Introduction
○ ○ ○

Possibility Trees and Multiplication Rule
○ ○ ○ ○ ○

Counting Elements of Disjoint Sets
○ ○ ○

The Pigeonhole Principle
○ ○ ○

Lecture #10: Counting and Probability I

Aaron Tan

10. Counting and Probability 1

9.1 Introduction

This lecture is based on Epp's book chapter 9.
Hence, the section numbering is according to the book.

- Random process, sample space, event and probability

9.2 Possibility Trees and the Multiplication Rule

- Possibility trees
- The multiplication/product rule
- Permutations

9.3 Counting Elements of Disjoint Sets

- The addition/sum rule
- The difference rule
- The inclusion/exclusion rule

9.4 The Pigeonhole Principle (PHP)

- Pigeonhole principle, general pigeonhole principle

Reference: Epp's Chapter 9 Counting and Probability

2

9.1 Introduction

Introduction

Counting

- **Counting schemes covered**

    - Basics of counting: product/multiplication rule and sum/addition rule

    - Permutations, with or without repetitions

    - Combinations, with or without repetitions

    - Inclusion-exclusion principle

    - Pigeonhole principle (Dirichlet drawer principle)

Introduction

Counting

- Counting has important applications in Computer Science
    - Determining complexity of algorithms
    - Computing discrete probability of some events
- Example: How many times is the routine this() called in the following algorithm?

```
process(n, m) {
    while (n>0) {
        for j ← 1 to m {
            this(j);
        }
        n = n/2;
    }
}
```

Tossing two coins

- 0, 1 or 2 heads?
- Does each of these events occur about 1/3 of the time?

1/4

No heads obtained

2/4

One head obtained

1/4

Two heads obtained

Introduction   Possibility Trees and Multiplication Rule   Counting Elements of Disjoint Sets   The Pigeonhole Principle
● ○ ○   ○ ○ ○ ○ ○   ○ ○ ○   ○ ○ ○

Introduction

Table 9.1.1. Relative frequencies.

Event	Tally	Frequency (Number of times the event occurred)	Relative Frequency (Fraction of times the event occurred)
2 heads obtained	⦀⦀ ⦀⦀ ⎮	11	22%
1 head obtained	⦀⦀ ⦀⦀ ⦀⦀ ⦀⦀ ⦀⦀ ⎮⎮	27	54%
0 heads obtained	⦀⦀ ⦀⦀ ⎮⎮	12	24%

Formalizing the analysis, we introduce:

- random process
- sample space
- event
- probability

Introduction

To say that a process is **random** means that when it takes place, one outcome from some set of outcomes is sure to occur, but it is impossible to predict with certainty which outcome that will be.

Definitions

A **sample space** is the set of all possible outcomes of a random process or experiment.
An **event** is a subset of a sample space.

Equally Likely Probability

Notation

For a finite set A, $|A|$ denotes the number of elements in A.

> Note: Epp uses $N(A)$ to denote the number of elements in A.
> For consistency with earlier topic (Sets), we shall use $|A|$.
> Also, we deal with finite sets here. Infinite sets will be
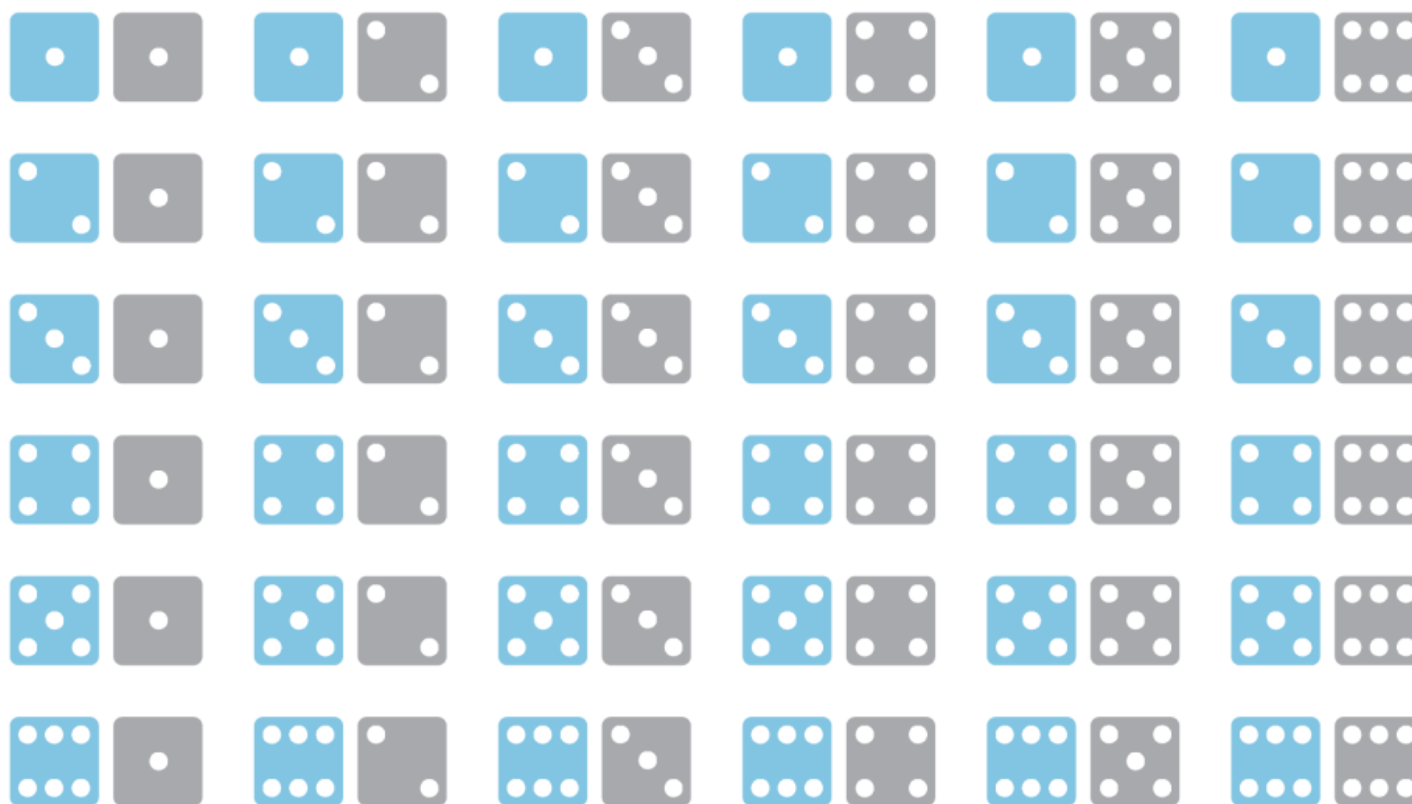> discussed another time.

Equally Likely Probability Formula

If S is a finite sample space in which all outcomes are equally likely and E is an event in S, then the **probability** of E, denoted $P(E)$, is

$$P(E) = \frac{\text{The number of outcomes in } E}{\text{The total number of outcomes in } S} = \frac{|E|}{|S|}$$

Equally Likely Probability

Rolling a Pair of Dice

Equally Likely Probability

A more compact notation identifies, say, ▢▢ with the notation 24, ▢▢ with 53, and so forth.

a.   Use the compact notation to write the sample space *S* of possible outcomes.

S = { 11, 12, 13, 14, 15, 16, 21, 22, 23, 24, 25, 26, 31, 32, 33, 34, 35, 36,
41, 42, 43, 44, 45, 46, 51, 52, 53, 54, 55, 56, 61, 62, 63, 64, 65, 66 }

b.   Use set notation to write the event *E* that the numbers showing face up have a sum of 6 and find the probability of this event.

E = { 15, 24, 33, 42, 51 }   *P*(*E*) = 5/36

11

Counting the Elements of a List

Counting the Elements of a List

Some counting problems are as simple as counting the elements of a list.

Eg: How many integers are there from 5 through 12?

list:	5	6	7	8	9	10	11	12
	↕	↕	↕	↕	↕	↕	↕	↕
count:	1	2	3	4	5	6	7	8

Theorem 9.1.1 The Number of Elements in a List

If m and n are integers and $m \leq n$, then there are

$$n - m + 1$$

integers from m to n inclusive.

Counting the Elements of a List

a. How many 3-digit integers (from 100 to 999 inclusive) are divisible by 5?

QUIZ

© Can Stock Photo

100 101 102 103 104 **105** 106 107 108 109 **110** … 994 **995** 996 997 998 999

5×**20**   5×**21**   5×**22**   5×**199**

From Theorem 9.1.1, there are 199 – 20 + 1 = 180 such integers.
Hence, there are **180** 3-digit integers that are divisible by 5.

b. What is the probability that a randomly chosen 3-digit integer is divisible by 5?

By Theorem 9.1.1, total number of integers from 100 through 999 = 999 – 100 + 1 = 900.

By part (a), 180 of these are divisible by 5.

Hence, answer = 180/900 = **1/5**.

13

9.2 Possibility Trees and the Multiplication Rule

Introduction
○ ○ ○

Possibility Trees and Multiplication Rule
● ○ ○ ○ ○

Counting Elements of Disjoint Sets
○ ○ ○

The Pigeonhole Principle
○ ○ ○

Possibility Trees

Possibility Trees

A **tree structure** is a useful tool for keeping systematic track of all possibilities in situations in which events happen in order.

