Combinatorics

Quinten Jin, Emre Kocaman

Massachusetts Institute of Technology

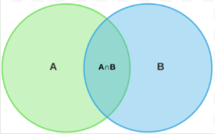
MIT PRIMES Circle 2024

Example

40 people speak French and 30 people speak German in a classroom. 20 people speak both. Everyone speaks 1 language at least. How many people are there?
Answer: 50
When we add 40 and 30, we get 70. However, we count the students who speak both two times. Thus, we have to subtract the number of students who speak both. When we subtract 20 from 70, we get 50. We can state our problem as $|A| + |B| - |A \cap B| = |A \cup B|$ where $|A|$ represents French speakers and $|B|$ represents German speakers. It can be shown in Venn Diagram as follows:

Example

40 people speak French and 30 people speak German in a classroom. 20 people speak both. Everyone speaks 1 language at least. How many people are there?
Answer: 50
When we add 40 and 30, we get 70. However, we count the students who speak both two times. Thus, we have to subtract the number of students who speak both. When we subtract 20 from 70, we get 50. We can state our problem as $|A| + |B| - |A \cap B| = |A \cup B|$ where $|A|$ represents French speakers and $|B|$ represents German speakers. It can be shown in Venn Diagram as follows:



Theorem (Principle of Inclusion and Exclusion)

PIE stands for the Principle of Inclusion and Exclusion and is used to avoid over counting and over subtracting. The generalized formula for PIE is:

$$\left|\bigcup_{i=1}^{n} S_k\right| = \sum_{I\subseteq[n]} (-1)^{(|I|+1)} \left|\bigcap_{i\in I} S_i\right|$$

Theorem (Principle of Inclusion and Exclusion)

PIE stands for the Principle of Inclusion and Exclusion and is used to avoid over counting and over subtracting. The generalized formula for PIE is:

$$\left|\bigcup_{i=1}^{n} S_k\right| = \sum_{I\subseteq[n]} (-1)^{(|I|+1)} \left|\bigcap_{i\in I} S_i\right|$$

The language problem we just gave is an example for PIE. In the formula, the left side is $| S_1 | \cup | S_2 | \cup | S_3 | \cdots | S_k |$. The right side indicates that if the number of intersecting sets is odd, we add the intersection. If the number of intersecting sets is even, we subtract the intersection. For example, the right side says that $| S_1 \cap S_2 |$ is subtracted and $| S_1 \cap S_2 \cap S_3 |$ is added.

Now, we are going to use PIE in a hard problem which seems irrelevant to PIE.

Example

$N = 3 \times 5 \times 7 \times 11$ How many numbers which are smaller than N have greatest common divisor as 1 with N? Answer: $2 \times 4 \times 6 \times 10$

Now, we are going to use PIE in a hard problem which seems irrelevant to PIE.

Example

$N = 3 \times 5 \times 7 \times 11$ How many numbers which are smaller than N have greatest common divisor as 1 with N? Answer: $2 \times 4 \times 6 \times 10$

Solution

1.$| S_1 |$ The numbers which are divisible by 3.

Now, we are going to use PIE in a hard problem which seems irrelevant to PIE.

Example

$N = 3 \times 5 \times 7 \times 11$ How many numbers which are smaller than N have greatest common divisor as 1 with N? Answer: $2 \times 4 \times 6 \times 10$

Solution

1.$| S_1 |$ The numbers which are divisible by 3.
2.$| S_2 |$ The numbers which are divisible by 5.

Now, we are going to use PIE in a hard problem which seems irrelevant to PIE.

Example

$N = 3 \times 5 \times 7 \times 11$ How many numbers which are smaller than N have greatest common divisor as 1 with N? Answer: $2 \times 4 \times 6 \times 10$

Solution

1.$| S_1 |$ The numbers which are divisible by 3.
2.$| S_2 |$ The numbers which are divisible by 5.
3.$| S_3 |$ The numbers which are divisible by 7.

Now, we are going to use PIE in a hard problem which seems irrelevant to PIE.

