BBM 205 Discrete Mathematics
Hacettepe University
http://web.cs.hacettepe.edu.tr/∼bbm205

Lecture 4: Counting, Pigeonhole Principle, Permutations, Combinations

Lecturer: Lale Özkahya

Resources:
Kenneth Rosen, "Discrete Mathematics and App."
http://www.inf.ed.ac.uk/teaching/courses/dmmr

Chapter Summary

- The Basics of Counting
- The Pigeonhole Principle
- Permutations and Combinations
- Binomial Coefficients and Identities
- Generalized Permutations and Combinations

Basic Counting: The Product Rule

Recall: For a set A, $|A|$ is the cardinality of A (# of elements of A).

For a pair of sets A and B, $A \times B$ denotes their cartesian product:

$$A \times B = \{(a, b) \mid a \in A \ \wedge b \in B\}$$

Product Rule

If A and B are finite sets, then: $|A \times B| = |A| \cdot |B|$.

Proof: Obvious, but prove it yourself by induction on |A|. □

Basic Counting: The Product Rule

Recall: For a set A, $|A|$ is the cardinality of A (# of elements of A).

For a pair of sets A and B, $A \times B$ denotes their cartesian product:

$$A \times B = \{(a, b) \mid a \in A \ \wedge b \in B\}$$

Product Rule

If A and B are finite sets, then: $|A \times B| = |A| \cdot |B|$.

Proof: Obvious, but prove it yourself by induction on |A|. □

general Product Rule

If $A_1, A_2, \ldots, A_m$ are finite sets, then

$$|A_1 \times A_2 \times \ldots \times A_m| = |A_1| \cdot |A_2| \cdot \ldots \cdot |A_m|$$

Proof: By induction on m, using the (basic) product rule. □

Product Rule: examples

Example 1: How many bit strings of length seven are there?

Product Rule: examples

Example 1: How many bit strings of length seven are there?

Solution: Since each bit is either 0 or 1, applying the product rule, the answer is $2^7 = 128$. $\square$

Product Rule: examples

Example 1: How many bit strings of length seven are there?

Solution: Since each bit is either 0 or 1, applying the product rule, the answer is $2^7 = 128$. $\square$

Example 2: How many different car license plates can be made if each plate contains a sequence of three uppercase English letters followed by three digits?

Product Rule: examples

Example 1: How many bit strings of length seven are there?

Solution: Since each bit is either 0 or 1, applying the product rule, the answer is $2^7 = 128$. □

Example 2: How many different car license plates can be made if each plate contains a sequence of three uppercase English letters followed by three digits?

Solution: $26 \cdot 26 \cdot 26 \cdot 10 \cdot 10 \cdot 10 = 17{,}576{,}000$. □

Counting Subsets

Number of Subsets of a Finite Set

A finite set, S, has $2^{|S|}$ distinct subsets.

Counting Subsets

Number of Subsets of a Finite Set

A finite set, S, has $2^{|S|}$ distinct subsets.

Proof: Suppose $S = \{s_1, s_2, \ldots, s_m\}$.
There is a one-to-one correspondence (bijection), between subsets of S and bit strings of length $m = |S|$.
The bit string of length $|S|$ we associate with a subset $A \subseteq S$ has a 1 in position i if $s_i \in A$, and 0 in position i if $s_i \notin A$, for all $i \in \{1, \ldots, m\}$.

$$\{s_2, s_4, s_5, \ldots, s_m\} \quad \cong \quad \underbrace{\boxed{0}\boxed{1}\boxed{0}\boxed{1}\boxed{1}\boxed{\ldots}\boxed{1}}_{m}$$

By the product rule, there are $2^{|S|}$ such bit strings. $\square$

Counting Functions

Number of Functions

For all finite sets A and B, the number of distinct functions, $f : A \to B$, mapping A to B is:

$$|B|^{|A|}$$

Proof: Suppose $A = \{a_1, \ldots, a_m\}$.
There is a one-to-one correspondence between functions $f : A \to B$ and strings (sequences) of length $m = |A|$ over an alphabet of size $n = |B|$:

$$(f : A \to B) \quad \cong \quad \begin{array}{|c|c|c|c|c|} \hline f(a_1) & f(a_2) & f(a_3) & \ldots & f(a_m) \\ \hline \end{array}$$

By the product rule, there are n^m such strings of length m. $\square$

Sum Rule

Sum Rule

If A and B are finite sets that are disjoint (meaning $A \cap B = \emptyset$), then

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

Proof. Obvious. (If you must, prove it yourself by induction on $|A|$.) $\square$

general Sum Rule

If $A_1, \ldots, A_m$ are finite sets that are pairwise disjoint, meaning $A_i \cap A_j = \emptyset$, for all $i, j \in \{1, \ldots, m\}$, then

$$|A_1 \cup A_2 \cup \ldots \cup A_m| = |A_1| + |A_2| + \ldots + |A_m|$$

Sum Rule: Examples

Example 1: Suppose variable names in a programming language can be either a single uppercase letter or an uppercase letter followed by a digit. Find the number of possible variable names.

Sum Rule: Examples

Example 1: Suppose variable names in a programming language can be either a single uppercase letter or an uppercase letter followed by a digit. Find the number of possible variable names.

Solution: Use the sum and product rules: $26 + 26 \cdot 10 = 286$. $\square$

Sum Rule: Examples

Example 1: Suppose variable names in a programming language can be either a single uppercase letter or an uppercase letter followed by a digit. Find the number of possible variable names.

Solution: Use the sum and product rules: $26 + 26 \cdot 10 = 286$. $\square$

Example 2: Each user on a computer system has a password which must be six to eight characters long.
Each character is an uppercase letter or digit.
Each password must contain at least one digit.
How many possible passwords are there?

Sum Rule: Examples

Example 1: Suppose variable names in a programming language can be either a single uppercase letter or an uppercase letter followed by a digit. Find the number of possible variable names.

Solution: Use the sum and product rules: $26 + 26 \cdot 10 = 286$. $\square$

Example 2: Each user on a computer system has a password which must be six to eight characters long.
Each character is an uppercase letter or digit.
Each password must contain at least one digit.
How many possible passwords are there?

Solution: Let P be the total number of passwords, and let P_6, P_7, P_8 be the number of passwords of lengths 6, 7, and 8, respectively.

- By the sum rule $P = P_6 + P_7 + P_8$.
- $P_6 = 36^6 - 26^6$; $P_7 = 36^7 - 26^7$; $P_8 = 36^8 - 26^8$.
- So, $P = P_6 + P_7 + P_8 = \sum_{i=6}^{8}(36^i - 26^i)$. $\square$

Subtraction Rule (Inclusion-Exclusion for two sets)

Subtraction Rule

For any finite sets A and B (not necessarily disjoint),

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

Proof: Venn Diagram:

|A| + |B| overcounts (twice) exactly those elements in $A \cap B$. □

Subtraction Rule: Example

Example: How many bit strings of length 8 either start with a 1 bit or end with the two bits 00?

Subtraction Rule: Example

Example: How many bit strings of length 8 either start with a 1 bit or end with the two bits 00?

Solution:

- Number of bit strings of length 8 that start with 1: $2^7 = 128$.
- Number of bit strings of length 8 that end with 00: $2^6 = 64$.
- Number of bit strings of length 8 that start with 1 <u>and</u> end with 00: $2^5 = 32$.

Applying the subtraction rule, the number is $128 + 64 - 32 = 160$. $\square$

The Pigeonhole Principle

Pigeonhole Principle

For any positive integer k, if $k+1$ objects (pigeons) are placed in k boxes (pigeonholes), then at least one box contains two or more objects.

Proof: Suppose no box has more than 1 object. Sum up the number of objects in the k boxes. There can't be more than k. Contradiction. $\square$

Pigeonhole Principle (rephrased more formally)

If a function $f : A \to B$ maps a finite set A with $|A| = k+1$ to a finite set B, with $|B| = k$, then f is not one-to-one.

(**Recall:** a function $f : A \to B$ is called **one-to-one** if $\forall a_1, a_2 \in A$, if $a_1 \neq a_2$ then $f(a_1) \neq f(a_2)$.)

Pigeonhole Principle: Examples

Example 1: At least two students registered for this course will receive exactly the same final exam mark. Why?

Pigeonhole Principle: Examples

Example 1: At least two students registered for this course will receive exactly the same final exam mark. Why?

Reason: There are at least 102 students registered for DMMR (suppose the actual number is 145), so, at least 102 objects. Final exam marks are integers in the range 0-100 (so, exactly 101 boxes). □

Generalized Pigeonhole Principle

Generalized Pigeonhole Principle (GPP)

If $N \geq 0$ objects are placed in $k \geq 1$ boxes, then at least one box contains at least $\left\lceil \frac{N}{k} \right\rceil$ objects.

Generalized Pigeonhole Principle

Generalized Pigeonhole Principle (GPP)

If $N \geq 0$ objects are placed in $k \geq 1$ boxes, then at least one box contains at least $\lceil \frac{N}{k} \rceil$ objects.

Proof: Suppose no box has more than $\lceil \frac{N}{k} \rceil - 1$ objects. Sum up the number of objects in the k boxes. It is at most

$$k \cdot (\left\lceil \frac{N}{k} \right\rceil - 1) < k \cdot ((\frac{N}{k} + 1) - 1) = N$$

Thus, there must be fewer than N. Contradiction.
(We are using the fact that $\lceil \frac{N}{k} \rceil < \frac{N}{k} + 1$.) $\square$

Exercise: Rephrase GPP as a statement about functions $f : A \to B$ that map a finite set A with $|A| = N$ to a finite set B, with $|B| = k$.

Generalized Pigeonhole Principle: Examples

Example 1: Consider the following statement:

"At least d students in this course were born in the same month." (1)

Suppose the actual number of students registered for DMMR is 145. What is the maximum number d for which it is certain that statement (1) is true?

Generalized Pigeonhole Principle: Examples

Example 1: Consider the following statement:

"At least d students in this course were born in the same month." (1)

Suppose the actual number of students registered for DMMR is 145. What is the maximum number d for which it is certain that statement (1) is true?

Solution: Since we are assuming there are 145 registered students in DMMR.
$\lceil \frac{145}{12} \rceil = 13$, so by GPP we know statement (1) is true for $d = 13$.
Statement (1) need not be true for $d = 14$, because if 145 students are distributed *as evenly as possible* into 12 months, the maximum number of students in any month is 13, with other months having only 12. □

Generalized Pigeonhole Principle: Examples

Example 1: Consider the following statement:

"At least d students in this course were born in the same month." (1)

Suppose the actual number of students registered for DMMR is 145. What is the maximum number d for which it is certain that statement (1) is true?

Solution: Since we are assuming there are 145 registered students in DMMR.
$\lceil \frac{145}{12} \rceil = 13$, so by GPP we know statement (1) is true for $d = 13$.
Statement (1) need not be true for $d = 14$, because if 145 students are distributed *as evenly as possible* into 12 months, the maximum number of students in any month is 13, with other months having only 12. □

(In **probability theory** you will learn that nevertheless it is highly probable, assuming birthdays are randomly distributed, that at least 14 of you (and more) were indeed born in the same month.)

GPP: more Examples

Example 2: How many cards must be selected from a standard deck of 52 cards to guarantee that at least thee cards of the same suit are chosen?

GPP: more Examples

Example 2: How many cards must be selected from a standard deck of 52 cards to guarantee that at least thee cards of the same suit are chosen?

Solution: There are 4 suits. (In a standard deck of 52 cards, every card has exactly one suit. There are no jokers.) So, we need to choose N cards, such that $\lceil \frac{N}{4} \rceil \geq 3$. The smallest integer N such that $\lceil \frac{N}{4} \rceil \geq 3$ is $2 \cdot 4 + 1 = 9$. $\square$

Permutations

Permutation

A **permutation** of a set S is an ordered arrangement of the elements of S.

In other words, it is a sequence containing every element of S exactly once.

Example: Consider the set $S = \{1, 2, 3\}$.

The sequence $(3, 1, 2)$ is one permutation of S.

There are 6 different permutations of S. They are:

$$(1, 2, 3)\ ,\ (1, 3, 2)\ ,\ (2, 1, 3)\ ,\ (2, 3, 1)\ ,\ (3, 1, 2)\ ,\ (3, 2, 1)$$

Permutations (an alternative view)

A permutation of a set S can alternatively be viewed as a bijection (a one-to-one and onto function), $\pi : S \to S$, from S to itself.

Specifically, if the finite set is $S = \{s_1, \ldots, s_m\}$, then by fixing the ordering $s_1, \ldots, s_m$, we can uniquely associate to each bijection $\pi : S \to S$ a sequence ordering $\{s_1, \ldots, s_m\}$ as follows:

$$(\pi : S \to S) \quad \cong \quad \begin{array}{|c|c|c|c|c|} \hline \pi(s_1) & \pi(s_2) & \pi(s_3) & \ldots & \pi(s_m) \\ \hline \end{array}$$

Note that π is a bijection if and only if the sequence on the right containing every element of S exactly once.

r-Permutation

r-Permutation

An r**-permutation** of a set S, is an ordered arrangement (sequence) of r distinct elements of S.

(For this to be well-defined, r needs to be an integer with $0 \leq r \leq |S|$.)

Examples:

There is only one 0-permutation of any set: the empty sequence ().

For the set $S = \{1, 2, 3\}$, the sequence $(3, 1)$ is a 2-permutation.

$(3, 2, 1)$ is both a permutation and 3-permutation of S (since $|S| = 3$).

There are 6 different different 2-permutations of S. They are:

$$(1, 2)\ ,\ (1, 3)\ ,\ (2, 1)\ ,\ (2, 3)\ ,\ (3, 1)\ ,\ (3, 2)$$

Question: How many r-permutations of an n-element set are there?

r-Permutations (an alternative view)

An r-permutation of a set S, with $1 \leq r \leq |S|$, can alternatively be viewed as a one-to-one function, $f : \{1, \ldots, r\} \to S$.

Specifically, we can uniquely associate to each one-to-one function $f : \{1, \ldots, r\} \to S$, an r-permutation of S as follows:

$$(f : \{1, \ldots, r\} \to S) \quad \cong \quad \begin{array}{|c|c|c|c|c|} \hline f(1) & f(2) & f(3) & \ldots & f(r) \\ \hline \end{array}$$

Note that f is one-to-one if and only if the sequence on the right is an r-permutation of S.

So, for a set S with $|S| = n$, the number of r-permutions of S, $1 \leq r \leq n$, is equal to the number of one-to-one functions:

$$f : \{1, \ldots, r\} \to \{1, \ldots, n\}$$

Formula for # of permutations, and # of r-permutations

Let $\mathbf{P}(n, r)$ denote the number of r-permutations of an n-element set.

$P(n, 0) = 1$, because the only 0-permutation is the empty sequence.

Theorem

For all integers $n \geq 1$, and all integers r such that $1 \leq r \leq n$:

$$P(n, r) = n \cdot (n-1) \cdot (n-2) \ldots (n-r+1) \;=\; \frac{n!}{(n-r)!}$$

Formula for # of permutations, and # of r-permutations

Let $\mathbf{P}(n, r)$ denote the number of r-permutations of an n-element set.

$P(n, 0) = 1$, because the only 0-permutation is the empty sequence.