Coin 1		H	T
Coin 2	H	T	H	T
Coin 3	H T	H T	H T	H T
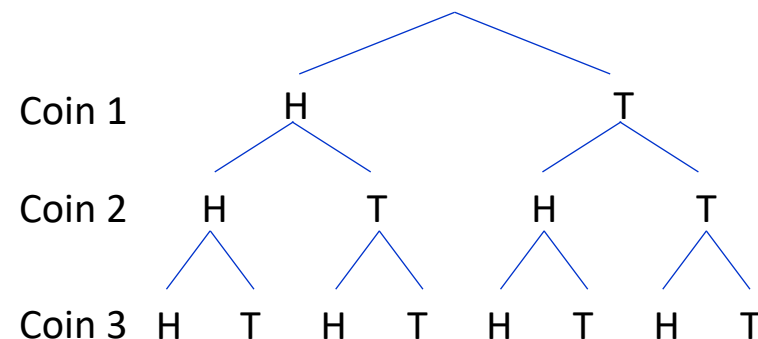
Example 1: Possibilities for Tournament Play

Teams *A* and *B* are to play each other repeatedly until one wins two games in a row, or a total of three games.

One way in which this tournament can be played is for *A* to win the first game, *B* to win the second, and *A* to win the third and fourth games. Denote this by writing *A–B–A–A.*

15

Introduction

○ ○ ○

Possibility Trees and Multiplication Rule

● ○ ○ ○ ○

Counting Elements of Disjoint Sets

○ ○ ○

The Pigeonhole Principle

○ ○ ○

Possibility Trees

Example 1: Possibilities for Tournament Play

a. How many ways can the tournament be played?

Possible ways are represented by the distinct paths from "root" (the start) to "leaf" (a terminal point) in the tree below.
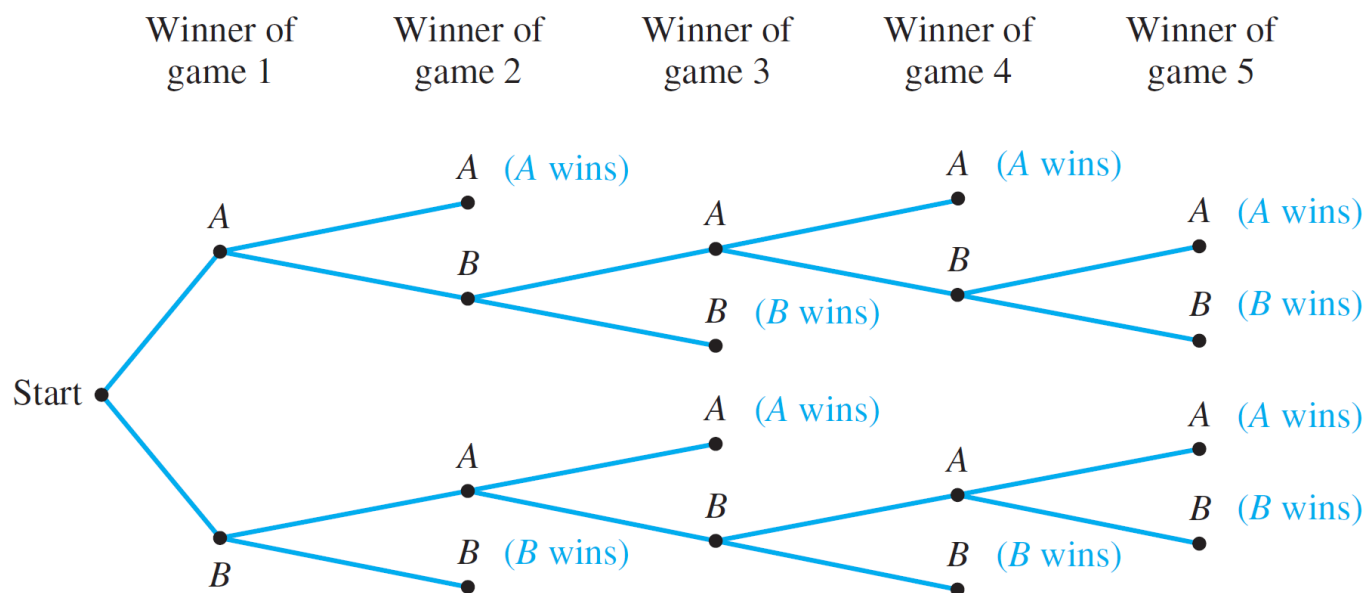

Figure 9.2.1 The Outcomes of a Tournament

Possibility Trees

Example 1: Possibilities for Tournament Play

a. How many ways can the tournament be played?

Ten paths from the root of the tree to its leaves ➜ ten possible ways for the tournament to be played.

1. A-A
2. A-B-A-A
3. A-B-A-B-A
4. A-B-A-B-B
5. A-B-B

6. B-A-A
7. B-A-B-A-A
8. B-A-B-A-B
9. B-A-B-B
10. B-B

Possibility Trees

Example 1: Possibilities for Tournament Play

b.   Assuming that all the ways of playing the tournament are equally likely, what is the probability that five games are needed to determine the tournament winner?

1.   A-A
2.   A-B-A-A
3.   A-B-A-B-A
4.   A-B-A-B-B
5.   A-B-B

6.   B-A-A
7.   B-A-B-A-A
8.   B-A-B-A-B
9.   B-A-B-B
10. B-B

Probability that 5 games are needed = 4/10 = **2/5**

18

Introduction
○ ○ ○

Possibility Trees and Multiplication Rule
○ ● ○ ○ ○ ○

Counting Elements of Disjoint Sets
○ ○ ○

The Pigeonhole Principle
○ ○ ○

The Multiplication/Product Rule

The Multiplication/Product Rule

Consider the following example.

Suppose a computer installation has four input/output units (*A*, *B*, *C*, and *D*) and three central processing units (*X*, *Y*, and *Z*).

Any input/output unit can be paired with any central processing unit. How many ways are there to pair an input/output unit with a central processing unit?

Introduction
○ ○ ○

Possibility Trees and Multiplication Rule
○ ● ○ ○ ○ ○

Counting Elements of Disjoint Sets
○ ○ ○

The Pigeonhole Principle
○ ○ ○

The Multiplication/Product Rule

Possibility tree:

The total number of ways to pair the two types of units…

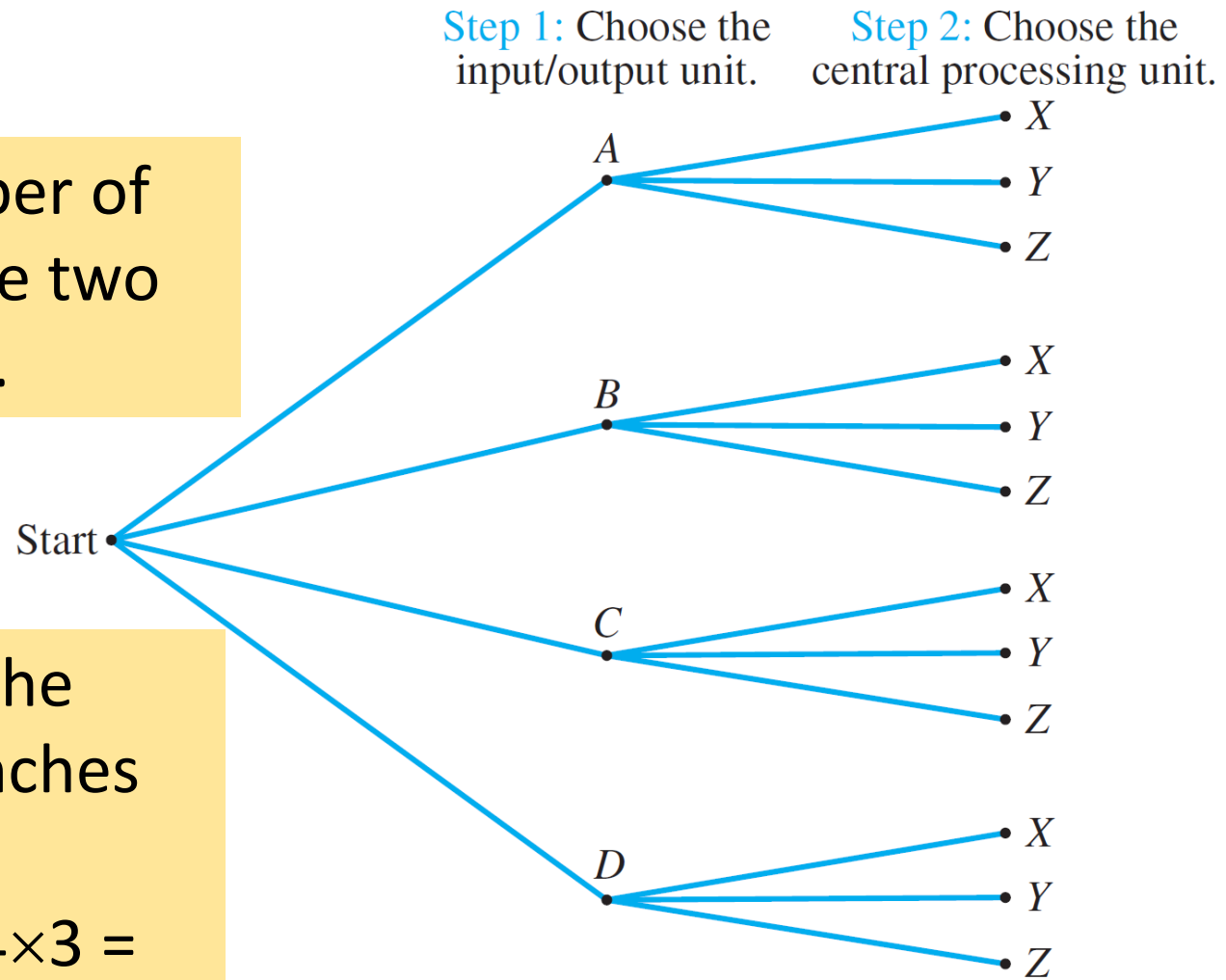is the same as the number of branches of the tree:

3 + 3 + 3 + 3 = 4×3 = **12**

Step 1: Choose the input/output unit.   Step 2: Choose the central processing unit.