Example

$N = 3 \times 5 \times 7 \times 11$ How many numbers which are smaller than N have greatest common divisor as 1 with N? Answer: $2 \times 4 \times 6 \times 10$

Solution

1.$| S_1 |$ The numbers which are divisible by 3.
2.$| S_2 |$ The numbers which are divisible by 5.
3.$| S_3 |$ The numbers which are divisible by 7.
4.$| S_4 |$ The numbers which are divisible by 11.

Now, we are going to use PIE in a hard problem which seems irrelevant to PIE.

Example

$N = 3 \times 5 \times 7 \times 11$ How many numbers which are smaller than N have greatest common divisor as 1 with N? Answer: $2 \times 4 \times 6 \times 10$

Solution

1.$\mid S_1 \mid$ The numbers which are divisible by 3.
2.$\mid S_2 \mid$ The numbers which are divisible by 5.
3.$\mid S_3 \mid$ The numbers which are divisible by 7.
4.$\mid S_4 \mid$ The numbers which are divisible by 11.
5.The number of alternatives we don't want is $\mid S_1 \mid + \mid S_2 \mid + \mid S_3 \mid + \mid S_4 \mid$ $-(\mid S_1 \cap S_2 \mid + \mid S_1 \cap S_3 \mid \cdots) + (\mid S_1 \cap S_2 \cap S_3 \cdots) - (\mid S_1 \cap S_2 \cap S_3 \cap S_4)$.

Now, we are going to use PIE in a hard problem which seems irrelevant to PIE.

Example

$N = 3 \times 5 \times 7 \times 11$ How many numbers which are smaller than N have greatest common divisor as 1 with N? Answer: $2 \times 4 \times 6 \times 10$

Solution

1.$\mid S_1 \mid$ The numbers which are divisible by 3.
2.$\mid S_2 \mid$ The numbers which are divisible by 5.
3.$\mid S_3 \mid$ The numbers which are divisible by 7.
4.$\mid S_4 \mid$ The numbers which are divisible by 11.
5.The number of alternatives we don't want is $\mid S_1 \mid + \mid S_2 \mid + \mid S_3 \mid + \mid S_4 \mid$ $-(\mid S_1 \cap S_2 \mid + \mid S_1 \cap S_3 \mid \cdots) + (\mid S_1 \cap S_2 \cap S_3 \cdots) - (\mid S_1 \cap S_2 \cap S_3 \cap S_4)$.
6.When we subtract them from all of the alternatives, we get our answer.

Now, we are going to use PIE in a hard problem which seems irrelevant to PIE.

Example

$N = 3 \times 5 \times 7 \times 11$ How many numbers which are smaller than N have greatest common divisor as 1 with N? Answer: $2 \times 4 \times 6 \times 10$

Solution

1.$\mid S_1 \mid$ The numbers which are divisible by 3.
2.$\mid S_2 \mid$ The numbers which are divisible by 5.
3.$\mid S_3 \mid$ The numbers which are divisible by 7.
4.$\mid S_4 \mid$ The numbers which are divisible by 11.
5.The number of alternatives we don't want is $\mid S_1 \mid + \mid S_2 \mid + \mid S_3 \mid + \mid S_4 \mid$ $-(\mid S_1 \cap S_2 \mid + \mid S_1 \cap S_3 \mid \cdots) + (\mid S_1 \cap S_2 \cap S_3 \cdots) - (\mid S_1 \cap S_2 \cap S_3 \cap S_4)$.
6.When we subtract them from all of the alternatives, we get our answer.
7.We could write it as
$N(1 - (\frac{1}{3} + \frac{1}{5} + \frac{1}{7} + \frac{1}{11}) + (\frac{1}{3\times5} + \frac{1}{3\times7} + \cdots) - (\frac{1}{3\times5\times7} + \cdots) + (\frac{1}{3\times5\times7\times11})) =$
$N(1 - \frac{1}{3})(1 - \frac{1}{5})(1 - \frac{1}{7})(1 - \frac{1}{11}))$.

Now, we are going to use PIE in a hard problem which seems irrelevant to PIE.

