# Collatz-Pascal Triangle

Reed Kelly

October 12, 2006

## 0.1 Introduction

The Collatz-Pascal triangle is a creation which combines a concept from the Collatz sequence with another concept from Pascal's Triangle. Pascal's Triangle is very well understood, while the Collatz sequence is a great mystery. By combining concepts from each, a new entity can be created which has the mystery of Collatz, while preserving a few aspects of Pascal's triangle.

Each subsequent term of the Collatz sequence is formed by conditionally dividing a term of the sequence by two if it is even and multiplying the term by three and adding one if it is. The concept of conditionally performing arithmetic operations based on some characteristic of previous terms leads to interesting results. Specifically, dividing terms by two if they are even leads to many interesting things.

Using the "conditional division by two" approach, we can modify the traditional Pascal triangle. This is what we will call the Collatz-Pascal Triangle.

## 0.2 Background

Before exploring the properties of the Collatz-Pascal Triangle, we will define Collatz Sequences and the Pascal Triangle a little more precisely.

### 0.2.1 Collatz Sequences

Define a function $f(n)$ on the integers as follows:

$$f(n) = \begin{cases} 3n+1 & \text{if } n \text{ is odd;} \\ n/2 & \text{if } n \text{ is even.} \end{cases} \tag{1}$$

A Collatz sequence is defined by choosing an initial integer $k$ and forming the sequence of iterates of the function 1.

$$(k,\ f(k),\ f(f(k)), \ldots,\ f^{(i)}(k),\ \ldots)$$

#### Collatz Conjecture

In 1937, Lothar Collatz conjectured that all such sequences eventually arrive at the number 1 and then repeat the sequence:

$$(1, 2, 4, 1, 2, 4, \ldots)$$

More formally,

**Conjecture 1**

$$\forall\ k \in \mathbb{N},\ k > 0,\ \exists\ i \in \mathbb{N} : (f^{(i)} = 1)$$

This conjecture has yet to be proven. See http://en.wikipedia.org/wiki/Collatz_conjecture for more information on the Collatz Conjecture.

Figure 1: Pascal's Triangle

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| 1 | | | | | | | | | |
| 1 | 1 | | | | | | | | |
| 1 | 2 | 1 | | | | | | | |
| 1 | 3 | 3 | 1 | | | | | | |
| 1 | 4 | 6 | 4 | 1 | | | | | |
| 1 | 5 | 10 | 10 | 5 | 1 | | | | |
| 1 | 6 | 15 | 20 | 15 | 6 | 1 | | | |
| 1 | 7 | 21 | 35 | 35 | 21 | 7 | 1 | | |
| 1 | 8 | 28 | 56 | 70 | 56 | 28 | 8 | 1 | |
| 1 | 9 | 36 | 84 | 126 | 126 | 84 | 36 | 9 | 1 |
| ... | | | | | | | | | |

### 0.2.2 Pascal's Triangle

Although I cannot do justice to the history and definition of Pascal's triangle, I will include a few points for those who are less familiar with it. It plays a central role in all of mathematics and if frequently referenced in the studies of combinatorics, probability and statistics, and number theory, to name a few. For further references, see Wikipedia's definition: http://en.wikipedia.org/wiki/Pascal's_triangle.

The history of Pascal's triangle goes very far back. It appears to have been known by the Indian author and mathematician, Pingala, in about 400 BC. Other writings indicate that the Yang Hui of China knew about the triangle and it's properties in the eleventh century.

In the seventeenth century, Blaise Pascal wrote extensively about the triangle. Since Pascal's paper, *Traité du triangle arithmétique*, describes so many of the triangle's properties, many mathematicians still refer to it by his name. To avoid confusion, I will continue to refer to it as Pascal's triangle.

