Pigeon Hole Principle and Principle of Inclusion-Exclusion

Benjamin Cosman, Patrick Lin and Mahesh Viswanathan

Fall 2020

TAKE-AWAYS

- The pigeon hole principle says that if $|A| > |B|$ then for any function $f : A \to B$ there are $a, b \in A$ such that $f(a) = f(b)$.
- The generalized pigeon hole principle is as follows. Let A be a set and B be an n-element set (say) $\{b_1, b_2, \ldots b_n\}$. Let $q_1, \ldots q_n$ be n natural numbers such that

  $$|A| > q_1 + q_2 + \cdots + q_n.$$

  For any function $f : A \to B$ there is an $i \in \{1, 2, \ldots n\}$ such that $|\{a \in A \mid f(a) = b_i\}| > q_i$.
- Observe that the (basic) pigeon hole principle is a special case of the generalized pigeon hole principle, where each $q_i = 1$.
- Another special case of the generalized pigeon hole principle is as follows. If $|A| > k|B|$ then for any function $f : A \to B$ there are $k+1$ elements $a_1, a_2, \ldots a_{k+1} \in A$ such that $f(a_i) = f(a_j)$ for any $i, j \in \{1, 2, \ldots k+1\}$.
- The principle of inclusion-exclusion says that for any sets $S_1, S_2, \ldots S_n$,

  $$\left|\bigcup_{i=1}^{n} S_i\right| = \sum_{\emptyset \neq I \subseteq \{1,2,\ldots n\}} (-1)^{|I|+1} \left|\bigcap_{i \in I} S_i\right|$$

- When $n = 2$ or $n = 3$, the principle of inclusion-exclusion specializes to the following equations.

  $$\begin{aligned} |A \cup B| &= |A| + |B| - |A \cap B| \\ |A \cup B \cup C| &= |A| + |B| + |C| - |A \cap B| - |A \cap C| - |B \cap C| \\ &\quad + |A \cap B \cap C| \end{aligned}$$

Pigeon Hole Principle

THE PIGEON HOLE PRINCIPLE is a simple, yet extremely powerful proof principle. Informally it says that if $n+1$ or more pigeons are placed in n holes, then some hole must have at least 2 pigeons. This is also known as the Dirichlet's drawer principle or Dirichlet's box principle after the mathematician Peter Gustav Dirichlet. Formally, we could restate it as follows.

Proposition 1 (Pigeon Hole Principle)**.** *If A and B are sets such that $|A| > |B|$ then for any function $f : A \to B$ there are $a, b \in A$ such that $f(a) = f(b)$.*

Proof. Observe that we have previously established that if $f : A \to B$ is injective then $|A| \leq |B|$. The contrapositive of this is that if $|A| > |B|$ then a function $f : A \to B$ is not injective. Or there must be two elements $a, b \in A$ such that $f(a) = f(b)$. □

The pigeon hole principle is used to establish many different results. For example, suppose you have a drawer with orange and blue colored socks, each of which can be worn on either foot, and you are pulling socks from the drawer without looking. How many socks do you need to pull before you are guaranteed to get a pair of the same color? Taking each sock to be a pigeon and the colors to be holes into which they are assigned, the pigeon hole principle guarantees that in any set of 3 socks there is a pair of the same color. While this is an extremely straightforward application of the pigeon hole principle, there are many non-trivial observations that follow from this principle.

Proposition 2. *In any set $S \subseteq \mathbb{Z}$ with $|S| = n$, there are $a, b \in S$ such that $a - b$ is a multiple of $n - 1$, i.e., $(n-1)|(a-b)$.*

Before proving the proposition, let us look at an example to understand what it is saying. Consider an example set of integers S that is (say) $\{4, 3, 1, 7, 8\}$. There are two elements, namely 4 and 8, such that $4 - 8 = -4$ is a multiple of $|S| - 1 = 5 - 1 = 4$.

Proof of Proposition 2. Consider the function $r : S \to \{0, 1, 2, \ldots n-2\}$ where $r(a) = \mathsf{rem}(a, n-1)$, i.e., r maps an number a to the remainder when a is divided by $n - 1$. Since $|S| = n > n - 1 = |\{0, 1, 2 \ldots n - 2\}|$, by the pigeon hole principle, there are $a, b \in S$ such that $r(a) = r(b)$. That is, $\mathsf{rem}(a, n-1) = \mathsf{rem}(b, n-1)$. Therefore, $a \equiv b \bmod (n-1)$ or $(n-1)|(a-b)$. □

Problem 1. A chess player trains for a championship by playing practice games over 77 days. She plays at least one game on any day,

and plays a total of at most 132 games. Prove that no matter what her schedule of games looks like, there is a period of consecutive days in which she plays *exactly* 21 games.

