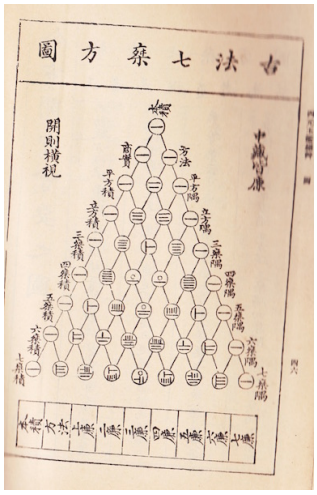
Mathematical Foundations of Computer Science

CS 499, Shanghai Jiaotong University, Dominik Scheder

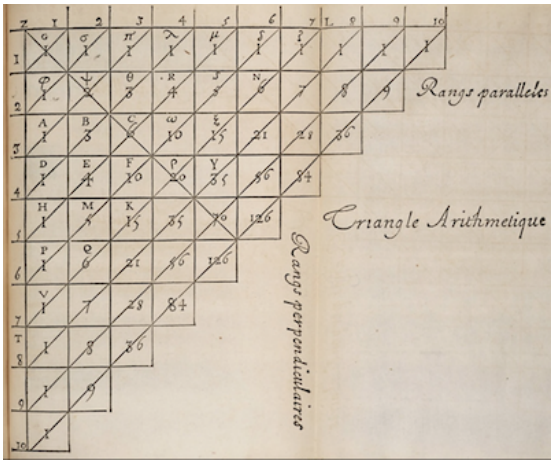
- Monday, 2018-03-19, homework handed out
- Sunday, 2018-03-25, 12:00: submit questions and first submissions. You'll get feedback until Wednesday.
- Sunday, 2018-03-29 (Wednesday), 18:00: submit your review of the other group's first submission.
- 2018-04-01: submit final solution.

4 Pascal's Triangle Modulo 2

Here are two early tables of the binomial coefficient:

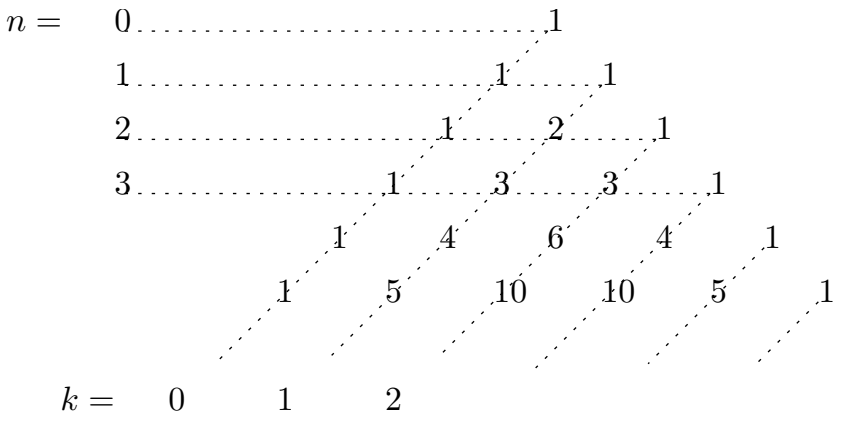


Yang Hui triangle from the book "Jade Mirror of the Four Unknowns" by Zhu Shijie, 1303 (Wikimedia)



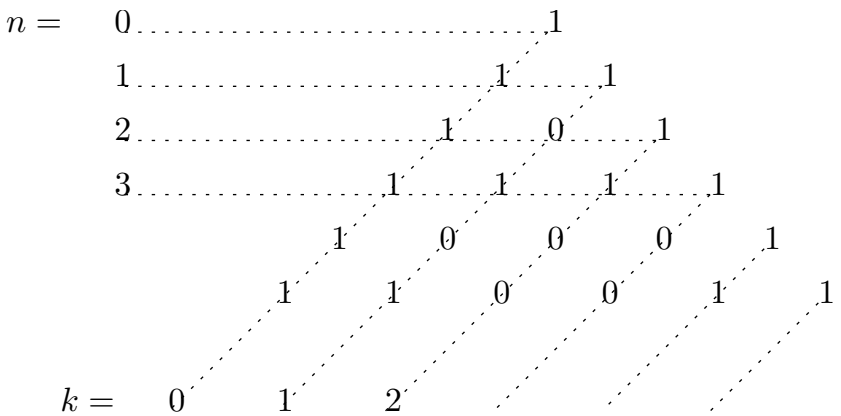
Blaise Pascals version of the triangle (Source: Wikimedia)

