A and B are well-formed, then AB is well-formed.

We say that the right parenthesis following A in the second rule *closes* the left parenthesis preceding A. □

Example 3 Find a sequence of $n = 4$ parentheses that is not well-formed and a sequence that is well-formed.

Solution: The sequence (( ))) (( ) is not well-formed since only one of the third and fourth left parentheses can be closed by the single right parenthesis that follows them. But the sequence (( )) ( ) ( ) is well-formed since each left parenthesis is closed by the first right parenthesis following it that does not close a left parenthesis between them. □

Clearly each well-formed sequence of parentheses must have an equal number of left parentheses and right parentheses. Further, in any initial string consisting of the first i parentheses of a well-formed sequence of $2n$ parentheses, there must be at least as many left parentheses as right parentheses. Thus, if we replace left parentheses by 1s and right parentheses by -1s in a well-formed sequence of parentheses, we obtain a sequence $(x_1, x_2, \ldots, x_{2n}) \in \{1, -1\}^{2n}$ with all partial sums $s_i \geq 0$, and hence a nonnegative path from $(0, 0)$ to $(2n, 0)$. Conversely, any nonnegative path from $(0, 0)$ to $(2n, 0)$ produces a well-formed sequence of parentheses. By Theorem 1 and this one-to-one correspondence, we have established the following result.

Theorem 2 The number of well-formed sequences of parentheses of length $2n$ is the Catalan number c_n. ■

Example 4 Find the well-formed sequences of parentheses of length $2n = 6$.

Solution: Since $n = 3$, by Theorem 2 there are exactly $c_3 = 5$ such sequences:

$$(\,) (\,) (\,), \quad (\,) ((\,)), \quad ((\,)) (\,), \quad ((\,) (\,)), \quad (((\,))).$$ □

Stack Permutations

In *Fundamental Algorithms*, Volume 1, of his classic series of books, *The Art of Computer Programming*, Donald Knuth posed the problem of counting the number of permutations of a particular type that arise in the computer.

A **stack** is a list which can only be changed by insertions or deletions at one distinguished end, called the **top** of the list. When characters are individually inserted and deleted from the stack, the last one inserted must be the first one deleted from the stack (*lifo*). An insertion to the top of the stack is called a **push**; a deletion from the top is called a **pop**. We can interpret a stack as a stack of plates, with a push representing the placement of a plate on top, and a pop corresponding to the removal of a plate from the top.

A sequence of pushes and pops is **admissible** if the sequence has an equal number n of pushes and pops, and at each stage the sequence has at least as many pushes as pops. If we identify pushes with 1s and pops with -1s, then an admissible sequence corresponds to a nonnegative path from the origin to the point $(2n, 0)$. Thus, by Theorem 1, the number of admissible sequences of pushes and pops of length $2n$ is the Catalan number c_n. When applied in a computer, an admissible sequence of pushes and pops of length $2n$ transforms an input string of length n to an output string with the same symbols, but in a possibly different order.

Suppose we have as initial input string the standard permutation $123\ldots n$ of the set $N = \{1, 2, \ldots, n\}$ and an admissible sequence of pushes and pops of length 2n. Each push in the sequence transfers the last element of the input string to the top of the stack, and each pop transfers the element on top of the stack to the beginning of the output string. After the n pushes and n pops have been performed, the output string is a permutation of N called a **stack permutation**.

Example 5 Let $n = 4$ and consider the admissible sequence

$$(+, +, +, -, -, -, +, -, -),$$

where a plus sign represents a push and a minus sign represents a pop. Find the stack permutation produced by this sequence of pushes and pops.

Solution: We will denote the result of each operation of the admissible sequence of pushes and pops by a string of the form $\alpha[\sigma]\beta$, where α is the current input string, β the current output string, and $[\sigma]$ the current stack, with the first element of σ being the top of the stack.

Then the admissible sequence given proceeds as follows to produce the stack permutation 4132:

Sequence	$\alpha[\sigma]\beta$	**Operation**
	1234[ ]	
+	123[4]	Push 4
+	12[34]	Push 3
+	1[234]	Push 2
−	1[34]2	Pop 2
−	1[4]32	Pop 3
+	[14]32	Push 1
−	[4]132	Pop 1
−	[ ]4132	Pop 4.

□

A stack permutation of $123\dots n$ is defined as one produced from $1, 2, \dots, n$ by an admissible sequence of pushes and pops. But we have seen that there is a one-to-one correspondence between admissible sequences of pushes and pops of length $2n$ and nonnegative paths from the origin to the point $(2n, 0)$. The stack permutation can be found from its corresponding nonnegative path in the plane as follows.

Let i index the horizontal axis and s the vertical axis, and suppose the path passes through the points $(i, s_i),\ i = 0, 1, 2, \dots, 2n$. Let k be the maximum ordinate s_i of the path.
(1) Draw (segments of) the horizontal lines with equations $s = j$ for $j = 1, 2, \dots, k$. The region bounded by the lines $i = 0,\ i = 2n,\ s = j-1,\ s = j$ will be called box j.
(2) Label the increasing steps of the path from left to right with $n, n-1, \dots, 1$.
(3) For $j = 1, 2, \dots, k$, label each decreasing step in box j with the label of the last increasing step in box j that precedes it.
(4) The stack permutation is the sequence *from right to left* of labels on the n decreasing steps.

Since the path starts at the origin, never crosses the i-axis (with equation $s = 0$), and ends on the i-axis at $(2n, 0)$, each box j must contain an equal number of increasing and decreasing steps. Further, they must alternate increasing-decreasing from left to right in box j, starting with an increasing step. Thus the labeling of the increasing steps in step (2) and the decreasing steps in step (3) establishes a one-to-one correspondence between the n increasing steps and the n decreasing steps of the path, where corresponding steps have the same label.

The label l on the increasing step in box j represents the element pushed onto the stack at that term of the admissible sequence of pushes and pops.

Element l is not popped until the path first returns to box j by the decreasing step labeled l. The distinct labels on the steps between the two steps labeled l each occur twice, and represent the elements that were pushed, and hence must be popped, while l was on the stack.

Alternatively, in the well-formed sequence of parentheses corresponding to the nonnegative path, each increasing step is regarded as a left parenthesis, and the corresponding decreasing step is regarded as the right parenthesis that closes it.

