香港電腦奧林匹克競賽
Hong Kong Olympiad in Informatics

Mathematics in OI (II)

Isaac Wong {WongChun1234}

2026-05-02

Motivation

- In Mathematics in OI (I), we mainly dealt with topics in **number theory** and **modular arithmetic** with an emphasis on how to compute things **correctly** and **efficiently**.
- In today's session, we will shift our focus to **combinatorics** – the mathematics of counting and arranging.

Motivation

- As you can see, knowing how to code things is not enough. You can never be certain whether a statement is true or false.
- That’s why **proving** is an important aspect in mathematics.
- Today, we’ll focus on proving. Hopefully it helps in developing your critical reasoning skills, and develop better intuition.

Table of Contents

Factorial

$n!$ (n factorial)

Recursive definition: $n! = \begin{cases} 1 & \text{if } n = 0 \\ n \times (n-1)! & \text{otherwise} \end{cases}$

Combinatorial definition: $n!$ is the **number of permutations** of $1 \ldots n$

⚠ **Are these two definitions equivalent?**

Factorial

Aim: To show that the two definitions are equivalent.

Let $n!$ be the number of permutations of $1 \ldots n$. ← *combinatorial definition*

- If $n = 0$, there is 1 permutation of 0 objects.
  $\therefore n! = 1$
- Otherwise, consider any permutation of $1 \ldots n-1$:

  $$\hat{} \quad 1 \quad \hat{} \quad 2 \quad \hat{} \quad 3 \quad \hat{} \quad \ldots \quad \hat{} \quad n-1 \quad \hat{}$$

  There are n locations to insert the number $n \Rightarrow$ permutation of $1 \ldots n$.
  $\therefore n! = n \times (n-1)!$

Combining the two cases, $n! = \begin{cases} 1 & \text{if } n = 0 \\ n \times (n-1)! & \text{otherwise} \end{cases}$. ← *recursive definition*

Explicit Formula for $\binom{n}{r}$

Theorem

For $0 \leq r \leq n$, the following holds:

$$\binom{n}{r} = \frac{n!}{r!(n-r)!}$$

Explicit Formula for $\binom{n}{r}$

$$\binom{n}{r} = \frac{n!}{r!(n-r)!} \tag{1}$$

We show that $n! = \binom{n}{r} r!(n-r)!$.

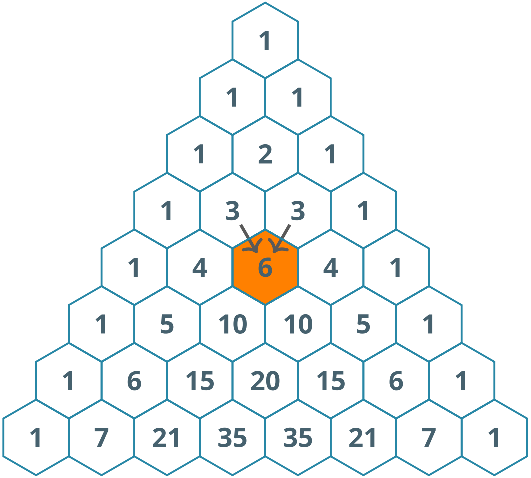
What are we counting?

The number of permutations of $1 \ldots n$.

- LHS: By definition, it is $n!$.
- RHS: Choose r elements out of n elements to be the first r elements.
  - For the first r elements, there are $r!$ ways to permute them.
  - For the next $n-r$ elements, there are $(n-r)!$ ways to permute them.

By multiplication principle, there are $\binom{n}{r} r!(n-r)!$ permutations.

Combinations and Pascal's Triangle



Pascal Triangle: The r-th column on the n-th row is $\binom{n}{k}$.

Pascal's Identity:

$$\binom{n}{r} = \binom{n-1}{r} + \binom{n-1}{r-1}$$

Let's use **two different methods** to prove this.

Equation (2) – Pascal's Identity

$$\binom{n}{r} = \binom{n-1}{r} + \binom{n-1}{r-1} \tag{2}$$

Method 1:

What are we counting?

The number of ways to choose r elements from $1 \ldots n$.

- LHS: By definition, it is $\binom{n}{r}$.
- RHS: Condition on whether n is chosen.
  - If no, we need to choose r elements from $1 \ldots (n-1)$. By definition, there are $\binom{n-1}{r}$ ways.
  - If yes, we need to choose $r-1$ elements from $1 \ldots (n-1)$. By definition, there are $\binom{n-1}{r-1}$ ways.

Equation (2) – Pascal's Identity

$$\binom{n}{r} = \binom{n-1}{r} + \binom{n-1}{r-1} \tag{2}$$

Method 2: Use the explicit formula for $\binom{n}{r}$.

$$\begin{aligned}
\binom{n-1}{r} + \binom{n-1}{r-1} &= \frac{(n-1)!}{r!(n-1-r)!} + \frac{(n-1)!}{(r-1)!(n-r)!} \\
&= \frac{(n-1)!}{(r-1)!(n-r-1)!} \cdot \left(\frac{1}{r} + \frac{1}{n-r}\right) \\
&= \frac{(n-1)!}{(r-1)!(n-r-1)!} \cdot \frac{n}{r(n-r)} \\
&= \frac{n!}{r!(n-r)!}
\end{aligned}$$

Combinatorial Proof and Algebraic Proof

Notice that Method 1 and Method 2 are very different approaches.