The above statement is a surprising observation. It says that even though we have no idea of what the chess player's match schedule looks like, there is a period of time when she plays exactly 21 games! This illustrates the power of the pigeon hole principle which can be used to prove statements like this, that establish the existence of pattern in an arbitrary situation.

Proof. Take a_i to denote the total number of games played in the first i days. Given that the chess player plays at least one game each day, and plays no more than 132 games over the 77 days, we say that

$$1 \leq a_1 < a_2 < \cdots < a_{77} \leq 132.$$

Let us define b_i to be $a_i + 21$. Using the previous sequence of inequalities, we can say that

$$22 \leq b_1 = a_1 + 21 < b_2 = a_2 + 21 < \cdots < b_{77} = a_{77} + 21 \leq 153.$$

Observe that we have 154 numbers in the sequence $a_1, a_2, \ldots a_{77}$, $b_1, b_2, \ldots b_{77}$. Each of these numbers takes a value between 1 and 153. Thus, by the pigeon hole principle, there are two numbers in this sequence that are the same. Observe that none of the a_is are equal, and none of the b_is are equal. Thus, it must be the case that there are j, k such that $a_j = b_k = a_k + 21$. Note that k must be less than i since the a_is form a strictly increasing sequence of numbers. Therefore, the chess player plays exactly 21 games in total on the days $k + 1, k + 2, \ldots j$. □

Theorem 3 (Chinese Remainder Theorem)**.** *Let $m, n \in \mathbb{N}$ be coprime, i.e., $\mathsf{gcd}(m, n) = 1$. Let a, b be any integers such that $0 \leq a < m$ and $0 \leq b < n$. Then there is $x \in \mathbb{N}$ such that $x < mn$, $\mathsf{rem}(x, m) = a$ and $\mathsf{rem}(x, n) = b$.*

Before presenting the proof of this important result in number theory, let us look at an example to understand what it is saying. Consider coprime numbers 6 (m) and 11 (n). Suppose we take a to be 5 and b to be 10, the theorem says that there is a number x that leaves a remainder of 5 when divided by 6 and a remainder of 10 when divided by 11. What is such a number? 65; $\mathsf{rem}(65, 6) = 5$ and $\mathsf{rem}(65, 11) = 10$.

Proof of Theorem 3. Consider the set of n numbers $S = \{a + im \mid 0 \leq i < n\} = \{a, a + m, a + 2m, \ldots a + (n - 1)m\}$. Since $a < m$, we have $a + (n - 1)m < nm$; thus, every element of S is $< mn$. We will show

that one of the numbers in S is our desired number x, i.e., it leaves a remainder of a when divided by m and a remainder of b when divided by n.

First observe that for any i, $\mathsf{rem}(a + im, m) = a$. Thus, every number in S leaves a remainder of a when divided by m. All that is left to show is that one of the numbers in S leaves a remainder of b when divided by n.

Suppose (for contradiction), for all $\ell \in S$, $\mathsf{rem}(\ell, n) \neq b$. Consider the function $r : S \to \{0, 1, 2, \ldots n-1\}$ defined as $r(\ell) = \mathsf{rem}(\ell, n)$. Since (by assumption) $b \notin \mathsf{rng}(r)$, $|\mathsf{rng}(r)| \leq n - 1$. Observe that $|S| = n$. Thus, by the pigeon hole principle, there are $i < j$ such that $r(a + im) = \mathsf{rem}(a + im, n) = \mathsf{rem}(a + jm, n) = r(a + jm)$. This means that $a + jm \equiv a + im \bmod n$ or $n|((a + jm) - (a + im))$. Simplifying we have, $n|(j - i)m$. Since m and n are coprime, it must be the case that $n|(j - i)$. But this is impossible since $0 < j - i < n$. □

Applications of the Chinese Remainder Theorem

The Chinese Remainder theorem is an important result in number theory that has applications in computer science. Our statement of the theorem is weaker than what the proof entails. One can observe that the proof shows that, for m, n, a, b as given in the theorem, there is a *unique* number $x < mn$ such that $x \equiv a \bmod m$ and $x \equiv b \bmod n$. Therefore, one way to interpret the Chinese Remainder Theorem is that it says that the (large) number x can be represented by the pair of (small) numbers (a, b). Many number theoretic algorithms exploit this interpretation and use the "chinese remainder theorem representation" of numbers to compute large numbers.

