Pigeonhole principle, inclusion-exclusion

CS 206: Discrete Structures II

Fall 2020

Pigeonhole principle

If you have two colors of socks in a drawer, how many do you have to pick to guarantee getting a pair of the same color?

What if you have three colors?

Pigeonhole principle

If there are more pigeons than pigeonholes, at least one pigeonhole must have more than one pigeon.

- If f is a function from A to B
- and $|A| > |B|$
- then there exist $a_1, a_2 \in A$ such that $f(a_1) = f(a_2)$

(Recall that if f is injective, this implies $|A| \leq |B|$)

Generalized pigeonhole principle

If

- $f : A \to B$
- $|A| > k \cdot |B|$
- then f maps at least $k+1$ elements of A to the same element of B

Example

How many people in Boston have the same number of hairs on their head?

- a person has at most 200,000 hairs
- Boston has a population of about 500,000

Example

Given 90 25-digit numbers, do any subsets have the same sum?

Aside: non-constructive proofs

Are there irrational numbers a and b such that a^b is rational?

Ramsey theory

$R(3) > 5$

$R(3) \leq 6$ by pigeonhole

Pigeonhole examples

- Given 367 people, at least 2 must have the same birthday
- Given 45000 US college students, at least 900 must be from the same state
- Given a 27-word passage, two words must start with the same letter
- Given 5 numbers in $[8]$, two must sum to 9

Pigeonhole examples

At a party of n people, there must be at least 2 people who have the same number of friends.

Two cases:

- everyone has at least 1 friend.
- someone has 0 friends.

Lossless compression

Some compression algorithms sometimes produce a compressed file that's larger than the original.

Does there exist a lossless compression scheme that doesn't do this?

Hash collisions

When you download a file, sometimes a SHA hash is given to verify your download is correct.

But the original file size is much larger

- e.g., SHA-256 uses only 256 bits

So some files must map to the same hash value!

Hash collisions

Suppose you try to download firefox.zip, but someone tricks you into downloading a corrupted firefox.zip.

Fortunately:

- 2^{256} is still a very large space
- if $f \mapsto h$, hard to find another $f' \mapsto h$

Suppose we have 60 math and 200 CS students.

How many students are there in total?

What if some students can double major?

Inclusion-exclusion

Suppose we have 60 math, 200 CS, and 40 physics students.

How many students are there in total?

What if some students can double/triple major?

Inclusion-exclusion

Suppose we have 60 math, 200 CS, and 40 physics students, and

- 10 math-CS double majors
- 7 CS-physics double majors
- 5 math-physics double majors
- 2 triple majors

Inclusion-exclusion

How many integers in $[100]$ are not divisible by 2, 3, or 5?

Inclusion-exclusion

For two sets:

$$|S_1 \cup S_2| = |S_1| + |S_2| - |S_1 \cap S_2|$$

Inclusion-exclusion

For three sets:

$$
\begin{aligned}
|S_1 \cup S_2 \cup S_3| = {} & |S_1| + |S_2| + |S_3| \\
& - |S_1 \cap S_2| - |S_1 \cap S_3| - |S_2 \cap S_3| \\
& + |S_1 \cap S_2 \cap S_3|
\end{aligned}
$$

Inclusion-exclusion

In general:

$$
\begin{aligned}
|S_1 \cup \cdots \cup S_n| &= \sum_{i=1}^{n} |S_i| \\
&- \sum_{1 \leq i < j \leq n} |S_i \cap S_j| \\
&+ \sum_{1 \leq i < j < k \leq n} |S_i \cap S_j \cap S_k| \\
&- \ldots \\
&+ (-1)^{(n-1)} \left| \bigcap_{i=1}^{n} S_i \right|
\end{aligned}
$$

Inclusion-exclusion

The same as a one-liner:

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

Derangement

At graduation, everyone throws their caps in the air. Does anyone get their same cap back?

A derangement is a permutation of a set such that no element is in its original position.

For a set of n elements, there are $n!$ permutations.

How many are derangements?

Let's count the non-derangements. Treat each permutation as a sequence.

Let S_i be the set of sequences where the ith element is in its original position.

Then all non-derangements are $\bigcup_{i=1}^{n} S_i$.

Derangement

By inclusion-exclusion,

$$\left|\bigcup_{i=1}^{n} S_i\right| = \sum_{i=1}^{n} |S_i| - \sum_{i<j} |S_i \cap S_j| + \cdots$$

Note that $S_i \cap S_j$ has two elements in their original position, but the number of ways is independent of which two elements.

$$= \binom{n}{1}|S_1| - \binom{n}{2}|S_1 \cap S_2| + \cdots + (-1)^{(p-1)}\binom{n}{p}|S_1 \cap \cdots \cap S_p| + \cdots$$

Derangement

$$= \binom{n}{1}|S_1| - \binom{n}{2}|S_1 \cap S_2| + \cdots + (-1)^{(p-1)}\binom{n}{p}|S_1 \cap \cdots \cap S_p| + \cdots$$

And $S_1 \cap \cdots \cap S_p$ has p elements in their original position. The other $n - p$ elements can be in any of their $(n-p)!$ possible orders.

$$= \binom{n}{1}(n-1)! - \binom{n}{2}(n-2)! + \cdots + (-1)^{(p-1)}\binom{n}{p}(n-p)! + \cdots$$

Derangement

$$
\begin{aligned}
&= \binom{n}{1}(n-1)! - \binom{n}{2}(n-2)! + \cdots + (-1)^{(p-1)}\binom{n}{p}(n-p)! + \cdots \\
&= \sum_{p=1}^{n}(-1)^{(p-1)}\binom{n}{p}(n-p)! \\
&= \sum_{p=1}^{n}(-1)^{(p-1)}\frac{n!}{p!(n-p)!}(n-p)! \\
&= \sum_{p=1}^{n}(-1)^{(p-1)}\frac{n!}{p!}
\end{aligned}
$$

Derangement

So the number of derangements is all permutations minus this:

$$n! - \sum_{p=1}^{n} (-1)^{(p-1)} \frac{n!}{p!}$$

Derangement

Consider permutations of three elements $\{a, b, c\}$:

Permutation	Elements in original position	Derangement?
abc		
acb		
bac		
bca		
cab		
cba		

Derangement

$$
\begin{aligned}
n! - \sum_{p=1}^{n} (-1)^{(p-1)} \frac{n!}{p!} &= 3! - \sum_{p=1}^{3} (-1)^{(p-1)} \frac{3!}{p!} \\
&= 6 - \left((-1)^0 \frac{3!}{1!} + (-1)^1 \frac{3!}{2!} + (-1)^2 \frac{3!}{3!} \right) \\
&= 6 - (6 - 3 + 1) \\
&= 2
\end{aligned}
$$