- **Method 1 (Combinatorial Proof)**: "Tells a story" and through the metaphor of the story proves the result.
  - Question asked: What are we counting?
- **Method 2 (Algebraic Proof)**: Manipulate formulas – Simplifications, substitutions, applying other equations.
  - Question asked: How can we make LHS equal to RHS?

Equation (3) – Sum of a row

$$\sum_{r=0}^{n} \binom{n}{r} = 2^n \qquad (3)$$

Let’s try both combinatorial proof and algebraic proof.

Equation (3) – Sum of a row

$$\sum_{r=0}^{n} \binom{n}{r} = 2^n \tag{3}$$

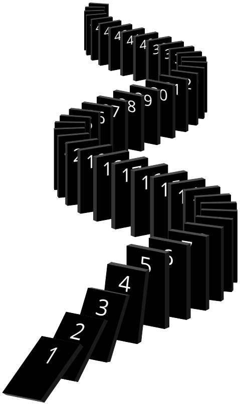
Combinatorial Proof:

What are we counting?

The number of subsets of $1 \ldots n$.

- LHS: Condition on the size of the subset. There are $\binom{n}{r}$ subsets of size r.
- RHS: By definition, it is 2^n.

Sidetrack: What is Induction?

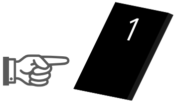


"Mathematical induction proves that we can knock down as many dominoes as we like, by proving that we can knock down the first domino (the basis) and each domino knocked down causes the next one to fall (the induction)."

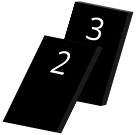
Sidetrack: What is Induction?

Goal: Prove that some statement about the integer n is true for all $n \geq n_0$.

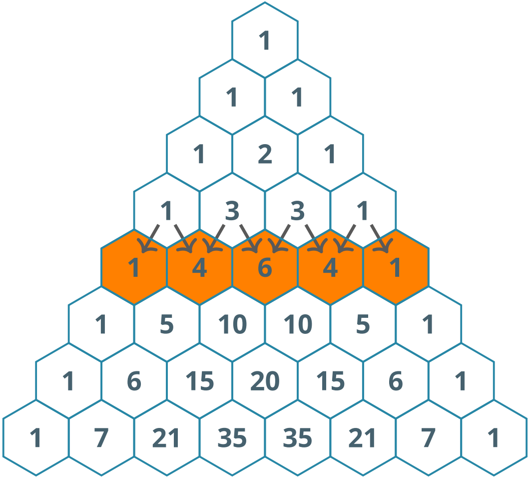
Basis step: Prove the statement when n has its smallest value n_0.



Induction step: Prove that if the statement is true for the value $n-1$, then it is true for the value n.



Equation (3) – Sum of a row



What is the relationship between $\sum_{r=0}^{k}\binom{k}{r}$ and $\sum_{r=0}^{k+1}\binom{k+1}{r}$?

Observation:

$$\sum_{r=0}^{k+1}\binom{k+1}{r} = \sum_{r=0}^{k+1}\left(\binom{k}{r}+\binom{k}{r-1}\right)$$

$$= 2\sum_{r=0}^{k}\binom{k}{r}$$

Equation (3) – Sum of a row

Algebraic Proof: Induct on n.

- Let $P(n)$ be the proposition $\sum_{r=0}^{n} \binom{n}{r} = 2^n$.
- **Basis step**: When $n = 0$, LHS $=$ RHS $= 1$. $\therefore P(0)$ is true.
- **Induction step**: Assume $P(k)$ is true for some $k \geq 0$. Then

$$\sum_{r=0}^{k+1} \binom{k+1}{r} = \sum_{r=0}^{k+1} \left(\binom{k}{r} + \binom{k}{r-1} \right) = 2 \sum_{r=0}^{k} \binom{k}{r} = 2 \times 2^k = 2^{k+1}$$

  $\therefore P(k+1)$ is true.
- By mathematical induction, $P(n)$ is true for all $n \geq 0$.

Equation (4) – Symmetry

$$\binom{n}{r} = \binom{n}{n-r} \tag{4}$$

Combinatorial Proof:

What are we counting?

The number of ways to choose r elements in $1 \ldots n$.

- LHS: By definition, it is $\binom{n}{r}$.
- RHS: Pick $(n-r)$ elements to **exclude**. There are $\binom{n}{n-r}$ ways to do so.

Equation (4) – Symmetry

$$\binom{n}{r} = \binom{n}{n-r} \tag{4}$$

Algebraic Proof: Use the explicit formula for $\binom{n}{r}$.

$$\begin{aligned}\binom{n}{r} &= \frac{n!}{r!(n-r)!} \\ &= \frac{n!}{(n-r)!(n-(n-r))!} \\ &= \binom{n}{n-r}\end{aligned}$$

Equation (5) – Weighted sum of a row

$$\sum_{r=0}^{n} r \times \binom{n}{r} = n \times 2^{n-1} \tag{5}$$

Combinatorial Proof:

What are we counting?

From n people, the number of ways to form a committee with a chairperson.
(Yes, you can use everyday language to reason about mathematics!)

- LHS: Form a committee first, then elect chairperson.
- RHS: Choose chairperson first, then pick the rest of the committee.

Equation (5) – Weighted sum of a row

Algebraic Proof: Let $P(n)$ be the proposition $\sum_{r=0}^{n} r \times \binom{n}{r} = n \times 2^{n-1}$.