Theorem

For all integers $n \geq 1$, and all integers r such that $1 \leq r \leq n$:

$$P(n, r) = n \cdot (n-1) \cdot (n-2) \ldots (n-r+1) \;=\; \frac{n!}{(n-r)!}$$

Proof. There are n different choices for the first element of the sequence. For each of those choices, there are $n-1$ remaining choices for the second element. For every combination of the first two choices, there are $n-2$ choices for the third element, and so forth. □

Corollary: the number of permutations of an n element set is:

$$n! \;=\; n \cdot (n-1) \cdot (n-2) \ldots \cdot 2 \cdot 1 = P(n, n)$$

Example: a simple counting problem

Example: How many permutations of the letters ABCDEFGH contain the string ABC as a (consecutive) substring?

Example: a simple counting problem

Example: How many permutations of the letters ABCDEFGH contain the string ABC as a (consecutive) substring?

Solution: We solve this by noting that this number is the same as the number of permutations of the following six objects:
ABC, D, E, F, G, and H. So the answer is:

$$6! = 720.$$

$\square$

Combinations

r-Combinations

An r**-combination** of a set S is an unordered collection of r elements of S. In other words, it is simply a subset of S of size r.

Example: Consider the set $S = \{1, 2, 3, 4, 5\}$.

The set $\{2, 5\}$ is a 2-combination of S.

There are 10 different 2-combinations of S. They are:

$\{1, 2\}$, $\{1, 3\}$, $\{1, 4\}$, $\{1, 5\}$,
$\{2, 3\}$, $\{2, 4\}$, $\{2, 5\}$,
$\{3, 4\}$, $\{3, 5\}$,
$\{4, 5\}$

Question: How many r-combinations of an n-element set are there?

Formula for the number of r-combinations

Let $\mathbf{C}(n, r)$ denote the number of r-combinations of an n-element set. Another notation for $C(n, r)$ is: $\binom{n}{r}$

These are called **binomial coefficients**, and are read as "n choose r".

Theorem

For all integers $n \geq 1$, and all integers r such that $0 \leq r \leq n$:

$$C(n, r) \doteq \binom{n}{r} = \frac{n!}{r! \cdot (n-r)!} = \frac{n \cdot (n-1) \cdot \ldots \cdot (n-r+1)}{r!}$$

Formula for the number of r-combinations

Let $\mathbf{C}(n, r)$ denote the number of r-combinations of an n-element set. Another notation for $C(n, r)$ is: $\binom{n}{r}$

These are called **binomial coefficients**, and are read as "n choose r".

Theorem

For all integers $n \geq 1$, and all integers r such that $0 \leq r \leq n$:

$$C(n, r) \doteq \binom{n}{r} = \frac{n!}{r! \cdot (n-r)!} = \frac{n \cdot (n-1) \cdot \ldots \cdot (n-r+1)}{r!}$$

Proof. We can see that $P(n, r) = C(n, r) \cdot P(r, r)$. (To get an r-permutation: first choose r elements, then order them.) Thus

$$C(n, r) = \frac{P(n, r)}{P(r, r)} = \frac{n!/(n-r)!}{r!/(r-r)!} = \frac{n!}{r! \cdot (n-r)!}$$

□

Combinations: examples

Example:

1. How many different 5-card poker hands can be dealt from a deck of 52 cards?
2. How many different 47-card poker hands can be dealt from a deck of 52 cards?

Solutions:

1. $$\binom{52}{5} = \frac{52!}{5! \cdot 47!} = \frac{52 \cdot 51 \cdot 50 \cdot 49 \cdot 48}{5 \cdot 4 \cdot 3 \cdot 2 \cdot 1} = 2,598,960$$
2. $$\binom{52}{47} = \frac{52!}{47! \cdot 5!} = \frac{52 \cdot 51 \cdot 50 \cdot 49 \cdot 48}{5 \cdot 4 \cdot 3 \cdot 2 \cdot 1} = 2,598,960$$

Question: Why are these numbers the same?

Combinations: an identity

Theorem

For all integers $n \geq 1$, and all integers r, $1 \leq r \leq n$:

$$\binom{n}{r} = \binom{n}{n-r}$$

Combinations: an identity

Theorem

For all integers $n \geq 1$, and all integers r, $1 \leq r \leq n$:

$$\binom{n}{r} = \binom{n}{n-r}$$

Proof:

$$\binom{n}{r} = \frac{n!}{r! \cdot (n-r)!} = \frac{n!}{(n-r)! \cdot (n-(n-r))!} = \binom{n}{n-r}$$

□

We can also give a **combinatorial proof**: Suppose $|S| = n$. A function, f, that maps each r-element subset A of S to the $(n-r)$-element subset $(S - A)$ is a bijection.
Any two finite sets having a bijection between them must have exactly the same number of elements. □

Binomial Coefficients

Consider the polynomial in two variables, x and y, given by:

$$(x+y)^n = \underbrace{(x+y)\cdot(x+y)\ldots(x+y)}_{n}$$

By multiplying out the n terms, we can expand this polynomial and write it in a standard sum-of-monomials form:

$$(x+y)^n = \sum_{j=0}^{n} c_j x^{n-j} y^j$$

Question: What are the coefficients c_j? (These are called binomial coefficients.)

Examples:

$$(x+y)^2 = x^2 + 2xy + y^2$$

$$(x+y)^3 = x^3 + 3x^2y + 3xy^2 + y^3$$

The Binomial Theorem

Binomial Theorem

For all $n \geq 0$:

$$(x+y)^n = \sum_{j=0}^{n} \binom{n}{j} x^{n-j} y^j = \binom{n}{0} x^n + \binom{n}{1} x^{n-1} y + \ldots + \binom{n}{n} y^n$$

The Binomial Theorem

Binomial Theorem

For all $n \geq 0$:

$$(x+y)^n = \sum_{j=0}^{n} \binom{n}{j} x^{n-j} y^j = \binom{n}{0} x^n + \binom{n}{1} x^{n-1} y + \ldots + \binom{n}{n} y^n$$

Proof: What is the coefficient of $x^{n-j}y^j$?
To obtain a term $x^{n-j}y^j$ in the expansion of the product

$$(x+y)^n = \underbrace{(x+y)(x+y)\ldots(x+y)}_{n}$$

we have to choose exactly $n-j$ copies of x and (thus) j copies of y.
How many ways are there to do this? Answer: $\binom{n}{j} = \binom{n}{n-j}$. □

Corollary: $\sum_{j=0}^{n} \binom{n}{j} = 2^n$.

Proof: By the binomial theorem, $2^n = (1+1)^n = \sum_{j=0}^{n} \binom{n}{j}$. □

Pascal's Identity

Theorem (Pascal's Identity)

For all integers $n \geq 0$, and all integers r, $0 \leq r \leq n+1$:

$$\binom{n+1}{r} = \binom{n}{r-1} + \binom{n}{r}$$

Pascal's Identity

Theorem (Pascal's Identity)

For all integers $n \geq 0$, and all integers r, $0 \leq r \leq n+1$:

$$\binom{n+1}{r} = \binom{n}{r-1} + \binom{n}{r}$$

Proof: Suppose $S = \{s_0, s_1, \ldots, s_n\}$. We wish to choose s subset $A \subseteq S$ such that $|A| = r$. We can do this in two ways. We can either:
(I) choose a subset A such that $s_0 \in A$, or
(II) choose a subset A such that $s_0 \notin A$.

There are $\binom{n}{r-1}$ sets of the first kind,
and there are $\binom{n}{r}$ sets of the second kind.
So, $\binom{n+1}{r} = \binom{n}{r-1} + \binom{n}{r}$. $\square$

r-Combinations with repetition (with replaced)

Sometimes, we want to choose r elements with repetition allowed from a set of size n. In how many ways can we do this?

Example: How many different ways are there to place 12 colored balls in a bag, when each ball should be either Red, Green, or Blue?

Let us first formally phrase the general problem.
A **multi-set** over a set S is an unordered collection (bag) of copies of elements of S with possible repetition. The size of a multi-set is the number of copies of all elements in it (counting repetitions).
For example, if $S = \{\mathrm{R}, \mathrm{G}, \mathrm{B}\}$, then the following two multi-sets over S both have size 4:

$$[\mathrm{G}, \mathrm{G}, \mathrm{B}, \mathrm{B}] \qquad [\mathrm{R}, \mathrm{G}, \mathrm{G}, \mathrm{B}]$$

Note that *ordering doesn't matter* in multi-sets, so $[\mathrm{R}, \mathrm{R}, \mathrm{B}] = [\mathrm{R}, \mathrm{B}, \mathrm{R}]$.

Definition: an r-Combination with repetition (r-comb-w.r.) from a set S is simply a multi-set of size r over S.

Formula for # of r-Combinations with repetition

Theorem

For all integers $n, r \geq 1$, the number of r-combs-w.r. from a set S of size n is:

$$\binom{n+r-1}{r} = \binom{n+r-1}{n-1}$$

Formula for # of r-Combinations with repetition

Theorem

For all integers $n, r \geq 1$, the number of r-combs-w.r. from a set S of size n is:

$$\binom{n+r-1}{r} = \binom{n+r-1}{n-1}$$

Proof: Each r-combination with repetition can be associated uniquely with a string of length $n+r-1$ consisting of of $n-1$ bars and r stars, and vice versa.

The bars partition the string into n different segments, and the number of stars in each segment denotes the number of copies of the corresponding element of S in the multi-set.

For example, for $S = \{\mathrm{R}, \mathrm{G}, \mathrm{B}, \mathrm{Y}\}$, then with the multiset

$$[\mathrm{R}, \mathrm{R}, \mathrm{B}, \mathrm{B}] \qquad \text{we associate the string} \qquad \star\star||\star\star|$$

How many strings of length $n+r-1$ with $n-1$ bars and r stars are there? Answer: $\binom{n+r-1}{r} = \binom{n+r-1}{n-1}$. □

Example

Example

How many different solutions in non-negative integers x_1, x_2, and x_3, does the following equation have?

$$x_1 + x_2 + x_3 = 11$$

Example

Example

How many different solutions in non-negative integers x_1, x_2, and x_3, does the following equation have?

$$x_1 + x_2 + x_3 = 11$$

Solution: We have to place 11 "pebbles" into three different "bins", x_1, x_2, and x_3.
This is equivalent to choosing an 11-comb-w.r. from a set of size 3, so the answer is

$$\binom{11+3-1}{11} = \binom{13}{2} = \frac{13 \cdot 12}{2 \cdot 1} = 78.$$

□

Permutations with indistinguishable objects

Question: How many different strings can be made by reordering the letters of the word "SUCCESS"?

Theorem: The number of permutations of n objects, with n_1 indistinguishable objects of Type 1, n_2 indistinguishable objects of Type 2, ..., and n_k indistinuishable objects of Type k, is: $$\frac{n!}{n_1! \cdot n_2! \ldots n_k!}$$

Permutations with indistinguishable objects

Question: How many different strings can be made by reordering the letters of the word "SUCCESS"?

Theorem: The number of permutations of n objects, with n_1 indistinguishable objects of Type 1, n_2 indistinguishable objects of Type 2, ..., and n_k indistinuishable objects of Type k, is: $\dfrac{n!}{n_1! \cdot n_2! \ldots n_k!}$

Proof: First, the n_1 objects of Type 1 can be placed among the n positions in $\binom{n}{n_1}$ ways. Next, the n_2 objects of Type 2 can be placed in the remaining $n - n_1$ positions in $\binom{n-n_1}{n_2}$ ways, and so on... We get:

$$\binom{n}{n_1} \cdot \binom{n-n_1}{n_2} \cdot \binom{n-n_1-n_2}{n_3} \ldots \binom{n-n_1-n_2-\ldots-n_{k-1}}{n_k} =$$

$$\frac{n!}{n_1!(n-n_1)!} \frac{(n-n_1)!}{n_2!(n-n_1-n_2)!} \ldots \frac{(n-n_1-\ldots-n_{k-1})!}{n_k!0!} = \frac{n!}{n_1!n_2!\ldots n_k!}.$$

Multinomial Coefficients

Multinomial coefficients

For integers $n, n_1, n_2, \ldots, n_k \geq 0$, such that $n = n_1 + n_2 + \ldots + n_k$, let:

$$\binom{n}{n_1, n_2, \ldots, n_k} = \frac{n!}{n_1! n_2! \ldots n_k!}$$

Multinomial Theorem

For all $n \geq 0$, and all $k \geq 1$:

$$(x_1 + x_2 + \ldots + x_k)^n = \sum_{\substack{0 \leq n_1, n_2, \ldots, n_k \leq n \\ n_1 + n_2 + \ldots + n_k = n}} \binom{n}{n_1, n_2, \ldots, n_k} x_1^{n_1} x_2^{n_2} \ldots x_k^{n_k}$$

Note: the Binomial Theorem is the special case of this where $k = 2$.

Question: In how many ways can the elements of a set S, $|S| = n$, be partitioned into k distinguishable boxes, such that Box 1 gets n_1 elements, ..., Box k gets n_k elements? **Answer:** $\binom{n}{n_1, n_2, \ldots, n_k}$. □

BBM 205 Discrete Mathematics
Hacettepe University
http://web.cs.hacettepe.edu.tr/∼bbm205

Lecture 5: Arithmetic Modulo m, Primes and Greatest Common Divisors

Lecturer: Lale Özkahya

Resources:
Kenneth Rosen, "Discrete Mathematics and App."
http://www.math.caltech.edu/ 2016-17/1term/ma006a/#notes

Our Problem: Encryption

- Alice needs to send bob a message.
- Eve can read the communications.
- Alice *encrypts* the message.

Classic Cryptography

- Alice and Bob exchange some information in advance, in a secure way.

Example: Atbash Cipher

- Replace each letter with a symbol, according to the sequence (*key*):

A	B	C	D	E	F	G	H	I	J	K	L	M
Z	Y	X	W	V	U	T	S	R	Q	P	O	N

"My hovercraft is full of eels"

"Nb slevixizug rh ufoo lu vvoh"

Other Historical Ciphers

- Scytale transposition cipher, used by the Spartan military.

- The Enigma machine in World War II.

- Cipher runes.

The Internet

- **Problem.** When performing a secret transaction over the internet, we cannot securely exchange information in advance.

Public-key Cryptography

- **Idea.** Use a *public key* which is used for *encryption* and a *private key* used for *decryption*.
- Bob generates both keys. Keeps the private key and publishes the public one.

Public-key Cryptography

- **Idea.** Use a *public key* which is used for *encryption* and a *private key* used for *decryption*.
- Alice encrypts her message with Bob's public key and sends it.

Public-key Cryptography

- Eve has the public key and the encrypted message.
- We need an action that is easy to do (*encrypt* using a public key) but very difficult to reverse (*decrypt* using a public key).

Public-key Cryptography

- Eve has the public key and the encrypted message.
- We need an action that is easy to do (*encrypt* using a public key) but very difficult to reverse (*decrypt* using a public key).
- **Bad example.** The public key is the number k. We encrypt a number a as $a \cdot k$. The adversary can divide by k…

Public-key Cryptography

- Eve has the public key and the encrypted message.
- We need an action that is easy to do (*encrypt* using a public key) but very difficult to reverse (*decrypt* using a public key).

Prime factorization

Integers

- We consider the set of integers

$$\mathbb{Z} = \{\dots, -2, -1, 0, 1, 2, \dots\}.$$

- The expression "$a \in \mathbb{Z}$" means that a is *in* the set $\mathbb{Z}$.
- For example, we have

$$1 \in \mathbb{Z}, \qquad 10^2 \in \mathbb{Z}.$$

- On the other hand

$$0.3 \notin \mathbb{Z}, \qquad \sqrt{2} \notin \mathbb{Z}.$$

Division

- Given two integers $a, b \in \mathbb{Z}$, we say that a divides b (or $a|b$) if there exists $s \in \mathbb{Z}$ such that $b = sa$.
- True or false:

$3|12$ ✓ $\qquad 12|3$ ✗

$3|-15$ ✓ $\qquad -3|3$ ✓

$-7|0$ ✓ $\qquad 0|-7$ ✗

$0|0$ ✓

Our First Proof

- **Claim.** If $a|b$ an $b|c$ then $a|c$.
- **Proof.**
  - There exists $s \in \mathbb{Z}$ such that $b = as$.
  - There exists $t \in \mathbb{Z}$ such that $c = bt$.
  - Therefore, $c = ast$.
  - Setting $r = st$, we have $c = ar$.