Figure 9.2.2 Pairing Objects Using a Possibility Tree

20

The Multiplication/Product Rule

Theorem 9.2.1 The Multiplication/Product Rule

If an operation consists of k steps and

the first step can be performed in n_1 ways,

the second step can be performed in n_2 ways

(regardless of how the first step was performed),

:

the kth step can be performed in n_k ways

(regardless of how the preceding steps were performed),

Then the entire operation can be performed in

$$n_1 \times n_2 \times n_3 \times \ldots \times n_k \text{ ways.}$$

21

Introduction
○ ○ ○

Possibility Trees and Multiplication Rule
○ ● ○ ○ ○ ○

Counting Elements of Disjoint Sets
○ ○ ○

The Pigeonhole Principle
○ ○ ○
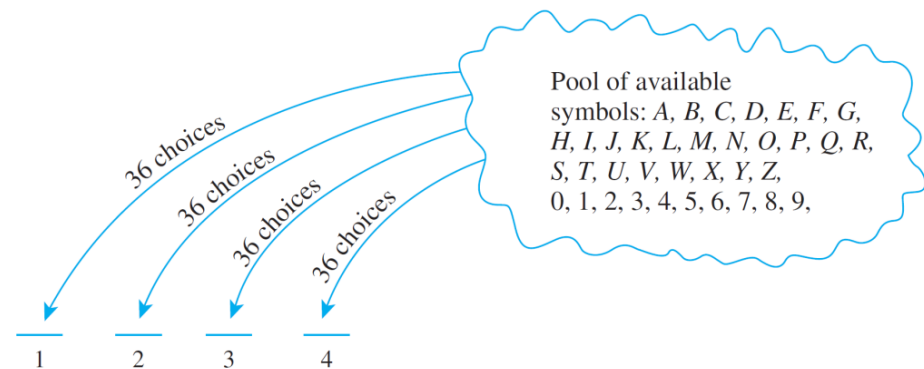
The Multiplication/Product Rule

Example 2: No. of Personal Identification Numbers (PINs)

A typical PIN is a sequence of any four symbols chosen from the 26 letters in the alphabet and the ten digits, with repetition allowed. Examples: CARE, 3387, B32B, and so forth.

How many different PINs are possible?

You can think of forming a PIN as a four-step operation to fill in each of the four symbols in sequence.

36 choices
36 choices
36 choices
36 choices

Pool of available symbols: $A, B, C, D, E, F, G, H, I, J, K, L, M, N, O, P, Q, R, S, T, U, V, W, X, Y, Z, 0, 1, 2, 3, 4, 5, 6, 7, 8, 9,$

1   2   3   4

Hence, by the **multiplication rule**, there are: $36\times36\times36\times36 = 36^4 =$ **1,679,616** PINs in all.

22

Introduction
○ ○ ○

Possibility Trees and Multiplication Rule
○ ● ○ ○ ○

Counting Elements of Disjoint Sets
○ ○ ○

The Pigeonhole Principle
○ ○ ○

The Multiplication/Product Rule

Example 3: No. of PINs without Repetition

Now, suppose that repetition is not allowed.

a. How many different PINs are there?

Step 1: Choose the 1st symbol.	36 ways
Step 2: Choose the 2nd symbol.	35 ways
Step 3: Choose the 3rd symbol.	34 ways
Step 4: Choose the 4th symbol.	33 ways

Hence, by the **multiplication rule**, there are:

$36 \times 35 \times 34 \times 33$ = **1,413,720** PINs in all with no repeated symbol.

b. If all PINs are equally likely, what is the probability that a PIN chosen at random contains no repeated symbols?

1,679,616 PINS in all.
1,413,720 PINs with no repeated symbol.

Hence, probability that a PIN chosen at random contains no repeated symbols:

$$\frac{1,413,720}{1,679,616} \cong \mathbf{0.8417}$$

23

The Multiplication/Product Rule

Power set

> ### Theorem 5.2.4 (Sets)
>
> Suppose A is a finite set. Then $|\wp(A)| = 2^{|A|}$.

Let $A = \{a_1, a_2, \cdots, a_n\}$. Then $|A| = n$.

A subset of A is formed by picking or dropping a_i from A, for $i = 1, 2, \ldots, n$.

Examples:

$$\{\, a_1 \quad a_2 \quad a_3 \quad a_4 \quad \cdots \quad a_n \,\} \implies \{a_1, a_3, a_n\}$$
$$\checkmark \quad \times \quad \checkmark \quad \times \quad \times \quad \checkmark$$

$$\{\, a_1 \quad a_2 \quad a_3 \quad a_4 \quad \cdots \quad a_n \,\} \implies \emptyset$$
$$\times \quad \times \quad \times \quad \times \quad \times \quad \times$$

Since there are 2 choices (pick or drop) for every element a_i and there are n elements, by the multiplication rule there are 2^n ways of forming a subset.

24

Introduction
○ ○ ○

Possibility Trees and Multiplication Rule
○ ○ ● ○ ○

Counting Elements of Disjoint Sets
○ ○ ○

The Pigeonhole Principle
○ ○ ○

When the Multiplication Rule is Difficult or Impossible to Apply

When the Multiplication Rule is Difficult/Impossible to Apply

Example 4: Consider the following problem:

Three officers – a president, a treasurer, and a secretary – are to be chosen from among four people: Ann, Bob, Cyd, and Dan. Suppose that, for various reasons, Ann cannot be president and either Cyd or Dan must be secretary. How many ways can the officers be chosen?

Is this correct?

It is natural to try to solve this problem using the multiplication rule. A person might answer as follows:

There are three choices for president (all except Ann), three choices for treasurer (all except the one chosen as president), and two choices for secretary (Cyd or Dan).

Therefore, by the **multiplication rule**, $3\times3\times2$ = **18**

Introduction
○○ ○

Possibility Trees and Multiplication Rule
○○ ● ○○

Counting Elements of Disjoint Sets
○○ ○

The Pigeonhole Principle
○○ ○

When the Multiplication Rule is Difficult or Impossible to Apply
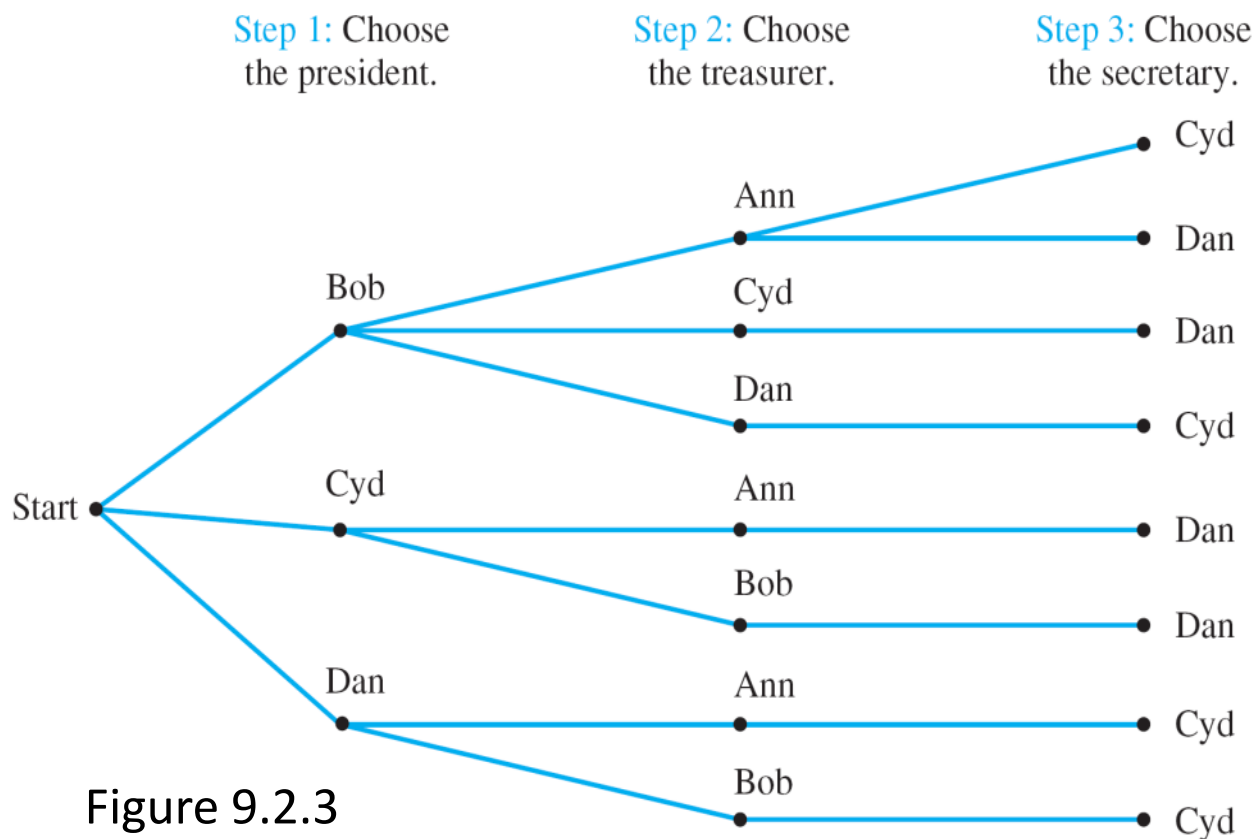
It is incorrect. The number of ways to choose the secretary varies depending on who is chosen for president and treasurer.