Basis Step: When $n = 0$, LHS = RHS $= 0$. $\therefore P(0)$ is true.

Induction Step: Assume $P(k)$ is true for some $k \geq 0$. Then

$$\sum_{r=0}^{k+1} r \times \binom{k+1}{r} = \sum_{r=0}^{k+1} r \times \left(\binom{k}{r-1} + \binom{k}{r}\right)$$

$$= \sum_{r=0}^{k+1} (r-1) \times \binom{k}{r-1} + \sum_{r=0}^{k+1} \binom{k}{r-1} + \sum_{r=0}^{k+1} r \times \binom{k}{r}$$

$$= k \times 2^{k-1} + 2^k + k \times 2^{k-1} = k \times 2^k + 2^k = (k+1) \times 2^k$$

$\therefore P(k+1)$ is true.
By mathematical induction, $P(n)$ is true for all $n \geq 0$.

Equation (6) – Binomial Theorem

$$\sum_{r=0}^{n} \binom{n}{r} a^r b^{n-r} = (a+b)^n \tag{6}$$

- Think about the number of ways to form the term $a^r b^{n-r}$ from the expansion of $(a+b)(a+b)\ldots(a+b)$.
- $a = b = 1$:

$$\sum_{r=0}^{n} \binom{n}{r} = 2^n$$

- $a = -1, b = 1$:

$$\sum_{r=0}^{n} (-1)^r \binom{n}{r} = 0$$

Table of Contents

Fibonacci Numbers (A000045)

(0,)1, 1, 2, 3, 5, 8, 13, 21, ...

Recursive Definition:

- $F_0 = 0, F_1 = 1$
- $F_{n+2} = F_n + F_{n+1}$

Combinatorial Definition:

- F_n counts the number of ways to tile a $1 \times (n-1)$ board with squares and dominoes.
- $F_0 = 0$ because, well, a $1 \times (-1)$ board does not exist.

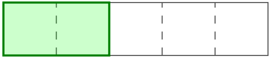
⚠ **Are these two definitions equivalent?**

Fibonacci Numbers (A000045)

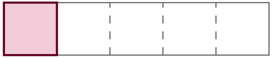
Aim: To show that the two definitions are equivalent.

$F_n =$ number of ways to tile a $1 \times (n-1)$ board with squares and dominoes.

- $F_1 = 1$ since a 1×0 board can only be tiled with nothing.
- F_{n+2}: We tile a $1 \times (n+1)$ board with squares and dominoes.
  - Start with a domino: We're left with a $1 \times (n-1)$ board – F_n ways.



  - Start with a square: We're left with a $1 \times n$ board – F_{n+1} ways.



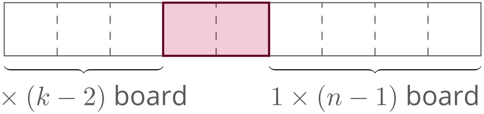
$\therefore F_0 = 0$, $F_1 = 1$ and $F_{n+2} = F_n + F_{n+1}$.

Fibonacci Numbers – Equation (7)

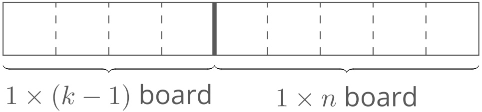
$$F_{n+k} = F_k F_{n+1} + F_{k-1} F_n \quad (k \geq 1) \tag{7}$$

The number of square-domino tilings of a $1 \times (n+k-1)$ board.

- LHS: By definition, it is F_{n+k}.
- RHS: Condition on whether a domino covers cells $(k-1)$ and k.
  - If yes, there are $F_{k-1} \times F_n$ ways to tile the rest.



  - If no, there are $F_k \times F_{n+1}$ ways to tile the rest.



Fibonacci Numbers – Equation (7)

$$F_{n+k} = F_k F_{n+1} + F_{k-1} F_n \quad (k \geq 1) \tag{7}$$

Putting $k = n$ and $k = n + 1$, respectively, we get

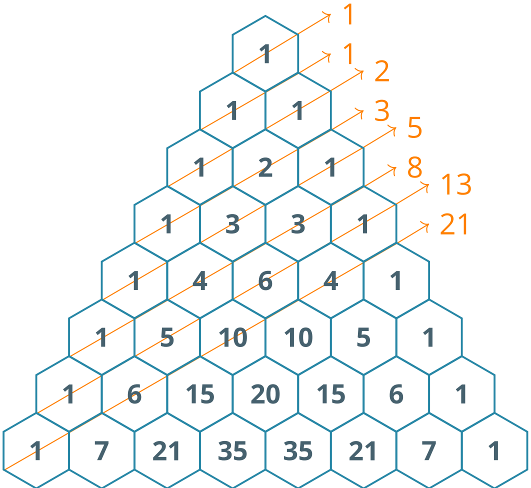
$$\begin{aligned} F_{2n} &= F_n F_{n+1} + F_{n-1} F_n \\ F_{2n} &= F_n F_{n+1} + (F_{n+1} - F_n) F_n \\ \mathbf{F_{2n}} &= \mathbf{2F_n F_{n+1} - F_n F_n} \end{aligned}$$

and

$$\mathbf{F_{2n+1}} = \mathbf{F_{n+1} F_{n+1} + F_n F_n},$$

which give an $O(\log n)$ algorithm to compute F_n mod M.