Pascal's triangle can be represented as an infinite array: $A = \{a_{i\,j}\}$ for all $i \in \mathbb{Z}$, and $j \in \mathbb{N}_0 = \{0, 1, 2, ...\}$, where $a_{0\,0} = 1$, and $a_{0\,j} = 0$ for all $j \in \mathbb{Z} \setminus \{0\}$, and

$$a_{i\,j} = a_{i-1\,j-1} + a_{i-1\,j}, \text{ for all } i \in \mathbb{N},\ j \in \mathbb{Z} \tag{2}$$

Given this representation, the first 10 rows of the array are shown in Figure 1 (omitting 0s).

### 0.2.3 Properties of Pascal's Triangle

Some of the well known properties of Pascal's Triangle include the following. There are many more properties that have been written about, but this is just a few.

### Binomial Coefficients

The process of adding adjacent entries of the previous row to form the terms of the next row is the same process that occurs to the coefficients of a polynomial in $x$ and $y$ that is multiplied by $(x + y)$. It is also the same process that occurs when you a ask how many ways can you get m heads from n independent coin tosses. This process leads to terms of the form:

$$a_{mn} = \binom{n}{m} = \frac{n!}{m!(n-m)!} \tag{3}$$

These terms are called the binomial coefficients. They are also written as shown in equation 3 and as $nCm$. A row of Pascal's triangle is referred to as a binomial distribution. They are used in probability, number theory and many other areas of mathematics.

### Symmetry

You may have noticed from Figure 1 that the left and right side of the non-zero numbers in the rows are the same. This is because, for row $n$, and $j \leq n$, we have $a_{nj} = a_{n\,(n-j)}$.

This can be seen from either the definition or from the fact that the terms are binomial coefficients.

From the definition, we can construct a proof by induction using the initial and inductive steps: 4 and 5.

For row 0, we have the following.

$$a_{0\,0} = a_{0\,(0-0)} \tag{4}$$

Assume that all the rows are symmetric up to row $n - 1$.

$$a_{n\,j} = a_{n-1\,j-1} + a_{n-1\,j} = a_{n-1\,(n-1-(j-1))} + a_{n-1\,(n-1-j)} = a_{n\,(n-j)} \tag{5}$$

### Row Sums

If you add the terms of the $n$th row from Pascal's Triangle, you get $2^n$. The 0th row is just $(1)$, so the sum is just $2^0$. The sum of the 4th row $(1, 4, 6, 4, 1)$ is $16 = 2^4$. The reason for this can be seen by considering the coefficients of the polynomial $(1 + x)n$. As mentioned before, the coefficients correspond to the rows of Pascal's Triangle. The value of this polynomial at $x = 1$ is simply the sum of the coefficients. It is also $(1 + 1)^n = 2^n$.

Again, an inductive proof can show this another way. We know that it is true for row 0 that $2^0 = 1$, which is the sum of the entries in row 0. In creating each row, each non-zero term is added to two of the terms of the next row. Therefore, the sum of the terms in each row is twice the sum of the terms in the previous row.

### Relationship to the Normal Distribution

If you divide the terms of a row of Pascal's triangle by $2^n$, the result is the range of a probability density function (p.d.f.). This is the p.d.f. of the binomial distribution. To be the range of a p.d.f., the sum of the values of a discrete function must sum to 1.

Very informally, the limit of this distribution as the row number goes to infinity is a normal distribution with the same mean and variance.

## 0.3 Collatz-Pascal Triangle

There is no single best choice for a Collatz-Pascal triangle. Unlike Pascal's triangle, choosing $(...0, 0, 1, 0, 0, ...)$ as an initial row results in a trivial triangle. Rather, we can define a Collatz-Pascal triangle that requires the choice of an initial row and then explore the set of all such triangles.