For instance, if Bob is chosen for president and Ann for treasurer, then there are two choices for secretary: Cyd and Dan.

But if Bob is chosen for president and Cyd for treasurer, then there is just one choice for secretary: Dan.

When the Multiplication Rule is Difficult or Impossible to Apply

The clearest way to see all the possible choices is to construct the possibility tree, as shown below.

How many ways?

8 ways.

Figure 9.2.3

Introduction
○ ○ ○

Possibility Trees and Multiplication Rule
○ ○ ● ○ ○

Counting Elements of Disjoint Sets
○ ○ ○

The Pigeonhole Principle
○ ○ ○

When the Multiplication Rule is Difficult or Impossible to Apply

Example 5: A More Subtle Use of the Multiplication Rule

Reorder the steps for choosing the officers in the previous example so that the total number of ways to choose officers can be computed using the multiplication rule.

Step 1: Choose the secretary.
Step 2: Choose the president.
Step 3: Choose the treasurer.

2 ways: Either Cyd or Dan.

2 ways: Neither Ann nor the person chosen in step 1 may be chosen, but either of the other two may.

2 ways: Either of the 2 persons not chosen in steps 1 and 2 may be chosen.

Hence, total number of ways = 2 ×2 ×2 = **8**

28

When the Multiplication Rule is Difficult or Impossible to Apply

A possibility tree illustrating this sequence of choices:

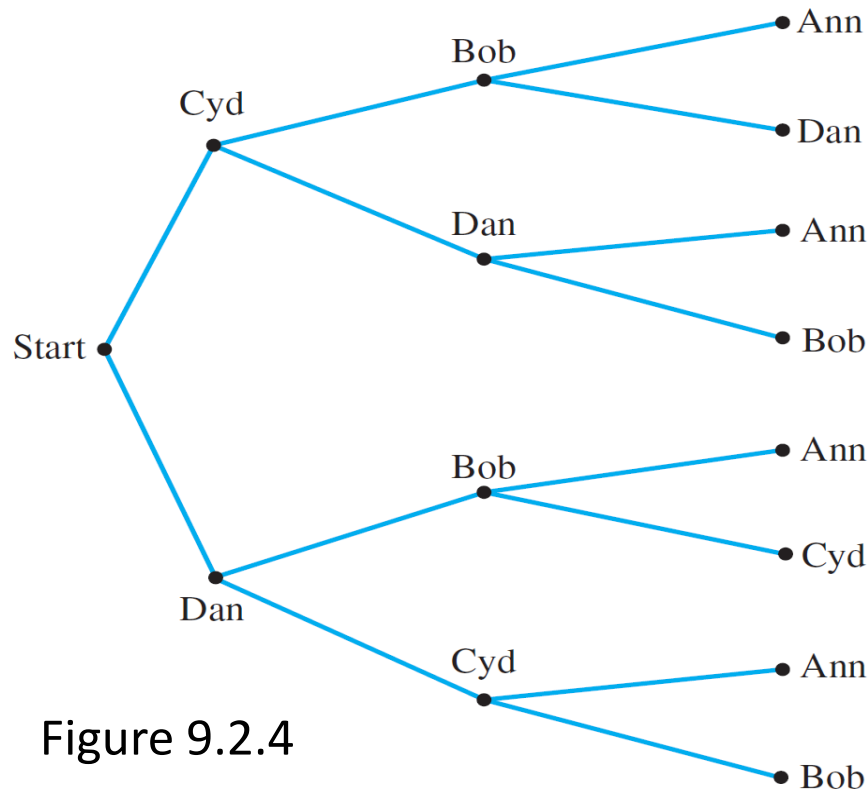Step 1: Choose the secretary.   Step 2: Choose the president.   Step 3: Choose the treasurer.

Figure 9.2.4

29

QUIZ

© Can Stock Photo

Given 26 scrabble tiles with letters 'A' to 'Z', what is the probability of drawing "ICANDOIT" if…

a. you are not allowed to return the tile after it is drawn.

0

b. you are allowed to return the tile after it is drawn.

$$(1/26)^8$$

Summary

> ## Principle of Product (or the Multiplication Principle)
>
> If there are m ways of doing something and n ways of doing another thing, then there are mn ways of performing both actions.

> ## Principle of Sum (or the Addition Principle)
>
> If we have m ways of doing something and n ways of doing another thing and we cannot do both at the same time, then there are $m+n$ ways to choose one of these actions.

Permutations

Permutations

A **permutation** of a set of objects is an ordering of the objects in a row. For example, the set of elements *a*, *b*, and *c* has six permutations.

$$abc \quad acb \quad cba \quad bac \quad bca \quad cab$$

In general, given a set of *n* objects, how many permutations does the set have?

Introduction
○ ○ ○

Possibility Trees and Multiplication Rule
○ ○ ○ ● ○

Counting Elements of Disjoint Sets
○ ○ ○

The Pigeonhole Principle
○ ○ ○

Permutations

Permutations

Imagine forming a permutation as an n-step operation:

Step 1: Choose an element to write first.   ← n ways

Step 2: Choose an element to write second.   ← $n-1$ ways

Step 3: Choose an element to write third.   ← $n-2$ ways

            :

Step n: Choose an element to write nth.   ← 1 way

By the **multiplication rule**, there are

$$n \times (n-1) \times (n-2) \times \ldots \times 2 \times 1 = \textbf{\textit{n!}}$$

ways to perform the entire operation.

Permutations

Theorem 9.2.2 Permutations

The number of permutations of a set with n ($n \geq 1$) elements is $\boldsymbol{n!}$

Note

0! is defined to be 1.
(See https://www.khanacademy.org/math/precalculus/x9e81a4f98389efdf:prob-comb/x9e81a4f98389efdf:combinatorics-precalc/v/zero-factorial-or-0)

Permutations

Example 6 – Permutations of the Letters in a Word

a.   How many ways can the letters in the word *COMPUTER* be arranged in a row?

All the eight letters in the word *COMPUTER* are distinct. Hence, **8!** = **40320**.

b.   How many ways can the letters in the word *COMPUTER* be arranged if the letters *CO* must remain next to each other (in order) as a unit?

7! = **5040**.

c.   If letters of the word *COMPUTER* are randomly arranged in a row, what is the probability that the letters *CO* remain next to each other (in order) as a unit?

$$\frac{7!}{8!} = \frac{1}{8}$$

Introduction
○ ○ ○

Possibility Trees and Multiplication Rule
○ ○ ○ ○ ●

Counting Elements of Disjoint Sets
○ ○ ○

The Pigeonhole Principle
○ ○ ○

Permutations of Selected Elements

Permutations of Selected Elements

Given the set {*a*, *b*, *c*}, there are six ways to select two letters from the set and write them in order.

$$ab \quad ac \quad ba \quad bc \quad ca \quad cb$$

Each such ordering of two elements of {*a*, *b*, *c*} is called a **2-*permutation*** of {*a*, *b*, *c*}?.

Definition

An ***r*-permutation** of a set of ***n* elements** is an ordered selection of *r* elements taken from the set.
The number of *r*-permutations of a set of *n* elements is denoted **P(*n*, *r*)**.

Other notations: $_nP_r$, P_r^n

36

Introduction
○ ○ ○

Possibility Trees and Multiplication Rule
○ ○ ○ ○ ●

Counting Elements of Disjoint Sets
○ ○ ○

The Pigeonhole Principle
○ ○ ○

Permutations of Selected Elements

Theorem 9.2.3 r-permutations from a set of n elements

If n and r are integers and $1 \leq r \leq n$, then the number of r-permutations of a set of n elements is given by the formula

$$P(n, r) = n(n - 1)(n - 2) \ldots (n - r + 1) \quad \text{first version}$$

or, equivalently,

$$P(n, r) = \frac{n!}{(n - r)!} \quad \text{second version}$$

Introduction
○ ○ ○

Possibility Trees and Multiplication Rule
○ ○ ○ ○ ●

Counting Elements of Disjoint Sets
○ ○ ○

The Pigeonhole Principle
○ ○ ○

Permutations of Selected Elements

a. Evaluate $P(5, 2)$.

$P(5, 2) = 5! / (5 − 2)! = 5 \times 4 = $ **20**

b. How many 4-permutations are there of a set of seven objects?

$P(7, 4) = 7! / (7 − 4)! = 7 \times 6 \times 5 \times 4 = $ **840**

c. How many 5-permutations are there of a set of five objects?

$P(5, 5) = 5! / (5 − 5)! = 5! = $ **120**

Introduction
○ ○ ○

Possibility Trees and Multiplication Rule
○ ○ ○ ○ ○

Counting Elements of Disjoint Sets
● ○ ○

The Pigeonhole Principle
○ ○ ○

9.3 Counting Elements of Disjoint Sets

Introduction
○ ○ ○

Possibility Trees and Multiplication Rule
○ ○ ○ ○ ○

Counting Elements of Disjoint Sets
● ○ ○

The Pigeonhole Principle
○ ○ ○

Counting Elements of Disjoint Sets: The Addition/Sum Rule

Counting Elements of Disjoint Sets: The Addition/Sum Rule

The **addition rule** (or **sum rule**) states that the number of elements in a union of mutually disjoint finite sets equals the sum of the number of elements in each of the component sets.