Example

$N = 3 \times 5 \times 7 \times 11$ How many numbers which are smaller than N have greatest common divisor as 1 with N? Answer: $2 \times 4 \times 6 \times 10$

Solution

1.$\mid S_1 \mid$ The numbers which are divisible by 3.
2.$\mid S_2 \mid$ The numbers which are divisible by 5.
3.$\mid S_3 \mid$ The numbers which are divisible by 7.
4.$\mid S_4 \mid$ The numbers which are divisible by 11.
5.The number of alternatives we don't want is $\mid S_1 \mid + \mid S_2 \mid + \mid S_3 \mid + \mid S_4 \mid$ $-(\mid S_1 \cap S_2 \mid + \mid S_1 \cap S_3 \mid \cdots) + (\mid S_1 \cap S_2 \cap S_3 \cdots) - (\mid S_1 \cap S_2 \cap S_3 \cap S_4)$.
6.When we subtract them from all of the alternatives, we get our answer.
7.We could write it as
$N(1 - (\frac{1}{3} + \frac{1}{5} + \frac{1}{7} + \frac{1}{11}) + (\frac{1}{3\times5} + \frac{1}{3\times7} + \cdots) - (\frac{1}{3\times5\times7} + \cdots) + (\frac{1}{3\times5\times7\times11})) =$
$N(1 - \frac{1}{3})(1 - \frac{1}{5})(1 - \frac{1}{7})(1 - \frac{1}{11}))$.
8. We get $2 \times 4 \times 6 \times 10$.

Theorem (Catalan Numbers)

Let's say there are n (+1)'s and n (-1)'s. When these are arranged and added from left to right, the sum is always zero or bigger. That means that the number of (-1)'s is never larger than the number of (+1)'s. C_n *is the number of arrangements which satisfy that. The closed formula for* C_n *is* $C_n = \frac{1}{n+1}\binom{2n}{n}$.

Theorem (Catalan Numbers)

Let's say there are n (+1)'s and n (-1)'s. When these are arranged and added from left to right, the sum is always zero or bigger. That means that the number of (-1)'s is never larger than the number of (+1)'s. C_n is the number of arrangements which satisfy that. The closed formula for C_n is $C_n = \frac{1}{n+1}\binom{2n}{n}$.

For example, let's say that there are 2 (-1)'s and 2 (+1)'s.
1.(+1),(+1),(-1),(-1) is okay.
2.(+1),(-1),(+1),(-1) is okay.
3.(-1),(-1),(+1),(+1) is NOT okay.
4.(+1),(-1),(-1), (+1) is NOT okay.

Theorem (Catalan Numbers)

Let's say there are n (+1)'s and n (-1)'s. When these are arranged and added from left to right, the sum is always zero or bigger. That means that the number of (-1)'s is never larger than the number of (+1)'s. C_n is the number of arrangements which satisfy that. The closed formula for C_n is $C_n = \frac{1}{n+1}\binom{2n}{n}$.

For example, let's say that there are 2 (-1)'s and 2 (+1)'s.
1.(+1),(+1),(-1),(-1) is okay.
2.(+1),(-1),(+1),(-1) is okay.
3.(-1),(-1),(+1),(+1) is NOT okay.
4.(+1),(-1),(-1), (+1) is NOT okay.
Catalan numbers are the numbers of arrangements which are OKAY.

Example

There are 3 open parenthesis and 3 closed parenthesis. In how many ways can these parenthesis be arranged that no parenthesis are left open? Answer: 5

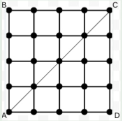
Example

There are 3 open parenthesis and 3 closed parenthesis. In how many ways can these parenthesis be arranged that no parenthesis are left open? Answer: 5

Here, open parenthesis is same with $+1$ and closed parenthesis is same with -1.
For example: $(+1)$,(-1),$(+1)$,(-1),$(+1)$,(-1) is same with ( ) ( ) ( ).
Thus, all we have to do is placing n with 3 in the closed formula. $\frac{1}{4}\binom{6}{3} = 5$.

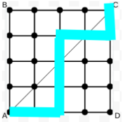
Example

In a 4×4 map, a car has to go from A to C. It can touch the diagonal but cannot pass it. It can only move to right and up. How many ways are there?



