

# The Pill Problem, Lattice Paths and Catalan Numbers

Margaret Bayer  
University of Kansas  
Lawrence, KS 66045-7594  
bayer@ku.edu

Keith Brandt  
Rockhurst University  
Kansas City, MO 64110  
Keith.Brandt@Rockhurst.edu

## Introduction

In 1991, Knuth and McCarthy [9] posed the following problem in the *American Mathematical Monthly*:

A certain pill bottle contains  $m$  large pills and  $n$  small pills initially, where each large pill is equivalent to two small ones. Each day the patient chooses a pill at random, if a small pill is selected, (s)he eats it; otherwise (s)he breaks the selected pill and eats one half, replacing the other half, which thenceforth is considered to be a small pill.

(a) What is the expected number of small pills remaining when the last large pill is selected?

(b) On which day can we expect the last large pill to be selected?

When the solution was published [6], the *Monthly* editors commented that the origins of the problem were not clear. It turns out that several of those who submitted solutions had seen the problem before (at MIT and Michigan State, for example). In 2003, Brennan and Prodinger [3] studied the problem further and considered some variations, such as breaking the whole pills into more than two pieces.

In this paper we study a related question. The different sequences of pill selections are represented as paths in a binary tree that we call the *pill tree*. We count the vertices in the pill tree as a function of the initial numbers of large and small pills. In what follows we shall use the terms *whole* and *half* pills for large and small terms, to emphasize the size relationship. We

observe some connections among the pill tree, lattice paths and Catalan numbers.

The pill tree arises naturally in the work of Brandt and Waite [2] on the following probability question. What is the probability  $P_k(w, h)$  that a whole pill is selected on the  $k$ th day, given that the bottle starts with  $w$  whole pills and  $h$  half pills? In [2] a recurrence relation is given for  $P_k(w, h)$ . The size of the pill tree shows the difficulty of implementing the recursion efficiently. Brandt and Waite study various storage techniques (arrays, trees, etc.) to eliminate redundant calculations and thereby improve the performance of their implementation.

## The Pill Tree

For  $w$  and  $h$  nonnegative integers, the *pill tree*  $PT(w, h)$  is a labeled binary rooted tree with root labeled  $\langle w, h \rangle$ . A node labeled  $\langle u, v \rangle$  has *left child*  $\langle u - 1, v + 1 \rangle$  (if  $u > 0$ ) and *right child*  $\langle u, v - 1 \rangle$  (if  $v > 0$ ). A node labeled  $\langle u, v \rangle$  represents a bottle containing  $u$  whole pills and  $v$  half pills. The paths from root to leaf describe all possible sequences of configurations of pills in the bottle. A step down to the left represents choosing a whole pill; a step down to the right represents choosing a half pill. The root is the initial pill configuration, and the leaves all represent the empty configuration  $\langle 0, 0 \rangle$ . For example,  $PT(2, 1)$  is given in Figure 1.

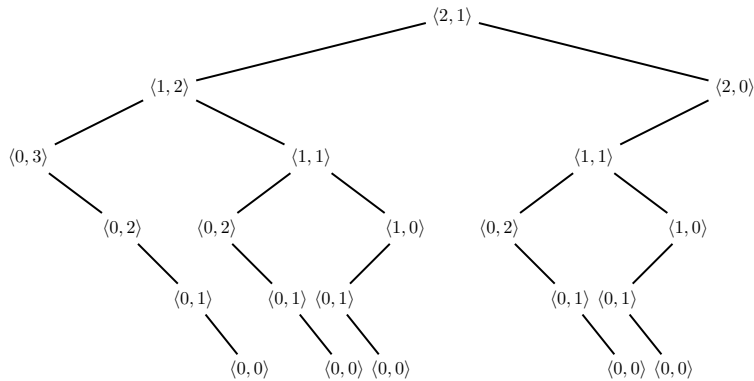


Figure 1: The Pill Tree  $PT(2, 1)$

Let  $T(w, h)$  be the number of nodes in the pill tree with initial configuration  $\langle w, h \rangle$ . The function  $T$  has an initial condition and recurrence relation

similar to the recurrence of Brandt and Waite for  $P_k(w, h)$ . For  $w, h \geq 0$ ,

$$T(w, h) = \begin{cases} h + 1 & \text{if } w = 0 \\ 1 + T(w - 1, 1) & \text{if } h = 0, w > 0 \\ 1 + T(w - 1, h + 1) + T(w, h - 1) & \text{otherwise} \end{cases} \quad (1)$$

Table 1 gives  $T(w, h)$  for some small values of  $w$  and  $h$ .