Example 6 Draw the nonnegative path produced by the admissible sequence in Example 5, and find the corresponding stack permutation by labeling the steps.

Solution: The sequence of ordinates of the path is the partial sum sequence

$$\mathbf{s} = (0, 1, 2, 3, 2, 1, 2, 1, 0)$$

computed from the admissible sequence (slope sequence)

$$(+, +, +, -, -, -, +, -, -)$$

given in Example 5. The nonnegative path, with maximum ordinate $k = 3$, is displayed in Figure 3, along with (segments of) the lines $s = 1, 2, 3$. The $n = 4$ increasing steps are labeled from left to right as $4, 3, 2, 1$. The decreasing steps are then labeled as shown in accordance with (3). When read from right to left, the labels on the decreasing steps produce the stack permutation 4132. □

Figure 3. Stack permutation 4132 from a nonnegative path with n = 4.

A stack permutation can only be produced by a unique admissible sequence, and therefore by a unique nonnegative path. We can recover the admissible sequence, and hence the path, from the stack permutation as follows. Starting with the stack permutation, precede it by the empty input string α and the empty stack $[\sigma]$ as in Example 5. Begin with an empty push-pop sequence.

Then, given a current content $\alpha[\sigma]\beta$ of the input, stack, and output, respectively, with at least one of σ and β nonempty, proceed repeatedly as follows:

> If the stack is empty or both of σ and β are nonempty, with the first element b of β less than the first element s of σ, transfer b to the left of σ (b was just popped) and add $-$ to the right of the push-pop sequence.

> If the output is empty or both of σ and β are nonempty, with the first element b of β greater than the first element s of σ, transfer s to the right of α (s was just pushed) and add $+$ to the right of the push-pop sequence.

When the input string is $\alpha = 12\ldots n$, σ and β are empty and the admissible sequence that produced the stack permutation is the push-pop sequence constructed.

Example 7 Find the the admissible sequence of pushes and pops that produces the stack permutation 4213.

Solution:

$\alpha[\sigma]\beta$	**Operation**	**Sequence**
[ ]4213		
[4]213	4 Popped	$-$
[24]13	2 Popped	$-$
[124]3	1 Popped	$-$
1[24]3	1 Pushed	$+$
12[4]3	2 Pushed	$+$
12[34]	3 Popped	$-$
123[4]	3 Pushed	$+$
1234[ ]	4 Pushed	$+$

The admissible sequence, obtained by reading the sequence in the third column upward, is $(+,+,-,+,+,-,-,-)$. □

It follows from the foregoing that there is a one-to-one correspondence between the set of stack permutations of N and the set of nonnegative paths, or between the set of stack permutations of N and set of admissible sequences. By Theorem 1 we have the following theorem.

Theorem 3 The number of stack permutations of an n-element set is the Catalan number c_n. ■

Well-parenthesized Products

Consider an algebraic structure S with a binary operation,which we will refer to as multiplication. Then, as usual, we can denote the product of $x, y \in S$ by xy. Let us further assume that the operation is *not commutative*, so that $xy \neq yx$ in general. The product $x_1x_2 \cdots x_{n+1}$ of $n+1$ elements of S in that order is well-defined provided that multiplication is associative, i.e. that $(xy)z = x(yz)$ for all $x, y, z \in S$. But let us suppose that multiplication in S is *not associative*. Then a product $x_1x_2 \cdots x_{n+1}$ is defined only after parentheses have been suitably inserted to determine recursively pairs of elements to be multiplied. However, we will refer to the sequence $x_1x_2 \cdots x_{n+1}$ without parentheses simply as a product.

We shall determine the number of ways a product $x_1x_2 \cdots x_{n+1}$ can be parenthesized. The assumption that the binary operation is noncommutative and nonassociative allows us to interpret this number as the maximum number of different elements of S that can be obtained by parenthesizing the product. However, we could instead consider the x_is to be real numbers and the binary operation to be ordinary multiplication. In this case we are seeking the number of ways that the product $x_1x_2 \cdots x_{n+1}$ can be computed by successive multiplications of exactly two numbers each time. This was Catalan's original formulation of the problem.

Example 8 Find the distinct ways to parenthesize the product $x_1x_2x_3x_4$.

Solution: Whenever a left parenthesis is closed by a right parenthesis, and we have carried out any products defined by closed pairs of parentheses nested between them, we must have a product of exactly two elements of S. The well-parenthesized sequences for the product $x_1x_2x_3x_4$ are found to be

$$((x_1(x_2x_3))x_4), \quad (x_1((x_2x_3)x_4)), \quad ((x_1x_2)(x_3x_4)),$$

$$(x_1(x_2(x_3x_4))), \quad (((x_1x_2)x_3)x_4). \qquad \square$$

Note that we used $n = 3$ pairs of parentheses in each parenthesized product in Example 8. Although not necessary, it is convenient to include outer parentheses, where the left parenthesis is first and the right parenthesis last.

We will now formally define what is meant by parenthesizing a product.

Definition 3 A product is *well-parenthesized* if it can be obtained recursively by a finite sequence of the following rules:

Each single term $x \in S$ is well-parenthesized.

If A and B are well-parenthesized, then (AB) is well-parenthesized. $\square$

Note that a well-parenthesized product, other than a single element of S, includes outer parentheses. Thus, (xy) is well-parenthesized, but xy is not. We can then use mathematical induction to prove that n pairs of parentheses must be added to $x_1x_2\cdots x_{n+1}$ to form a well-parenthesized product.

If the $n+1$ variables $x_1, x_2, \ldots, x_{n+1}$ are deleted from a well-parenthesized product, the n pairs of parentheses that remain must be a well-formed sequence of parentheses. But not every well-formed sequence of n pairs of parentheses can arise in this way. For example,

$$((\quad)(\quad)(\quad))$$

is a well-formed sequence of $n = 4$ pairs of parentheses, but since the operation is binary, the outer pair of parentheses would call for the undefined product of the three elements of S that are to be computed within the inner parentheses.

Example 9 Show by identifying A and B at each step that

$$((x_1(x_2x_3))(x_4x_5))$$

is obtained from the product $x_1x_2x_3x_4x_5$ by the recursive definition, and hence is a well-parenthesized product.