The pigeon hole principle can be generalized as follows.

Proposition 4 (Generalized Pigeon Hole Principle)**.** *Let A be a set and B be an n-element set (say) $\{b_1, b_2, \ldots b_n\}$. Let $q_1, \ldots q_n \in \mathbb{N}$ be such that $|A| > q_1 + q_2 + \cdots + q_n$. For any function $f : A \to B$, there is an $i \in \{1, 2, \ldots n\}$ such that $|\{a \in A \mid f(a) = b_i\}| > q_i$.*

Proof. For an element $b \in B$, let us define $f^{-1}(b) = \{a \in A \mid f(a) = b\}$. Observe that since each element gets mapped to exactly one value in B, we have $A \subseteq \cup_{i=1}^{n} f^{-1}(b_i)$ and for any $j \neq k$, $f^{-1}(b_j) \cap f^{-1}(b_k) = \emptyset$.

We need to show that for some i, $|f^{-1}(b_i)| > q_i$. Suppose (for contradiction) this is not true, i.e., for every i, $|f^{-1}(b_i)| \leq q_i$. Based on the observation in the previous paragraph, and using the sum rule,

we have

$$|A| \leq |\cup_{i=1}^{n} f^{-1}(b_i)| = \sum_{i=1}^{n} |f^{-1}(b_i)| \leq \sum_{i=1}^{n} q_i.$$

But this contradicts the assumption that $|A| > \sum_{i=1}^{n} q_i$. □

Observe that the (simple) pigeon hole principle (Proposition 1) follows from Proposition 4 — simply take each $q_i = 1$. There is another special case of Proposition 4 that arises in many situations. If we take each q_i to be some (fixed) number k, we see that Proposition 4 says that if the cardinality of the domain of a function f is greater than k times the cardinality of its co-domain, then there are at least $k+1$ elements that are mapped to the same value by f. Since this form is very useful, we state it explicitly.

Proposition 5. *If $|A| > k|B|$ then for any function $f : A \to B$ there are $k+1$ elements $a_1, a_2, \ldots a_{k+1} \in A$ such that $f(a_i) = f(a_j)$ for any $i, j \in \{1, 2, \ldots k+1\}$.*

We conclude our discussion on the pigeon hole principle by looking at some applications of the generalized pigeon hole principle.

Example 6. Suppose we draw cards from a standard 52 card deck. How many cards must we draw to ensure that we get 3 cards of one suit? How many cards must we draw if we instead want 3 cards of the hearts suit? The answer to these two questions is different and can be computed using the generalized pigeon hole principle.

Let us consider the first question. Consider the function that maps each drawn card to its suit. The co-domain of this function has 4 elements. Taking each q_i to be 2 in Proposition 4, we see that if we draw more than $2 \times 4 = 8$ cards then we can ensure that there are 3 cards in one suit. Thus, we need to draw at least 9 cards.

To answer the second question, once again consider the function that maps each drawn card to its suit. Since we may draw all 13 of the spades, all 13 of the diamonds, all 13 of the clubs, before we get the 3 hearts cards, we need to draw $13 + 13 + 13 + 3 = 42$ cards to ensure that we get 3 cards of the hearts suit.

Problem 2. Two people will be said to be *acquaintances* if they have met before, and they will be said to be *strangers* if they have never met before. Prove that in any group of 6 people, there is either a group of 3 people who are mutual acquantainces (i.e., any two in this group of 3 have met before) or there is a group of 3 mutual strangers (i.e., no two in this group of 3 have met before).

Proof. Let a be one of the 6 people. Let K denote the set of people a has met (among the other 5 people) and let S be the set of people a

has not met before. By the generalized pigeonhole principle, either $|K| \geq 3$ or $|S| \geq 3$.

Case $|K| \geq 3$: Let b, c, d denote 3 of the people a is acquanited with. If any pair among b, c, d are acquainted then a, along with the pair form a group of mutual acquaintances. On the other hand, if none among b, c, d have met before, then they form a group of mutual strangers.