Choose the first row as a sequence of integers $c_{0j}$ for all $j \in \mathbb{Z}$. A Collatz-Pascal triangle can be represented as an infinite array:

$$C = \{c_{i\,j}\} \text{ for all } i \in \mathbb{Z}, \text{and } j \in \mathbb{N}_0 = \{0, 1, 2, ...\}, \text{ such that}$$

$$\text{for all } j \in \mathbb{Z},\ c_{0j} \text{ is the chosen initial row, and}$$

$$\text{for all } i \in \mathbb{N} \text{ and } j \in \mathbb{Z},\ c_{i\,j} = \begin{cases} c_{i-1\,j-1} + c_{i-1\,j} & \text{if } c_{i-1\,j-1} + c_{i-1\,j} \text{ is odd;} \\ (c_{i-1\,j-1} + c_{i-1\,j})/2 & \text{if } c_{i-1\,j-1} + c_{i-1\,j} \text{ is even.} \end{cases} \quad (6)$$

If there are an infinite number of non-zero entries in the first row, the array is more difficult to analyze. Simplifications include: limiting the non-zero terms of the initial row to the non-negative indices $\mathbb{N}_0 = \{0, 1, 2, ...\}$, and limiting the non-zero terms to a finite set of non-negative indices. Unless mentioned otherwise, a Collatz-Pascal triangle will refer to the latter case.

A good starting row is $(...0, 0, 4, 0, 0, ...)$. The first 10 rows of this triangle are shown in Figure 2. Henceforth, this instance of a Collatz-Pascal triangle will be referred to as *CPT4*.

### 0.3.1 Properties of a Collatz-Pascal Triangle

The Collatz-Pascal triangle has some properties that are understood, but many others that have not been proven. The following are a couple of the know properties of the Collatz-Pascal triangle.

#### Symmetry of Collatz-Pascal Triangles

If the initial sequence is symmetric, then the generated Collatz-Pascal triangle is symmetric as well. This follows, because the Collatz-Pascal recursion step 6 is itself a symmetric function. This applies to initial sequences that do not have a finite basis (infinite non-zero elements).

Figure 2: CPT4

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| 4 | | | | | | | | | |
| 2 | 2 | | | | | | | | |
| 1 | 2 | 1 | | | | | | | |
| 1 | 3 | 3 | 1 | | | | | | |
| 1 | 2 | 3 | 2 | 1 | | | | | |
| 1 | 3 | 5 | 5 | 3 | 1 | | | | |
| 1 | 2 | 4 | 5 | 4 | 2 | 1 | | | |
| 1 | 3 | 3 | 9 | 9 | 3 | 3 | 1 | | |
| 1 | 2 | 3 | 6 | 9 | 6 | 3 | 2 | 1 | |
| 1 | 3 | 5 | 9 | 15 | 15 | 9 | 5 | 3 | 1 |
| ... | | | | | | | | | |

#### Repeating Columns

Any Collatz-Pascal triangle starting with a row that is 0 for all indices less than some index, the first non-zero column eventually becomes a constant. This can easily been seen as follows. During each iteration, the entry in the previous row and same column is added to zero and then divided by 2 if possible. This results in the entries of the column being divided by 2 until an odd number is returned. Once an odd number appears, then all subsequent numbers in that column are the same odd number.

It can also be proven that every column of such a Collatz-Pascal triangle will eventually repeat. The period of repetition may grow the further the column is from the initial column.

### 0.3.2 Sequences Related to a Collatz-Pascal triangle

Based on *CPT4*, some interesting subsequences can be defined. Some of these are analogous to subsequences of Pascal's triangle, but others can be defined.

#### Center of Collatz-Pascal triangle

The terms: $CC = c_{2i\,i}$ for all $i \in \mathbb{N}_0$, form a mostly increasing sequence. Here are some of the terms of the sequence:

4, 2, 3, 5, 9, 15, 27, 25, 47, 89, 107, 119, 241, 545, 699, 1471, 3313, 4288, 15661, 31739, 30813, 35143, 92101, 123614, 384815, 788429, 1532363, 2995379, 6281191, 13569969

**Conjecture 2** *The terms of $CC$ have no maximum value.*

#### Periods of Repetition

Since all the columns eventually repeat, one can ask what the periods of repetition are. Obviously, the zeroth column has a period of 1. Since the first