Relating Fibonacci Numbers and $\binom{n}{r}$



$$F_{n+1} = \sum_{r=0}^{\lfloor n/2 \rfloor} \binom{n-r}{r}$$

Exercise: Let's prove this using a combinatorial proof and an algebraic proof.

Relating Fibonacci Numbers and $\binom{n}{r}$

$$F_{n+1} = \sum_{r=0}^{\lfloor n/2 \rfloor} \binom{n-r}{r} \qquad (8)$$

Combinatorial Proof:

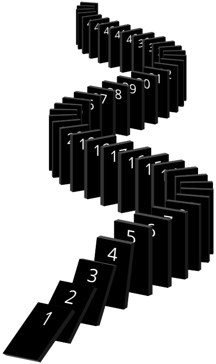
What are we counting?

The number of square-domino tilings of a $1 \times n$ board.

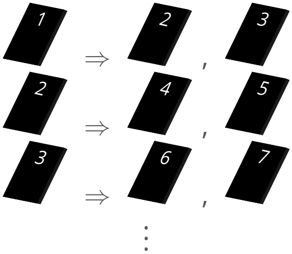
- LHS: By definition, it is F_{n+1}.
- RHS: Condition on the number of dominoes. For tilings with k dominoes, a total of $n-k$ pieces will be used, and so there are $\binom{n-k}{k}$ ways to tile the board.

Sidetrack: Weak vs Strong Mathematical Induction

Weak MI: Knock the dominoes one by one. $1 \to 2 \to 3 \to \cdots$.



Strong MI: Knock the dominoes in **any creative order**, e.g. $P(n) \to P(2n)$ & $P(2n+1)$.



Relating Fibonacci Numbers and $\binom{n}{r}$

Algebraic Proof: Let $P(n)$ be the proposition $F_{n+1} = \sum_{r=0}^{\lfloor n/2 \rfloor} \binom{n-r}{r}$.

Basis step: When $n = 0$ and $n = 1$, LHS = RHS $= 1$. $\therefore P(0)$ and $P(1)$ are true.

Induction step: Assume $P(k)$ and $P(k+1)$ are true for some $k \geq 0$. Then

$$\begin{aligned} F_{k+3} &= F_{k+1} + F_{k+2} \\ &= \sum_{r=0}^{\lfloor k/2 \rfloor} \binom{k-r}{r} + \sum_{r=0}^{\lfloor (k+1)/2 \rfloor} \binom{k+1-r}{r} \\ &= \sum_{r=0}^{\lfloor (k+2)/2 \rfloor} \binom{k+1-r}{r-1} + \sum_{r=0}^{\lfloor (k+2)/2 \rfloor} \binom{k+1-r}{r} = \sum_{r=0}^{\lfloor (k+2)/2 \rfloor} \binom{k+2-r}{r} \end{aligned}$$

$\therefore P(k+2)$ is true.

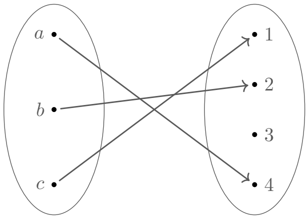
By **strong** mathematical induction, $P(n)$ is true for all $n \geq 0$.

A Brief Summary: Proof Techniques

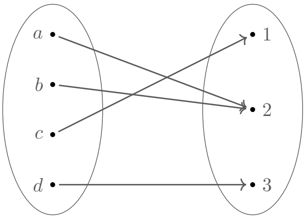
- **Combinatorial Proof**: "Tells a story" and through the metaphor of the story proves the result.
  - Question asked: What are we counting?
- **Algebraic Proof**: Manipulate formulas – Simplifications, substitutions, applying other equations.
  - Question asked: How can we make LHS equal to RHS?
- **Induction**: A tool for us to prove a statement for all $n \geq n_0$. It can be used in **BOTH** combinatorial and algebraic proofs.
  - Weak Mathematical Induction: Prove that $P(n_0)$ is true and $P(k) \to P(k+1)$ for all $k \geq n_0$.
  - Strong Mathematical Induction: Prove in any creative order!

Injections and Surjections

Given two sets X, Y and a function $f : X \to Y$.



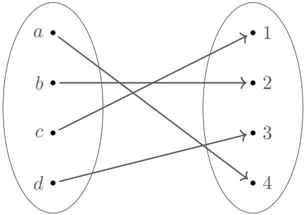
Injection (One-to-one)
Each element in Y is mapped to **at most** one element from X



Surjection (Onto)
Each element in Y is mapped to **at least** one element from X

Bijections

What if a function $f : X \to Y$ is **both** an injection and a surjection?



Bijection

Each element in Y is mapped to **exactly** one element from X

Important property: If we can find a bijection between two finite sets X and Y, then $|X| = |Y|$ (i.e. X and Y has the same cardinality).

Bijective Proofs

Bijective Proof can be used to show that two sets have the same cardinality.

1. Define a (bijective) function between the sets X and Y.
2. Prove that your function is a bijection.
   ℹ A function is a bijection if and only if an **inverse function** exists.
   Therefore, we can prove that the function is a bijection by providing the inverse function.
3. Conclude that $|X| = |Y|$.

Bijective Proof – An example

Example: Prove that the number of subsets of $\{1, 2, \ldots, n\}$ without two consecutive integers is F_{n+2}.