Here is my version of "Pascal's triangle", indicating that rows are indexes by n and "columns" by k:

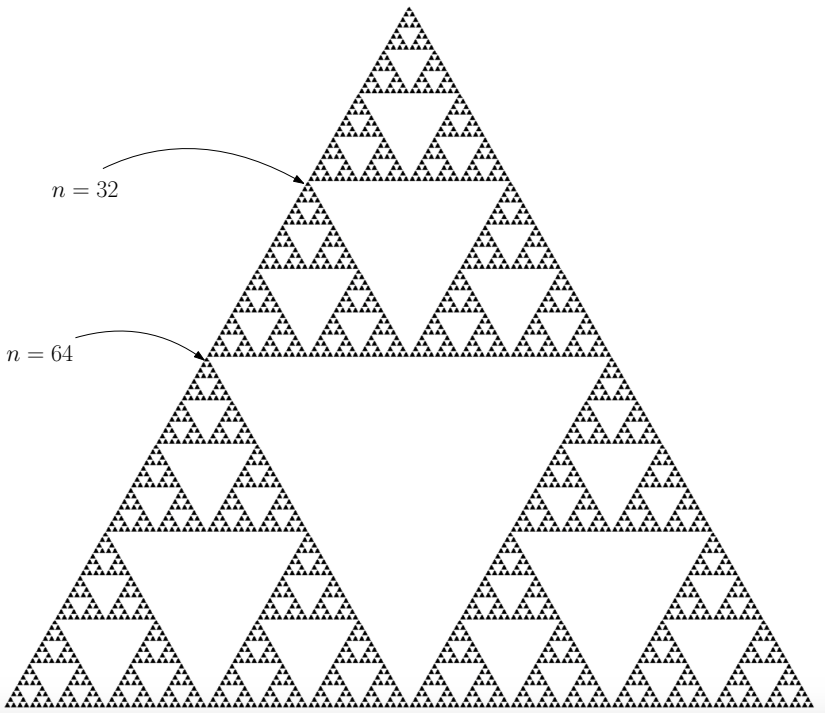


4.1 Lucas Theorem: $\binom{n}{k} \mod 2$

Something interesting happens when we take the triangle modulo 2, that is, we replace even numbers by 0 and odd numbers by 1:



If we draw a black dot for every 1 and look at a larger section of this triangle, we get the following pattern, known as the Sierpinski triangle:



Note the amazing recursive structure. This suggests we should be able to compute $\binom{n}{k} \mod 2$ without actually computing $\binom{n}{k}$, by somehow employing this structure. In fact, here is a cool result by Édouard Lucas, which we state here in a simpler, more special version:

The set $\mathbb{N}_0$ comes equipped with a partial ordering $\preceq$, in which $x \preceq y$ if for every i, the i^{th} least significant bit of x is at most that of y. Put in a simpler way, we write x and y as bit strings in binary. If their length differ, we put a bunch of 0's in front of the smaller number to make both strings of equal length d. Then we simply compare those strings using the usual partial ordering $\preceq$ on $\{0,1\}^d$. For example, $3 \preceq 7$ since $011 \preceq 111$, and $5 \preceq 23$ since $00101 \preceq 10111$, but $7 \not\preceq 8$ since $0111 \not\preceq 1000$.

Theorem 4.1. *Let $n, k \in \mathbb{N}_0$. Then $\binom{n}{k}$ is odd if $k \preceq n$ and even otherwise.*

Note that this theorem lets us compute $\binom{n}{k} \mod 2$ quickly for numbers n, k having millions of digits, whereas no computer on Earth has the memory to evaluate the formula

$$\binom{n}{k} = \frac{n \cdot (n-1) \cdot (n-2) \cdot \cdots \cdot (n-k+2) \cdot (n-k+1)}{k \cdot (k-1) \cdot (k-2) \cdot \cdots \cdot 2 \cdot 1}$$

for values that large. Let me now walk you through a proof of this theorem.

Definition 4.2. *For a natural number $n \in \mathbb{N}$, let $|n|_1$ be the number of 1's in the binary representation of n. For example, $|1|_1 = |2|_1 = |4|_1 = 1$ but $|3|_1 = 2$ and $|7|_1 = 3$.*

Definition 4.3. *For a natural number $a \in \mathbb{N}$ define $f(a)$ as the number of times the factor 2 appears in a. Formally,*

$$f(a) := \max\{k \mid 2^k \text{ divides } a\} .$$

For example, $f(24) = 3$ since 8 divides 24 but 16 does not.

