# Discrete Structures for Computer Science

**William Garrison**

bill@cs.pitt.edu

6311 Sennott Square

Lecture #20: Inclusion/Exclusion, Pigeonhole Principle

University of Pittsburgh

Based on materials developed by Dr. Adam Lee

# Today's Topics

- Inclusion/exclusion principle

- The pigeonhole principle

# Sometimes when counting a set, we count the same item more than once

For instance, if something can be done $n_1$ ways or $n_2$ ways, but some of the $n_1$ ways are the same as some of the $n_2$ ways.

In this case $n_1 + n_2$ is an overcount of the ways to complete the task!

What we really want to do is count the $n_1 + n_2$ ways to complete the task and then subtract out the common solutions.
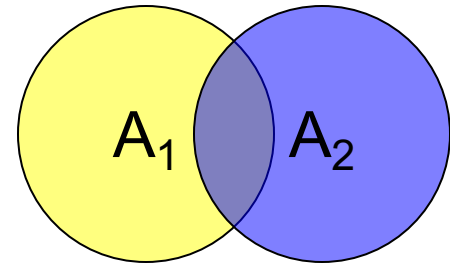
This is called the inclusion/exclusion principle.

# We can formulate this concept using set theory

Suppose that a task T can be completed using a solution drawn from one of two classes: $A_1$ and $A_2$

As in the sum rule, we can define the solution set for the task T as $S = A_1 \cup A_2$

Then $|S| = |A_1 \cup A_2|$

$\qquad = |A_1| + |A_2| - |A_1 \cap A_2|$

*Do you remember this from earlier this semester?*

# Counting Bit Strings

***Example:*** How many bit strings of length 8 start with a 1 or end with 00?

***Solution:***

- $2^7 = 128$ 8-bit strings start with a 1
- $2^6 = 64$ 8-bit strings end with 00
- $2^5 = 32$ 8-bit strings start with a 1 and end with 00

So, we have 128 + 64 – 32 = 160 ways to construct an 8-bit string that starts with a 1 or ends with 00.

# Job Applications

*Example:* A company receives 350 applications. Suppose 220 of these people majored in CS, 147 majored in business, and 51 were CS/business double-majors. How many applicants majored in neither CS nor business?

*Solution:*

- Let C be the set of CS majors, B be the set of business majors
- $|C \cup B| = |C| + |B| - |C \cap B|$
- $\quad\quad = 220 + 147 - 51$
- $\quad\quad = 316$

So of the 350 applications, 350 – 316 = 34 applications neither majored in CS nor business.

# The pigeonhole principle is an incredibly simple concept that is extremely useful!

**The pigeonhole principle:** If k is a positive integer and k+1 objects are placed in k boxes, then at least one box contains at least two objects.



Example:  k = 4

# The pigeonhole principle is also easy to prove

The pigeonhole principle:  If k is a positive integer and k+1 objects are placed in k boxes, then at least one box contains at least two objects.

*Proof:*  Assume that each of the k boxes contains at most 1 item.  This means that there are at most k items, which is a contradiction of our assumption that we have k+1 items, so at least one box must contain more than one item.  ❏

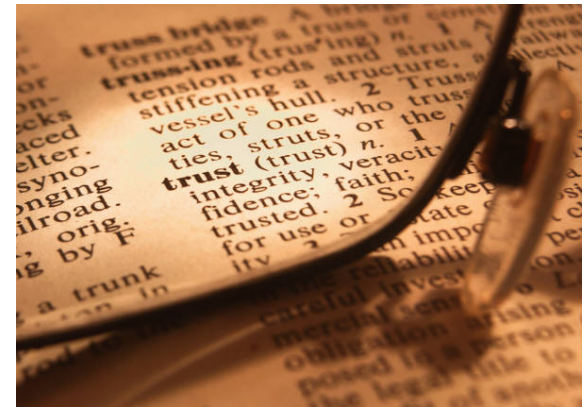# Examples

**_Example:_** Among any group of 367 people there are at least two with the same birthday, since there are only 366 possible birthdays.

**_Example:_** Among any 27 English words, at least two will start with the same letter.

# The pigeonhole principle can be used to prove a number of interesting results

*Claim:* Every integer n has a multiple whose decimal representation contains only 1s and 0s

*Proof:*

*This number has n+1 ones*

- Let n be a positive integer
- Consider the set of n+1 integers S = {1, 11, 111, ..., 111...1}
- Note that when any integer x is divided by n, there are n possible remainders (0 through n-1)
- Since S contains n+1 elements, at least two elements of S have the same remainder when divided by n (call them x and y, with x > y)
- Since x ≡ y (mod n), n | (x – y), and thus na = (x – y)
- Finally, we note that x – y contains only 0s and 1s  ❏

# In-class exercises

**Problem 1:** How many strings of three decimal digits begin with a 1 or end with a 4?

**Problem 2:** How many bit strings of length 10 begin with three 0s or end with two 0s?

**Problem 3:** If a student can get either an A, B, C, D, or F on a test, how many students are needed to ensure that at least two get the same grade?