Consider the following sets:

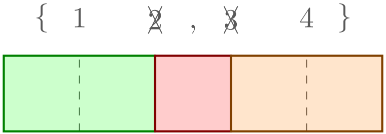
- X: The number of subsets of $\{1, 2, \ldots, n\}$ without two consecutive integers.
- Y: The number of ways to tile a $1 \times (n+1)$ board with squares and dominoes.

Since we want to prove that $|X| = |Y|$, it's a good idea to try establishing a bijection between the two sets.

Bijective Proof – An example

Given a way to tile a $1 \times (n+1)$ board with squares and dominoes:

- Consider the n gaps between the cells.
- If there is a gap between cell i and cell $i+1$, exclude i from the set.
- Otherwise, include i in the set.
- This forms a subset of $\{1, 2, \ldots, n\}$ without two consecutive integers.

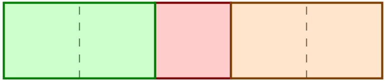


Bijective Proof – An example

Inverse function: Given a subset of $\{1, 2, \ldots, n\}$ without two consecutive integers.

- For each $1 \leq i \leq n$, if i is in the set, tile cells i and $i + 1$ with a domino.
- For all the other cells, tile them with squares.
- This forms a unique square-domino tiling for a $1 \times (n + 1)$ board.

$$\{\ 1 \quad , \quad 4\ \}$$



Bijective Proof – Another example

Example: The number of ways to tile a $1 \times n$ rectangle with rectangles/squares with odd area is F_n.

Consider the following sets:

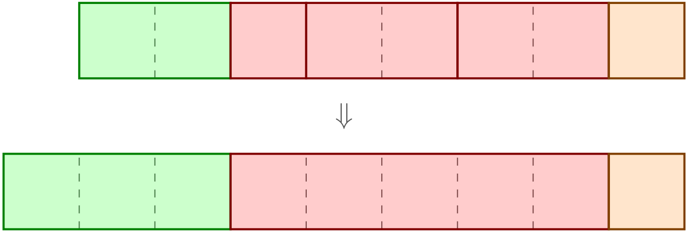
- X: The number of ways to tile a $1 \times n$ rectangle with rectangles/squares with odd area.
- Y: The number of ways to tile a $1 \times (n-1)$ board with squares and dominoes.

Since we want to prove that $|X| = |Y|$, it's a good idea to try establishing a bijection between the two sets.

Bijective Proof – Another example

Given a way to tile a $1 \times (n-1)$ board with squares and dominoes:

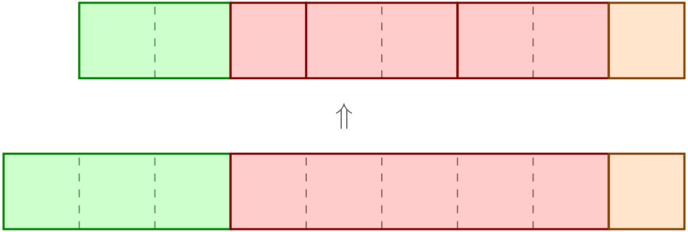
- Add a square in front of the board.
- Merge all dominoes to the square in front of it.
- All merged segment has odd area.
- This uniquely tiles a $1 \times n$ board with rectangle/squares with odd area.



Bijective Proof – Another example

Inverse function: Given a way to tile a $1 \times n$ rectangle with rectangles/squares with odd area.

- For each rectangle/square with odd area, split it into one square and multiple dominoes.
- Remove the first square.
- This forms a unique square-domino tiling for a $1 \times (n-1)$ board.



Catalan Numbers (A000108)

(1,)1, 2, 5, 14, 42, ...

What does this sequence C_n describe?

Many Things!

- C_n is the number of valid bracket sequences of length $2n$.
- C_n is the number of triangulations of a convex $(n+2)$-gon.
- C_n is the number of "ordered" rooted binary trees with n nodes.
- The list goes on...

Let's Take A Look At Small n's

Definition

Define C_n to be the number of valid bracket sequences of length $2n$.

Valid means the (and) brackets form matching pairs.

- $n = 0$: (empty sequence)
- $n = 1$: ()
- $n = 2$: ()(), (())
- $n = 3$: ()()(), ()(()), (())(), (()()), ((()))
- $n = 4$: ()()()(), ()()(()), ()(())(), ()(()()), ()((())), (())()(), (())(()), (()())(), ((()))(), (()()()), (()(())), ((())()), ((()())), (((())))

Recurrence Formula

$$C_n = \sum_{k=1}^{n} C_{k-1} \times C_{n-k} \tag{9}$$

What are we counting?

The number of valid bracket sequences of length $2n$.

- LHS: By definition, it is C_n.
- RHS: Condition on the position of) which closes the first bracket (. For $1 \le k \le n$, there are $C_{k-1} \times C_{n-k}$ valid bracket sequences with) positioned at $2k$ (note that the position must be even).

This gives an $O(n^2)$ algorithm to compute $C_n \bmod M$.

Explicit Formula For C_n

$$C_n = \frac{1}{n+1}\binom{2n}{n} \tag{10}$$

Reformulate the Catalan model as follows:

- You start from $(0, 0)$ and want to reach $(2n, 0)$.
- From (x, y) you can move “up” to $(x + 1, y + 1)$ or “down” to $(x + 1, y - 1)$.
- C_n is the number of “good” paths – paths that do not go below the x-axis.