Exercise 4.4. Find a closed formula for $f(n!)$ in terms of n and $|n|_1$.

Exercise 4.5. Find a closed formula for $f\left(\binom{n}{k}\right)$ in terms of $n, k, |n|_1$, and so on.

Exercise 4.6. Prove Theorem 4.1. With our new notation, prove that $f\left(\binom{n}{k}\right)$ is 0 if $k \preceq n$ and at least 1 if $k \not\preceq n$.

4.2 Almost Empty Rows

One feature of the Sierpinski triangle is that some rows are almost empty. For example, row 64 has a black dot at the very left and the very right, and only white space in between. This is because

Theorem 4.7. *Let $d \in \mathbb{N}_0$ and $0 < k < 2^d$. Then $\binom{2^d}{k}$ is even.*

Although this theorem follows easily from Lucas' Theorem, I want you to think about an alternative proof. Intuitively, if some number is even, then one suspects it can be proved by "pairing things up" perfectly. After all, if you can prove that in a set S, every element can be "married" to another element, you have partitioned S into couples and thus $|S|$ must be even. So let's see whether there is a proof of Theorem 4.7 along these lines. This is also valuable because it lets you practice with notions of sets and functions.

Consider the set $\{0,1\}^d$. You can view this as the set of all binary strings of length d. This set has size $2^d = n$. For $1 \leq i \leq d$ and $x \in \{0,1\}^d$ let $f_i(x)$ be x with the i^{th} position flipped. For example, $f_3(11011) = 11111$.

Exercise 4.8. Show that f_i is an involution without a fixed point. That is, $f(f(x)) = x$ and $f(x) \neq x$ for all $x \in \{0,1\}^d$.

Proof. If x has n bits, write down x as $x = a_1a_2a_3...a_n$.
$\forall\ j = 1, 2, 3, ...n,\ a_j = 0 or 1$.
If a_j is flipped, record it as $\bar{a}_j$.
Then $\forall\ i = 1, 2, 3, ...n,\ f_i(x) = a_1a_2a_3...\bar{a}_i...a_n$
Then $f(f_i(x)) = a_1a_2a_3...\bar{\bar{a}}_i...a_n = a_1a_2a_3...a_i...a_n = x$.
If there is x_k making $f(x) = x$,.
Then for an i, $x = a_1a_2a_3...a_n = a_1a_2a_3...\bar{a}_i...a_n$.
$a_i = \bar{a}_i$.
It is impossible.
Above all,f_i is an involution without a fixed point.

□

Let $S \subseteq \{0,1\}^d$. We define $f_i(S)$ as the set arising from applying f_i to every element of S. Formally,

$$f_i(S) := \{f_i(x) \mid x \in S\}\ .$$

Given a set $S \subseteq \{0,1\}^d$, we call an index $i \in [n]$ *active* for S if $f_i(S) \neq S$.

Exercise 4.9. Let $d = 3$ and $S = \{000, 100\}$. Which of the indices $1, 2, 3$ are active?

$f_1(S) = \{100, 000\} = S$
$f_2(S) = \{010, 110\} = S$
$f_1(S) = \{001, 101\} = S$
2,3 are active.

Exercise 4.10. Let $S \subseteq \{0,1\}^d$. Show that if $S \neq \emptyset$ and $S \neq \{0,1\}^d$ then S has at least one active index.

Proof. Soppose S has no active index and $S \neq \emptyset$.
Then $\forall i \in \{1, 2, 3...d\}$, $f_i(S) = S$.
$S = \{s_1, s_2...s_n\}$
$s_j = a_1a_2...a_d$
So $f_i(s_j) = a_1a_2...\bar{a}_ja_n \in S$.
So take a sequence s from S, $\forall i$, when the i^th position is flipped,the new $s' \in S$.
Then through change every bit of s and its child sequences, we can get every sequences of $\{0,1\}^d$.
$s = \{0,1\}^d$.
So if $S \neq \emptyset$ and $S \neq \{0,1\}^d$ then S has at least one active index.
□

Given $S \subseteq \{0,1\}^d$, define $f(S)$ as follows: if $S = \emptyset$ or $S = \{0,1\}^d$ define $f(S) = S$. Otherwise, let $f(S) := f_i(S)$ where i is the smallest active index of S (which exists by the previous exercise).