Our Second Proof

- **Claim.** If $a|b$ and $b|a$ then $a = \pm b$.
- **Proof.**
  - There exists $s \in \mathbb{Z}$ such that $a = sb$.
  - There exists $t \in \mathbb{Z}$ such that $b = ta$.
  - That is, $a = sta$.
  - $st = 1$ so either $s = t = 1$ or $s = t = -1$.

Prime Numbers

- A *natural number* is an integer that is non-negaive. The set of natural numbers:
$$\mathbb{N} = \{0,1,2,3, \dots\}.$$
- A number of $\mathbb{N} \setminus \{0,1\}$ is said to be *prime* if its only positive divisors are one at itself.

Proof by Induction

- **Claim (Prime decomposition).** Every natural number $n \geq 2$ is either a prime or a product of primes.
- **Proof.**
  - Induction basis: The claim holds for 2.
  - Induction step: Assume that the claim holds for every natural number smaller than n.
  - If n is a prime, the claim holds for n.
  - Otherwise, we can write $n = ab$.
  - By the induction hypothesis, both a and b are either primes or a product of primes.
  - Thus, n is a product of primes.

Proof by Contradiction

- **Claim.** There exist infinitely many prime numbers.
  - **Proof.** Assume, for contradiction, that there exists a finite set of primes
    $$P = \{p_1, p_2, p_3, \dots, p_n\}.$$
  - The number $p_1 p_2 \cdots p_n + 1$ is not prime, since it is not in P.
  - The number $p_1 p_2 \cdots p_n + 1$ is prime, since it cannot be divided by any of the primes of P.
  - *Contradiction! So there must be infinitely many primes!*

More Division Properties

- **Claim.** Given two numbers $a, b \in \mathbb{N}$, there are unique $q, r \in \mathbb{N}$ such that $r < b$ and
  $$a = qb + r.$$
- We say that q and r are the *quotient* and the *remainder* of dividing a with b.
- We write $r = a \bmod b$.
- *Proof by algorithm!*

Our First Algorithm

- **Input.** Two numbers $a, b \in \mathbb{N}$.
- **Output.** Two number $q, r \in \mathbb{N}$ such that $a = qb + r$ and $r < b$.

- $q \leftarrow 0$ and $n \leftarrow a$.
- While $n \geq b$:
  - $n \leftarrow n - b$.
  - $q \leftarrow q + 1$.
- $r \leftarrow n$

$a = 12 \quad b = 5$

$q = 0 \quad n = 12 \quad r = ?$
$q = 1 \quad n = 7 \quad r = ?$
$q = 2 \quad n = 2 \quad r = ?$
$r = 2$

Greatest Common Divisor

- We say that d is a *common divisor* of a and b (where $a, b, d \in \mathbb{N}$) if $d|a$ and $d|b$.
- The *greatest common divisor* of a and b, denoted $\mathrm{GCD}(a, b)$, is a common divisor c of a and b, such that
  - If $d|a$ and $d|b$ then $d \leq c$.
  - Equivalently, if $d|a$ and $d|b$ then $d|c$.

Examples: GCD

- What is $\mathrm{GCD}(18,42)$? *6*
- What is $\mathrm{GCD}(50,100)$? *50*
- What is $\mathrm{GCD}(6364800, 1491534000)$?
  - $\mathrm{GCD}(2^7 \cdot 3^2 \cdot 5^2 \cdot 13 \cdot 17, 2^4 \cdot 3^7 \cdot 5^3 \cdot 11 \cdot 31)$?

    $= 2^4 \cdot 3^2 \cdot 5^2 = 3600.$
- What can we do when dealing with numbers that are too large to factor?

GCD Property

- **Claim.** If $a = bq + r$ then

$$\mathrm{GCD}(a, b) = \mathrm{GCD}(b, r)$$

- **Example.**

$$66 = 21 \cdot 3 + 3$$

$$\Downarrow$$

$$GCD(66,21) = GCD(21,3) = 3.$$

Computing GCD: General approach

- **Problem.** Compute $\mathrm{GCD}(a, b)$.
  - Find $q_1, r_1 \in \mathbb{Z}$ such that $a = q_1 b + r_1$.
  - Since $\mathrm{GCD}(a, b) = \mathrm{GCD}(b, r_1)$, it suffices to compute the latter.
  - Find $q_2, r_2 \in \mathbb{Z}$ such that $b = q_2 r_1 + r_2$.
  - Since $\mathrm{GCD}(b, r_1) = \mathrm{GCD}(r_1, r_2)$, it suffices to compute the latter.
  - …
  - Continue until obtaining a zero remainder (then the divider is the required GCD).

The Euclidean Algorithm

- **Input.** Two numbers $a, b \in \mathbb{N}$.
- **Output.** $\mathrm{GCD}(a, b)$.

- $r \leftarrow a \bmod b$.
- While $r \neq 0$:
  - $a \leftarrow b$.
  - $b \leftarrow r$.
  - $r \leftarrow a \bmod b$.
- Output b.

$a = 78 \quad b = 45$

$a = 78$	$b = 45$	$r = 33$
$a = 45$	$b = 33$	$r = 12$
$a = 33$	$b = 12$	$r = 9$
$a = 12$	$b = 9$	$r = 3$
$a = 9$	$b = 3$	$r = 0$

Proof of GCD Property

- **Claim.** If $a = bq + r$ then

$$\mathrm{GCD}(a, b) = \mathrm{GCD}(b, r)$$

- **Proof.**
  - Since $r = a - bq$, every common divisor of a and b is also a divisor of r. Thus,
    $$\mathrm{GCD}(a, b) | \mathrm{GCD}(b, r)$$
  - Since $a = bq + r$, every common divisor of b and r is also a common divisor of a. Thus,
    $$\mathrm{GCD}(b, r) | \mathrm{GCD}(a, b).$$

The End

- The Voynich manuscript:

Warm-up: The Fibonacci Numbers

- *Fibonacci numbers*:

$$F_0 = F_1 = 1 \qquad F_i = F_{i-1} + F_{i-2}.$$

$$1,1,2,3,5,8,13,21,34, \ldots$$

- How many rounds of the algorithm are required to compute $GCD(F_n, F_{n-1})$?
  - Round 1: $r = F_n - F_{n-1} = F_{n-2}$.
  - Round 2: $r = F_{n-1} - F_{n-2} = F_{n-3}$.
  - …
  - **Round n**: $r = F_1 - F_0 = 0$.

More GCDs

- **Theorem.** For any $a, b \in \mathbb{N}$, there exist $s, t \in \mathbb{Z}$ such that

$$\mathrm{GCD}(a, b) = as + bt.$$

$$\mathrm{GCD}(18,27) = 9 \qquad -1 \cdot 18 + 1 \cdot 27 = 9$$

$$\mathrm{GCD}(25,65) = 5 \qquad 8 \cdot 25 - 3 \cdot 65 = 5$$

The Extended Euclidean Algorithm

- Build a matrix: First two rows are $(a, 1,0)$ and $(b, 0,1)$.
- Every other row is obtained by subtracting the two rows above it, to obtain the next value of b.

$$\begin{pmatrix} 78 & 1 & 0 \\ 45 & 0 & 1 \\ 33 & 1 & -1 \\ 12 & -1 & 2 \\ 9 & 3 & -5 \\ 3 & -4 & 7 \end{pmatrix}$$

$$a = 78 \qquad b = 45$$

$$\begin{array}{ccc} a = 78 & b = 45 & r = 33 \\ a = 45 & b = 33 & r = 12 \\ a = 33 & b = 12 & r = 9 \\ a = 12 & b = 9 & r = 3 \\ a = 9 & b = 3 & r = 0 \end{array}$$

The Extended Euclidean Algorithm

- Build a matrix: First two rows are $(a, 1,0)$ and $(b, 0,1)$.
- Every other row is obtained by subtracting the two rows above it, to obtain the next value of b.

$$\begin{pmatrix} 78 & 1 & 0 \\ 45 & 0 & 1 \\ 33 & 1 & -1 \\ 12 & -1 & 2 \\ 9 & 3 & -5 \\ 3 & -4 & 7 \end{pmatrix}$$

In every step, we have

$$a = qb + r,$$

and then

$$a \leftarrow b, \qquad b \leftarrow r.$$

If R_i denotes the i'th row:

$$R_i = R_{i-2} - qR_{i-1}.$$

The Extended Euclidean Algorithm

- Build a matrix: First two rows are $(a, 1,0)$ and $(b, 0,1)$.
- Every other row is obtained by subtracting the two rows above it, to obtain the next value of b.

$$\begin{pmatrix} 78 & 1 & 0 \\ 45 & 0 & 1 \\ 33 & 1 & -1 \\ 12 & -1 & 2 \\ 9 & 3 & -5 \\ 3 & -4 & 7 \end{pmatrix}$$

$33 = 2 \cdot 12 + 9$
so $R_5 = R_3 - 2R_4$

Proof by Algorithm!

- **Theorem.** If $c = \mathrm{GCD}(a, b)$, then there exist $s, t \in \mathbb{Z}$ such that
$$as + bt = c.$$

$$\begin{pmatrix} 78 & 1 & 0 \\ 45 & 0 & 1 \\ 33 & 1 & -1 \\ 12 & -1 & 2 \\ 9 & 3 & -5 \\ 3 & -4 & 7 \end{pmatrix} \Rightarrow \begin{aligned} 78 &= 1 \cdot 78 + 0 \cdot 45 \\ 45 &= 0 \cdot 78 + 1 \cdot 45 \\ 33 &= 1 \cdot 78 - 1 \cdot 45 \\ 12 &= -1 \cdot 78 + 2 \cdot 45 \\ 9 &= 3 \cdot 78 - 5 \cdot 45 \\ 3 &= -4 \cdot 78 + 7 \cdot 45 \end{aligned}$$

Algorithm Correctness

- **Proof Sketch.**
  - Induction basis. Trivial for the first two rows.
  - Induction step.

$$\begin{array}{c} R_i \\ R_{i+1} \\ R_{i+2} \end{array} = \begin{pmatrix} s_1 & s_2 & s_3 \\ t_1 & t_2 & t_3 \\ u_1 & u_2 & u_3 \end{pmatrix}$$

$$s_1 = a \cdot s_2 + b \cdot s_3,$$
$$t_1 = a \cdot t_2 + b \cdot t_3,$$

Induction hypothesis

$$u_1 = s_1 - qt_1 = a(s_2 - qt_2) + b(s_3 - qt_3)$$
$$= a \cdot u_2 + b \cdot u_3.$$

Scales Problem

- We need to verify the weights of various objects by using scales.
- We have an unlimited amount of weights in two different integer sizes - a and b.
- For which values of a and b can we measure every possible integer weight?
- **Answer.** Whenever

$$GCD(a, b) = 1.$$

Number Theory

- **Number theory**: the study of integers.
- Some famous theorems:
  - **Euclid.** There are infinitely many prime numbers.
  - **"Fermat's last theorem".** The equation $x^n + y^n = z^n$ has no integer solutions when $n > 2$.
  - **Lagrange 1770.** Every natural number can be represented as the sum of four integer squares.

$15 = 1^2 + 1^2 + 2^2 + 3^2$ $\quad$ $110 = 10^2 + 3^2 + 1^2 + 0^2$

Number Theory (2)

- A couple of famous open problems:
  - **Twin prime conjecture.** There are infinitely many pairs of prime numbers that differ by two (5 and 7, 17 and 19, 41 and 43, …).
  - **Goldbach's conjecture.** Every even integer greater than 2 can be expressed as the sum of two primes.

Congruences

- **Recall.** The remainder of dividing a by m can be written as

$$r = a \bmod m.$$

- If also $r = b \bmod m$, we say that "a is *congruent* to b modulo m", and write

$$a \equiv b \bmod m.$$

  - Equivalently, $m|(a-b)$.

- The numbers $3, 10, 17, 73, 1053$ are all congruent modulo 7.

Congruence Classes

- If $m = 2$, numbers are congruent if they have the same parity.
- If $m = 3$, there are three distinct classes of numbers

$$0 \equiv 3 \equiv 6 \equiv 9 \equiv \cdots \bmod 3$$
$$1 \equiv 4 \equiv 7 \equiv 10 \equiv \cdots \bmod 3$$
$$2 \equiv 5 \equiv 8 \equiv 11 \equiv \cdots \bmod 3$$

- In general, we have exactly m *equivalence classes* of numbers.

Congruency is Transitive

- **Claim 1.** If $a \equiv b \; mod \; m$ and $b \equiv c \; mod \; m$ then $a \equiv c \; mod \; m$.

$$5 \equiv 55 \; mod \; 10.$$
$$55 \equiv 95 \; mod \; 10$$
$$\Downarrow$$
$$5 \equiv 95 \; mod \; 10$$

Congruency is Transitive

- **Claim 1.** If $a \equiv b \; mod \; m$ and $b \equiv c \; mod \; m$ then $a \equiv c \; mod \; m$.
- **Proof.** If $m|(a-b)$ and $m|(b-c)$ then $m|(a-c)$ since

$$a - c = (a - b) + (b - c).$$

Congruency and Addition

- **Claim 2.** If $a \equiv b \; mod \; m$ and $c \equiv d \; mod \; m$, then

$$a + c \equiv b + d \; mod \; m.$$

$$3 \equiv 15 \; mod \; 12$$
$$2 \equiv 26 \; mod \; 12$$

$$3 + 2 \equiv 15 + 26 \bmod 12$$

Congruency and Addition

- **Claim 2.** If $a \equiv b \; mod \; m$ and $c \equiv d \; mod \; m$, then

$$a + c \equiv b + d \; mod \; m.$$

- **Proof.** If $m|(a - b)$ and $m|(\mathrm{c} - \mathrm{d})$ then $m|\big((a + c) - (b + d)\big)$.

Congruency and Multiplication

- **Claim 3.** If $a \equiv b \bmod m$ and $c \equiv d \bmod m$, then

$$ac \equiv bd \bmod m.$$

$$3 \equiv 15 \bmod 12$$
$$2 \equiv 26 \bmod 12$$

$$\Downarrow$$

$$3 \cdot 2 \equiv 15 \cdot 26 \bmod 12$$

Congruency and Multiplication

- **Claim 3.** If $a \equiv b \bmod m$ and $c \equiv d \bmod m$, then

$$ac \equiv bd \bmod m.$$

- **Proof.** We have

$$\begin{aligned} ac - bd &= (ac - cb) + (cb - bd) \\ &= c(a - b) + b(c - d). \end{aligned}$$

That is, $m|(ac - bd)$.