Ideas:

- We count, instead, the number of bad paths.
- Then $C_n = \binom{2n}{n} -$ (number of bad paths).
- Therefore, for the formula to hold, we expect there to be $\binom{2n}{n+1}$ bad paths.

Explicit Formula For C_n

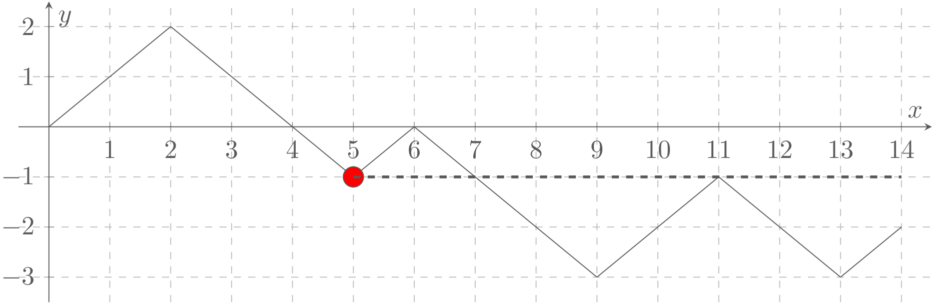
Establish a bijection between:

- Bad paths from $(0, 0)$ to $(2n, 0)$, and
- Paths from $(0, 0)$ to $(2n, -2)$

Explicit Formula For C_n

Given a bad path from $(0, 0)$ to $(2n, 0)$:

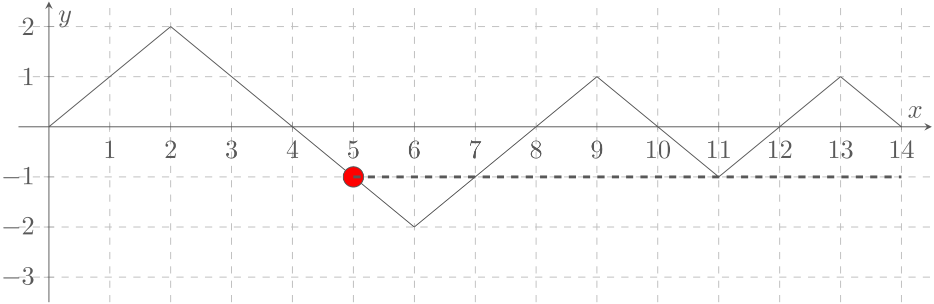
- We take the portion of the path **after** the first visit below x-axis.
- Reflect it about $y = -1 \Rightarrow$ a path from $(0, 0)$ to $(2n, -2)$.



Explicit Formula For C_n

Inverse function: Given a path from $(0, 0)$ to $(2n, -2)$:

- We take the portion of the path **after** the first visit below x-axis.
- Reflect it about $y = -1 \Rightarrow$ a bad path from $(0, 0)$ to $(2n, 0)$.



Reflection Trick – Exercise

Exericse: Codeforces 26D, La Salle-Pui Ching Programming Challenge 2017 Problem L (Let Me Count The Ways)

Further Reading: Young Tableaux

Table of Contents

Introduction to Probability Space

A (finite) probability space consists of two parts:

1. A set $\Omega = \{a_1, \ldots, a_n\}$ of outcomes
2. Associated probabilities $p_1, \ldots, p_n$ of each outcome ($\sum_i p_i = 1$)

Probabilities can be abstract (based on idealised model) or empirical (based on statistical evidence).
Usually we consider uniform probability spaces, i.e. $p_1 = \cdots = p_n = \frac{1}{n}$.

Examples:

- Dice roll
- Coin toss
- A sequence of N coin tosses
- Drawing a ball from a bag of colored balls

Event Probability

An event is (a short description of) a subset of Ω.
For example:

- $\Omega_1 :=$ sequences of 3 coin tosses, $A_1 :=$ seqs with two consecutive heads.
  Then, $A_1 = \{HHH, HHT, THH\}$.
  $P(A_1) = \frac{3}{8}$.
- $\Omega_2 :=$ dice roll, $A_2 :=$ odd outcomes.
  Then, $A_2 = \{1, 3, 5\}$.
  $P(A_2) = \frac{3}{6} = \frac{1}{2}$.

For uniform probability spaces,

$$P(A) = \frac{\text{number of outcomes in } A}{\text{number of possible outcomes}},$$

so we just need to count :)

Conditional Probability

For two events A and B with $P(B) > 0$, **define** conditional probability as

$$P(A|B) := \frac{P(A \cap B)}{P(B)}.$$

For DP on probabilities, we often use the fact that

$$\underbrace{P(A \cap B)}_{\text{new state}} = \underbrace{P(A|B)}_{\text{transition}} \cdot \underbrace{P(B)}_{\text{previous state}}$$

Exercise: Suppose we pick an integer x in $[1, 20]$ uniformly at random.
Calculate $P(x$ is odd | x is prime) and $P(x$ is prime | x is odd).

Random Variable, Expectation, Variance

A random variable is a function $f : \Omega \longrightarrow \mathbb{R}$.
It assigns a value of each of $a_1, \ldots, a_n$.

Examples: slot machine, dice roll

More Definitions

Expectation of f : $\mathbb{E}[f] := \sum_{i=1}^{n} f(a_i)p_i.$

Variance of f : $Var(f) := \mathbb{E}[(f - \mathbb{E}[f])^2].$

$\mathbb{E}[f]$ is the average value of f, weighted by probability (think lottery).
$Var(f)$ measures how f deviates from its average.