Answer: 336

In order not to pass the diagonal, the number of vertical moves shouldn't pass the number of horizontal moves. There should be 4 vertical and 4 horizontal moves. The vertical moves should be same as -1 and horizontal moves should be $+1$. For example, $(+1)$,$(+1)$,(-1),(-1),(-1),$(+1)$,$(+1)$,(-1) is equivalent to:



The Domino Tiling Problem

Determine how many ways to cover 2×100 board using 1×2 dominos.

The Domino Tiling Problem

Determine how many ways to cover 2×100 board using 1×2 dominos.

Solution

We define $F(n)$ to be the number of ways to cover a $2 \times n$ board. To solve the tiling problem, we can consider two primary cases:

The Domino Tiling Problem

Determine how many ways to cover 2×100 board using 1×2 dominos.

Solution

We define $F(n)$ to be the number of ways to cover a $2 \times n$ board. To solve the tiling problem, we can consider two primary cases:

- **Case 1: Horizontal Placement.** Placing a horizontal domino covers two consecutive cells in one row, requiring another horizontal domino directly beneath it. This reduces the problem to $F(n-2)$.

The Domino Tiling Problem

Determine how many ways to cover 2×100 board using 1×2 dominos.

Solution

We define $F(n)$ to be the number of ways to cover a $2 \times n$ board. To solve the tiling problem, we can consider two primary cases:

- **Case 1: Horizontal Placement.** Placing a horizontal domino covers two consecutive cells in one row, requiring another horizontal domino directly beneath it. This reduces the problem to $F(n-2)$.

- **Case 2: Vertical Placement.** Placing a vertical domino covers one cell in each row, leaving a smaller $2 \times (n-1)$ board, hence, $F(n-1)$.

Recurrence Relation

The recursion for this problem is derived as:

$$F(n) = F(n-1) + F(n-2)$$

Recurrence Relation

The recursion for this problem is derived as:

$$F(n) = F(n-1) + F(n-2)$$

This relation resembles the Fibonacci sequence, with initial conditions:

- $F(0) = 1$ (empty board)
- $F(1) = 1$ (single vertical domino)

Recurrence Relation

The recursion for this problem is derived as:

$$F(n) = F(n-1) + F(n-2)$$

This relation resembles the Fibonacci sequence, with initial conditions:

- $F(0) = 1$ (empty board)
- $F(1) = 1$ (single vertical domino) 

Consequently, $F(n)$ follows the Fibonacci sequence shifted by one position.

Challenge

Prove why

$$F_n = \frac{1}{\sqrt{5}}\left[\left(\frac{1+\sqrt{5}}{2}\right)^n - \left(\frac{1-\sqrt{5}}{2}\right)^n\right]$$

Challenge

Prove why

$$F_n = \frac{1}{\sqrt{5}}\left[\left(\frac{1+\sqrt{5}}{2}\right)^n - \left(\frac{1-\sqrt{5}}{2}\right)^n\right]$$

Hints: (1. Use the recursive formula $F_n = F_{n-1} + F_{n-2}$.
2. Find a geometric sequence which satisfies the recursive formula. (A geometric sequence is $a, ar, ar^2, ar^3 \cdots$) Then solve for $X^n = X^{n-1} + X^{n-2}$).

Generating Closed Forms

Now we are going to take a look at how to create an explicit formula from a recursive formula.

Generating Closed Forms

Now we are going to take a look at how to create an explicit formula from a recursive formula.

Example

$X_n = 4X_{n-1} - 3X_{n-2}$ $X_0 = 3$ and $X_1 = 7$ What is the explicit formula?

Generating Closed Forms

Now we are going to take a look at how to create an explicit formula from a recursive formula.

Example

$X_n = 4X_{n-1} - 3X_{n-2}$ $X_0 = 3$ and $X_1 = 7$ What is the explicit formula?

1. We should find a geometric sequence which satisfies the recursive formula so we write $X^n = 4X^{n-1} - 3X^{n-2}$.
2. $(x-1)(x-3) = 0$. $x = 1$ and $x = 3$.
3. Since there are two possible geometric sequences, we algebraically combine them and set up the equations as $3 = A + B$ and $7 = A + 3B$. (3 is X_0 and 7 is X_1)
4.We find $A = 1$ and $B = 2$. Thus, the explicit formula of this recursive formula is $X_n = 1 + 2(3)^n$.