Relatively Prime Numbers

- Two integers $m, n \in \mathbb{Z}$ are ***relatively prime*** if $\mathrm{GCD}(m, n) = 1$.
- **Claim 4.** If a and m are relatively prime, then there exists $b \in \mathbb{Z}$ such that

$$ab \equiv 1 \; mod \; m.$$

$$\mathrm{GCD}(6,17) = 1$$

$$\Downarrow$$

$$6 \cdot 3 = 1 \; mod \; 17$$

Relatively Prime Numbers

- Two integers $m, n \in \mathbb{Z}$ are ***relatively prime*** if $\mathrm{GCD}(m, n) = 1$.
- **Claim 4.** If a and m are relatively prime, then there exists $b \in \mathbb{Z}$ such that

$$ab \equiv 1 \; mod \; m.$$

- **Proof.** There exist $s, t \in \mathbb{Z}$ such that $as + mt = 1$. Taking $b = s$, we have

$$m|(ab + mt - 1) \quad \Rightarrow \quad m|(ab - 1).$$

A Cancellation Law

- **Claim 5.** If k, m are relatively prime, and
$$ak \equiv bk \; mod \; m,$$
then $a \equiv b \; mod \; m$.

$$\mathrm{GCD}(5,9) = 1$$
$$1 \cdot 5 \equiv 10 \cdot 5 \; mod \; 9$$
$$\Downarrow$$
$$1 \equiv 10 \; mod \; 9$$

A Cancellation Law

- **Claim 5.** If k, m are relatively prime and
$$ak \equiv bk \; mod \; m,$$
then $a \equiv b \; mod \; m$.
- **Proof.**
  - By Claim 4 there exist $s \in \mathbb{Z}$ such that $ks \equiv 1 \; mod \; m$.

$$a \equiv a \cdot 1 \equiv aks \equiv bks \equiv b \cdot 1 \equiv b \; mod \; m.$$

A Cancellation Law

- **Claim 5.** If k, m are relatively prime and
$$ak \equiv bk \; mod \; m,$$
then $a \equiv b \; mod \; m$.

- What happens when $\mathrm{GCD}(k, m) \neq 1$?

$$k = 4 \qquad m = 8$$
$$2 \cdot k \equiv 4 \cdot k \equiv 0 \; mod \; 8$$

Latin Squares

- **Claim 6.** Let $a, b, m \in \mathbb{Z}$ and let a, m be relatively prime. Then there is a *unique* x $(mod \; m)$ such that $ax \equiv b \; mod \; m$.

$$m = 11, \qquad a = 5, \qquad b = 6$$
$$5 \cdot x \equiv 6 \; mod \; 11$$
$$\Downarrow$$
$$x = 10.$$

Latin Squares

- **Claim 6.** Let $a, b, m \in \mathbb{Z}$ and let a, m be relatively prime. Then there is a *unique* x $(mod\ m)$ such that $ax \equiv b\ mod\ m$.
- **Proof.** By Claim 4 there exists $s \in \mathbb{Z}$ such that $as \equiv 1\ mod\ m$.
- Thus, $x = sb$ is one valid solution.
- Assume, for contradiction, that there are two distinct solutions x, x'.
  - Then $ax \equiv ax'\ mod\ m$.

$$x = sax = sax' = x'.$$

Problem: Large Powers

- **Problem.** Compute $3^{100}\ mod\ 7$.

Problem: Large Powers

- **Problem.** Compute $3^{100}\ mod\ 7$.
- **Modest beginning.**

$$3^1 \equiv 3\ mod\ 7$$
$$3^2 \equiv 3 \cdot 3^1 \equiv 2\ mod\ 7$$
$$3^3 \equiv 3 \cdot 3^2 \equiv 6\ mod\ 7$$
$$3^4 \equiv 3 \cdot 3^3 \equiv 4\ mod\ 7$$
$$3^5 \equiv 3 \cdot 3^4 \equiv 5\ mod\ 7$$
$$3^6 \equiv 3 \cdot 3^5 \equiv 1\ mod\ 7$$
$$3^7 \equiv 3 \cdot 3^6 \equiv 3\ mod\ 7$$

Problem: Large Powers

$$3^1 \equiv 3\ mod\ 7$$
$$3^2 \equiv 3 \cdot 3^1 \equiv 2\ mod\ 7$$
$$3^3 \equiv 3 \cdot 3^2 \equiv 6\ mod\ 7$$
$$3^4 \equiv 3 \cdot 3^3 \equiv 4\ mod\ 7$$
$$3^5 \equiv 3 \cdot 3^4 \equiv 5\ mod\ 7$$
$$3^6 \equiv 3 \cdot 3^5 \equiv 1\ mod\ 7$$
$$3^7 \equiv 3 \cdot 3^6 \equiv 3\ mod\ 7$$
$$3^8 \equiv 3 \cdot 3^7 \equiv 2\ mod\ 7$$
$$\ldots$$
$$3^{100} \equiv 3^{4+6\cdot 16} \equiv 3^4 \cdot 1 \equiv 4\ mod\ 7$$

Reminder: Some Congruent Properties

- **Addition.** If $a \equiv b \bmod m$ and $c \equiv d \bmod m$, then $a + c \equiv b + d \bmod m$.
- **Products.** If $a \equiv b \bmod m$ and $c \equiv d \bmod m$, then $ac \equiv bd \bmod m$.
- **Cancellation.** If $\mathrm{GCD}(k, m) = 1$ and $ak \equiv bk \bmod m$, then $a \equiv b \bmod m$.
- **Inverse.** If $\mathrm{GCD}(a, m) = 1$, then there exists $b \in \mathbb{Z}$ such that $ab \equiv 1 \bmod m$.

Warm-up: Division by Nine

- **Claim.** A number $a \in \mathbb{N}$ is divisible by 9 if and only if the sum of its digits is divisible by 9.
- Is 123456789 divisible by 9?

$$1 + 2 + 3 + 4 + 5 + 6 + 7 + 8 + 9 = 45$$
$$4 + 5 = 9$$

✓

Warm-up: Division by Nine (2)

- **Claim.** A number $a \in \mathbb{N}$ is divisible by 9 if and only if the sum of its digits is divisible by 9.
  - **Proof.** Write a as $a_k a_{k-1} \cdots a_1 a_0$ where a_i is the $(i+1)$'th rightmost digit of a.

$$a - (a_0 + a_1 + \cdots + a_k) =$$
$$(a_0 \cdot 10^0 + a_1 \cdot 10^1 + a_2 \cdot 10^2 + \cdots) - (a_0 + \cdots + a_k)$$
$$= a_1 \cdot 9 + a_2 \cdot 99 + a_3 \cdot 999 + \cdots$$

  - That is, $9 | a - (a_0 + a_1 + \cdots + a_k)$

Warm-up: Division by Nine (3)

- **Claim.** A number $a \in \mathbb{N}$ is divisible by 9 if and only if the sum of its digits is divisible by 9.
  - **Proof.** Write a as $a_k a_{k-1} \cdots a_1 a_0$ where a_i is the $(i-1)$'th rightmost digit of a.
  - We have: $9 | a - (a_0 + a_1 + \cdots + a_k)$.
  - Equivalently,

$$a \equiv (a_0 + a_1 + \cdots + a_k) \; mod \; 9.$$

Casting Out Nines

- **Problem.** Is the following correct?
  $54{,}321 \cdot 98{,}765 = 5{,}363{,}013{,}565.$
- If this is correct, then

$54{,}321 \cdot 98{,}765 \equiv 5{,}363{,}013{,}565 \; mod \; 9.$

$$5 + 4 + 3 + 2 + 1 \equiv 6 \; mod \; 9$$
$$9 + 8 + 7 + 6 + 5 \equiv 2 \; mod \; 9$$
$$5 + 3 + 6 + 3 + 0 + 1 + 3 + 5 + 6 + 5 \equiv 1 \; mod \; 9.$$

$$6 \cdot 2 \not\equiv 1 \; mod \; 9$$

X

Casting Out Nines (cont.)

- Is the *casting out nines* technique always correct in verifying whether $a \cdot b = c$?
  - If the calculation $mod \; 9$ is wrong, the original calculation must be wrong.
  - If the calculation $mod \; 9$ is correct, the original calculation might still be wrong!

$$1 \cdot 2 \equiv 11 \; mod \; 9.$$

Casting Out Nines Crank

- In the 1980's, a crank wrote a 124-page book explaining *the law of conservations of numbers* that he "developed for 24 years".
- This law "was perfected with 100% effectiveness".
- The book is basically 124 pages about the casting out nines trick. It does not mention that the method can sometimes fail.

Fermat's Little Theorem

- **Theorem.** For any prime p and integer a,

$$a^p \equiv a \bmod p.$$

- Examples:

$$15^7 \equiv 15 \equiv 1 \bmod 7$$

$$20^{53} \equiv 20 \bmod 53$$

$$2^{1009} \equiv 2 \bmod 1009$$

Pierre de Fermat

Fermat's Little Theorem

- **Theorem.** For any prime p and integer a,

$$a^p \equiv a \bmod p.$$

- **Proof.** By induction on a:
  - We now prove only the case of $a \geq 0$.
  - Induction basis: Obviously holds for $a = 0$.
  - Induction step: Assume that the claim holds for a. In a later lecture we prove
    $$(a+b)^p \equiv a^p + b^p \bmod p.$$
  - Thus:
    $$(a+1)^p \equiv a^p + 1 \equiv a + 1 \bmod p.$$

A Corollary

- **Corollary.** If $a \in \mathbb{N}$ is not divisible by a prime p then $a^{p-1} \equiv 1 \bmod p$.
- **Proof.**
  - We have $\mathrm{GCD}(a, p) = 1$.
  - Fermat's little theorem: $a^p \equiv a \bmod p$.
  - Combine with cancelation property: If $\mathrm{GCD}(k, m) = 1$ and $ak \equiv bk \bmod m$, then $a \equiv b \bmod m$.

Euler's Totient Function

- *Euler's totient $\varphi(n)$* is defined as follows: Given $n \in \mathbb{N} \setminus \{0\}$, then

$$\varphi(n) = |\{x \mid 1 \leq x < n \text{ and } \mathrm{GCD}(x, n) = 1\}|.$$

- In more words: $\varphi(n)$ is the number of natural numbers $1 \leq x \leq n$ such that x and n are relatively prime.

$$\varphi(12) = |\{1,5,7,11\}| = 4$$

Leonhard Euler

The Totient of a Prime

- **Observation.** If p is a prime number, then
$$\varphi(p) = p - 1.$$

The first thousand values of $\varphi(n)$:

Euler's Theorem

- **Theorem.** For any pair $a, n \in \mathbb{N}$ such that $GCD(a, n) = 1$, we have
$$a^{\varphi(n)} \equiv 1 \; mod \; n.$$

- This is a generalization of the claim $a^{p-1} \equiv 1 \; mod \; p$ (when p is prime).

The RSA Algorithm

- Public key cryptosystem.
- Discovered in 1977 by *R*ivest, *S*hamir, and *A*dleman.
- Still extremely common!

Ron Rivest

Adi Shamir

Leonard Adleman

RSA Public and Private Keys

1. Choose two LARGE primes p, q (say, 500 digits).
2. Set $n = pq$.
3. Compute $\varphi(n)$, and choose $1 < e < \varphi(n)$ such that $\mathrm{GCD}(e, \varphi(n)) = 1$.
4. Find d such that $de \equiv 1 \; mod \; \varphi(n)$.

Public key. n and e.

Private information. p, q, and d.

Preparing for Secure Communication

- Bob generates p, q, n, d, e, and transmits only e and n.

Encrypting a Message

- Alice wants to send Bob the number $m < n$ without Eve deciphering it.
- Alice uses n, e to calculate $X = m^e \bmod n$, and sends X to Bob.

Decrypting a Message

- Bob receives message $X = m^e \bmod n$ from Alice. Then he calculates:

$$X^d \bmod n \equiv m^{ed} \bmod n$$
$$\equiv m^{1+k\cdot\varphi(n)} \bmod n \equiv m \bmod n.$$

$de \equiv 1 \bmod \varphi(n)$

Euler's Theorem: $m^{\varphi(n)} \equiv 1 \bmod n$

Slightly cheating since the theorem requires $GCD(m, n) = 1$

RSA in One Slide

- Bob wants to generate keys:
  - Arbitrarily chooses primes p and q.
    sets $n = pq$ and finds $\varphi(n)$.
  - Chooses e such that $\mathrm{GCD}(\varphi(n), e) = 1$.
  - Find d such that $de \equiv 1 \; mod \; \varphi(n)$.
- Alice wants to pass bob m.
  - Receives n, e from Bob.
  - Returns $X \equiv m^e \; mod \; n$.
- Bob receives X.

  Calculates $X^d \; mod \; n$.

Example: RSA (with small numbers)

- Bob wants to generate keys:
  - Arbitrarily chooses primes $p = 61$ and $q = 53$.
    $n = 61 \cdot 52 = 3233$. $\varphi(3233) = 3120$.
  - Chooses $e = 17$ ($GCD(3120,17) = 1$).
  - For $de \equiv 1 \; mod \; 3120$, we have $d = 2753$.
- Alice wants to pass bob $m = 65$.
  - Receives n, e from Bob. Returns
    $m^e = 65^{17} \equiv 2790 \; mod \; 3233$.
- Bob receives $X \equiv 2790 \; mod \; 3233$.

  Calculates $X^d = 2790^{3233} \equiv 65 \; mod \; 3233$.

Some Details

- Bob needs to:
  - Find two large primes p, q.
  - Calculate n, d, e.
- Alice needs to
  - Use n, e to calculate $X = m^e \ mod \ n$.
- Eve must not be able to
  - Use n, e, X to find m.
- Bob needs to:
  - Use n, d, X to find m.

That is: Easy to compute a large power $mod \ n$. Hard to compute a large "root" $mod \ n$.

Taking Large Roots

- Eve has n, e, and Alice's message $X \equiv m^e \ mod \ n$.
- If Eve can compute $X^{1/e} \ mod \ n$, she can read the message! (i.e., if she can factor n).
- So far nobody knows how to compute this in a reasonable time.
- Or do they?

Computing a Large Power

- **Problem.** How can we compute

$$65^{2^{4000}} \; mod \; 9721?$$

- A naïve approach:

$$65^2 \equiv 4225 \; mod \; 9721$$
$$65^3 \equiv 65 \cdot 65^2 \equiv 2437 \; mod \; 9721$$
$$65^4 \equiv 65 \cdot 65^3 \equiv 2869 \; mod \; 9721$$
$$\ldots$$

- This approach requires 2^{4000} (about $1.3 \cdot 10^{1204}$) steps. *Impossible!*

Computing a Large Power – Fast!

- **Problem.** How can we compute

$$65^{2^{4000}} \; mod \; 9721?$$

$$65^2 \equiv 4225 \; mod \; 9721$$
$$65^4 \equiv 65^2 \cdot 65^2 \equiv 2869 \; mod \; 9721$$
$$65^8 \equiv 65^4 \cdot 65^4 \equiv 7195 \; mod \; 9721$$
$$65^{16} \equiv 65^8 \cdot 65^8 \equiv 3700 \; mod \; 9721$$
$$\ldots$$

Only 4000 steps. *Easy!*

A Small Technical Issue

- What if we calculate a^b where b is not a power of two?
- We calculate $a, a^2, a^4, a^8, a^{16}, a^{32}, \ldots$
- Every number can be expressed as a sum of distinct powers of 2.

$$57 = 32 + 16 + 8 + 1$$

$$\Downarrow$$

$$a^{57} = a^{32}a^{16}a^{8}a$$

What is Left to Do?

- Bob wants to generate keys:
  - Arbitrarily chooses primes p and q. ?
    $n = pq$ ✓ find $\varphi(n)$. ?
  - Chooses e such that $\mathrm{GCD}(\varphi(n), e) = 1$.
  - Find d such that $de \equiv 1 \; mod \; \varphi(n)$. ?
- Alice wants to pass bob m.
  - Receives n, e from Bob.
  - Returns $X \equiv m^e \; mod \; n$. ✓
- Bob receives X.