A Remark

Event probability is just the expected value of certain **indicator functions**.
Just choose $f(x) := 1$ if $x \in A$ and $f(x) := 0$ otherwise.
Then $\mathbb{E}[f] = P(A)$.
We denote this function by 1_A.

Example 1

There is a bag with two red balls and a black ball. Pick a ball uniformly at random from the bag. If it is red, you get 3 coins. If it is black, you get 9 coins. What is the expected number of coins you will get?

Definition of Expectation

Expectation of f : $$\mathbb{E}[f] := \sum_{i=1}^{n} f(a_i)p_i.$$

Example 1

There is a bag with two red balls and a black ball. Pick a ball uniformly at random from the bag. If it is red, you get 3 coins. If it is black, you get 9 coins. What is the expected number of coins you will get?

$$\begin{aligned}\mathbb{E}[X] &= \tfrac{2}{3} \times 3 + \tfrac{1}{3} \times 9 \\ &= 2 + 3 \\ &= 5.\end{aligned}$$

However, there is another way of calculating it.

Linearity of Expectation

Theorem (Linearity of Expectation)

Suppose random variable f is written as sum $f_1 + \cdots + f_k$. Then,

$$\mathbb{E}[f] = \sum_{i=1}^{k} \mathbb{E}[f_i].$$

(Think about: do we have "linearity of variance"?) NO.

This simple theorem is very powerful!

Core Idea: Break down a quantity into small, easy-to-calculate parts.

Example 1 (With Linearity of expectation)

There is a bag with two red balls and a black ball. Pick a ball uniformly at random from the bag. If it is red, you get 3 coins. If it is black, you get 9 coins. What is the expected number of coins you will get?

Theorem (Linearity of Expectation)

Suppose random variable f is written as sum $f_1 + \cdots + f_k$. Then,

$$\mathbb{E}[f] = \sum_{i=1}^{k} \mathbb{E}[f_i].$$

Example 1 (With Linearity of expectation)

There is a bag with two red balls and a black ball. Pick a ball uniformly at random from the bag. If it is red, you get 3 coins. If it is black, you get 9 coins. What is the expected number of coins you will get?

Let X_r and X_b be the coins you get from picking a red ball and a black ball respectively.

$$\begin{aligned} \mathbb{E}[X] &= \mathbb{E}[X_r] + \mathbb{E}[X_b] \\ &= \frac{2}{3} \times 3 + \frac{1}{3} \times 9 \\ &= 2 + 3 \\ &= 5. \end{aligned}$$

Looks the same in this case, but will help us massively in more complicated examples.

Example 2

Suppose we pick a permutation $p = [p_1, \ldots, p_n]$ of $1..n$ uniformly at random. What is the expected number of fixed points of p?

Let X be the number of fixed points of p. Then $X = X_1 + \cdots + X_n$ where X_i is indicator function for $p_i = i$.
By linearity of expectation,

$$\begin{aligned} \mathbb{E}[X] &= \sum_{i=1}^{n} \mathbb{E}[X_i] \\ &= \sum_{i=1}^{n} P(p_i = i) \\ &= \sum_{i=1}^{n} \frac{1}{n} \\ &= 1. \end{aligned}$$

Example 3 (HKOJ S232)

Dr Jones partitions N students into K teams randomly. M pairs of students are enemies. What is the expected number of pairs of enemies that are placed in the same team?

Let X be the number of pairs of enemies that are placed in the same team. Then $X = X_1 + \cdots + X_M$ where X_i is the indicator function for "the i-th pair of enemy is placed in the same team".
By linearity of expectation,

$$\begin{aligned} \mathbb{E}[X] &= \sum_{i=1}^{M} \mathbb{E}[X_i] \\ &= \sum_{i=1}^{M} P(i\text{-th pair of enemy is placed in the same team}) \\ &= \sum_{i=1}^{M} \frac{1}{K} \\ &= \frac{M}{K} \end{aligned}$$

Example 4

Generate a sequence of length n, where each element is chosen independently and uniformly from $1..k$.
What is the expected number of distinct elements?

Let X be the number of distinct elements. Then $X = X_1 + \cdots + X_k$ where X_i is indicator function for "number i appears in the sequence".
By linearity of expectation,

$$\begin{aligned}
\mathbb{E}[X] &= \textstyle\sum_{i=1}^{k} \mathbb{E}[X_i] \\
&= \textstyle\sum_{i=1}^{k} P(\text{number } i \text{ appears in the sequence}) \\
&= \textstyle\sum_{i=1}^{k} (1 - P(\text{number } i \text{ does not appear in the sequence})) \\
&= \textstyle\sum_{i=1}^{k} (1 - (\frac{k-1}{k})^n) \\
&= k(1 - (\frac{k-1}{k})^n)
\end{aligned}$$

Example 2 (Modified)

Consider all permutations $p = [p_1, \ldots, p_n]$ of $1..n$. What is the **total number** of fixed points?

Expectation and Counting

As $\mathbb{E}[F]$ is the "Average" of all possibilities, by linearity of expectation,

$$\begin{aligned}
\sum_{i=1}^{n} f(a_i) &= n \times \mathbb{E}[f] \\
&= n \times \sum_{i=1}^{k} \mathbb{E}[f_i] \\
&= \sum_{i=1}^{k} n \times \mathbb{E}[f_i] \\
&= \sum_{i=1}^{k} n \times \sum_{j=1}^{n} \frac{f_i(a_j)}{n} \\
&= \sum_{i=1}^{k} \sum_{j=1}^{n} f_i(a_j)
\end{aligned}$$

This is known as the **contribution** technique where we count the "contribution" for every function f_i.