Solution:

$x_1x_2x_3x_4x_5$	
$x_1(x_2x_3)x_4x_5$	$A = x_2, \quad B = x_3$
$(x_1(x_2x_3))x_4x_5$	$A = x_1, \quad B = (x_2x_3)$
$(x_1(x_2x_3))(x_4x_5)$	$A = x_4, \quad B = x_5$
$((x_1(x_2x_3))(x_4x_5))$	$A = (x_1(x_2x_3)), \quad B = (x_4x_5).$

□

We will now show that there is a one-to-one correspondence between well-parenthesized products of $x_1, x_2, \cdots, x_{n+1}$ and nonnegative paths from the origin to the point $(2n, 0)$. A well-parenthesized product forms a string $\mathbf{p}$ of length $3n+1$ with three types of characters: n left parentheses, $n+1$ variables, and n right parentheses. But the slope sequence of the nonnegative paths from the origin to the point $(2n, 0)$ has just two different numbers, 1 and -1, with n of each. To obtain a sequence from $\mathbf{p}$ with a structure similar to that of the slope sequence, we form a string $\mathbf{q} = S(\mathbf{p})$ of length $2n$ by deleting the last variable x_{n+1} and the n right parentheses.

Example 10 Find the string $\mathbf{q}$, if $\mathbf{p} = ((x_1(x_2x_3))(x_4x_5))$ is the well-parenthesized product in Example 8.

Solution: Deleting x_5 and the right parentheses from $\mathbf{p}$ gives

$$\mathbf{q} = ((x_1(x_2x_3(x_4.$$ □

Let us examine the properties of this string $\mathbf{q} = S(\mathbf{p})$ obtained from a well-parenthesized product $\mathbf{p}$. We will then show that $\mathbf{p}$ can be obtained from a string $\mathbf{q}$ with these properties. By the way that $\mathbf{q}$ was defined, we first note the following.

Lemma 1 The string $\mathbf{q} = S(\mathbf{p})$ has n left parentheses and the n variables $x_1, x_2, \ldots, x_n$ in that order. ■

Since $\mathbf{p}$ cannot have a left parenthesis between x_n and x_{n+1}, the last character of the string $\mathbf{q}$ must be x_n. This is in fact implied by a more general property that $\mathbf{q}$ satisfies, analogous to the property satisfied by nonnegative paths, which we state in the following lemma.

Lemma 2 Let $\mathbf{q} = S(\mathbf{p})$, where $\mathbf{p}$ is a well-parenthesized product of the variables $x_1, x_2, \ldots, x_{n+1}$. For $i \leq 2n$, the number of left parentheses in the string $q_1q_2\cdots q_i$ is at least as large as the number of variables.

Proof: We will prove the lemma by induction on n. If $n = 1$, then we must have $\mathbf{p} = (x_1x_2)$, so $\mathbf{q} = (x_1$ and the conclusion holds.

Suppose $n \geq 2$ and that the conclusion holds whenever $\mathbf{q}$ is obtained from a well-parenthesized product $\mathbf{p}$ of $k \leq n$ variables. Let $\mathbf{p}$ be a well-parenthesized product of $n + 1$ variables. From the recursive definition, $\mathbf{p} = (AB)$ is the product of two nonempty well-parenthesized products, A and B. The first left parenthesis in $\mathbf{p}$ is placed before AB, so the number of left parentheses up to each character of $\mathbf{p}$ preceding B is one more than in A. But if x_k is the last variable in A, then it does not appear in $S(A)$, but does appear following x_{k-1} in $\mathbf{q} = S(\mathbf{p})$. Thus the difference between the number of left parentheses and the number of variables in $\mathbf{q}$ exceeds the corresponding difference in $S(A)$ by one until x_k is reached, where it becomes zero. The differences at each character of B are then the same as they are at that character in $\mathbf{p}$. Since A is a well-parenthesized product of $k \leq n$ variables, it then follows by the inductive hypothesis that the number of left parentheses in $q_1q_2\cdots q_i$ for $i \leq 2n$ is at least as large as the number of variables. ■

We say that a string $\mathbf{q}$ satisfying Lemmas 1 and 2 is **suitable**.

Theorem 4 There is a one-to-one correspondence between the set of suitable strings of length $2n$ and the set of well-parenthesized products of length $3n+1$.

Proof: We will show that the function S is a one-to-one correspondence. Let $\mathbf{q}$ be a suitable string. We need to show that we can reconstruct the well-parenthesized product $\mathbf{p}$ such that $\mathbf{q} = S(\mathbf{p})$. First we adjoin x_{n+1} to the right of $\mathbf{q}$ and call the new string $\mathbf{q}' = \mathbf{q}x_{n+1}$. Then by Lemma 1 we have the following.

Lemma 1′ The string $\mathbf{q}'$ has n left parentheses and the $n+1$ variables $x_1, x_2, \ldots, x_{n+1}$ in that order.

Since $\mathbf{q}$ and $\mathbf{q}'$ agree in the first $2n$ positions, we will denote the character of $\mathbf{q}'$ in position i by q_i for $i \le 2n$. Then $\mathbf{q}'$ satisfies the conclusion of Lemma 2.

Let us prove the theorem by mathematical induction. If $n = 1$ then $\mathbf{q}' = (x_1x_2$, so we must have $\mathbf{p} = (x_1x_2)$.

Assume that $n \ge 2$ and the theorem is true with $n-1$ replacing n. By Lemmas 1 and 2 the last two characters of $\mathbf{q}$ are either $x_{n-1}x_n$ or $(x_n$, so the last three characters of $\mathbf{q}'$ are either $x_{n-1}x_nx_{n+1}$ or $(x_nx_{n+1}$. By Lemma 1′, in either case $\mathbf{q}'$ will have three consecutive characters of the form $(x_jx_{j+1}$ for some $j \ge 1$. Let j_1 be the minimum such j. When right parentheses are inserted in $\mathbf{q}'$ to form a well-parenthesized product, a right parenthesis must immediately follow $(x_{j_1}x_{j_1+1}$, so $\mathbf{p}$ would contain $(x_{j_1}x_{j_1+1})$. Replace $(x_{j_1}x_{j_1+1})$ by a new variable y_1 in $\mathbf{q}'$ to form a string $\mathbf{q_1}$. Then $\mathbf{q_1}$ has $n-1$ left parentheses and n variables. Further, $\mathbf{q_1}$ satisfies the conclusion of Lemma 1′ with $n-1$ replacing n. Then by our inductive hypothesis, the well-parenthesized product $\mathbf{p_1}$ can be recovered. Substituting $(x_{j_1}x_{j_1+1})$ for y_1 in $\mathbf{p_1}$ gives a well-parenthesized product $\mathbf{p}$ such that $\mathbf{q} = S(\mathbf{p})$. Since S is one-to-one and can be inverted, it is a one-to-one correspondence. ∎

Example 11 Recover the well-parenthesized product $\mathbf{p}$ from the suitable string $\mathbf{q} = (x_1((x_2((x_3x_4x_5$ so that $\mathbf{q} = S(\mathbf{p})$.