  Calculates $X^d \; mod \; n$. ✓

Reminder: Euler's Totient Function

- *Euler's totient $\varphi(n)$* is defined as follows: Given $n \in \mathbb{N}$, then

$$\varphi(n) = |\{x \mid 1 \leq x < n \text{ and } \mathrm{GCD}(x, n) = 1\}|.$$

- In more words: $\varphi(n)$ is the number of natural numbers $1 \leq x \leq n$ such that x and n are coprime.

$$\varphi(12) = |\{1,5,7,11\}| = 4.$$

Reminder #2: The RSA Algorithm

- Bob wants to generate keys:
  - Arbitrarily chooses primes p and q. ?
    $n = pq$ ✓ find $\varphi(n)$. ?
  - Chooses e such that $\mathrm{GCD}(\varphi(n), e) = 1$. ?
  - Find d such that $de \equiv 1 \; mod \; \varphi(n)$. ?
- Alice wants to pass bob m.
  - Receives n, e from Bob.
  - Returns $X \equiv m^e \; mod \; n$. ✓
- Bob receives X.

  Calculates $X^d \; mod \; n$. ✓

Finding $\varphi(n)$

- **Problem.** Given $n = pq$, where p, q are large primes, find $\varphi(n)$.
  - We need the number of elements in $\{1,2, \dots, n\}$ that are not multiplies of p or q.
  - There are $\frac{n}{p} = q$ numbers are divisible by p.
  - There are $\frac{n}{q} = p$ numbers are divisible by q.
  - Only $n = pq$ is divided by both.
  - Thus: $\varphi(n) = n - p - q + 1$.

The RSA Algorithm

- Bob wants to generate keys:
  - Arbitrarily chooses primes p and q. ?
    $n = pq$ ✓ find $\varphi(n)$. ✓
  - Chooses e such that $\mathrm{GCD}(\varphi(n), e) = 1$. ?
  - Find d such that $de \equiv 1 \; mod \; \varphi(n)$. ?
- Alice wants to pass bob m.
  - Receives n, e from Bob.
  - Returns $X \equiv m^e \; mod \; n$. ✓
- Bob receives X.

  Calculates $X^d \; mod \; n$. ✓

Choose e s.t. $\mathrm{GCD}(\varphi(n), e) = 1$

- **Problem.** Given $n = pq$, where p, q are large primes, find $e \in \mathbb{N}$ such that $\mathrm{GCD}(\varphi(n), e) = 1$.
  - We can choose arbitrary numbers until we find one that is relatively prime to $\varphi(n)$.
  - For the "worst" values of $\varphi(n)$, a random number is good with probability $1/\log\log n$.

The RSA Algorithm

- Bob wants to generate keys:
  - Arbitrarily chooses primes p and q. ?
    $n = pq$ ✓ find $\varphi(n)$. ✓
  - Chooses e such that $\mathrm{GCD}(\varphi(n), e) = 1$. ✓
  - Find d such that $de \equiv 1 \; mod \; \varphi(n)$. ?
- Alice wants to pass bob m.
  - Receives n, e from Bob.
  - Returns $X \equiv m^e \; mod \; n$. ✓
- Bob receives X.

  Calculates $X^d \; mod \; n$. ✓

Find d such that $de \equiv 1 \; mod \; \varphi(n)$

- **Recall.** Since $\mathrm{GCD}(e, \varphi(n)) = 1$ then there exist $s, t \in \mathbb{Z}$ such that
$$se + t\varphi(n) = 1.$$
- That is, $se \equiv 1 \; mod \; \varphi(n)$.
- We can find s, t by the ***extended Euclidean algorithm*** from lecture 2.

The RSA Algorithm

- Bob wants to generate keys:
  - Arbitrarily chooses primes p and q. ?
    $n = pq$ ✓ find $\varphi(n)$. ✓
  - Chooses e such that $\mathrm{GCD}(\varphi(n), e) = 1$. ✓
  - Find d such that $de \equiv 1 \; mod \; \varphi(n)$. ✓
- Alice wants to pass bob m.
  - Receives n, e from Bob.
  - Returns $X \equiv m^e \; mod \; n$. ✓
- Bob receives X.

  Calculates $X^d \; mod \; n$. ✓

Quantum Computing

- A *bit* of a computer contains a value of either 0 or 1.
- A quantum computer contains *qubits*, which can be in superpositions of states.
- Theoretically, a quantum computer can easily factor numbers and decipher almost any known encryption.

Should We Stop Ordering Things Online?

SEARCH

The New York Times

TECHNOLOGY

Microsoft Makes Bet Quantum Computing Is Next Breakthrough

By JOHN MARKOFF JUNE 23, 2014

The Washington Post

EDGE

NSA seeks to build quantum computer that could crack most types of encryption

Finding Large Primes

- Let n be a LARGE integer (e.g., 2^{4000}).
- *The prime number theorem.* The probability of a random $p \in \{1, \dots, n\}$ being prime is about $1/\log n$.
- If we randomly choose numbers from $\{1, \dots, n\}$, we expect to have about $\log n$ iterations before finding a prime.
  - *But how can we check whether our choice is a prime or not?!*

Primality Testing

- Given a LARGE $q \in \mathbb{Z}$, how can we check whether q is prime?
- **The naïve approach.** Go over every number in $\{2, \dots, \sqrt{q}\}$ and check whether it divides q.
  - *But we chose our numbers to be too large for a computer to go over all of them!*

Recall: Fermat's Little Theorem

- For any prime p and integer a relatively prime to p, we have

$$a^p \equiv a \bmod p.$$

- Pick a random integer a and check whether $a^q \equiv a \bmod q$.
  - If not, q is not a prime!
  - If yes, ???

Pierre de Fermat

Example: Fermat Primality Testing

- Is $n = 355207$ prime?

$$2^{355207} \equiv 84927 \bmod 355207.$$

- n is not prime since $2^n \not\equiv 2 \bmod n$.

- We can try 1000 different values of a and see if $a^n \equiv a \bmod n$ for each of them.

Carmichael Numbers

- A number $q \in \mathbb{N}$ is said to be a *Carmichael number* if it is not prime, but still satisfies $a^q \equiv a \bmod q$ for every a that is relatively prime to q.
  - The smallest such number is 561.
  - Very rare – about one in 50 trillion in the range $1 - 10^{21}$.

R. D. Carmichael

Example: Carmichael Numbers

- **Claim.** Let $k \in \mathbb{N} \setminus \{0\}$ such that $6k+1, 12k+1$, and $18k+1$ are primes. Then
  $$n = (6k+1)(12k+1)(18k+1)$$
  is a Carmichael number.
- **Example.**
  - For $k = 1$, we have that 7,13,19 are primes.
  - $7 \cdot 13 \cdot 19 = 1729$ is a Carmichael number.

Proof

- We need to prove that for any a that is relatively prime to n, we have
$$a^n \equiv a \bmod n.$$
- **Recall.** Since $GCD(a, n) = 1$, this is equivalent to $a^{n-1} \equiv 1 \bmod n$.
- We rewrite $n = 1296k^3 + 396k^2 + 36k + 1$.
- For any such a, we have
$$a^{n-1} = a^{1296k^3+396k^2+36k}$$
$$= \left(a^{6k}\right)^{216k^2+66k+6}.$$

Proof (cont.)

- For any a relatively prime to n, we have
$$a^{n-1} = \left(a^{6k}\right)^{216k^2+66k+6}.$$
- **Recall.** If $a \in \mathbb{N}$ is not divisible by a prime p then $a^{p-1} \equiv 1 \bmod p$.
- Since a and $6k - 1$ are relatively prime
$$a^{n-1} \equiv 1^{216k^2+66k+6} \equiv 1 \bmod 6k + 1.$$
- Similarly, we have $a^{n-1} \equiv 1 \bmod 12k + 1$ and $a^{n-1} \equiv 1 \bmod 18k + 1$.
- Since $a^{n-1} - 1$ is divisible by the three pairwise coprime integers $6k + 1, 12k + 1$, and $18k + 1$, it is also divisible by their product n. That is, $a^{n-1} \equiv 1 \bmod n$.

Miller–Rabin Primality Test

- The *Miller–Rabin primality test* works on *every number*.

Gary Miller

Michael Rabin

Root of Unity

- **Claim.** For any prime p, the only numbers $a \in \{1, \dots, p\}$ such that $a^2 \equiv 1 \; mod \; p$ are 1 and $p - 1$.

- **Example.** The solutions to
$$a^2 \equiv 1 \; mod \; 1009$$
are exactly the numbers satisfying
$$a \equiv 1 \text{ or } 1008 \; mod \; 1009.$$

Root of Unity

- **Claim.** For any prime p, the only numbers $a \in \{1, \dots, p\}$ such that $a^2 \equiv 1 \; mod \; p$ are 1 and $p - 1$.
- **Proof.**

$$a^2 \equiv 1 \; mod \; p$$

$$a^2 - 1 \equiv 0 \; mod \; p$$

$$(a + 1)(a - 1) \equiv 0 \; mod \; p$$

- That is, either $p|(a + 1)$ or $p|(a - 1)$.

Roots of Unity Properties

- Given a prime $p > 2$, we write

$$p - 1 = 2^s d$$

  where d is odd and $s \geq 1$.
- **Claim.** For any ***odd*** prime p and any $1 < a < p$, one of the following holds.
  - $a^d \equiv 1 \; mod \; p$.
  - There exists $0 \leq r < s$ such that

$$a^{2^r d} \equiv -1 \; mod \; p.$$

Roots of Unity Properties (2)

- **Claim.** For any ***odd*** prime p and any $1 < a < p$, one of the following holds.
  - $a^d \equiv 1 \; mod \; p$.
  - There exists $0 \leq r < s$ such that $a^{2^r d} \equiv -1 \; mod \; p$.
- **Proof.**
  - By Fermat's little theorem $a^{p-1} \equiv 1 \; mod \; p$.
  - Consider $a^{(n-1)/2}, a^{(n-1)/4}, \ldots, a^{(n-1)/2^s}$. By the previous claim, each such root is $\pm 1 \; mod \; n$.
  - If all of these roots equal 1, we are in the first case. Otherwise, we are in the second.

Composite Witnesses

- Given a composite (non-prime) ***odd*** number n, we again write $n - 1 = 2^s d$ where d is odd and $s \geq 1$.
- We say that $a \in \{2,3,4, \ldots, n-2\}$ is a ***witness*** for n if
  - $a^d \not\equiv 1 \; mod \; n$.
  - For every $0 \leq r < s$, we have $a^{2^r d} \not\equiv -1 \; mod \; n$.

Example: Composite Witness

- **Problem.** Prove that 91 is not a prime.

$$90 = 2 \cdot 45.$$

$$2^{45} \equiv 57 \; mod \; 91.$$

- 2 is a witness that 91 is not a prime.

There are Many Witnsses

- Given an odd composite n, the probability of a number $\{2, \dots, n-2\}$ being a witness is at least ¾.
- Given an odd $n \in \mathbb{N}$, take i numbers and check if they are witnesses.
  - If we found a witness, n is composite.
  - If we did not find a witness, n is prime with probability at least

The End

I'VE DISCOVERED A WAY TO GET COMPUTER SCIENTISTS TO LISTEN TO ANY BORING STORY.

BBM 205 Discrete Mathematics
Hacettepe University
http://web.cs.hacettepe.edu.tr/∼bbm205

Lecture 6: Recursion: Definitions, Solving recursive equations

Lecturer: Lale Özkahya

Resources:
Kenneth Rosen, "Discrete Mathematics and App."
http://www1.cs.columbia.edu/ zeph/3203s04/lectures.html

Recursively Defined Sequences

EG: Recall the Fibonacci sequence:

$\{f_n\} = 0,1,1,2,3,5,8,13,21,34,55,\ldots$

Recursive definition for $\{f_n\}$:

INITIALIZE: $f_0 = 0$, $f_1 = 1$

RECURSE: $f_n = f_{n-1}+f_{n-2}$ for $n > 1$.

The recurrence relation is the recursive part

$f_n = f_{n-1}+f_{n-2}$. Thus a recurrence relation for a sequence consists of an equation that expresses each term in terms of lower terms.

Q: Is there another solution to the Fibonacci recurrence relation?

Recursively Defined Sequences

A: Yes, for example could give a different set of ***initial conditions*** such as f_0=1, f_1= -1 in which case would get the sequence

$\{f_n\}$ = 1,-1,0,-1,-2,-3,-5,-8,-13,-21,…

Q: How many solutions are there to the Fibonacci recursion relation?

Recursively Defined Sequences

A: Infinitely many solutions as each pair of integer initial conditions (a,b) generates a unique solution.

Recurrence Relations for Counting

Often it is very hard to come up with a closed formula for counting a particular set, but coming up with recurrence relation easier.

EG: Geometric example of counting the number of points of intersection of n lines.

Q: Find a recurrence relation for the number of bit strings of length n which contain the string 00.

Recurrence Relations for Counting

A: a_n= #(length n bit strings containing 00):

I. If the first n-1 letters contain 00 then so does the string of length n. As last bit is free to choose get contribution of $2a_{n-1}$

II. Else, string must be of the form u00 with u a string of length n-2 not containing 00 and **not ending in 0** (why not?). But the number of strings of length n-3 which don't contain 00 is the total number of strings minus the number that do. Thus get contribution of $2^{n-3}-a_{n-3}$

Solution: $a_n = 2a_{n-1} + 2^{n-3} - a_{n-3}$

Q: What are the initial conditions:

Recurrence Relations for Counting

A: Need to give enough initial conditions to avoid ensure well-definedness. The smallest n for which length is well defined is n=0. Thus the smallest n for which $a_n = 2a_{n-1} + 2^{n-3} - a_{n-3}$ makes sense is n=3. Thus need to give a_0, a_1 and a_2 explicitly.

$a_0 = a_1 = 0$ (strings to short to contain 00)

$a_2 = 1$ (must be 00).

Note: example 6 on p. 313 gives the simpler recursion relation $b_n = b_{n-1} + b_{n-2}$ for strings which do ***not*** contain two consecutive 0's.

Financial Recursion Relation

Most savings plans satisfy certain recursion relations.

Q: Consider a savings plan in which \$10 is deposited per month, and a 6%/year interest rate given with payments made every month. If P_n represents the amount in the account after *n* months, find a recurrence relation for P_n.

Financial Recursion Relation

A: $P_n = (1+r)\cdot P_{n-1} + 10$

where $r = 1 + 6\%/12 = 1.005$

Partition Function

A ***partition*** of a set is a way of grouping all the elements disjointly.

EG: All the partitions of {1,2,3} are:

1. { {1,2,3} }
2. { {1,2}, {3} }
3. { {1,3}, {2} }
4. { {2,3}, {1} }
5. { {1},{2},{3} }

The partition function p_n counts the number of partitions of $\{1,2,\ldots,n\}$. Thus $p_3 = 5$.

Partition Function

Let's find a recursion relation for the partition function. There are *n* possible scenarios for the number of members on *n*'s team:

0: *n* is all by itself ☹ (e.g. {{1,2},{3}})

1: *n* has 1 friend (e.g. {{1},{2,3}})

2: *n* has 2 friends (e.g. {{1,2,3}})

…

n-1: *n* has *n*-1 friends on its team.

By the sum rule, we need to count the number of partitions of each kind, and then add together.

Partition Function

Consider the *k'* th case:

k : *n* has *k* friends

There are *C* (*n*-1,*k*) ways of choosing fellow members of *n*'s team.