$w \backslash h$	0	1	2	3	4	5
0	1	2	3	4	5	6
1	3	7	12	18	25	33
2	8	21	40	66	100	143
3	22	63	130	231	375	572
4	64	195	427	803	1376	2210
5	196	624	1428	2805	5016	8398
6	625	2054	4860	9877	18276	31654

Table 1: Number of nodes in the pill tree  $PT(w, h)$

## What are these numbers?

How can mathematicians, faced with a table like this, find a formula for, or at least better understand, the numbers? We can turn to the *On-Line Encyclopedia of Integer Sequences* (OEIS) [4]. There we find that the numbers in column 0 are partial sums of Catalan numbers. The Catalan numbers  $C_n$  are defined as  $C_n = \frac{1}{n+1} \binom{2n}{n}$  for  $n \geq 1$ ; the first few terms in the sequence are 1, 2, 5, 14, 42. An equivalent and commonly used formula is  $C_n = \binom{2n}{n} - \binom{2n}{n-1}$ . In the 1750s Euler gave these as the numbers of triangulations of convex polygons. Eugene Catalan rediscovered them in 1838, using them to count well-formed sequences of parentheses. Since then, these numbers have turned up in a vast number of settings. Stanley [10] stresses this in his exercise 6.19 where he says, “Show that the Catalan numbers  $C_n = \frac{1}{n+1} \binom{2n}{n}$  count the number of elements of the 66 sets  $S_i \dots$  given below.” Stanley’s web page addendum [11] extends this list to over 200. A brief history with citations for the original references can be found in [10, p. 212]. For a comprehensive introduction to Catalan numbers see Koshy [7].

Returning to our table, we notice that each entry in the  $h = 1$  column is one less than the entry in the  $h = 0$  column one row down; this is easily confirmed in general by the middle line of equation (1). The entries of the  $h = 2$  column are Catalan numbers minus 2. Now it seems that we run out of luck: neither we nor the OEIS recognizes the sequence of numbers in the  $h = 3$  column. But what if these numbers are the partial sums of some sequence, as were the entries in the  $h = 0$  column? We ask OEIS about the sequence of differences from the  $h = 3$  column: 4, 14, 48, 165, 572, 2002, 7072. Again these numbers are related to Catalan numbers. The OEIS does not help to identify the sequences in the rows of the table, but it might be interesting to investigate these further. The connections between the finite sequences in the columns and the Catalan numbers lead to the following theorem.

**Theorem 1** *For  $w \geq 1$  the number of nodes in the pill tree is*

- a.  $T(w, 0) = \sum_{i=1}^{w+1} C_i$
- b.  $T(w, 1) = \sum_{i=2}^{w+2} C_i$
- c.  $T(w, 2) = C_{w+3} - 2$

Because of the recursion (1), each part of the theorem implies the others.

## Catalan numbers and lattice paths

To understand the role of the Catalan numbers in the pill tree, we focus on the application of Catalan numbers to counting lattice paths.

The integer lattice is the set of points in the Cartesian plane having integer coordinates. We think of this as a grid with length one vertical and horizontal line segments connecting adjacent lattice points. A *lattice path* is a sequence of connected rightward and upward segments going from  $(0, 0)$  to a point  $(c, d)$ . Moving along a rightward edge means adding  $(1, 0)$  to the lattice point; moving along an upward edge means adding  $(0, 1)$ . The whole path must contain exactly  $c + d$  edges,  $c$  horizontal (rightward) edges and  $d$  vertical (upward), and these  $c$  and  $d$  edges can be in any order. The number of lattice paths from  $(0, 0)$  to  $(c, d)$  is  $\binom{c+d}{d}$ , because the  $d$  horizontal edges can be chosen to be any of the  $c + d$  edges of the path.