Theorem 9.3.1 The Addition/Sum Rule

Suppose a finite set A equals the union of k distinct mutually disjoint subsets $A_1, A_2, …, A_k$. Then

$$|A| = |A_1| + |A_2| + … + |A_k|.$$

Introduction
○ ○ ○

Possibility Trees and Multiplication Rule
○ ○ ○ ○ ○

Counting Elements of Disjoint Sets
● ○ ○

The Pigeonhole Principle
○ ○ ○ ○

Counting Elements of Disjoint Sets: The Addition/Sum Rule

Example 7 – Counting Passwords with 3 or fewer Letters

A computer access password consists of from one to three letters chosen from the 26 letters in the alphabet with repetitions allowed. How many different passwords are possible?

The set of all passwords can be partitioned into subsets consisting of those of length 1, length 2, and length 3:

Figure 9.3.1
Set of all passwords
of length ≤ 3

passwords of length 1 | passwords of length 2 | passwords of length 3

Introduction
○ ○ ○

Possibility Trees and Multiplication Rule
○ ○ ○ ○ ○

Counting Elements of Disjoint Sets
● ○ ○

The Pigeonhole Principle
○ ○ ○

Counting Elements of Disjoint Sets: The Addition/Sum Rule
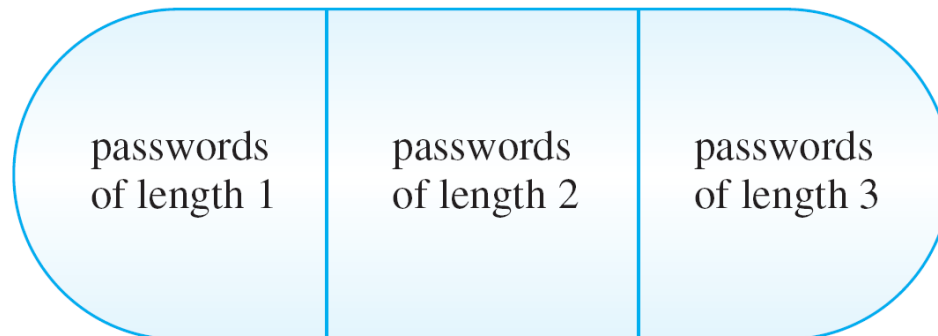
Example 7 – Counting Passwords with 3 or fewer Letters

By the **addition rule**, the total number of passwords equals the sum of the number of passwords of length 1, length 2, and length 3.

Number of passwords of length 1 = 26

Number of passwords of length 2 = 26^2

Number of passwords of length 3 = 26^3

Hence, total number of passwords = $26 + 26^2 + 26^3$

$= $ **18,278**.

The Difference Rule

The Difference Rule

An important consequence of the addition rule is the fact that if the number of elements in a set A and the number in a subset B of A are both known, then the number of elements that are in A and not in B can be computed.

Theorem 9.3.2 The Difference Rule
If A is a finite set and $B \subseteq A$, then $$\lvert A \setminus B \rvert = \lvert A \rvert - \lvert B \rvert.$$

Introduction
○ ○ ○

Possibility Trees and Multiplication Rule
○ ○ ○ ○ ○

Counting Elements of Disjoint Sets
○ ● ○

The Pigeonhole Principle
○ ○ ○

The Difference Rule
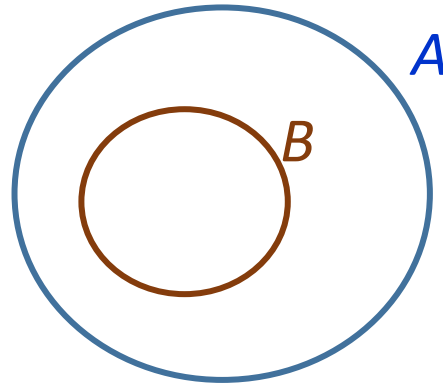
The difference rule holds for the following reason: If $B \subseteq A$, then the two sets B and $A \setminus B$ have no elements in common and $B \cup (A \setminus B) = A$.

Hence, by the addition rule,

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

Subtracting $|B|$ from both sides gives the equation

$$|A \setminus B| = |A| - |B|.$$

44

The Difference Rule

Example 8 – Counting PINs with Repeated Symbols

A typical PIN (personal identification number) is a sequence of any four symbols chosen from the 26 letters in the alphabet and the ten digits, with repetition allowed.

a.   How many PINs contain repeated symbols?

There are 36^4 = 1,679,616 PINs when repetition is allowed, and there are $36 \times 35 \times 34 \times 33$ = 1,413,720 PINs when repetition is not allowed.

By the **difference rule**, there are

$$1,679,616 - 1,413,720 = \mathbf{265,896}$$

PINS that contain at least one repeated symbol.

Introduction
○ ○ ○

Possibility Trees and Multiplication Rule
○ ○ ○ ○ ○

Counting Elements of Disjoint Sets
○ ● ○

The Pigeonhole Principle
○ ○ ○

The Difference Rule

Example 8 – Counting PINs with Repeated Symbols

b. If all PINs are equally likely, what is the probability that a randomly chosen PIN contains a repeated symbol?

There are 1,679,616 PINs in all, and by part (a) 265,896 of these contain at least one repeated symbol.

Thus, the probability that a randomly chosen PIN contains a repeated symbol is

$$\frac{265,896}{1,679,616} \cong \mathbf{0.158}$$

Introduction
○ ○ ○

Possibility Trees and Multiplication Rule
○ ○ ○ ○ ○

Counting Elements of Disjoint Sets
○ ● ○

The Pigeonhole Principle
○ ○ ○

The Difference Rule

Example 8 – Counting PINs with Repeated Symbols

An alternative solution to part (b) is based on the observation that if S is the set of all PINs and A is the set of all PINs with no repeated symbol, then $S \setminus A$ is the set of all PINs with at least one repeated symbol.

It follows that:

$$P(S \setminus A) = \frac{|S \setminus A|}{|S|} = \frac{|S| - |A|}{|S|}$$

$$= \frac{|S|}{|S|} - \frac{|A|}{|S|}$$

$$= 1 - P(A)$$

$$P(A) = \frac{1{,}413{,}720}{1{,}679{,}616}$$

$$\cong 0.842$$

$$P(S \setminus A) \cong 1 - 0.842$$

$$\cong \mathbf{0.158}$$

The Difference Rule

Probability of the Complement of an Event

This solution illustrates a more general property of probabilities: that the probability of the complement of an event is obtained by subtracting the probability of the event from the number 1.

> **Formula for the Probability of the Complement of an Event**
>
> If S is a finite sample space and A is an event in S, then
> $$P(\bar{A}) = 1 - P(A)$$

The Inclusion/Exclusion Rule

The Inclusion/Exclusion Rule

The addition rule says how many elements are in a union of sets <u>if the sets are mutually disjoint</u>. Now consider the question of how to determine the number of elements in a union of sets when **some of the sets overlap**.

For simplicity, begin by looking at a union of two sets *A* and *B*, as shown in Figure 9.3.5.

Figure 9.3.5

49
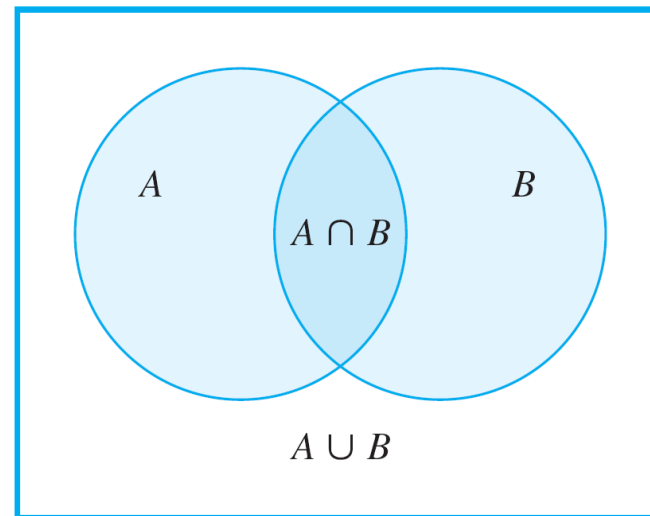
The Inclusion/Exclusion Rule

To get an accurate count of the elements in $A \cup B$, it is necessary to subtract the number of elements that are in both A and B, i.e. $A \cap B$,

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

Theorem 9.3.3 The Inclusion/Exclusion Rule for 2 or 3 Sets

If A, B, and C are any finite sets, then
$$|A \cup B| = |A| + |B| - |A \cap B|$$
and
$$|A \cup B \cup C| = |A| + |B| + |C| - |A \cap B|$$
$$- |A \cap C| - |B \cap C| + |A \cap B \cap C|$$

The Inclusion/Exclusion Rule

Example 9 – Counting Elements of a General Union

a. How many integers from 1 through 1,000 are multiples of 3 or multiples of 5?

Let A = the set of all integers in [1, 1000] that are multiples of 3.
Let B = the set of all integers in [1, 1000] that are multiples of 5.

Then $A \cup B$ = the set of all integers in [1, 1000] that are multiples of 3 or 5.

Then $A \cap B$ = the set of all integers in [1, 1000] that are multiples of both 3 and 5
 = the set of all integers in [1, 1000] that are multiples of 15.

Every multiple of 3 in [1, 1000] has the form $3k$, for some integer k in [1, 333].