Solution: We form $\mathbf{q}' = (x_1((x_2((x_3x_4x_5x_6$ by adding x_6 to the right of $\mathbf{q}$. Then locate at each stage j the first occurrence of a left parenthesis immediately followed by two variables, and replace this string of length three by the new variable y_j equal to this string with a right parenthesis added on the right as a fourth character.

$$\begin{aligned}
\mathbf{q}' &= (x_1((x_2((x_3x_4x_5x_6 & & \\
&= (x_1((x_2(y_1x_5x_6 & \qquad y_1 &= (x_3x_4) \\
&= (x_1((x_2y_2x_6 & y_2 &= (y_1x_5) \\
&= (x_1(y_3x_6 & y_3 &= (x_2y_2) \\
&= (x_1y_4 & y_4 &= (y_3x_6) \\
&= y_5 & y_5 &= (x_1y_4)
\end{aligned}$$

Then, on setting $\mathbf{p} = y_5$ and successively substituting for the new variables $y_5, y_4, \ldots, y_1$ their values on the right we obtain

$$\begin{aligned}\mathbf{p} &= y_5\\ &= (x_1 y_4)\\ &= (x_1(y_3 x_6))\\ &= (x_1((x_2 y_2)x_6))\\ &= (x_1((x_2(y_1 x_5))x_6))\\ &= (x_1((x_2((x_3 x_4)x_5))x_6)).\end{aligned}$$ □

Corollary 1 The number of well-parenthesized products of $n+1$ variables is the Catalan number c_n.

Proof: By Theorem 4, each well-parenthesized sequence from $x_1 x_2 \cdots x_{n+1}$ corresponds to a suitable string $\mathbf{q}$ of length $2n$ with n left parentheses and the n variables $x_1, x_2, \ldots, x_n$. Define a sequence $\mathbf{z} = (z_1, z_2, \ldots, z_{2n})$ by

$$z_i = \begin{cases} 1 & \text{if } q_i \text{ is a left parenthesis} \\ -1 & \text{if } q_i \text{ is a variable } x_j. \end{cases}$$

Then it follows from Lemma 2 that the partial sums s_i of $\mathbf{z}$ are nonnegative. Thus, corresponding to $\mathbf{q}$ is a nonnegative path from $(0,0)$ to $(2n,0)$. Consequently, by Theorem 1, the number of well-parenthesized products is c_n. ■

Full Binary Trees

Recall (see Section 9.1 of *Discrete Mathematics and Its Applications*, Fifth Edition) that a **full binary tree** is a rooted tree in which each internal vertex has exactly two children. Thus, a full binary tree with n internal vertices has $2n$ edges. Since a tree has one more vertex than it has edges, a full binary tree T with n internal vertices has $2n+1$ vertices, and thus $n+1$ leaves. Suppose we label the leaves of T as they are encountered along a transversal (preorder, postorder, or inorder; see Section 9.3 of *Discrete Mathematics and Its Applications*) with $x_1, x_2, \ldots, x_{n+1}$. Then T recursively defines a well-parenthesized product of $x_1, x_2, \cdots, x_{n+1}$ by the following rule.

> Labeling rule: If v is an internal vertex with left child a and right child b, having labels A and B, respectively, then label v with (AB).
>
> The label on the root of the tree will be the well-parenthesized product.

Conversely, given a well-parenthesized product of $n+1$ variables $x_1, x_2, \ldots, x_{n+1}$, a labeled full binary tree is determined by first labeling the root with the well-parenthesized product, then moving from the outer parentheses inward by adding two children labeled A and B to each vertex v with label (AB). The leaves of the tree will be labeled with the variables $x_1, x_2, \ldots, x_{n+1}$ in the order encountered by a traversal. Consequently, there is a one-to-one correspondence between the well-parenthesized products of $n+1$ variables and the full binary trees with $n+1$ leaves and n internal vertices. By Theorem 4 we therefore have the following result.

Theorem 5 The number of full binary trees with n internal vertices is the Catalan number c_n. ■

Example 12 Draw and label the full binary tree defined by the well-parenthesized product $((1(23))(45))$.

Solution: The full binary tree is shown in Figure 4. □

Figure 4. Full binary tree obtained from ((1(23))(45)).

Triangulations of a Convex Polygon

In this section we consider a geometric interpretation of the Catalan numbers. An n**-gon** ($n \geq 3$) in the plane is a polygon P with n vertices and n sides. Let $v_0, v_1, \ldots, v_{n-1}$ be the vertices (in counterclockwise order). Let us denote by v_iv_j the **line segment** joining v_i and v_j. Then the n sides of P are $s_i = v_{i-1}v_i$ for $1 \leq i \leq n-1$ and $s_0 = v_nv_0$. A **diagonal** of P is a line segment v_iv_j joining two nonadjacent vertices of P. An n-gon P is **convex** if every diagonal lies wholly in the interior of P.

Let D be a set of diagonals, no two of which meet in the interior of a convex n-gon P, that partitions the interior of P into triangles. The sides of the triangles are either diagonals in D or sides of P. The set T of triangles obtained in this way is called a **triangulation** of P. Three triangulations of a convex hexagon are shown in Figure 5.

Figure 5. Three triangulations of a convex hexagon.

We shall next determine the number of diagonals needed to form a triangulation.

Lemma 3 A triangulation of a convex n-gon has $n-2$ triangles determined by $n-3$ diagonals.