Catalan numbers count a restricted class of lattice paths. First, assume the final point lies on the line  $y = x$ . Second, assume the lattice path never goes above the line  $y = x$ ; that is, each lattice point in the path is of the form  $(s, t)$ , with  $s \geq t$ . Such a path is called a *ballot path*, sometimes referred to as a *Dyck path*. The number of ballot paths from  $(0, 0)$  to  $(n, n)$  is the  $n$ th Catalan number  $C_n$ . (For a proof, see [7, p. 259] or your favorite combinatorics textbook.) For example,  $C_3 = 5$ , and Figure 2 shows the five ballot paths from  $(0, 0)$  to  $(3, 3)$ .

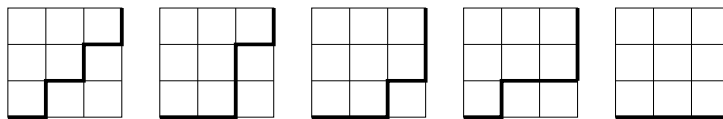


Figure 2: Ballot Paths

Counting lattice paths with various restrictions goes back a long way. For a detailed history see [5]. Before the 1960s interest was primarily among statisticians. Indeed, the book, *Lattice Path Counting and Applications*, by S. G. Mohanty [8], was published in the series Probability and Mathematical Statistics. In the past thirty years lattice paths have become a standard topic in combinatorics.

Why the term “ballot path” for a lattice path that does not go above the line  $y = x$ ? The following “ballot problem” dates from the 1880s [1]:

Candidate A wins an election with  $a$  votes over opponent B, who receives  $b$  votes. The votes are counted one at a time. What is the probability that throughout the count, A stays ahead of B?

The vote count can be represented by a lattice path, where a rightward edge is drawn each time a vote for A is counted, and an upward edge is drawn each time a vote for B is counted. To compute the probability that A stays ahead of B we need to count a set of restricted lattice paths: those lattice paths from  $(0, 0)$  to  $(a, b)$  that stay strictly below the line  $y = x$ . (What we have called a “ballot path” is slightly different, but it is not too hard to make the conversion: appending a horizontal segment at the beginning and a vertical segment at the end changes a path that does not go above the diagonal to one that stays strictly below the diagonal.)

It is time to return to the pill problem! It turns out that in the special case of  $h = 0$ , the pill tree  $PT(w, 0)$  represents ballot paths from  $(0, 0)$  to

$(w, w)$ . More generally we show a connection between pill sequences and lattice paths.

In the pill tree  $PT(w, h)$ , a node labeled  $\langle u, v \rangle$  represents a bottle having  $u$  whole pills and  $v$  half pills. We reach  $\langle u, v \rangle$  by selecting whole pills  $s = w - u$  times and selecting half pills  $t = (h + w - u) - v$  times. Note that  $t$  can also be interpreted as the total reduction in the number of whole or half pills. The pill configuration labeled  $\langle u, v \rangle$  in the pill tree can alternatively be identified by the pair  $(s, t)$ . To empty the bottle, whole pills are selected  $w$  times and half pills are selected  $w + h$  times. Thus, when the nodes of the pill tree are relabeled, each path in the tree from the root (now labeled  $(0, 0)$ ) to a leaf (now labeled  $(w, w + h)$ ) represents a lattice path from  $(0, 0)$  to  $(w, w + h)$ . See Figure 3 for the pill tree  $PT(2, 1)$  relabeled as the lattice path tree.

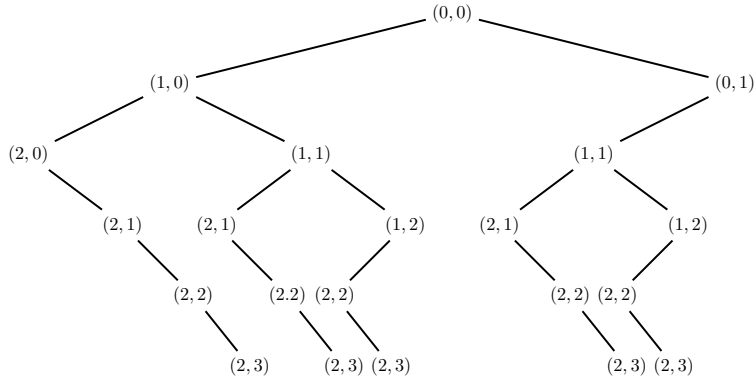


Figure 3: The Lattice Path Tree

When  $h = 0$ , the only half pills available in the process are those that came from whole pills, so  $s \geq t$ . This is also clear from the form  $(s, t) = (w - u, w - (u + v))$ . In this case the lattice paths corresponding to pill sequences are ballot paths.