Figure 3: Fourth Column-3

| | | | |
|---|---|---|---|
| 0 | 1 | 2 | 3 |
| 0 | 1 | 3 | 5 |
| 0 | 1 | 2 | 4 |
| 0 | 1 | 3 | 3 |
| 0 | 1 | 2 | 3 |
| 0 | 1 | 3 | 5 |
| 0 | 1 | 2 | 4 |
| 0 | 1 | 3 | 3 |
| ... | | | |

Figure 4: Fourth Column-5

| | | | |
|---|---|---|---|
| 0 | 1 | 2 | 5 |
| 0 | 1 | 3 | 7 |
| 0 | 1 | 2 | 5 |
| 0 | 1 | 3 | 7 |
| ... | | | |

column depends only on the starting term and the zeroth column, the first column eventually repeats the sequence $(2, 3)$, so its period is 2. For *CPT4*, some of the terms of this sequence are:

1, 2, 4, 8, 24, 24, 72, 72, 72, 1080, 4320, 25920

### 0.3.3 Minimal Repetition Sequence

If we consider arbitrary initial sequences that start with (0,1), we can consider try to characterize all the repeating columns. As was stated above, all the columns eventually repeat. If we assume that the first column is all zeros, then the second column would have to be all ones. Assuming that the subsequent terms of the first row are all non-negative integers, then there are two possibilities for the eventual repeating cycles in the third column: (1,1,...) and (2,3,2,3,...). If the third column repeats with ones (1,1,...), then the subsequent possibilities are the same as they were with the second column being (1,1,...). Therefore, we can focus our attention on the case where the third column is (2,3,2,3,...). Note that there are no other repeating sequences that are possible for the third column.

With the third column repeating (2,3,2,3,...), there are three possibilities for the fourth column. Following the 2 in the row, they start with 3, 5, and 6. Figures 3, 4, and 5.

With each additional column, more choices seem to occur. Since every column eventually repeats, the subsequent columns must themselves repeat. This

Figure 5: Fourth Column-6

| | | | |
|---|---|---|---|
| 0 | 1 | 2 | 6 |
| 0 | 1 | 3 | 4 |
| 0 | 1 | 2 | 7 |
| 0 | 1 | 3 | 9 |
| 0 | 1 | 2 | 6 |
| 0 | 1 | 3 | 4 |
| 0 | 1 | 2 | 7 |
| 0 | 1 | 3 | 9 |
| ... | | | |

can lead to many more questions about the growth in number and size of the repeating columns.

If we start with the row (0,1,2,3) as in Figure 3, we can consider this minimal in a way. It is not as small as (1,1,1,...), but 3 is the least next term for the first row after (0,1,2). Continuing in this way, we can ask what is the minimal row that starts a repeating sequence. More precisely, we can use a greedy strategy to choose the least term in the row each time. This results in the following sequence:

0, 1, 2, 3, 4, 11, 21, 27, 28, 58, 220, 649, 1042, 1427, 2265, 3520, 7009, 12500, 16525, 26425, 68235, 234717,...

Note that computing the last term (234717) took about 4 hours on a dual-core pentium system using a Perl script.

## 0.4 Shape

While the columns of the array eventually repeat, the terms toward the middle of the triangle appear to grow exponentially. To see that, the image in Figure 6 is a chronographic map of the log of terms from *CPT4*.

This image was produced by tilting the triangle so that the zeroth column goes off to the left at a 45 degree angle and the right most non-zero entries are off to the right at a 45 degree angle form the top. This has the effect of making the triangle appear symmetric about the middle of the page.

Also, after computing the entries of the triangle, the logarithm was taken, followed by getting the floor to make an integer and computing a color modulo 8. This produces a kind of contour map of the logarithm of the triangle's entries.

One thing that can be noticed from the graph is that the colored lines on the edges seem to be evenly spaced. If this is true, then the growth of the average value of the columns of the triangle is exponential.

Figure 6: Chronographic Map of Log(CPT4)