Hence, $|A|$ = 333

Every multiple of 5 in [1, 1000] has the form $5k$, for some integer k in [1, 200].

Hence, $|B|$ = 200

Every multiple of 15 in [1, 1000] has the form $15k$, for some integer k in [1, 66].

Hence, $|A \cap B|$ = 66

$$|A \cup B| = |A| + |B| - |A \cap B|$$
$$= 333 + 200 - 66 = 467$$

Thus, **467** integers from 1 through 1000 are multiples of 3 or multiples of 5.

51

The Inclusion/Exclusion Rule

b. How many integers from 1 through 1,000 are neither multiples of 3 nor multiples of 5?

There are 1000 integers in [1, 1000].
By part (a), 467 of these are multiples of 3 or multiples of 5.

Thus, by the **difference rule**, there are 1000 – 467 = **533** that are neither multiples of 3 nor multiples of 5.

Note that the solution to part (b) hid a use of De Morgan's law.

The number of elements that are neither in A nor in B is $|\bar{A} \cap \bar{B}|$.
By De Morgan's law, $\bar{A} \cap \bar{B} = \overline{(A \cup B)}$.

So $|\overline{A \cup B}|$ was then calculated using the set difference rule:
$$|\overline{A \cup B}| = |U| - |A \cup B|$$
where the universe U was the set of all integers in [1, 1000].

Introduction
○ ○ ○

Possibility Trees and Multiplication Rule
○ ○ ○ ○ ○

Counting Elements of Disjoint Sets
○ ○ ○

The Pigeonhole Principle
● ○ ○

9.4 The Pigeonhole Principle (PHP)

The Pigeonhole Principle: Introduction

The Pigeonhole Principle: Introduction

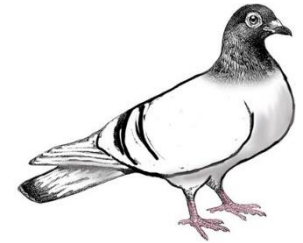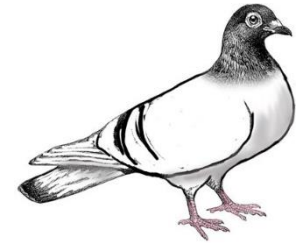If n pigeons fly into m pigeonholes and $n > m$, then at least one hole must contain two or more pigeons.
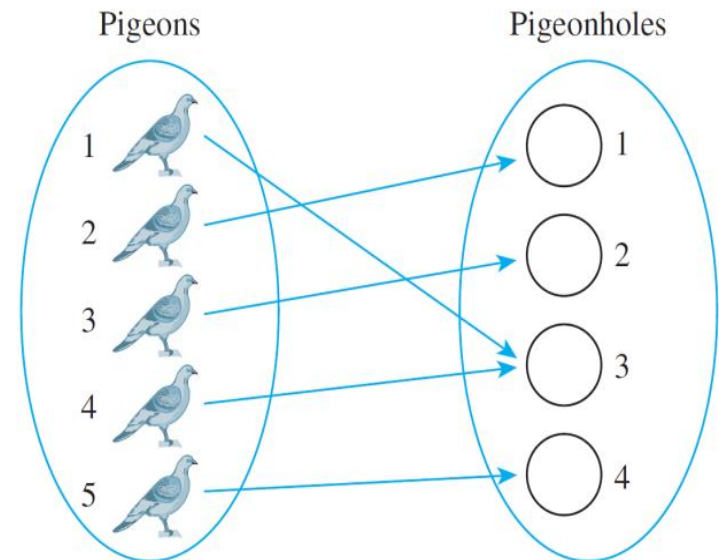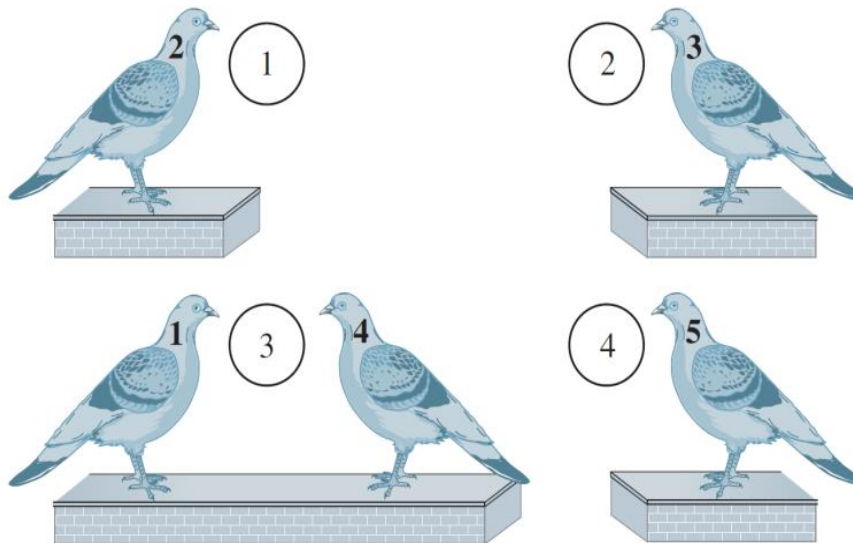
Figure 9.4.1 $n = 5$ and $m = 4$
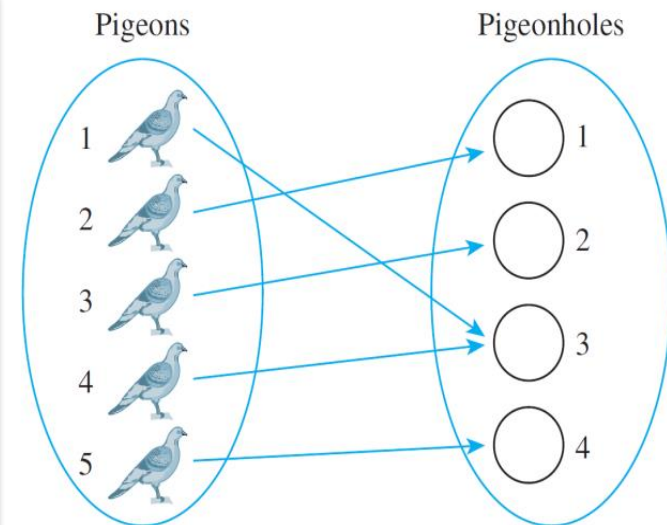
The Pigeonhole Principle: Introduction

The pigeonhole principle is sometimes called the *Dirichlet box principle* because it was first stated formally by J. P. G. L. Dirichlet (1805–1859).

Mathematical formulation:

Pigeonhole Principle (PHP)

A function from one finite set to a smaller finite set cannot be one-to-one: There must be at least 2 elements in the domain that have the same image in the co-domain.

Epp uses the terms one-to-one for injective, and onto for surjective.
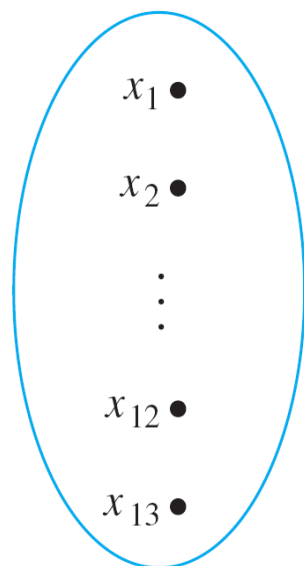
The Pigeonhole Principle: Introduction

Example 10 – Applying the Pigeonhole Principle

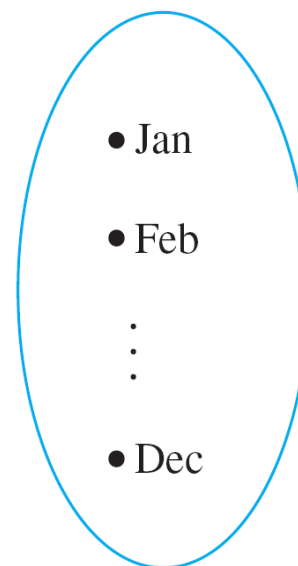a. In a group of six people, must there be at least two who were born in the same month? No.

In a group of 13 people, must there be at least two who were born in the same month? Why?

13 people (pigeons)   12 months (pigeonholes)

x_1 •

x_2 •

B

$B(x_i)$ = birth month of x_i

• Jan

• Feb

⋮

x_{12} •

x_{13} •

• Dec

Yes. At least 2 people must have been born in the same month.

56

The Pigeonhole Principle: Introduction

Example 10 – Applying the Pigeonhole Principle

b. Among the population of Singapore, are there at least two people with the same number of hairs on their heads? Why?

Population of Singapore: 5.47m (June 2014).

Hairs on head: average up to 150,000; no more than 300,000.

Define a function H from the set of people in Singapore $\{x_1, x_2, \ldots x_p\}$ to the set $\{0, 1, 2, \ldots 300000\}$ as shown in the next slide.
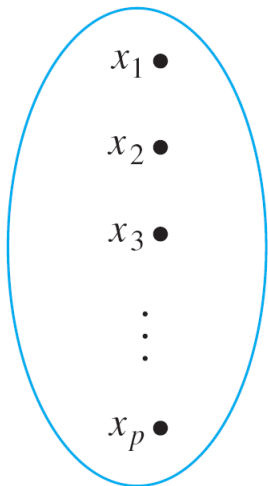
The Pigeonhole Principle: Introduction
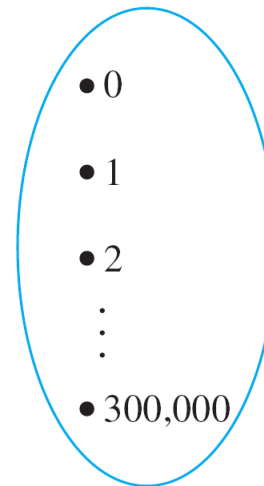
Example 10 – Applying the Pigeonhole Principle

b. Among the population of Singapore, are there at least two people with the same number of hairs on their heads? Why?