Furthermore, there are p_{n-k-1} ways of partitioning the rest of the *n* elements. Using the product and sum rules we get:

$$p_n = \sum_{i=0}^{n-1} p_i \cdot C(n-1, n-1-i) =$$

$$p_0 \cdot C(n-1, n-1) + \cdots + p_{n-1} \cdot C(n-1, 0)$$

Solving Recurrence Relations

We will learn how to give closed solutions to certain kinds of recurrence relations. Unfortunately, most recurrence relations cannot be solved analytically.

EG: If you can find a closed formula for partition function tell me!

However, recurrence relations can all be solved quickly by using *dynamic programming*.

Closed Solutions by Telescoping

We've already seen technique in the past:

1) Plug recurrence into itself repeatedly for smaller and smaller values of *n*.
2) See the pattern and then give closed formula in terms of initial conditions.
3) Plug values into initial conditions getting final formula.

Telescoping also called ***back-substitution***

Telescope Example

Find a closed solution to $a_n = 2a_{n-1}$, $a_0 = 3$:

$a_n = 2a_{n-1}$

Telescope Example

Find a closed solution to $a_n = 2a_{n-1}$, $a_0 = 3$:

$$a_n = 2a_{n-1} = 2^2 a_{n-2}$$

Telescope Example

Find a closed solution to $a_n = 2a_{n-1}$, $a_0 = 3$:

$$a_n = 2a_{n-1} = 2^2 a_{n-2} = 2^3 a_{n-3}$$

Telescope Example

Find a closed solution to $a_n = 2a_{n-1}$, $a_0 = 3$:

$$a_n = 2a_{n-1} = 2^2 a_{n-2} = 2^3 a_{n-3} = \ldots$$

Telescope Example

Find a closed solution to $a_n = 2a_{n-1}$, $a_0 = 3$:

$$a_n = 2a_{n-1} = 2^2 a_{n-2} = 2^3 a_{n-3} = \ldots = 2^i a_{n-i}$$

Telescope Example

Find a closed solution to $a_n = 2a_{n-1}$, $a_0 = 3$:

$$a_n = 2a_{n-1} = 2^2 a_{n-2} = 2^3 a_{n-3} = \ldots \quad = 2^i a_{n-i} = \ldots$$

Telescope Example

Find a closed solution to $a_n = 2a_{n-1}$, $a_0 = 3$:

$$a_n = 2a_{n-1} = 2^2 a_{n-2} = 2^3 a_{n-3} = \ldots = 2^i a_{n-i} = \ldots = 2^n a_0$$

Telescope Example

Find a closed solution to $a_n = 2a_{n-1}$, $a_0 = 3$:

$$a_n = 2a_{n-1} = 2^2 a_{n-2} = 2^3 a_{n-3} = \ldots = 2^i a_{n-i} = \ldots = 2^n a_0$$

Plug in $a_0 = 3$ for final answer:

$$a_n = 3 \cdot 2^n$$

Blackboard Exercise for 5.1

- 5.1.21: Give a recurrence relation for the number of ways to climb *n* stairs if the climber can take one or two stairs at a time.

Closed Solutions by Telescoping

The only case for which telescoping works with a high probability is when the recurrence give the next value in terms of a single previous value.

There is a class of recurrence relations which *can* be solved analytically in general. These are called *linear recurrences* and include the Fibonacci recurrence.

Linear Recurrences

The only case for which telescoping works with a high probability is when the recurrence gives the next value in terms of a single previous value. But…

There is a class of recurrence relations which *can* be solved analytically in general. These are called *linear recurrences* and include the Fibonacci recurrence.

Begin by showing how to solve Fibonacci:

Solving Fibonacci

Recipe solution has 3 basic steps:

1) Assume solution of the form $a_n = r^n$
2) Find all possible r's that seem to make this work. Call these[1] r_1 and r_2. Modify assumed solution to ***general solution*** $a_n = Ar_1^n + Br_2^n$ where A,B are constants.
3) Use initial conditions to find A,B and obtain specific solution.

Solving Fibonacci

1) Assume exponential solution of the form $a_n = r^n$:

Plug this into $a_n = a_{n-1} + a_{n-2}$:

$$r^n = r^{n-1} + r^{n-2}$$

Notice that all three terms have a common r^{n-2} factor, so divide this out:

$$r^n / r^{n-2} = (r^{n-1} + r^{n-2}) / r^{n-2} \rightarrow r^2 = r + 1$$

This equation is called the ***characteristic equation*** of the recurrence relation.

Solving Fibonacci

2) Find all possible r's that solve characteristic

$$r^2 = r + 1$$

Call these r_1 and r_2.[1] General solution is $a_n = Ar_1^n + Br_2^n$ where A,B are constants.

Quadratic formula[2] gives:

$$r = (1 \pm \sqrt{5})/2$$

So $r_1 = (1+\sqrt{5})/2$, $r_2 = (1-\sqrt{5})/2$

General solution:

$$a_n = A\,[(1+\sqrt{5})/2]^n + B\,[(1-\sqrt{5})/2]^n$$

Solving Fibonacci

3) Use initial conditions $a_0 = 0$, $a_1 = 1$ to find A,B and obtain specific solution.

$0 = a_0 = A\,[(1+\sqrt{5})/2]^0 + B\,[(1-\sqrt{5})/2]^0 = A + B$

$1 = a_1 = A\,[(1+\sqrt{5})/2]^1 + B\,[(1-\sqrt{5})/2]^1 =$
$A(1+\sqrt{5})/2 + B\,(1-\sqrt{5})/2 = (A+B)/2 + (A-B)\sqrt{5}/2$

First equation give $B = -A$. Plug into 2nd:

$1 = 0 + 2A\sqrt{5}/2$ so $A = 1/\sqrt{5}$, $B = -1/\sqrt{5}$

Final answer:

(CHECK IT!) $$a_n = \frac{1}{\sqrt{5}}\left(\frac{1+\sqrt{5}}{2}\right)^n - \frac{1}{\sqrt{5}}\left(\frac{1-\sqrt{5}}{2}\right)^n$$

Linear Recurrences with Constant Coefficients

Previous method generalizes to solving "*linear recurrence relations with constant coefficients*":

DEF: A recurrence relation is said to be ***linear*** if a_n is a linear combination of the previous terms plus a function of n. I.e. no squares, cubes or other complicated function of the previous a_i can occur. If in addition all the coefficients are constants then the recurrence relation is said to have ***constant coefficients***.

Linear Recurrences with Constant Coefficients

Q: Which of the following are linear with constant coefficients?

1. $a_n = 2a_{n-1}$
2. $a_n = 2a_{n-1} + 2^{n-3} - a_{n-3}$
3. $a_n = a_{n-1}^2$
4. Partition function:

$$p_n = \sum_{i=0}^{n-1} p_i \cdot C(n-1, n-1-i)$$

Linear Recurrences with Constant Coefficients

A:

1. $a_n = 2a_{n-1}$: YES
2. $a_n = 2a_{n-1} + 2^{n-3} - a_{n-3}$: YES
3. $a_n = a_{n-1}^2$: NO. Squaring is not a linear operation. Similarly $a_n = a_{n-1}a_{n-2}$ and $a_n = \cos(a_{n-2})$ are non-linear.
4. Partition function:

NO. This is linear, but coefficients are not constant as $C(n-1, n-1-i)$ is a non-constant function of n.

$$p_n = \sum_{i=0}^{n-1} p_i \cdot C(n-1, n-1-i)$$

Homogeneous Linear Recurrences

To solve such recurrences we must first know how to solve an easier type of recurrence relation:

DEF: A linear recurrence relation is said to be ***homogeneous*** if it is a linear combination of the previous terms of the recurrence *without* an additional function of *n*.

Q: Which of the following are homogeneous?

1. $a_n = 2a_{n-1}$
2. $a_n = 2a_{n-1} + 2^{n-3} - a_{n-3}$
3. Partition function: $p_n = \sum_{i=0}^{n-1} p_i \cdot C(n-1, n-1-i)$

Linear Recurrences with Constant Coefficients

A:

1. $a_n = 2a_{n-1}$: YES
2. $a_n = 2a_{n-1} + 2^{n-3} - a_{n-3}$: No. There's an extra term $f(n) = 2^{n-3}$
3. Partition function:

YES. No terms appear not involving the previous p_i

$$p_n = \sum_{i=0}^{n-1} p_i \cdot C(n-1, n-1-i)$$

Homogeneous Linear Recurrences with Const. Coeff.'s

The 3-step process used for the Fibonacci recurrence works well for general homogeneous linear recurrence relations with constant coefficients. There are a few instances where some modification is necessary.

Homogeneous -Complications

1) *Repeating roots* in characteristic equation. Repeating roots imply that don't learn anything new from second root, so may not have enough information to solve formula with given initial conditions. We'll see how to deal with this on next slide.
2) *Non-real number roots* in characteristic equation. If the sequence has periodic behavior, may get complex roots (for example $a_n = -a_{n-2}$)[1]. We won't worry about this case (in principle, same method works as before, except use complex arithmetic).

Complication: Repeating Roots

EG: Solve $a_n = 2a_{n-1} - a_{n-2}$, $a_0 = 1$, $a_1 = 2$

Find characteristic equation by plugging in $a_n = r^n$:

$$r^2 - 2r + 1 = 0$$

Since $r^2 - 2r + 1 = (r-1)^2$ the root $r = 1$ repeats.

If we tried to solve by using general solution

$$a_n = Ar_1^n + Br_2^n = A1^n + B1^n = A + B$$

which forces a_n to be a constant function (→←).

SOLUTION: Multiply second solution by n so general solution looks like:

$$a_n = Ar_1^n + Bnr_1^n$$

Complication: Repeating Roots

Solve $a_n = 2a_{n-1} - a_{n-2}$, $a_0 = 1$, $a_1 = 2$

General solution: $a_n = A1^n + Bn1^n = A + Bn$

Plug into initial conditions

$1 = a_0 = A + B \cdot 0 \cdot 1^0 = A$

$2 = a_0 = A \cdot 1^1 + B \cdot 1 \cdot 1^1 = A + B$

Plugging first equation $A = 1$ into second:

$2 = 1 + B$ implies $B = 1$.

Final answer: $a_n = 1 + n$

(CHECK IT!)

The Nonhomogeneous Case

Consider the Tower of Hanoi recurrence (see Rosen p. 311-313) $a_n = 2a_{n-1}+1$.

Could solve using telescoping. Instead let's solve it methodically. Rewrite:

$$a_n - 2a_{n-1} = 1$$

1) Solve with the RHS set to 0, i.e. solve the homogeneous case.
2) Add a particular solution to get general solution. I.e. use rule:

General Nonhomogeneous	=	General homogeneous	+	Particular Nonhomogeneous

The Nonhomogeneous Case

$$a_n - 2a_{n-1} = 1$$

1) Solve with the RHS set to 0, i.e. solve

$$a_n - 2a_{n-1} = 0$$

Characteristic equation: $r - 2 = 0$

so unique root is $r = 2$. General solution to homogeneous equation is

$$a_n = A \cdot 2^n$$

The Nonhomogeneous Case

2) Add a particular solution to get general solution for $a_n - 2a_{n-1} = 1$. Use rule:

There are little tricks for guessing particular nonhomogeneous solutions. For example, when the RHS is constant, the guess should also be a constant.[1]

So guess a particular solution of the form $b_n = C$.

Plug into the original recursion:

$1 = b_n - 2b_{n-1} = C - 2C = -C$. Therefore $C = -1$.

General solution: $a_n = A \cdot 2^n - 1$.

The Nonhomogeneous Case

Finally, use initial conditions to get closed solution. In the case of the Towers of Hanoi recursion, initial condition is:

$$a_1 = 1$$

Using general solution $a_n = A \cdot 2^n - 1$ we get:

$1 = a_1 = A \cdot 2^1 - 1 = 2A - 1$.

Therefore, $2 = 2A$, so $A = 1$.

Final answer: $a_n = 2^n - 1$

More Complicated

EG: Find the general solution to recurrence from the bit strings example: $a_n = 2a_{n-1} + 2^{n-3} - a_{n-3}$

1) Rewrite as $a_n - 2a_{n-1} + a_{n-3} = 2^{n-3}$ and solve homogeneous part:

Characteristic equation: $r^3 - 2r + 1 = 0$.

Guess root $r = \pm 1$ as integer roots divide.

$r = 1$ works, so divide out by $(r - 1)$ to get

$r^3 - 2r + 1 = (r - 1)(r^2 + r - 1)$.

More Complicated

$r^3 - 2r + 1 = (r - 1)(r^2 + r - 1)$.

Quadratic formula on $r^2 + r - 1$:

$$r = (-1 \pm \sqrt{5})/2$$

So $r_1 = 1$, $r_2 = (-1+\sqrt{5})/2$, $r_3 = (-1-\sqrt{5})/2$

General homogeneous solution:

$$a_n = A + B\,[(-1+\sqrt{5})/2]^n + C\,[(-1-\sqrt{5})/2]^n$$

More Complicated

2) Nonhomogeneous particular solution to $a_n - 2a_{n-1} + a_{n-3} = 2^{n-3}$

Guess the form $b_n = k\,2^n$. Plug guess in:

$$k\,2^n - 2k\,2^{n-1} + k\,2^{n-3} = 2^{n-3}$$

Simplifies to: $k = 1$.

So particular solution is $b_n = 2^n$

Final answer:

$a_n = A + B\,[(-1+\sqrt{5})/2]^n + C\,[(-1-\sqrt{5})/2]^n + 2^n$

General Nonhomogeneous	=	General homogeneous	+	Particular Nonhomogeneous

Blackboard Exercise for 5.2

- Solve the following recurrence relation in terms of a_1 assuming n is odd:

$$a_n = (n-1)a_{n-2}$$

BBM 205 Discrete Mathematics
Hacettepe University
http://web.cs.hacettepe.edu.tr/∼bbm205

Lecture 7: Graph Terminology and Graph Isomorphism

Lecturer: Lale Özkahya

Resources:
Kenneth Rosen, "Discrete Mathematics and App."
http://www.inf.ed.ac.uk/teaching/courses/dmmr
http://www.cs.nthu.edu.tw/ wkhon/math16.html

What is a Graph ?

A graph consists of a nonempty set V of vertices and a set E of edges, where each edge in E connects two (may be the same) vertices in V.

- Let G be a graph associated with a vertex set V and an edge set E

  We usually write G = (V, E) to indicate the above relationship

Examples

- Furthermore, if each edge connects two different vertices, and no two edges connect the same pair of vertices, then the graph is a simple graph
- Which of the above is a simple graph ?

Directed Graph

- Sometimes, we may want to specify a direction on each edge

  Example : Vertices may represent cities, and edges may represent roads (can be one-way)

- This gives the directed graph as follows :

A directed graph G consists of a nonempty set V of vertices and a set E of directed edges, where each edge is associated with an ordered pair of vertices. We write G = (V, E) to denote the graph.

Examples

Test Your Understanding

- Suppose we have a simple graph G with n vertices

  What is the maximum number of edges G can contain, if

  (i) G is an undirected graph ?

  (ii) G is a directed graph ?

Terminology (Undirected Graph)

- Let e be an edge that connects vertices u and v

  We say (i) e is incident with u and v

  (ii) u and v are the endpoints of e ;

  (iii) u and v are adjacent (or neighbors)

  (iv) if u = v, the edge e is called a loop

- The degree of a vertex v, denoted by deg(v), is the number of edges incident with v, except that a loop at v contributes twice to the degree of v

Example

- What are the degrees and neighbors of each vertex in the following graph ?