**Theorem 2** *The number of pill sequences that start with  $w$  whole pills and no half pills is the Catalan number  $C_w$ .*

## Counting nodes in the pill tree

We found the number of lattice paths, or, equivalently, the number of leaves in the lattice path tree. But we want the total number of nodes in the pill

tree/lattice path tree. For the moment we still restrict ourselves to the case  $h = 0$ , and we work with the lattice path version of the pill tree. To count all the nodes in the lattice path tree, we count how many times a fixed label  $(s, t)$  (with  $0 \leq t \leq s \leq w$ ) occurs in the tree. For each node labeled  $(s, t)$ , there is a unique path from the root to that node, representing a lattice path from  $(0, 0)$  to  $(s, t)$  that does not go above the line  $y = x$ , and all such lattice paths are represented by paths in the tree. (Such a path can be extended in one or more ways to form a ballot path from  $(0, 0)$  to  $(w, w)$ .) Write  $C(s, t)$  for the number of such lattice paths. A formula for  $C(s, t)$  comes later, but we do not need it now, because our goal is a formula for the sum of such numbers for all pairs  $(s, t)$  with  $0 \leq t \leq s \leq w$ .

**Lemma 3** For all  $s \geq 0$ ,  $\sum_{t=0}^s C(s, t) = C_{s+1}$ .

**Proof:** Divide all allowable lattice paths from  $(0, 0)$  to  $(s + 1, s + 1)$  into  $s + 1$  categories, depending on which lattice point with  $x$ -coordinate  $s + 1$  they reach first. (Here “allowable” means not going above the line  $y = x$ .) Any path first reaching an  $x$ -coordinate of  $s + 1$  at the point  $(s + 1, t)$ , where  $0 \leq t \leq s$ , must have previously passed through the point  $(s, t)$ . The number of such lattice paths is  $C(s, t)$ . Summing these numbers  $C(s, t)$  gives the total number of lattice paths from  $(0, 0)$  to  $(s + 1, s + 1)$  that do not go above the line  $y = x$ , that is,  $C_{s+1}$ .  $\square$

**Proof of Theorem 1:** The number of nodes in the pill tree  $PT(w, 0)$  is the sum of  $C(s, t)$  for all pairs  $(s, t)$  with  $0 \leq t \leq s \leq w$ . Thus, by Lemma 3,

$$T(w, 0) = \sum_{s=0}^w C_{s+1} = \sum_{i=1}^{w+1} C_i.$$

The recursion (1) gives  $T(w, 1) = T(w + 1, 0) - 1$  and  $T(w, 2) = T(w + 1, 1) - T(w + 1, 0) - 1$ . Thus

$$T(w, 1) = \sum_{i=1}^{w+2} C_i - 1 = \sum_{i=2}^{w+2} C_i$$

and

$$T(w, 2) = \sum_{i=2}^{w+3} C_i - \sum_{i=1}^{w+2} C_i - 1 = C_{w+3} - 2.$$

$\square$

Can we use lattice paths to compute  $T(w, h)$  for  $h > 2$ ? The lattice points  $(s, t)$  no longer stay at or below the line  $y = x$ . For  $s = w - u$  and  $t = (w + h) - (u + v)$ , we know only that  $s + h \geq t$ , i.e., that the lattice points  $(s, t)$  do not go above the line  $y = x + h$ .

These and other variations of ballot paths also have a long history (see [5]). The following result was rediscovered by various people. We take it from [8, p. 3]. Write  $C_h(s, t)$  for the number of lattice paths from  $(0, 0)$  to  $(s, t)$  that do not go above the line  $y = x + h$ . ( $C_0(s, t)$  is what we called  $C(s, t)$  before.)

**Theorem 4** For  $s \geq 0$ ,  $h \geq 0$ , and  $t \leq s + h$ ,

$$C_h(s, t) = \binom{s+t}{s} - \binom{s+t}{s+h+1}.$$

The number of nodes in the pill tree  $P(w, h)$  is then the sum of the numbers  $C_h(s, t)$  over all pairs  $(s, t)$  satisfying  $0 \leq s \leq w$  and  $0 \leq t \leq s + h$ .