Proof: We shall argue by induction on n. If $n=3$ there is one triangle and no diagonals, while if $n=4$ there are two triangles and one diagonal.

Assume that $n \geq 5$ and that the conclusion holds for triangulations of a convex m-gon with $3 \leq m < n$. Let P be a convex n-gon and let T be a triangulation of P with diagonal set D. Since $n > 3$ there must be at least one diagonal in D, say v_0v_k, where $2 \leq k \leq n-2$. The diagonal v_0v_k of P then serves jointly as a side of the convex $(k+1)$-gon P_1 with vertices $v_0, v_1, \ldots, v_k$ and the convex $(n-k+1)$-gon $P_2$with vertices $v_0, v_k, v_{k+1} \ldots, v_{n-1}$. Triangulations T_1 and T_2 of these two convex polygons are defined by subsets D_1 and D_2, respectively, of the set D of diagonals of P other than v_0v_k. Since $k+1 \leq n$ and $n-k+1 \leq n$, we may apply the inductive hypothesis. Then T_1 and T_2 have $k-1$ and $n-k-1$ triangles defined by $k-2$ and $n-k-2$ diagonals, respectively. Adding the numbers of triangles gives $(k-1)+(n-k-1) = n-2$ triangles in T. We add 1 (for v_0v_k) to the sum $(k-2)+(n-k-2) = n-4$ of the numbers of diagonals to get $n-3$ diagonals in D. ■

It is convenient now to assume that P is a convex $(n+2)$-gon for $n \geq 1$, with vertices $v_0, v_1, \ldots, v_{n+1}$ and sides $s_0, s_1, \ldots, s_{n+1}$. Then, by Lemma 3, every triangulation of P is defined by $n-1$ diagonals and has n triangles.

Let t_n be the number of triangulations of a convex $(n+2)$-gon P. Then clearly $t_1 = 1$ and $t_2 = 2$.

Example 13 Draw all the triangulations of a convex pentagon to show that $t_3 = 5$.

Solution: Since $n = 3$, each triangulation has three triangles defined by two diagonals. If $v_i v_j$ is one of the diagonals, the other diagonal must meet v_i or v_j. Thus each of the two diagonals must join the common endpoint to a nonadjacent vertex. We find the five triangulations shown in Figure 6. □

Figure 6. The triangulations of a convex pentagon.

Consider the string $s_1 s_2 \cdots s_{n+1}$ formed by taking the $n+1$ sides of P other than s_0 in order around P. We shall show that the product $s_1 s_2 \cdots s_{n+1}$ is well-parenthesized by a triangulation of P. One side, s_0, is excluded above in order that the remaining sides form an open path in the plane. The sides, taken in the order of the path, then form a sequence analogous to the product of $n+1$ variables considered earlier.

Let T be a triangulation of P defined by a set D of $n-1$ diagonals, where $n \geq 4$. Each side s_i of P is a side of exactly one of the n triangles of T. There are two types of triangles in T that contain sides of P:

(i) The three sides of an **outer triangle** are two adjacent sides s_i, s_{i+1} of P and the diagonal $v_{i-1}v_{i+1}$ of D.

(ii) The three sides of an **inner triangle** are a side $s_i = v_{i-1}v_i$ of P and the two diagonals $v_{i-1}v_j$, $v_i v_j$ of D for some vertex v_j, with $j \neq i-2, i+1$.

For example, hexagons (a) and (b) in Figure 5 each have two outer triangles, while (c) has three. In order to establish a one-to-one correspondence between triangulations and well-parenthesized products, we must show that any triangulation has an outer triangle not having s_0 as a side.

Lemma 4 If $n \geq 2$, every triangulation T of a convex $(n+2)$-gon P has at least two outer triangles.

Proof: Suppose that at most one of the n triangles of T has two sides of P as sides. Let n_i be the number of sides of P that are sides of the ith triangle. Then $n_i = 1$ for at least $n-1$ triangles and $n_i = 2$ for at most one. When we sum these n numbers n_i, we conclude that P has at most $n+1$ sides. But P has $n+2$ sides, so we have a contradiction. Thus P must have at least two outer triangles. ■

Theorem 6 The number of triangulations of a convex $(n + 2)$-gon is the Catalan number c_n.

Proof: By Corollary 1, it will suffice to establish a one-to-one correspondence between triangulations of a convex $(n+2)$-gon P and well-parenthesized products of the $n+1$ sides of P other than s_0.

Let $n \geq 2$ and assume such a one-to-one correspondence exists for $(n+1)$-gons. Let T be a triangulation of the convex $(n+2)$-gon P. Since every side of P is a side of exactly one triangle of T, and by Lemma 4 there are at least two outer triangles, there must be an outer triangle in T that does not have s_0 as a side. This triangle has vertices v_{i-1}, v_i, v_{i+1} for some i, $1 \leq i \leq n$, so has as sides the two sides s_i, s_{i+1} of P and the diagonal $v_{i-1}v_{i+1}$ in D. Label this diagonal with $(s_i s_{i+1})$. If we delete vertex v_i and replace sides s_i, s_{i+1} by the diagonal labeled $(s_i s_{i+1})$, we have a convex $(n+1)$-gon P'. By the inductive hypothesis we can establish a one-to-one correspondence between triangulations T' of P' and well-parenthesized products of the sides of P' other than s_0. But one of the sides of P' is labeled with the well-parenthesized product $(s_i s_{i+1})$ of two sides of P. Thus the well-parenthesized product of the sides of P' represents a well-parenthesized product of the sides of P.

Conversely, each innermost pair of parentheses in a well-parenthesized product of the sides of P other than s_0 indicates that the two sides within that pair are in an outer triangle. Then the diagonal completing the outer triangle must be included in D. Each closing of parentheses acts in this way to add diagonals that complete outer triangles on the reduced polygon until a triangulation of P is obtained. ■

Summary of Objects Counted by the Catalan Numbers

The one-to-one correspondences we have established between the sets of triangulations of a convex $(n+2)$-gon, well-parenthesized products of $n+1$ variables, well-formed sequences of n pairs of parentheses, stack permutations of $12 \cdots n$, and nonnegative paths from the origin to the point $(2n, 0)$ are illustrated in Figure 7 for the case $n = 4$.