People in Singapore
(pigeons)

Possible number of hairs on
a person's head (pigeonholes)

H

$H(x_i) =$ the number of
hairs on x_i's head

$x_1 \bullet$

$x_2 \bullet$

$x_3 \bullet$

$\vdots$

$x_p \bullet$

$\bullet \, 0$

$\bullet \, 1$

$\bullet \, 2$

$\vdots$

$\bullet \, 300{,}000$

Yes. At least 2 people must have the same number of hairs on their heads.

58

Introduction
○ ○ ○

Possibility Trees and Multiplication Rule
○ ○ ○ ○ ○

Counting Elements of Disjoint Sets
○ ○ ○

The Pigeonhole Principle
○ ● ○

Application to Decimal Expansions of Fractions

Application to Decimal Expansions of Fractions

One important consequence of the pigeonhole principle is the fact that *the decimal expansion of any rational number either terminates or repeats*.

A terminating decimal: 3.625

A repeating decimal: $2.38\overline{246}$ (2.38246246246246…)

A rational number can be written as a fraction *a/b, b*≠0.

When one integer is divided by another, it is the pigeonhole principle (together with the quotient-remainder theorem) that guarantees that such a repetition of remainders (and hence decimal digits) must always occur.

59

Introduction
○ ○ ○

Possibility Trees and Multiplication Rule
○ ○ ○ ○ ○

Counting Elements of Disjoint Sets
○ ○ ○

The Pigeonhole Principle
○ ● ○

Application to Decimal Expansions of Fractions

Consider a/b, where for simplicity assume that a and b are positive. The decimal expansion of a/b is obtained by dividing a by b as illustrated here for a = 3 and b = 14.

Let r_0 = a and let r_1, r_2, r_3, . . . be the successive remainders obtained in the long division of a by b.

By the quotient-remainder theorem (Theorem 8.1.16: Division Theorem), each remainder must be between 0 and $b-1$. (Here, b is 14, and so the remainders are from 0 to 13.)

If some remainder r_i = 0, then the division terminates and a/b has a terminating decimal expansion. If no r_i = 0, then the division process and hence the sequence of remainders continues forever.
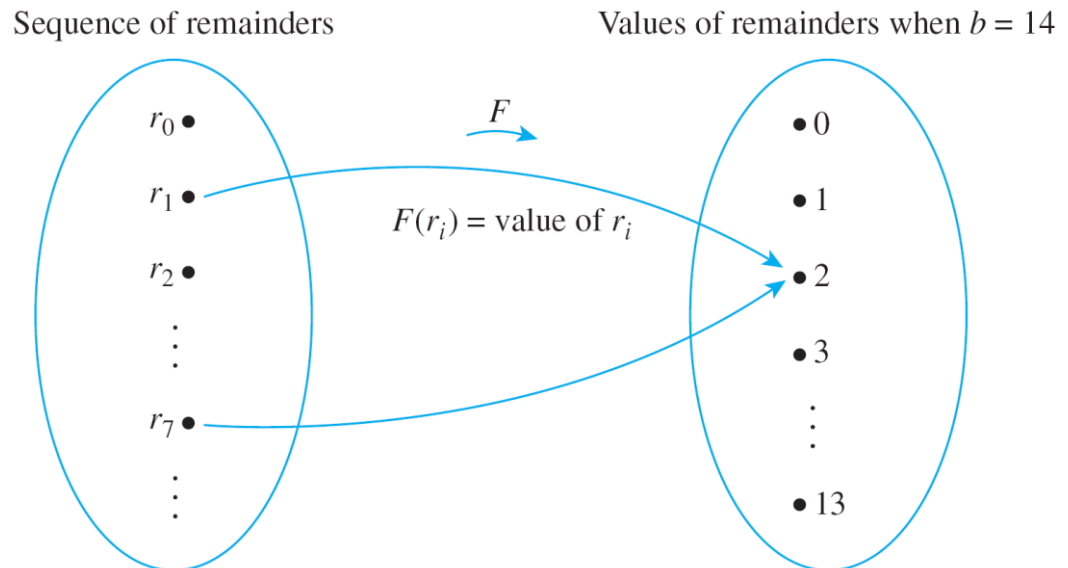
$$.2\,1\,4\,2\,8\,5\,7\,1\,4\,2\,8\,5\,7\ldots$$

$$14\,|\,\overline{③0\,0\,0\,0\,0\,0\,0\,0\,0\,0\,0\,0\,0\,0\,0\,0}$$

2 8	$r_0 = 3$
②0	$r_1 = 2$
1 4	
⑥0	$r_2 = 6$
5 6	
④0	$r_3 = 4$
2 8	
①②0	$r_4 = 12$
1 1 2	
⑧0	$r_5 = 8$
7 0	
①⓪0	$r_6 = 10$
9 8	
②0	$r_7 = 2 = r_1$
1 4	
⑥0	$r_8 = 6 = r_2$
5 6	
④0	$r_9 = 4 = r_3$
⋮	⋮

Application to Decimal Expansions of Fractions

By the PHP, since there are more remainders than values that the remainders can take, some remainder value must repeat: $r_j = r_k$, for some indices j and k with $j < k$.

It follows that the decimal digits obtained from the divisions between r_j and r_{k-1} repeat forever.

Sequence of remainders

Values of remainders when $b = 14$

F

$F(r_i)$ = value of r_i

r_0

r_1

r_2

r_7

0

1

2

3

13

In the case of 3/14, the repetition begins with $r_7 = 2 = r_1$ and the decimal expansion repeats the quotients obtained from the divisions from r_1 through r_6 forever:

$$3/14 = 0.2\overline{142857}$$

61

Introduction

○ ○ ○

Possibility Trees and Multiplication Rule

○ ○ ○ ○ ○

Counting Elements of Disjoint Sets

○ ○ ○

The Pigeonhole Principle

○ ○ ●

Generalized Pigeonhole Principle

Generalized Pigeonhole Principle

If *n* pigeons fly into *m* pigeonholes and, for some positive integer *k*, *k* < *n/m*, then at least one pigeonhole contains *k* + 1 or more pigeons.

Generalized Pigeonhole Principle

For any function *f* from a finite set *X* with *n* elements to a finite set *Y* with *m* elements and for any positive integer *k*, if *k* < *n/m*, then there is some *y* ∈ *Y* such that *y* is the image of at least *k* + 1 distinct elements of *X*.

Generalized Pigeonhole Principle

Example 11 – Applying the General Pigeonhole Principle

Show how the generalized pigeonhole principle implies that in a group of 85 people, at least 4 must have the same last initial ('A', 'B', …, 'Z').
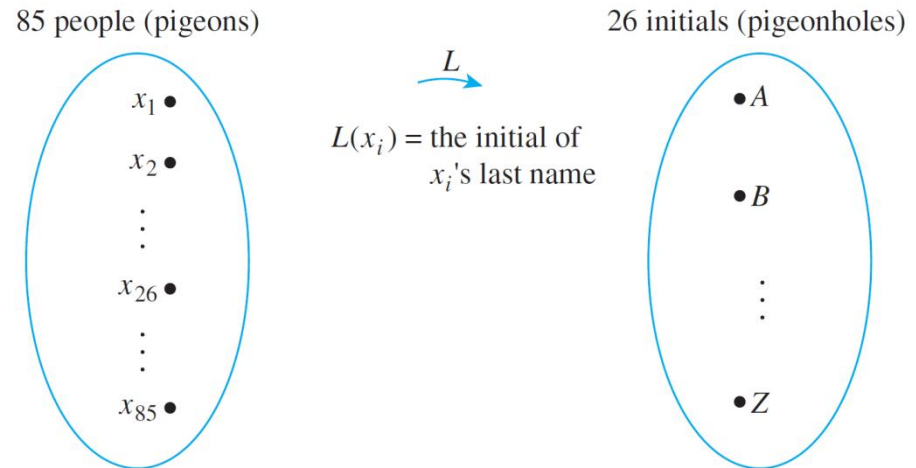
In this example the pigeons are the 85 people and the pigeonholes are the 26 possible last initials of their names. Note that

$$3 < 85/26 \cong 3.27$$

Generalized Pigeonhole Principle

Example 11 – Applying the General Pigeonhole Principle

Consider the function *L* from people to initials defined by the following arrow diagram.

85 people (pigeons)

26 initials (pigeonholes)

L

$L(x_i) =$ the initial of x_i's last name

$x_1 \bullet$

$x_2 \bullet$

$\vdots$

$x_{26} \bullet$

$\vdots$

$x_{85} \bullet$

$\bullet A$

$\bullet B$

$\vdots$

$\bullet Z$

Since 3 < 85/26, the generalized pigeonhole principle states that some initial must be the image of at least four (3 + 1) people.

Thus, at least 4 people have the same last initial. ∎

Introduction
○ ○ ○

Possibility Trees and Multiplication Rule
○ ○ ○ ○ ○

Counting Elements of Disjoint Sets
○ ○ ○

The Pigeonhole Principle
○ ○ ● ○

Generalized Pigeonhole Principle

Consider the following contrapositive form of the generalized pigeonhole principle. You may find it natural to use the contrapositive form of the generalized pigeonhole principle in certain situations.