Case $|S| \geq 3$: Let b, c, d denote 3 people that a has not met. If any pair among b, c, d are strangers, then a along with the pair form a group of 3 mutual strangers. If that is not the case then b, c, d all are acquainted, and so they form a group of 3 mutual acquaintances. □

Ramsey Theory

Problem 2 is a special case of a result due to Ramsey that started a sub-field within combinatorics called Ramsey Theory that tries to "find ordered regularity among disorder" — find regular sub-structures in any large object. The specific result of Ramsey states that for any ℓ and k, there is a number $R(\ell, k)$ such that in any group of size at least $R(\ell, k)$, there is either a group of size ℓ of mutual acquaintances, or a group of size k of mutual strangers. The special case here says that $6 \geq R(3, 3)$.

Principle of Inclusion-Exclusion

THE PRINCIPLE OF INCLUSION-EXCLUSION is a way to calculate the cardinality of a set that is expressed as a union of other sets. It is a generalization of the sum rule of counting that states that the cardinality of the union of disjoint sets is the sum of the cardinalities of the individual sets. Before presenting the general principle of inclusion-exclusion, we begin by looking at a couple of special cases that you maybe familiar with.

Proposition 7. *For any sets A, B, $|A \cup B| = |A| + |B| - |A \cap B|$.*

Proof. Observe that, $A \cup B = A \cup (B \setminus A)$. Notice that $A \cap (B \setminus A) = \emptyset$. Thus by the sum rule,

$$|A \cup B| = |A| + |B \setminus A|.$$

Next observe that $B = (B \setminus A) \cup (A \cap B)$ and that $(B \setminus A) \cap (A \cap B) = \emptyset$. Therefore, again by the sum rule we have

$$|B| = |B \setminus A| + |A \cap B|.$$

This means that $|B \setminus A| = |B| - |A \cap B|$. Putting it all together, we have

$$|A \cup B| = |A| + |B \setminus A| = |A| + |B| - |A \cap B|.$$

□

For 3 sets A, B, C, we can generalize Proposition 7 to get

$$|A \cup B \cup C| = |A| + |B| + |C| - |A \cap B| - |A \cap C| - |B \cap C| + |A \cap B \cap C|.$$

We leave the proof of this result as an exercise for the reader. In general, the principle of inclusion-exclusion says that the cardinality of the union of n sets is the sum of the cardinalities of the individual sets, minus the cardinality of their pairwise intersection, plus the cardinality of their 3-way intersections, and so on, with the plus and minus signs alternating with intersections of larger number of sets. Or informally,

$$|S_1 \cup S_2 \cup \cdots \cup S_n| = \sum_{1 \le i \le n} |S_i| - \sum_{1 \le i < j \le n} |S_i \cap S_j| + \sum_{1 \le i < j < k \le n} |S_i \cap S_j \cap S_k| \cdots$$

This is stated fomally below.

Theorem 8 (Principle of Inclusion-Exclusion)**.** *For any sets $S_1, S_2, \ldots S_n$,*

$$\left| \bigcup_{i=1}^{n} S_i \right| = \sum_{\emptyset \neq I \subseteq \{1,2,\ldots n\}} (-1)^{|I|+1} \left| \bigcap_{i \in I} S_i \right|.$$

Theorem 8 can be proved by induction using Proposition 7 to establish the base case and to carry out the induction step. We skip the proof here. Let us conclude our discussion by using Theorem 8 to prove a result.

Example 9. How many integers between 1 and 100 are multiples of either 2 or 3? We will Proposition 7 to establish this result.

Let us define a couple of sets.

$$A_2 = \{1 \le i \le 100 \mid 2|i\} \qquad A_3 = \{1 \le i \le 100 \mid 3|i\}$$

Observe that we need to compute $|A_2 \cup A_3|$. We will use the principle of inclusion-exclusion for this. We know,

$$|A_2| = \left\lfloor \frac{100}{2} \right\rfloor = 50 \qquad |A_3| = \left\lfloor \frac{100}{3} \right\rfloor = 33.$$

[1] Notice that

$$A_2 \cap A_3 = \{1 \le i \le 100 \mid 6|i\}.$$

That is, $A_2 \cap A_3$ is the set of all numbers that are multiples of 6. Thus, $|A_2 \cap A_3| = \lfloor \frac{100}{6} \rfloor = 16$. Thus, by the principle of inclusion-exclusion,

$$|A_2 \cup A_3| = |A_2| + |A_3| - |A_2 \cap A_3| = 50 + 33 - 16 = 67.$$

[1] For a real number $r \in \mathbb{R}$, $\lfloor r \rfloor$ is largest integer that is $\le r$. On the other hand, $\lceil r \rceil$ is the smallest integer $\ge r$.