# There is a more general form of the pigeonhole principle that is even more useful

The generalized pigeonhole principle:  If N objects are placed into k boxes, then there is at least one box containing at least ⌈N/k⌉ items.

*Proof:*

- Assume that no box contains more than ⌈N/k⌉ - 1 objects
- Note that ⌈N/k⌉ < (N/k) + 1
- So, k (⌈N/k⌉ - 1) < k((N/k + 1) – 1) = N
- This contradicts our assumption that we had N objects  ❑

# Example

What is the minimum number of students needed to guarantee that at least six students receive the same grade, if possible grades are A, B, C, D, and F?

## Solution:

- Need the smallest integer N such that $\lceil N/5 \rceil = 6$
- With 25 students, it would be possible (though maybe unlikely) to have 5 students get each possible grade
- By adding a 26th student, we guarantee that at least 6 students get one possible grade
- So, the smallest such N is $5 \times 5 + 1 = 26$ ❑

# From the casino…

How many cards must be drawn from a standard 52-card deck to guarantee that three cards of the same suit are drawn?

*Solution:*

- Let's make 4 piles: one for each suit
- We want to have $\lceil N/4 \rceil \geq 3$
- We can do this using $4 \times 2 + 1 = 9$ cards

Note: We don't need 9 cards to end up with three from the same suit—if we did, we could never get a flush in poker!

# We can't always use the pigeonhole principle directly

How many cards would we need to draw to ensure that we picked at least three hearts?



In the worst case, we would need to draw every club, spade, and diamond before getting three hearts...

So, to guarantee three hearts, we need to draw 3 ×13 + 3 = 42 cards!

# Ma Bell…

What is the least number of area codes needed to guarantee that the 25 million phones in some state can be assigned distinct 10-digit phone numbers of the form NXX-NXX-XXXX?

## *Solution:*

- The product rule tells us that there are 8 million phone numbers of the form NXX-XXXX
- Think of phones as objects and phone numbers as boxes
- By the generalized pigeonhole principle, we know that some "box" contains at least ⌈25,000,000/8,000,000⌉ = 4 "objects"
- This means that we need 4 area codes to ensure that each phone gets a unique 10-digit number   ❏

# This has been easy so far, right?

Unfortunately, life isn't always easy!

Sometimes, we need to be clever when we are defining our "boxes" or assigning objects to them

*For example…*

# Sports!

During a month with 30 days, a baseball team plays at least one game per day, but no more than 45 games total. Show that there must be some period of consecutive days in which exactly 14 games are played.

*Solution:*

- Let $a_j$ be the number of games played on or before the $j^{th}$ day of the month. Note that the sequence $\{a_j\}$ is <span style="color:red">strictly increasing</span>.
- Note also that $\{a_j + 14\}$ is also an increasing sequence
- Now, consider $a_1, a_2, …, a_{30}, a_1 + 14, a_2 + 14, …, a_{30} + 14$
- There are 60 terms in this sequence, all $\leq (45 + 14) = 59$
- By the pigeonhole principle, at least two terms are equal
- Note: Each $a_j$ for $j = 1, 2, …, 30$ is distinct, as is each $a_j + 14$
- This means there exists some $a_i$ that is equal to some $a_j + 14$, so 14 games were played from day $j + 1$ to day $i$ ❑

# Number theory

Show that among any n+1 positive integers not exceeding 2n, there must be an integer that divides one of the other integers.

*Proof:*

- Call our $n + 1$ positive integers $a_1, a_2, \ldots, a_{n+1}$
- Write each $a_i$ as $2^{k_i} q_i$, where $q_i$ is an odd positive integer and $k_i$ is non-negative (i.e., $k_i$ might be zero)
- Note that there are $n$ odd positive integers less than $2n$
- By the pigeonhole principle, at least two of $q_1, q_2, \ldots, q_{n+1}$ must be equal
- This means we have some $a_i = 2^{k_i} q_i$ and some $a_j = 2^{k_j} q_i$
- If $k_i < k_j$, then $a_i \mid a_j$. If $k_i > k_j$, then $a_j \mid a_i$  ❑

# In-class exercises

**Problem 4:** What is the minimum number of students, each of whom comes from one of the 50 states, who must be enrolled in a university to guarantee that there are at least 100 who come from the same state?

**Problem 5:** A drawer contains a dozen brown socks and a dozen black socks, all unmatched. How many socks must be drawn to find a matching pair? How many socks must be drawn to find a pair of black socks?

**Problem 6:** Let A be some subset of {1, 2, …, 50} where |A|=10. Show that there are two subsets of A that have the same sum.

# Final Thoughts

- The inclusion/exclusion principle is useful when we need to avoid <span style="color:red">overcounting</span>

- The pigeonhole principle and its generalized form are useful for solving many types of counting problems

- Next time:
  - Permutations and combinations (Section 6.3)