The method we used in our example is the same with the one used in proof of the Fibonacci explicit formula.

Generating Functions

Picking numbers from two boxes, and the two boxes has follows:

- Box 1: $\{1, 2, 3\}$
- Box 2: $\{2, 3, 4\}$

Generating Functions

Picking numbers from two boxes, and the two boxes has follows:

- Box 1: $\{1, 2, 3\}$
- Box 2: $\{2, 3, 4\}$

The sums obtained by choosing one number from each box are:

$$1 + 2 = 3$$
$$1 + 3 = 4$$
$$1 + 4 = 5$$
$$2 + 2 = 4$$
$$2 + 3 = 5$$
$$2 + 4 = 6$$
$$3 + 2 = 5$$
$$3 + 3 = 6$$
$$3 + 4 = 7$$

Generating Functions

Picking numbers from two boxes, and the two boxes has follows:

- Box 1: $\{1, 2, 3\}$
- Box 2: $\{2, 3, 4\}$

The sums obtained by choosing one number from each box are:

$$1 + 2 = 3$$
$$1 + 3 = 4$$
$$1 + 4 = 5$$
$$2 + 2 = 4$$
$$2 + 3 = 5$$
$$2 + 4 = 6$$
$$3 + 2 = 5$$
$$3 + 3 = 6$$
$$3 + 4 = 7$$

Thus, the distribution of sums is: $[3, 4, 4, 5, 5, 5, 6, 6, 7]$.

Introduction to generating functions: A generating function is a formal power series that encodes a sequence $a_0, a_1, a_2, \ldots$ as follows:

Introduction to generating functions: A generating function is a formal power series that encodes a sequence $a_0, a_1, a_2, \ldots$ as follows:

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

Introduction to generating functions: A generating function is a formal power series that encodes a sequence $a_0, a_1, a_2, \ldots$ as follows:

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

Generating functions for the same problem:

Introduction to generating functions: A generating function is a formal power series that encodes a sequence $a_0, a_1, a_2, \ldots$ as follows:

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

Generating functions for the same problem: The generating function for choosing numbers from Box 1 is:

$$G_1(x) = x^1 + x^2 + x^3$$

Introduction to generating functions: A generating function is a formal power series that encodes a sequence $a_0, a_1, a_2, \ldots$ as follows:

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

Generating functions for the same problem:The generating function for choosing numbers from Box 1 is:

$$G_1(x) = x^1 + x^2 + x^3$$

The generating function for choosing numbers from Box 2 is:

$$G_2(x) = x^2 + x^3 + x^4$$

Introduction to generating functions: A generating function is a formal power series that encodes a sequence $a_0, a_1, a_2, \ldots$ as follows:

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

Generating functions for the same problem:The generating function for choosing numbers from Box 1 is:

$$G_1(x) = x^1 + x^2 + x^3$$

The generating function for choosing numbers from Box 2 is:

$$G_2(x) = x^2 + x^3 + x^4$$

$$G(x) = G_1(x) \cdot G_2(x) = (x^1 + x^2 + x^3) \cdot (x^2 + x^3 + x^4) = x^1 \cdot (x^2 + x^3 + x^4) + x^2 \cdot (x^2 + x^3 + x^4) + x^3 \cdot (x^2 + x^3 + x^4)$$

Introduction to generating functions: A generating function is a formal power series that encodes a sequence $a_0, a_1, a_2, \ldots$ as follows:

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

Generating functions for the same problem:The generating function for choosing numbers from Box 1 is:

$$G_1(x) = x^1 + x^2 + x^3$$

The generating function for choosing numbers from Box 2 is:

$$G_2(x) = x^2 + x^3 + x^4$$

$$G(x) = G_1(x) \cdot G_2(x) = (x^1 + x^2 + x^3) \cdot (x^2 + x^3 + x^4) = x^1 \cdot (x^2 + x^3 + x^4) + x^2 \cdot (x^2 + x^3 + x^4) + x^3 \cdot (x^2 + x^3 + x^4)$$

Simplifying this product gives:

$$G(x) = x^3 + 2x^4 + 3x^5 + 2x^6 + x^7$$