The side s_0 in the hexagon that does not correspond to a variable in the corresponding well-parenthesized product of $n+1$ variables is shown as a dashed line segment. For clarity, each of the other sides, s_i, which corresponds to a variable in the corresponding well-parenthesized product, is labeled simply i.

The Generating Function of the Catalan Numbers

We started by defining the Catalan number c_n by means of a formula, and we

Figure 7. One-to-one correspondences with $n = 4$ between the sets of triangulations of a convex hexagon, well-parenthesized products, well-formed sequences of parentheses, stack permutations, and nonnegative paths.

then showed by a combinatorial argument that it enumerates nonnegative paths in the plane. We subsequently found one-to-one correspondences between several different types of combinatorial structures, starting with the nonnegative paths. It followed that the number of structures of each type must be equal to the number c_n of nonnegative paths. Having established the one-to-one correspondences, the same conclusion would follow if we showed (combinatorially or otherwise) that the number of structures of any particular type is given by

$$c_n = \frac{1}{n+1}\, C(2n, n). \tag{1}$$

In this section we will obtain (1) as the number of well-parenthesized products of $n+1$ variables using recurrence relations and generating functions (see Chapter 6 of *Discrete Mathematics and Its Applications*). In the next section we will investigate the behavior of the sequence $\{c_n\}$ for large values of n.

A sequence $\{a_n\} = a_0, a_1, \ldots$ satisfies a linear homogeneous recurrence relation of degree k with constant coefficients if each term a_n for $n \geq k$ can be computed recursively from the previous k terms by means of

$$a_n = C_1 a_{n-1} + C_2 a_{n-2} + \cdots + C_k a_{n-k} \tag{2}$$

for some constants C_i, $1 \leq i \leq k$, with $C_k \neq 0$. Any sequence satisfying (2) is completely determined by its k initial values.

The generating function of a sequence $\{a_n\}$ is the power series

$$A(x) = \sum_{n=0}^{\infty} a_n x^n.$$

If the sequence satisfies a linear homogeneous recurrence relation with constant coefficients of some fixed degree k, the generating function is a rational function of x. The polynomial in the numerator depends only on the k initial values, while the polynomial in the denominator depends only on the recurrence relation. Once the roots of the polynomial in the denominator are found (or estimated by an algorithm), the value of any term of the sequence can be obtained by means of partial fractions and known power series.

Suppose we define a_n to be the number of well-parenthesized products of n variables. We showed earlier that a_{n+1} is the Catalan number c_n given by (1). But let us find a_n using recurrence relations and generating functions. To form a well-parenthesized product of $n \geq 2$ variables $x_1, x_2, \cdots, x_n$ by the recursive definition, the outer parentheses would be the last ones added to form (AB), where A and B are well-parenthesized products, each having at least one of the variables. The outer parentheses then would enclose AB, where A is one of the a_i well-parenthesized products of the first i variables and B is one of the a_{n-i} well-parenthesized products of the last $n - i$ variables, for some i satisfying $1 \leq i \leq n-1$. Then, by the product rule and the sum rule, we obtain

$$a_n = a_1 a_{n-1} + a_2 a_{n-2} + \cdots + a_{n-1} a_1 = \sum_{i=1}^{n-1} a_i a_{n-i}. \tag{3}$$

Note that (3) is a homogeneous recurrence relation with constant coefficients. However, it is not linear, but *quadratic*. In addition, since the number of previous terms that determine a_n depends on n, and the degree of a recurrence relation is defined only when this number is some fixed integer k, the degree is not defined in (3).

If we define $a_0 = 0$, then the terms a_0a_n and a_na_0 can be added to the right-hand side of (3) without changing its value, and we obtain

$$a_n = a_0a_n + a_1a_{n-1} + \cdots + a_na_0 = \sum_{i=0}^{n} a_ia_{n-i}. \tag{4}$$

The right-hand side of (4) is zero for $n \le 1$, but for $n \ge 2$ it is the coefficient of x^n in the square of the generating function $A(x) = \sum_{i=0}^{\infty} a_nx^n$. Since $a_1 = 1$, it follows that

$$A^2(x) = A(x) - x,$$

which is a quadratic equation in $A(x)$. By the quadratic formula, we obtain as solutions

$$A(x) = \frac{1}{2}(1 \pm \sqrt{1-4x}). \tag{5}$$

Since $A(0) = a_0 = 0$, and when we set $x = 0$ the right-hand side of (5) with the plus sign is 1, we must take the minus sign. Thus the generating function of the sequence a_n is

$$A(x) = \frac{1}{2}(1 - \sqrt{1-4x}). \tag{6}$$

Using a generalization of the binomial theorem, it can be shown that

$$\sqrt{1-4x} = (1-4x)^{1/2} = 1 + \sum_{n=1}^{\infty} (-1)^n \frac{(\frac{1}{2})_n}{n!} 4^n x^n, \tag{7}$$

where $(\frac{1}{2})_n$ is the *falling factorial* function $(x)_n = x(x-1)\cdots(x-n+1)$ evaluated at $x = 1/2$. The terms in the sum in (7) can be simplified after writing 4^n as 2^n2^n, carrying out the multiplication of $(\frac{1}{2})_n = (\frac{1}{2})(\frac{1}{2}-1)\cdots(\frac{1}{2}-n+1)$ by $(-1)^n2^n$ termwise, writing the remaining factor 2^n as $2(2^{n-1})(n-1)!/(n-1)!$, and noting that $2^{n-1}(n-1)! = 2\cdot 4\cdots(2n-2)$. We then obtain

$$\sqrt{1-4x} = 1 - 2\sum_{n=1}^{\infty} \frac{1}{n}\, C(2n-2, n-1)\, x^n. \tag{8}$$

From (6) and (8) we obtain the generating function

$$A(x) = \sum_{n=1}^{\infty} \frac{1}{n}\, C(2n-2, n-1)\, x^n. \tag{9}$$

Thus the coefficient of x^{n+1} in (9) is $a_{n+1} = \frac{1}{n+1}\, C(2n, n)$, which by (1) is the Catalan number c_n. Similar methods could be used to show that the sizes of the other structures considered are Catalan numbers. It would suffice to show that the sequence satisfied the recurrence relation (4) and that the initial values agree, perhaps after a shift.