**Theorem 5** The number of nodes in the pill tree  $PT(w, h)$  is

$$T(w, h) = \sum_{s=0}^w \left[ \binom{2s+h+2}{s+1} - \binom{2s+h+2}{s} \right].$$

**Proof:** The number of nodes in the pill tree  $PT(w, h)$  is

$$\sum_{s=0}^w \sum_{t=0}^{s+h} C_h(s, t).$$

The inner sum works the same way as the sum of Lemma 3. That is,

$$\sum_{t=0}^{s+h} C_h(s, t) = C_h(s+1, s+h+1) = \binom{2s+h+2}{s+1} - \binom{2s+h+2}{s}.$$

□

Note: the summand  $\binom{2s+h+2}{s+1} - \binom{2s+h+2}{s}$  can also be written  $\frac{h+1}{s+h+2} \binom{2s+h+2}{s+1}$ .

This theorem gives part (a) of Theorem 1 directly. For  $h = 1$  note that

$$\binom{2s+3}{s+1} - \binom{2s+3}{s} = \binom{2s+4}{s+2} - \binom{2s+4}{s+1} = C_{s+2},$$

so we get part (b) of Theorem 1.



## Conclusion

Lattice paths come up in a variety of contexts, and so the lattice path tree can be interpreted in various ways. For example, a lattice path from  $(0, 0)$  to  $(n, m)$  gives a binary sequence of  $n$  minus ones and  $m$  plus ones, when an upward step is represented by 1 and a rightward step by  $-1$ . In the lattice path tree, moving to the left child appends a  $-1$  to the sequence, moving to the right child appends a 1. The number of nodes in the pill tree  $PT(w, h)$  is thus the number of sequences of plus and minus ones containing at most  $w$  minus ones and at most  $w + h$  plus ones, and having all partial sums (of initial sequences) at most  $h$ .

Yet again, the ubiquitous Catalan numbers have shown up in an unexpected place. Theorem 2 does not surprise us, as one can easily associate a pill sequence with a lattice path (or well-formed sequence of parentheses). But we found quite striking the values of  $T$  we observed in Table 1 (hence this paper). Theorem 1 reveals surprising connections between Catalan numbers and the lattice paths that are often used to define them.

## References

- [1] J. Bertrand. Calcul des probabilités. solution d'un problème. *C. R. Math. Acad. Sci. Paris*, 105:369, 1887.
- [2] K. Brandt and K. Waite. Using recursion to solve the pill problem. *J. Computing Sciences in College*, 24:136–141, 2009.
- [3] Charlotte A. C. Brennan and Helmut Prodinger. The pills problem revisited. *Quaest. Math.*, 26(4):427–439, 2003.
- [4] <http://oeis.org>. The on-line encyclopedia of integer sequences. 2011.
- [5] Katherine Humphreys. A history and survey of lattice path enumeration. *J. Statist. Plann. Inference*, 140:2237–2254, 2010.
- [6] Donald E. Knuth, John McCarthy, Walter Stromquist, Daniel H. Wagner, and Tim Hesterberg. E3429. *The American Mathematical Monthly*, 99(7):684, 1992.
- [7] Thomas Koshy. *Catalan numbers with applications*. Oxford University Press, Oxford, 2009.

- [8] Sri Gopal Mohanty. *Lattice path counting and applications*. Academic Press [Harcourt Brace Jovanovich Publishers], New York, 1979. Probability and Mathematical Statistics.
- [9] R. Padmanabhan, N. S. Mendelsohn, B. Wolk, Artin B. Boghossian, Donald E. Knuth, John McCarthy, Paul Erdos, Jeffrey Shallit, and Laszlo Toth. Elementary problems: E3427-e3432. *The American Mathematical Monthly*, 98(3):263–264, 1991.
- [10] Richard P. Stanley. *Enumerative combinatorics. Vol. 2*, volume 62 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, 1999. With a foreword by Gian-Carlo Rota and appendix 1 by Sergey Fomin.
- [11] Richard P. Stanley. Catalan addendum, 2013. <http://www-math.mit.edu/~rstan/ec/catadd.pdf>, 2013-05-25.

**Summary** We define the pill tree, which is a rooted, binary tree consisting of different sequences of whole and half pills from a problem posed by Knuth and McCarthy. We observe some connections among the pill tree, lattice paths, and Catalan numbers, and give an explicit formula for the number of nodes in the tree.