Generalized Pigeonhole Principle (Contrapositive Form)

For any function f from a finite set X with n elements to a finite set Y with m elements and for any positive integer k, if for each $y \in Y$, $f^{-1}(\{y\})$ has at most k elements, then X has at most km elements; in other words, $n \leq km$.

Generalized Pigeonhole Principle

For instance, the result of Example 11 can be explained as follows:

Suppose no 4 people out of the 85 had the same last initial. Then at most 3 would share any particular one.

By the generalized pigeonhole principle (contrapositive form), this would imply that the total number of people is at most $3 \times 26 = 78$. But this contradicts the fact that there are 85 people in all.

Hence at least 4 people share a last initial.

Introduction
○ ○ ○

Possibility Trees and Multiplication Rule
○ ○ ○ ○ ○

Counting Elements of Disjoint Sets
○ ○ ○

The Pigeonhole Principle
○ ○ ●

Generalized Pigeonhole Principle

Example 12 – Using the General Pigeonhole Principle (Contrapositive)

There are 42 students who are to share 12 computers. Each student uses exactly 1 computer, and no computer is used by more than 6 students. Show that at least 5 computers are used by 3 or more students.

Using an Argument by Contradiction: Suppose not. Suppose that 4 or fewer computers are used by 3 or more students. *[A contradiction will be derived.]* Then 8 or more computers are used by 2 or fewer students.

Divide the set of computers into two subsets: C_1 and C_2.

Generalized Pigeonhole Principle

Into C_1 place 8 of the computers used by 2 or fewer students; into C_2 place the computers used by 3 or more students plus any remaining computers (to make a total of 4 computers in C_2).
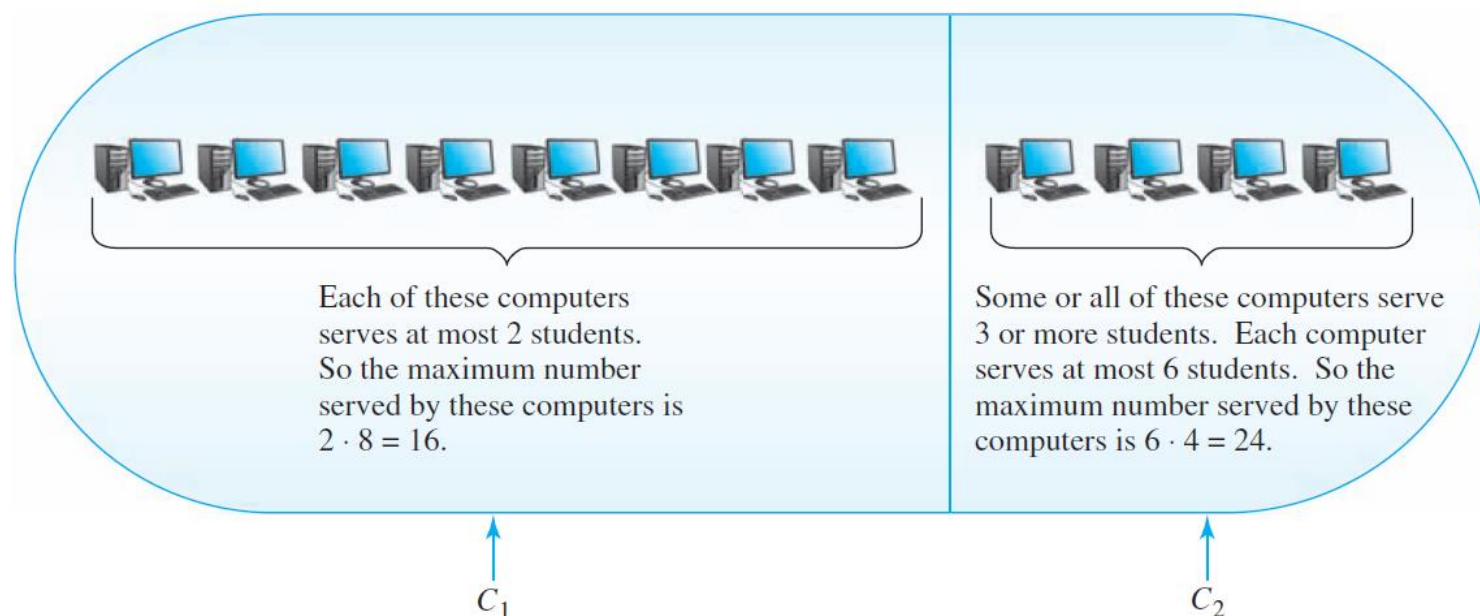

Each of these computers serves at most 2 students. So the maximum number served by these computers is $2 \cdot 8 = 16$.

Some or all of these computers serve 3 or more students. Each computer serves at most 6 students. So the maximum number served by these computers is $6 \cdot 4 = 24$.

C_1 C_2

Figure 9.4.3 The set of 12 computers

68

Generalized Pigeonhole Principle

Since at most 2 students are served by any one computer in C_1, by the generalized pigeonhole principle (contrapositive form), the computers in set C_1 serve at most 2 × 8 = 16 students.

Since at most 6 students are served by any one computer in C_2, by the generalized pigeonhole principle (contrapositive form), the computers in set C_2 serve at most 6 × 4 = 24 students.

Total number of students served by the computers is at most 24 + 16 = 40.

This contradicts that there are 42 students, each served by a computer.
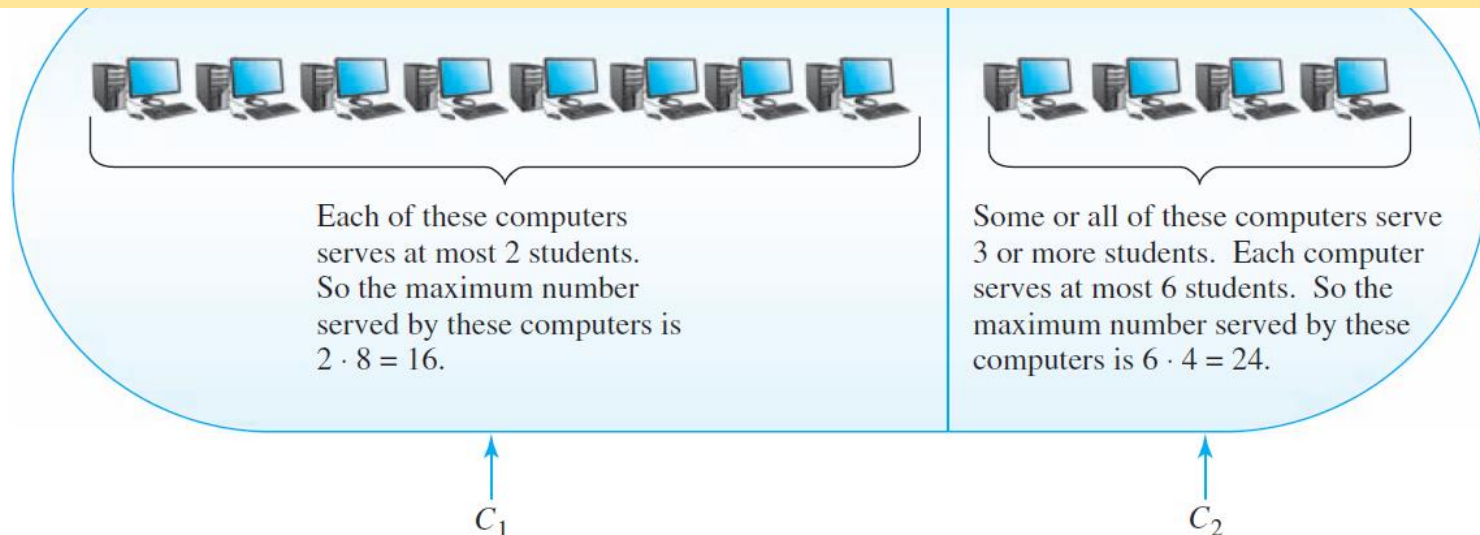
Therefore, the supposition is false.

Each of these computers serves at most 2 students. So the maximum number served by these computers is 2 · 8 = 16.

Some or all of these computers serve 3 or more students. Each computer serves at most 6 students. So the maximum number served by these computers is 6 · 4 = 24.

C_1

C_2

Figure 9.4.3 The set of 12 computers

69

Introduction

○ ○ ○

Possibility Trees and Multiplication Rule

○ ○ ○ ○ ○

Counting Elements of Disjoint Sets

○ ○ ○

The Pigeonhole Principle

○ ○ ●

Generalized Pigeonhole Principle

Using a Direct Argument: Let k be the number of computers used by 3 or more students. *[We must show that $k \geq 5$.]*

Because each computer is used by at most 6 students, these computers are used by at most 6k students (by the generalized PHP (contrapositive form)).

Each of the remaining 12 − k computers is used by at most 2 students.

Taken together, they are used by at most 2(12 − k) = 24 − 2k students (again, by the generalized PHP).

Thus the maximum number of students served by the computers is 6k + (24 − 2k) = 4k + 24.

Because 42 students are served by the computers, 4k + 24 ≥ 42.

Solving for k gives $k \geq 4.5$, and since k is an integer, this implies that $k \geq 5$.

70

Next week's lectures

More on Counting!

- Combinations
- *r*-Combinations
- Pascal's Formula and the Binomial Theorem
- Probability Axioms and Expected Value
- Conditional Probability, Bayes' Formula, and Independent Events

END OF FILE