Exercise 4.11. Show that f is an involution. That is, $f(f(S)) = S$. Furthermore, show that the only fixed points of f are $\emptyset$ and $\{0,1\}^d$.

Proof. From Ex 4.8, we know $\forall i = \{1, 2...d\}$,$\forall x \in S$, $f(f(x)) = x$.
Then $\forall s \in S$, $f(f(s)) = s$.
So $f(f(S)) = S$.
If f has fixed point, namely $f(S) = S$.
From Ex 4.10, we know that it is impossible that $f(S) = S$ unless $S = \emptyset$ or $S = \{0,1\}^d$.
Above all,$f(f(S)) = S$. And the only fixed points of f are $\emptyset$ and $\{0,1\}^d$.
□

Exercise 4.12. Let $\mathcal{S} = \binom{\{0,1\}^d}{k}$. This is a set of sets, and each set $S \in \mathcal{S}$ consists of exactly k strings from $\{0,1\}^d$. Prove the following statements:

1. f is a bijection from $\mathcal{S}$ to $\mathcal{S}$.
2. For $1 \leq k \leq 2^d - 1$, this bijection is an involution without fixed points.
3. $|\mathcal{S}|$ is even for $1 \leq k \leq 2^d - 1$.

Proof. □

1. For a sequence $\in S$, $f(s)$is a certain sequence.

   So for $S = \{s_1, s_2..., s_k\}$,$f(S) = \{f(s_1), f(s_2), ..., f(s_k)\}$is a certain set.

   Then f is surjective.

   If for $a, b \in \mathcal{S}$,$a \neq b$,$f(a) = f(b) = A$.

   From exercise above we know,$f(f(a)) = a$ and $f(f(b)) = b$,namely $f(A) = a$,$f(A) = b$

   which is contradictive with f is surjective.

   So f is bijective.

2. In 1 we have proved $f(f(S)) = S$,it is an involution.

   If f has fixed points, namely f(S)=S.

   From Ex 4.11, the only fixed points of f are $\emptyset$ and $\{0,1\}^d$.

   However $k \in [1, 2^d - 1]$, f has no fixed points.

   For $1 \leq k \leq 2^d - 1$, this bijection is an involution without fixed points.

3. Since $\mathcal{S}$ is bijective, it has a number of couples (S_i, S_j) in which $f(S_i) = S_j$.

   So $|\mathcal{S}|$ is even.

Exercise 4.13. Complete the proof of Theorem 4.7.

Proof. □

$|\mathcal{S}| = \binom{\{0,1\}^d}{k}$,then $|\mathcal{S}| = \binom{2^d}{k}$.
$|\mathcal{S}|$ is even.
Then $\binom{2^d}{k}$ is even.

***Exercise 4.14.** *Generalize the above "combinatorial" proof to show the following theorem:*

Theorem 4.15. *Let $n = p^d$ where p is a prime number. Then p divides $\binom{n}{k}$ unless $k = 0$ or $k = n$.*

Proof. □

1. Let $S \subseteq 0, 1, 2, 3...p - 1^d$.

   For $1 \leq i \leq d$ and $x \in \{0,1\}^d$ let $g_i(x)$ be x with the i^th position replaced with x'. 0 is replaced by 1, 1 by $2...p-2$ by $p-1$,$p-1$ by 0.

2. Let $S \subseteq \{0, 1...p-1\}^d$. We define $g_i(S)$ as the set arising from applying g_i to every element of S. Formally,

$$g_i(S) := \{g_i(x) \mid x \in S\} .$$

   Given a set $S \subseteq \{0,1\}^d$, we call an index $i \in [n]$ *active* for S if $g_i(S) \neq S$.

   Given a set $S \subseteq \{0, 1...p-1\}^d$, we call an index $i \in [n]$ *active* for S if $g_i(S) \neq S$.

3. Similar to Ex 4.12, let Q = $\binom{\{0,1...p-1\}^d}{k}$.

   Every p sequences are in a group, in which $g(s_1) = s_2$,$g_(s_2) = s_3...g_(s_{p-1}) = s_p$,$g(s_p) = s_1$.

4. Then $|Q|$ can be divided by p.

   Then p divides $\binom{n}{k}$ unless $k = 0$ or $k = n$.