Handshaking Theorem

- Let G = (V, E) be an undirected graph with m edges

Theorem:

$$\sum_{v \in V} \deg(v) = 2m$$

- Proof : Each edge e contributes exactly twice to the sum on the left side (one to each endpoint).

Corollary : An undirected graph has an even number of vertices of odd degree.

Terminology (Directed Graph)

- Let e be an edge that connects vertices from u to v

  We say (i) u = initial vertex, v = terminal vertex ;

  (ii) u is adjacent to v;

  (iii) v is adjacent from u;

  (iv) if u = v, the edge e is called a loop
- The in-degree of a vertex v, denoted by $\deg^{-}(v)$, is the number of edges with v as terminal vertex
- The out-degree of a vertex u, denoted by $\deg^{+}(u)$, is the number of edges with u as initial vertex

Example

- What are the in- and out-degrees of each vertex in the following graph ?

Handshaking Theorem

- Let G = (V, E) be directed graph with m edges

Theorem:

$$\sum_{v \in V} \deg^{-}(v) = \sum_{u \in V} \deg^{+}(u) = m$$

- Proof : Each edge e contributes exactly once to the in-degree and once to the out-degree

Some Special Simple Graphs

Definition: A complete graph on n vertices, denoted by K_n, is a simple graph that contains one edge between each pair of distinct vertices

Examples :

K_1 K_2 K_3 K_4 K_5

Some Special Simple Graphs

Definition: A cycle C_n, $n \geq 3$, is a graph that consists of n vertices $v_1, v_2, \ldots, v_n$ and n edges $\{v_1, v_2\}, \{v_2, v_3\}, \ldots, \{v_{n-1}, v_n\}, \{v_n, v_1\}$

Examples :

C_3 C_4 C_5 C_6

Some Special Simple Graphs

Definition: A wheel W_n, $n \geq 3$, is a graph that consists of a cycle C_n with an extra vertex that connects to each vertex in C_n

Examples :

W_3 W_4 W_5 W_6

Some Special Simple Graphs

Definition: An n-cube, denoted by Q_n, is a graph that consists of 2^n vertices, each representing a distinct n-bit string. An edge exists between two vertices ⇔ the corresponding strings differ in exactly one bit position.

Examples :

Some Special Simple Graphs

Definition: A bipartite graph is a graph such that the vertices can be partitioned into two sets V and W, so that each edge has exactly one endpoint from V, and one endpoint from W

Examples :

bipartite graphs

non-bipartite graphs

Some Special Simple Graphs

- Which of the following is a bipartite graph?

Check if a Graph is Bipartite

- The following is a very useful theorem :

Theorem: A simple graph is bipartite if and only if it is possible to assign one of two different colors to each vertex, so that no two adjacent vertices are assigned the same color

- Proof : If there is a way to color the vertices, the same way shows a possible partition of vertices. Conversely, if there is a way to partition the vertices, the same way gives a possible coloring.

Check if a Graph is Bipartite

- The above implies the following algorithm to check if a connected graph is bipartite :

Step 1 : Pick a vertex u. Color it with white ;

Step 2 : While there are uncolored vertices

(i) for each neighbor of a white vertex,

color it with black ;

(ii) for each neighbor of a black vertex,

color it with white ;

Step 3 : Report YES if each edge is colored properly.

Else, report NO ;

Some Special Simple Graphs

Definition: A complete bipartite graph $K_{m,n}$ is a bipartite graph with vertices partitioned into two subsets V and W of size m and n, respectively, such that there is an edge between each vertex in V and each vertex in W

Examples :

$K_{2,2}$ $K_{3,2}$ $K_{3,3}$

Subgraphs and Complements

If G = (V, E) is a graph, then G' = (V', E') is called a subgraph of G if $V' \subseteq V$ and $E' \subseteq E$.

- Which one is a subgraph of the leftmost graph G ?

Subgraphs and Complements

If G = (V, E) is a graph, then the subgraph of G induced by U $\subseteq$ V is a graph with the vertex set U and contains exactly those edges from G with both endpoints from U

Ex : Consider the graph on the right side

What is its subgraph induced by the vertex set { a, b, c, g } ?

Subgraphs and Complements

If G = (V, E) is a graph, then the complement of G, denoted by $\overline{G}$, is a graph with the same vertex set, such that

an edge e exists in $\overline{G}$ ⇔ e does not exist in G

Ex : Consider the graph on the right side

What is its complement ?

New Graphs from Old

Definition: The *union* of two simple graphs $G1 = (V1, E1)$ and $G2 = (V2, E2)$ is the simple graph with vertex set $V1 \bigcup V2$ and edge set $E1 \bigcup E2$. The union of $G1$ and $G2$ is denoted by $G1 \bigcup G2$.

Example:

Representing Graphs: Adjacency Lists

Definition: An *adjacency list* represents a graph (with no multiple edges) by specifying the vertices that are adjacent to each vertex.

Example:

TABLE 1 An Adjacency List for a Simple Graph.

Vertex	Adjacent Vertices
a	b, c, e
b	a
c	a, d, e
d	c, e
e	a, c, d

Example:

TABLE 2 An Adjacency List for a Directed Graph.

Initial Vertex	Terminal Vertices
a	b, c, d, e
b	b, d
c	a, c, e
d	
e	b, c, d

Representation of Graphs: Adjacency Matrices

Definition: Suppose that $G = (V, E)$ is a simple graph where $|V| = n$. Arbitrarily list the vertices of G as $v_1, v_2, \ldots, v_n$.

The *adjacency matrix*, **A**, of G, with respect to this listing of vertices, is the $n \times n$ 0-1 matrix with its (i, j)th entry = 1 when v_i and v_j are adjacent, and =0 when they are not adjacent.

- In other words: $A = [a_{ij}]$ and:

$$a_{ij} = \begin{cases} 1 & \text{if } \{v_i, v_j\} \text{ is an edge of } G, \\ 0 & \text{otherwise.} \end{cases}$$

Adjacency Matrices (*continued*)

Example: *The vertex ordering is is a, b, c, d.*

$$\begin{bmatrix} 0 & 1 & 1 & 1 \\ 1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix}$$

$$\begin{bmatrix} 0 & 1 & 1 & 0 \\ 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 \end{bmatrix}$$

A sparse graph has few edges relative to the number of possible edges. Sparse graphs are more efficient to represent using an adjacency list than an adjacency matrix. But for a dense graph, an adjacency matrix is often preferable.

Note: The adjacency matrix of an undirected graph is symmetric:

$$a_{ij} = a_{ji}, \forall i, j$$

Also, since there are no loops, each diagonal entry *is zero:* $a_{ii} = 0, \forall i$

Adjacency Matrices (*continued*)

- Adjacency matrices can also be used to represent graphs with loops and multi-edges.
- When multiple edges connect vertices *vi* and *vj*, (or if multiple loops present at the same vertex), the (*i*, *j*)th entry equals the number of edges connecting the pair of vertices.

Example: Adjacency matrix of a pseudograph, using vertex ordering *a*, *b*, *c*, *d:*

$$\begin{bmatrix} 0 & 3 & 0 & 2 \\ 3 & 0 & 1 & 1 \\ 0 & 1 & 1 & 2 \\ 2 & 1 & 2 & 0 \end{bmatrix}$$

Adjacency Matrices (*continued*)

- Adjacency matrices can represent directed graphs in exactly the same way. The matrix A for a directed graph $G = (V, E)$ has a 1 in its (i, j)th position if there is an edge from vi to vj, where $v1$, $v2$, ... vn is a list of the vertices.
  - In other words,

$$a_{ij} = 1 \quad if \quad (i, j) \in E$$
$$a_{ij} = 0 \quad if \quad (i, j) \notin E$$

- Note: the adjacency matrix for a directed graph need not be symmetric.

Graph Isomorphism

Graphs G = (V, E) and H = (U, F) are isomorphic if we can set up a bijection f : V → U such that

x and y are adjacent in G

⇔ f(x) and f(y) are adjacent in H

Ex : The following are isomorphic to each other :

Isomorphism of Graphs (*cont.*)

Example: Show that the graphs $G = (V, E)$ and $H = (W, F)$ are isomorphic.

Solution: The function f with $f(u1) = v1$, $f(u2) = v4$, $f(u3) = v3$, and $f(u4) = v2$ is a one-to-one correspondence between V and W.

Isomorphism of Graphs (*cont.*)

It is difficult to determine whether two graphs are isomorphic by brute force: there are $n!$ bijections between vertices of two n-vertex graphs.

Often, we can show two graphs are not isomorphic by finding a property that only one of the two graphs has. Such a property is called *graph invariant*:

- e.g., number of vertices of given degree, the degree sequence (list of the degrees), ...

Isomorphism of Graphs (*cont.*)

Example: Are these graphs are isomorphic?

Solution: No! Since *deg*(*a*) = 2 in *G*, *a* must correspond to *t*, *u*, *x*, or *y,* since these are the vertices of degree 2 in H. But each of these vertices is adjacent to another vertex of degree 2 in *H*, which is not true for *a* in *G*. So, G and H can not be isomorphic.

Isomorphism of Graphs (*cont.*)

Example: Determine whether these two graphs are isomorphic.

Solution: The function f is defined by: $f(u1) = v6$, $f(u2) = v3$, $f(u3) = v4$, $f(u4) = v5$, $f(u5) = v1$, and $f(u6) = v2$ is a bijection.

Algorithms for Graph Isomorphism

- The best algorithms known for determining whether two graphs are isomorphic have exponential worst-case time complexity (in the number of vertices of the graphs).
- However, there are algorithms with good time complexity in many practical cases.
- See, e.g., a publicly available software called NAUTY for graph isomorphism.

Applications of Graph Isomorphism

The question whether graphs are isomorphic plays an important role in applications of graph theory. For example:

Chemists use molecular graphs to model chemical compounds. Vertices represent atoms and edges represent chemical bonds. When a new compound is synthesized, a database of molecular graphs is checked to determine whether the new compound is isomorphic to the graph of an already known one.

Graph Isomorphism

The following graphs are isomorphic to each other. This graph is known as the Petersen graph :

Graph Isomorphism

How to show the following are not isomorphic ?

How to show the following are not isomorphic ?

BBM 205 Discrete Mathematics
Hacettepe University
http://web.cs.hacettepe.edu.tr/∼bbm205

Lecture 8: Connectivity, Euler and Hamilton Paths, Graph Coloring, Planar Graphs

Lecturer: Lale Özkahya

Resources:
Kenneth Rosen, "Discrete Mathematics and App."
http://www.inf.ed.ac.uk/teaching/courses/dmmr
http://www.cs.nthu.edu.tw/ wkhon/math16.html

What is a Path ?

A path is a sequence of edges that begins at a vertex, and travels from vertex to vertex along edges of the graph. The number of edges on the path is called the length of the path.

- Ex : Consider the graph on the right.
  w → x → y → z → x corresponds to a path of length 4

What is a Path ?

If a path begins and ends at the same vertex, the path is also called a circuit.

- Ex : Consider the graph on the right.

  w → x → y → z → w

  gives to a circuit of length 4

Paths and Isomorphism

Q: How to show that the following graphs are not isomorphic ?

A: One contains a circuit of length 3 (a triangle), while the other does not

Paths and Connected Components

An undirected graph is connected if there is a path between any pair of vertices. Otherwise, it is disconnected.

- Ex :

Connected Disconnected

Paths and Connected Components

A connected subgraph is a connected component if it is not contained in any other connected subgraphs.

- Ex :

1 connected component

3 connected components

Euler Paths and Circuits

- The above is a map of a Prussian city called Königsberg during the 18th century

Euler Paths and Circuits

- The Pregel River (blue part) divides the city into 4 parts : Two sides and two large islands

Euler Paths and Circuits

- Seven bridges connect the sides with the islands
- Can we start at some location, travel each bridge exactly once, and go back to the same location ?

Euler Paths and Circuits

- Euler first represents the four parts and the seven bridges by a graph shown on the right

  ➔ The problem will be equivalent to :

  Find a circuit that travels each edge exactly once

- Euler shows that there is NO such circuit

Euler Paths and Circuits

Definition : An Euler path in a graph is a path that contains each edge exactly once. If such a path is also a circuit, it is called an Euler circuit.

- Ex :

Euler path | Euler circuit

Euler Paths and Circuits

Theorem : A connected graph G has an Euler circuit ⇔ each vertex of G has even degree.

- Proof : [The "only if" case]

  If the graph has an Euler circuit, then when we walk along the edges according to this circuit, each vertex must be entered and exited the same number of times.

  Thus, the degree of each vertex must be even.

Euler Paths and Circuits

- Proof : [The "if" case]

  If each vertex has an even degree, we shall use induction (on the number of edges) to show that an Euler circuit exists.

  (Basis) When there is one edge, it must be a self-loop ➔ An Euler circuit exists.

  (Inductive) We start at a vertex x, and obtain a path without using any edge twice, until we end at a vertex without any more unused edge to travel ➔ This vertex must be x (why?)

Euler Paths and Circuits

- Proof : [The "if" case (continued)]

  Let C denote the above circuit.

  If we remove C from the graph, the degree of each vertex must still be even (why?). Further, each connected component with edges must share some vertex u with C, and has an Euler circuit C' (why?)

  ➔ We get an Euler circuit of the original graph, by walking on C until vertex u, then edges on C', then back to u, and the remaining edges on C

Euler Paths and Circuits

- Example on obtaining an Euler circuit :

Step 1:
Getting a circuit C by starting from a vertex x

Step 2:
Getting C' for each remaining component

Step 3:
Combining C and the C' of each component

Euler Paths and Circuits

Corollary : A connected graph G has an Euler path, but no Euler circuits ⇔ exactly two vertices of G has odd degree.

- Proof : [The "only if" case] The degree of the starting and ending vertices of the Euler path must be odd, and all the others must be even.

  [The "if" case] Let u and v be the vertices with odd degrees. Adding an edge between u and v will produce an Euler circuit ➔ Removal of this edge thus implies an Euler path in the graph

Hamilton Paths and Circuits

- The above is a regular dodecahedron (12-faced) with each vertex labeled with the name of a city

Hamilton Paths and Circuits

- Can we find a circuit (travelling along the edges) so that each city is visited exactly once ?

Hamilton Paths and Circuits

- The right graph is isomorphic to the dodecahedron, and it shows a possible way (in red) to travel

Hamilton Paths and Circuits

Definition : A Hamilton path in a graph is a path that visits each vertex exactly once. If such a path is also a circuit, it is called a Hamilton circuit.

- Ex :

Hamilton path | Hamilton circuit

Hamilton Paths and Circuits

- Which of the following have a Hamilton circuit or, if not, a Hamilton path ?