Asymptotic Behavior of the Catalan Numbers

Let us consider the behavior of the sequence $\{c_n\}$ of Catalan numbers for large values of n. In the previous section we let a_n be the number of well-parenthesized products of n variables and showed that the sequence $\{a_n\}$ satisfies the recurrence relation (3). Using the generating function $A(x)$ of the sequence $\{a_n\}$, we found that a_{n+1} is equal to the Catalan number c_n given by (1). If we substitute c_j for a_{j+1} in (3) and adjust the range of the index variable i, we see that the sequence $\{c_n\}$ satisfies the recurrence relation

$$c_n = c_0 c_{n-1} + c_1 c_{n-2} + \cdots + c_{n-1} c_0 = \sum_{i=0}^{n-1} c_i c_{n-i-1} \tag{10}$$

with $c_0 = 1$. This is a quadratic homogeneous recurrence relation with constant coefficients, but with the degree undefined. The asymptotic behavior of the sequence $\{c_n\}$ can be more easily found by showing that it satisfies a second homogeneous recurrence relation, one that is linear of degree one but with a *variable* coefficient. The linear recurrence relation could have been found earlier using Definition 1 of the Catalan numbers, but we will find it using the solution given by (1) to the quadratic recurrence relation (10).

A linear homogeneous recurrence relation of degree one with a constant coefficient has the form $a_n = Ca_{n-1}$. On iterating this recurrence relation $n-1$ times, we obtain the formula $a_n = C^n a_0$. Suppose that instead of making our original definition of c_n, which is the same as (1), we had defined c_n as the number of well-parenthesized products of $n+1$ variables. Then we would find, as in the last section, that the sequence $\{c_n\}$ satisfies the quadratic recurrence relation (10), and that the solution is given by (1). On using the formula for the binomial coefficient, it can be then be shown (see Exercise 6) that

$$c_n = \frac{4n-2}{n+1} c_{n-1}, \tag{11}$$

so that the constant coefficient C is replaced by a variable coefficient

$$C(n) = \frac{4n-2}{n+1} = 4 - \frac{6}{n+1}.$$

Clearly $C(n)$ increases to 4 as a limit as $n \to \infty$. However, this does not imply that c_n can be approximated by a constant multiple of 4^n. That would be the case if fact $C(n)$ was identically equal to the constant $C = 4$.

However, on using the familiar expression for a binomial coefficient involving three factorials, and replacing each of those factorials by an approximate value, we can approximate c_n, and use this approximation to find a function $f(n)$ with a relatively simple form such that $c_n = O(f(n))$. The simplest form of **Stirling's approximation** s_n of $n!$ is given by

$$s_n = \sqrt{2\pi n}\, e^{-n} n^n. \tag{12}$$

Using this approximation, it can be shown (see Exercise 7) that

$$c_n = O(n^{-3/2} 4^n). \tag{13}$$

Suggested Readings

1. K. Bogart, *Introductory Combinatorics*, Harcourt, Brace, Jovanovich, 1990.

2. L. Comptet, *Advanced Combinatorics*, D. Reidel, 1974.

3. W. Feller, *An Introduction to Probability Theory and Its Applications*, 2nd ed., Wiley, 1961.

4. F. Roberts, *Applied Combinatorics*, Prentice-Hall, 1984.

Exercises

In Exercises 1–4 find the structures that correspond to the given structures under the one-to-one correspondences established.

1. Given the sequence $(++-+-+--++--)$ of ± 1s with nonnegative partial sums, find or draw the corresponding
a) well-formed sequence of six pairs of parentheses
b) nonnegative path from the origin to $(12, 0)$
c) stack permutation of 123456.

2. Given the sequence $(++-+--++--)$ of ± 1s with nonnegative partial sums, find or draw the corresponding
a) well-parenthesized product of six variables
b) full binary tree with six leaves whose vertices are labeled with well-parenthesized products
c) triangulation of a convex septagon.

3. Given the following triangulation of a convex octagon, find or draw the corresponding
   a) well-parenthesized product of seven variables
   b) sequence of twelve ±1s with nonnegative partial sums
   c) nonnegative path from the origin to $(12, 0)$
   d) stack permutation of 123456.

4. Given the well-parenthesized product $((1(23))((45)6))$ of six variables (with x_i denoted by i), find or draw the corresponding
   a) sequence of ten ±1s with nonnegative partial sums
   b) stack permutation of 12345
   c) triangulation of a convex septagon.

5. Find the sequence of ten pushes $(+)$ and pops $(-)$ that produces the stack permutation 42135, and draw the corresponding nonnegative path.

6. Prove that the Catalan numbers c_n satisfy the recurrence relation (11).

7. Use the Stirling approximation (12) for $n!$ to prove that the Catalan number c_n satisfies (13).

In Exercises 8–10, suppose T is a triangulation of a convex $(n+2)$-gon P with diagonal set D. Let $v_0v_1 \cdots v_{n+1}$ be the vertices of P in order, and let s_i be the side $v_{i-1}v_i$ for $1 \le i \le n+1$, $s_0 = v_0v_{n+1}$. Denote by $\mathbf{p}$ the corresponding well-parenthesized product $s_1s_2 \cdots s_{n+1}$ of its sides other than s_0.

8. Prove that the diagonal v_iv_j is in the diagonal set D if and only if the product $s_{i+1}s_{i+2} \cdots s_j$ is well-parenthesized in $\mathbf{p}$.

9. Let $1 \le k \le n-1$. Prove that the nonnegative path corresponding to T meets the x-axis at the point $(2k, 0)$ if and only if D contains the diagonal v_kv_{n+1}.

10. Prove that the nonnegative path corresponding to T is positive if and only if D contains the diagonal v_0v_n.

Computer Projects

1. Given a positive integer n as input, find the Catalan number c_n using

a) the recurrence relation (10).

b) the recurrence relation (11).

2. Given positive integers n and N, find N random sequences $x_j = \{x_{ij}\}$ of n 1s and n -1s, compute the corresponding sequences $s_j = \{s_{ij}\}$ of partial sums $s_{ij} = x_{1j} + x_{2j} + \cdots + x_{ij}$, and let a, b be the number of positive, nonnegative sequences s_j, respectively. Compute the ratios

a) $a/C(2n, n)$.

b) $b/C(2n, n)$.

c) a/b.