Expectation and Counting (Visualization)

p[1]	p[2]	p[3]	p[4]	fixed pts
1	2	3	4	4
1	2	4	3	2
1	3	2	4	2
1	3	4	2	1
1	4	2	3	1
1	4	3	2	2
2	1	3	4	2
2	1	4	3	0
2	3	1	4	1
2	3	4	1	0
2	4	1	3	0
2	4	3	1	1

	p[1]	p[2]	p[3]	p[4]	fixed pts
	3	1	2	4	1
	3	1	4	2	0
	3	2	1	4	2
	3	2	4	1	1
	3	4	1	2	0
	3	4	2	1	0
	4	1	2	3	0
	4	1	3	2	1
	4	2	1	3	1
	4	2	3	1	2
	4	3	1	2	0
	4	3	2	1	0
SUM	6	6	6	6	24

Summing “vertically” down the four columns is much easier!

Example 5

Given an integer set $S = \{b_1, b_2, \ldots, b_n\}$. For each subset T of S, let $sum(T)$ denote the sum of elements of T. Calculate

$$\sum_{T \subseteq S} sum(T) \mod M.$$

Let

$$contain_i(T) := \begin{cases} 1, & \text{if } T \text{ contains } b_i; \\ 0, & \text{otherwise.} \end{cases}$$

Then $sum(T) = \sum_{i=1}^{n} b_i \times contain_i(T)$, and

$$\begin{aligned} \sum_{T \subseteq S} sum(T) &= \sum_{T \subseteq S} \sum_{i=1}^{n} b_i \times contain_i(T) \\ &= \sum_{i=1}^{n} b_i \left(\sum_{T \subseteq S} contain_i(T) \right) \\ &= \sum_{i=1}^{n} b_i 2^{n-1} \end{aligned}$$

Example 6

Given an array $a[1..n]$. For $1 \le l \le r \le n$, let $sum(l,r)$ denote $a[l] + a[l+1] + \cdots + a[r]$. Calculate

$$\sum_{1 \le l \le r \le n} sum(l,r).$$

Let

$$contain_p(l,r) := \begin{cases} 1, & \text{if } l \le p \le r; \\ 0, & \text{otherwise.} \end{cases}$$

Then $sum(l,r) = \sum_{p=1}^{n} a[p] \times contain_p(l,r)$, and

$$\begin{aligned} \sum_{1 \le l \le r \le n} sum(l,r) &= \sum_{1 \le l \le r \le n} \sum_{p=1}^{n} a[p] \times contain_p(l,r) \\ &= \sum_{p=1}^{n} a[p] \times p \times (n+1-p) \end{aligned}$$

Example 7

Given an integer set $S = \{s_1, s_2, \ldots, s_n\}$. For each nonempty subset T of S, let $diff(T)$ denote the difference between the largest and the smallest elements of T. Calculate

$$\sum_{\varnothing \neq T \subseteq S} diff(T) \mod M.$$

Sort the numbers. For $1 \leq p < n$, let

$$contain_p(T) := \begin{cases} 1, & \text{if } min\_index(T) \leq p < max\_index(T) \\ 0, & \text{otherwise.} \end{cases}$$

Then $diff(T) = \sum_{p=1}^{n-1} (s_{p+1} - s_p) \times contain_p(T)$, and

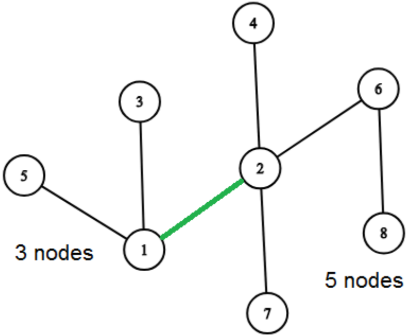
$$\begin{aligned} \sum_{\varnothing \neq T \subseteq S} diff(T) &= \sum_{\varnothing \neq T \subseteq S} \sum_{p=1}^{n-1} (s_{p+1} - s_p) \times contain_p(T) \\ &= \sum_{p=1}^{n-1} (s_{p+1} - s_p) \times \sum_{\varnothing \neq T \subseteq S} contain_p(T) \\ &= \sum_{p=1}^{n-1} (s_{p+1} - s_p) \times (2^p - 1) \times (2^{n-p} - 1) \end{aligned}$$

Example 8

Given a tree with N nodes. Let $l(x, y)$ be the length (number of edges) of the simple path from x to y. Calculate

$$\sum_{x,y} l(x, y).$$

For each edge, count the number of paths that pass through it.
It is (number of nodes on one side) $\times$ (number of nodes on the other side).
This can be done using a DFS on tree.



Further Reading

Proofs That Really Count (The Art of Combinatorial Proof)
by Arthur T. Benjamin and Jennifer J. Quinn

Sums and Expected Value – Part 1 (https://codeforces.com/blog/entry/62690)
Sums and Expected Value – Part 2 (https://codeforces.com/blog/entry/62792)
by Errichto

Concrete Mathematics: A Foundation for Computer Science
by Ronald L. Graham, Donald E. Knuth and Oren Patashnik

Table of Contents