3. Given a positive integer n, find a random sequence x of n 1s and n -1s that has a nonnegative sequence s of partial sums. Using s, produce the corresponding

a) well-formed sequence of parentheses.

b) stack permutation.

c) well-parenthesized product.

d) full binary tree (graphics).

e) triangulation of a convex polygon (graphics).

Exercise Solutions

1. a) Substituting left parentheses for 1 and right parentheses for -1 gives the well-formed sequence (( )( )( ))(( )).

b)

c) Starting with 123456[ ], and interpreting each + as a push and each $-$ as a pop, the sequence successively produces: 12345[6],1234[56], 1234[6]5, 123[46]5, 123[6]45, 12[36]45,12[6]345, 12[ ]6345, 1[2]6345, [12]6345, [2]16345, [ ]216345. The stack permutation is therefore 216345.

2. a) The string $\mathbf{q}$ of left parentheses together with the first five variables is ((1(23((45. Adding the last variable on the right and closing parentheses from the left gives the well-parenthesized product $\mathbf{p} = ((1(23))((45)6))$.

b) c)

3. a) ((1((2(34))(56)))7).
b) $(++-++-+--+--)$.
c)

d) 12345[6], 1234[56], 1234[6]5, 123[46]5, 12[346]5, 12[46]35, 1[246]35, 1[46]235, 1[6]4235, [16]4235, [6]14235, [ ]614235.

4. a) $(+++---++--+-)$.
b) 12345[6], 1234[56], 123[456], 123[56]4, 123[6]54, 123[ ]654, 12[3]654, 1[23]654, 1[3]2654, 1[ ]32654, [1]132654, [ ]132654.
c)

5.

[ ]42135		
[4]2135	4 Popped	$-$
[24]135	2 Popped	$-$
[124]35	1 Popped	$-$
1[24]35	1 Pushed	$+$
12[4]35	2 Pushed	$+$
12[34]5	3 Popped	$-$
123[4]5	3 Pushed	$+$
1234[ ]5	4 Pushed	$+$
1234[5]	5 Popped	$-$
12345[ ]	5 Pushed	$+$

The admissible sequence is therefore $(+-++-++---)$. The path is the following:

6. $c_n = \frac{1}{n+1}\frac{(2n)!}{(n!)^2} = \frac{2n(2n-1)}{(n+1)n^2}\frac{(2n-2)!}{((n-1)!)^2} = \frac{(4n-2)\cdot(2n-2)!}{(n+1)n\cdot((n-1)!)^2} = \frac{4n-2}{n+1}c_{n-1}$.

7. Substituting Stirling's approximation for the factorials in c_n gives $c_n = \frac{1}{n+1}\frac{(2n)!}{n!n!} \approx \frac{1}{n+1}\frac{\sqrt{4\pi n}\ (2n)^{2n}}{2\pi n\ n^{2n}} = \frac{2^{2n}}{(n+1)\sqrt{\pi n}} \approx \frac{4^n}{\sqrt{\pi}\ n^{3/2}} = O(n^{-3/2}\ 4^n)$.

8. If $v_iv_j \in D$, then the polygon P' with vertices $v_i, v_{i+1}, \ldots, v_j$ is triangulated by the set D' of diagonals of D that are also diagonals of P'. With v_iv_j serving as the excluded side of P', the product of sides $s_{i+1}s_{i+2}\cdots s_j$ is well-parenthesized into $\mathbf{p}'$ by recursively parenthesizing the two polygon sides that are sides of outer triangles, and then reducing the polygon. The order of reducing the polygons does not affect the well-parenthesized product finally obtained, so $\mathbf{p}'$ will be a subsequence of consecutive terms of $\mathbf{p}$. Conversely, if $s_{i+1}s_{i+2}\cdots s_j$ is well-parenthesized in p as $\mathbf{p}'$, then $\mathbf{p}'$ will appear on the diagonal joining the end vertices of the path formed by the sides in $\mathbf{p}'$. But that diagonal is v_iv_j.

9. Suppose that D contains the diagonal v_kv_{n+1}. Then after putting $i = k$, $j = n+1$ in Exercise 8, the product $s_{k+1}s_{k+2}\cdots s_{n+1}$ is well-parenthesized as a sequence $\mathbf{p}'$. Corresponding to $\mathbf{p}'$ is a nonnegative path from the origin to $(2(n-k), 0)$. Translating this path to the right $2k$ units identifies it with the segment of the path corresponding to T from $(2k, 0)$ to $(2n, 0)$. Conversely, suppose the nonnegative path corresponding to T meets the x-axis at the point $(2k, 0)$. On removing s_{n+1} and the n right parentheses from the corresponding well-parenthesized sequence $\mathbf{p}$ of the product $s_1s_2\cdots s_{n+1}$, we obtain a sequence $\mathbf{q}$ with n left parentheses interlaced with the product $s_1s_2\cdots s_n$. Since the path passes through the point $(2k, 0)$, the first $2k$ terms of $\mathbf{q}$ have k left parentheses interlaced with $s_1s_2\cdots s_k$. Hence the subsequence $\mathbf{q}'$ consisting of the last $2(n-k)$ terms of $\mathbf{q}$ interlaces $n-k$ left parentheses with $s_{k+1}s_{k+2}\cdots s_n$. Since the path is nonnegative, the segment corresponding to $\mathbf{q}'$ is nonnegative, and so corresponds to a triangulation of the $(n-k+2)$-gon with sides $s_{k+1}s_{k+2}\cdots s_{n+1}$ and the diagonal (of P) joining the end vertices of the path. But that diagonal is v_kv_{n+1}.

10. If D contains the diagonal v_0v_n, then $v_nv_{n+1}v_0$ is an outer triangle, so no diagonals on v_{n+1} can be in D. By Exercise 9 the corresponding path cannot meet the x-axis at any point $(2k, 0)$ for $1 \le k \le n-1$. Since the path starts on the x-axis, it can only return to the x-axis after an even number $2k$ of steps. Thus the path is positive. Conversely, if the path is positive, then D contains no diagonals v_kv_{n+1} for $1 \le k \le n-1$. Since every side of P is in exactly one triangle of T, s_n, s_0 must be in the same triangle. But the outer triangle $v_nv_{n+1}v_0$ is the only triangle that can contain both. Thus T contains the diagonal v_0v_n.