Hamilton Paths and Circuits

- Show that the n-dimensional cube Q_n has a Hamilton circuit, whenever $n \geq 2$

- Ex :

Q_2

Q_3

Hamilton Paths and Circuits

- Unlike Euler circuit or Euler path, there is no efficient way to determine if a graph contains a Hamilton circuit or a Hamilton path
  - ➔ The best algorithm so far requires exponential time in the worst case
- However, it is shown that when the degree of the vertices are sufficiently large, the graph will always contain a Hamilton circuit
  - ➔ We shall discuss two theorems in this form

Hamilton Paths and Circuits

- Before we give the two theorems, we show an interesting theorem by Bondy and Chvátal (1976)

  ➔ The two theorems will then become corollaries of Bondy-Chvátal theorem

- Let G be a graph with n vertices

Definition : The Hamilton closure of G is a simple graph obtained by recursively adding an edge between two vertices u and v, whenever

$$\deg(u) + \deg(v) \geq n$$

Hamilton Paths and Circuits

- Ex :

G Hamilton closure

- Ex :

G Hamilton closure

Hamilton Paths and Circuits

- Ex :

Hamilton closure

Hamilton Paths and Circuits

Theorem [Bondy and Chvátal (1976)] :

A graph G contains a Hamilton circuit $\Leftrightarrow$
its Hamilton closure contains a Hamilton circuit

- The "only if" case is trivial
- For the "if" case, we can prove it by contradiction
- However, we shall give the proof a bit later, as we are now ready to talk about the two corollaries

Hamilton Paths and Circuits

- Let G be a simple graph with $n \geq 3$ vertices

Corollary [Dirac (1952)] :

If the degree of each vertex in G is at least n/2, then G contains a Hamilton circuit

Corollary [Ore (1960)] :

If for any pair of non-adjacent vertices u and v,

$$\deg(u) + \deg(v) \geq n,$$

then G contains a Hamilton circuit

Hamilton Paths and Circuits

- Proof of Dirac's and Ore's Theorems :

It is easy to verify that

(i) if the degree of each vertex is at least n/2, or

(ii) if for any pair of non-adjacent vertices u and v,

$$\deg(u) + \deg(v) \geq n,$$

➔ G's Hamilton closure is a complete graph K_n

➔ When $n \geq 3$, K_n has a Hamilton circuit

➔ Bondy-Chvátal implies that there will be a Hamilton circuit in G

Hamilton Paths and Circuits

- Next, we shall give the proof of the “if case” of Bondy-Chvátal’s Theorem
- Proof (“if case”):

  Suppose on the contrary that

  (i) G does not have a Hamilton circuit, but

  (ii) G’s Hamilton closure has a Hamilton circuit.

  Then, consider the sequence of graphs obtained by adding one edge each time when we produce the Hamilton closure from G

Hamilton Paths and Circuits

- Proof ("if case" continued):

G
(no circuit)

Hamilton Closure
(has a circuit)

Hamilton Paths and Circuits

- Proof (“if case” continued):

  Let G’ be the first graph in the sequence that contains a Hamilton circuit

  Let { u, v } be the edge added to produce G’

Hamilton Paths and Circuits

- Proof ("if case" continued):

  Now, we show that the graph before G' must also contain a Hamilton circuit, which immediately will cause a contradiction.

  Consider the graph before adding { u, v } to G'.

  It must contain a Hamilton path from u to v (why?)

Hamilton Paths and Circuits

- Proof ("if case" continued):

  Also, since we are connecting u and v in G',

  $$\deg(u) + \deg(v) \geq n$$

  Consider all the nodes connected by u, and we mark their 'left' neighbors in red

Hamilton Paths and Circuits

- Proof ("if case" continued):

  Since

  (i) v does not connect to u nor itself, and

  (ii) deg(u) + deg(v) $\geq$ n

  ➔ v must connect to some red node (why?)

Hamilton Paths and Circuits

- Proof ("if case" continued):

➔ We get a Hamilton circuit, even without connecting u and v !

➔ This contradicts with the choice of G', and the theorem is thus correct

Graph Colouring

Suppose we have k distinct colours with which to colour the vertices of a graph. Let $[k] = \{1, \ldots, k\}$. For an undirected graph, $G = (V, E)$, an admissible vertex **k-colouring** of G is a function $c : V \to [k]$, such that for all $u, v \in V$, if $\{u, v\} \in E$ then $c(u) \neq c(v)$.

For an integer $k \geq 1$, we say an undirected graph $G = (V, E)$ is **k-colourable** if there exists a k-colouring of G.

The **chromatic number** of G, denoted $\chi(G)$, is the *smallest positive integer* k, such that G is k-colourable.

Some observations about Graph colouring

- Note that any graph G with n vertices in n-colourable.
- The ***n*-Clique**, K_n, i.e., the complete graph on n vertices, has chromatic number $\chi(K_n) = n$. All its vertices must get assigned different colours in any admissible colouring.
- The clique number, $\omega(G)$, of a graph G is the maximum positive integer $r \geq 1$, such that K_r is a subgraph of G.
- Note that for all graphs G, $\omega(G) \leq \chi(G)$: if G has an r-clique then it is not $(r-1)$-colorable.
- However, in general, $\omega(G) \neq \chi(G)$. For instance, The 5-cycle, C_5, has $\omega(C_5) = 2 < \chi(C_5) = 3$.

More observations about colouring

- As already mentioned, any bipartite graph is 2-colourable. Indeed, that is an equivalent definition of being bipartite.
- More generally, a graph G is k-colourable precisely if it is k-partite, meaning its vertices can be partitioned into k disjoint sets such that all edges of the graph are between nodes in different parts.

Algorithms/complexity of colouring graphs

To determine whether a n-vertex graph $G = (V, E)$ is k-colourable by "*brute force*", we could try all possible colourings of n nodes with k colours.

Difficulty: There are k^n such k-colouring functions $c : V \to [k]$.

Question: Is there an efficient (polynomial time) algorithm for determining whether a given graph G is k-colourable?

Algorithms/complexity of colouring graphs

To determine whether a n-vertex graph $G = (V, E)$ is k-colourable by "*brute force*", we could try all possible colourings of n nodes with k colours.

Difficulty: There are k^n such k-colouring functions $c : V \to [k]$.

Question: Is there an efficient (polynomial time) algorithm for determining whether a given graph G is k-colourable?

Answer: No, no generally efficient (polynomial time) algorithm is known, and even the problem of determining whether a given graph is 3-colourable is **NP-complete**. (Even approximating the chromatic number of a given graph is NP-hard.)

In practice, there are hueristic algorithms that do obtain good colourings for many classes of graphs.

Applications of Graph Colouring (many)

Final Exam Scheduling

- There are n courses, $\{1, \ldots, n\}$.
- Some courses have the same students registered for both, so their exams can't be scheduled at the same time.
- Let $G = (\{1, \ldots, n\}, E)$ be a graph such that $\{i, j\} \in E$ if and only if $i \neq j$ and courses i and j have a student in common.
- **Question:** What is the minimum number of exam time slots needed to schedule all n exams?
- **Answer:** This is precisely the chromatic number $\chi(G)$ of G.

  Furthermore, a k-colouring of G yields an *admissible schedule* of exams into k time slots, allowing all students to attend all their exams, as long as different "colors" are scheduled in disjoint time slots.

What is a Planar Graph ?

Definition : A planar graph is an undirected graph that can be drawn on a plane without any edges crossing. Such a drawing is called a planar representation of the graph in the plane.

- Ex : K_4 is a planar graph

Examples of Planar Graphs

- Ex : Other planar representations of K_4

Examples of Planar Graphs

- Ex : Q_3 is a planar graph

Examples of Planar Graphs

- Ex : $K_{1,n}$ and $K_{2,n}$ are planar graphs for all n

$K_{1,5}$

$K_{2,4}$

Euler's Planar Formula

Definition : A planar representation of a graph splits the plane into regions, where one of them has infinite area and is called the infinite region.

- Ex :

4 regions
(R_4 = infinite region)

2 regions
(R_2 = infinite region)

Euler's Planar Formula

- Let G be a connected planar graph, and consider a planar representation of G. Let

  V = # vertices, E = # edges, F = # regions.

Theorem : $V + F = E + 2$.

- Ex :

V = 4, F = 4, E = 6

V = 8, F = 6, E = 12

Euler's Planar Formula

- Proof Idea :
  - Add edges one by one, so that in each step, the subgraph is always connected
  - Use induction to show that the formula is always satisfied for each subgraph
  - For the new edge that is added, it either joins :

    (1) two existing vertices ➔ V~, F↑

    (2) one existing + one new vertex ➔ V~, F↑

Euler's Planar Formula

Euler's Planar Formula

- Let G be a connected simple planar graph with

  V = # vertices, E = # edges.

Corollary : If $V \geq 3$, then $E \leq 3V - 6$.

- Proof : Each region is surrounded by at least 3 edges (how about the infinite region?)

  ➔ $3F \leq$ total edges $= 2E$

  ➔ $E + 2 = V + F \leq V + 2E/3$

  ➔ $E \leq 3V - 6$

Euler's Planar Formula

Theorem : K_5 and $K_{3,3}$ are non-planar.

- Proof :

(1) For K_5, $V = 5$ and $E = 10$

➔ $E > 3V - 6$ ➔ non-planar

(2) For $K_{3,3}$, $V = 6$ and $E = 9$.

➔ If it is planar, each region is surrounded by at least 4 edges (why?)

➔ $F \leq \lfloor 2E/4 \rfloor = 4$

➔ $V + F \leq 10 < E + 2$ ➔ non-planar

Platonic Solids

Definition : A Platonic solid is a convex 3D shape that all faces are the same, and each face is a regular polygon

Platonic Solids

Theorem: There are exactly 5 Platonic solids

- Proof:

Let n = # vertices of each polygon

m = degree of each vertex

For a platonic solid, we must have

$$nF = 2E \quad \text{and} \quad Vm = 2E$$

Platonic Solids

- Proof (continued):

  By Euler's planar formula,

  $$2E/m + 2E/n = V + F = E + 2$$

  $$\Rightarrow \quad 1/m + 1/n = 1/2 + 1/E \qquad \ldots\ldots (*)$$

  Also, we need to have

  $$n \geq 3 \quad \text{and} \quad m \geq 3 \qquad \text{[from 3D shape]}$$

  but one of them must be $= 3$ [from (*)]

Platonic Solids

- Proof (continued):

➔ Either

(i) n = 3 (with m = 3, 4, or 5)

(ii) m = 3 (with n = 3, 4, or 5)

Map Coloring and Dual Graph

A Map M

Dual Graph of M

Map Coloring and Dual Graph

Observation: A proper color of M

⇔ A proper vertex color the dual graph

Proper coloring : Adjacent regions (or vertices) have to be colored in different colors

Five Color Theorem

- Appel and Haken (1976) showed that every planar graph can be 4 colored

  (Proof is tedious, has 1955 cases and many subcases)

- Here, we shall show that :

Theorem : Every planar graph can be 5 colored.

- The above theorem implies that every map can be 5 colored (as its dual is planar)

Five Color Theorem

- Proof :

  We assume the graph has at least 5 vertices. Else, the theorem will immediately follow.

  Next, in a planar graph, we see that there must be a vertex with degree at most 5.

  Else,

$$2E = \text{total degree} \geq 3V$$

  which contradicts with the fact $E \leq 3V - 6$.

Five Color Theorem

- Proof (continued) :

  Let v be a vertex whose degree is at most 5.

  Now, assume inductively that all planar graphs with n – 1 vertices can be colored in 5 colors

  ➔ Thus if v is removed, we can color the graph properly in 5 colors

  What if we add back v to the graph now ??

Five Color Theorem

- Proof (continued) :

  Case 1 : Neighbors of v uses at most 4 colors

there is a 5th color for v

Five Color Theorem

- Proof (continued) :

  Case 2 : Neighbors of v uses up all 5 colors

Can we save 1 color, by coloring the yellow neighbor in blue ?

Five Color Theorem

- Proof ("Case 2" continued):

  Can we color the yellow neighbor in blue ?

First, we check if the yellow neighbor can connect to the blue neighbor by a "switching" yellow-blue path

Five Color Theorem

- Proof ("Case 2" continued):

  Can we color the yellow neighbor in blue ?

If not, we perform "switching" and thus save one color for v

Five Color Theorem

- Proof ("Case 2" continued):

Can we color the yellow neighbor in blue ?

Else, they are connected
➔ orange and green cannot be connected by "switching path

Five Color Theorem

- Proof ("Case 2" continued):

We color the orange neighbor in green !

➔ we can perform "switching" (orange and green) to save one color for v

Kuratowski's Theorem

Definition : A subdivision operation on an edge { u, v } is to create a new vertex w, and replace the edge by two new edges { u, w } and { w, v }.

- Ex :

Kuratowski's Theorem

Definition : Graphs G and H are homeomorphic if both can be obtained from the same graph by a sequence of subdivision operations.

- Ex : The following graphs are all homeomorphic :

Kuratowski's Theorem

- In 1930, the Polish mathematician Kuratowski proved the following theorem :

Theorem :

Graph G is non-planar

$\Leftrightarrow$ G has a subgraph homeomorphic to K_5 or $K_{3,3}$

- The "if" case is easy to show (how?)
- The "only if" case is hard (I don't know either …)

Kuratowski's Theorem

- Ex : Show that the Petersen graph is non-planar.

Kuratowski's Theorem

- Proof :

Petersen Graph

Subgraph homeomorphic to $K_{3,3}$

Kuratowski's Theorem

- Ex : Is the following graph planar or non-planar ?

Kuratowski's Theorem

- Ans : Planar

BBM 205 Discrete Mathematics
Hacettepe University
http://web.cs.hacettepe.edu.tr/∼bbm205

Lecture 9a: Introduction to Discrete Probability

Lecturer: Lale Özkahya

Resources:
Kenneth Rosen, "Discrete Mathematics and App."
http://www.eecs70.org/

Key Points

- Uncertainty does not mean “nothing is known”
- Most real-world problems involve uncertainty
  - Predictions:
    - Will you get an A in CS70? Will the Raiders win the Super Bowl?
  - Strategy/Decision-making under uncertainty
    - Drop CS 70? How much to bet on blackjack? Buy a specific stock?
  - Engineering
    - Build a spam filter Improve wifi coverage. Control systems (Internet, airplane, robots, self-driving cars)
- How to best use ‘artificial’ uncertainty?
  - Play games of chance
  - Design randomized algorithms.
- Probability
  - Models knowledge about uncertainty: Mathematical discipline that allows you to **reason about uncertainty**.
  - Discovers best way to use that knowledge in making decisions

The Magic of Probability

Uncertainty: vague, fuzzy, confusing, scary, hard to think about.

Probability: A precise, unambiguous, simple(!) way to think about uncertainty.

Uncertainty = Fear

Probability = Serenity

Our mission: help you discover the serenity of Probability, i.e., enable you to think clearly about uncertainty.

Your cost: focused attention and practice, practice, practice.

Random Experiment: Flip one Fair Coin

Flip a fair coin: (*One flips or tosses a coin*)

- Possible outcomes: Heads (H) and Tails (T)
  (*One flip yields either 'heads' or 'tails'.*)
- Likelihoods: $H : 50\%$ and $T : 50\%$

Random Experiment: Flip one Fair Coin

Flip a fair coin:

What do we mean by the likelihood of tails is 50%?

Two interpretations:

- Single coin flip: 50% chance of 'tails' [subjectivist]

  Willingness to bet on the outcome of a single flip

- Many coin flips: About half yield 'tails' [frequentist]

  Makes sense for many flips

- Question: Why does the fraction of tails converge to the same value every time? Statistical Regularity! Deep!

Random Experiment: Flip one Fair Coin

Flip a fair coin: model

Physical Experiment

Probability Model

- The physical experiment is complex. (Shape, density, initial momentum and position, ...)
- The Probability model is simple:
  - A set Ω of outcomes: $\Omega = \{H, T\}$.
  - A probability assigned to each outcome: $Pr[H] = 0.5, Pr[T] = 0.5$.

Random Experiment: Flip one Unfair Coin

Flip an unfair (biased, loaded) coin:

- Possible outcomes: Heads (H) and Tails (T)
- Likelihoods: $H : p \in (0, 1)$ and $T : 1 - p$
- Frequentist Interpretation:

  Flip many times $\Rightarrow$ Fraction $1 - p$ of tails
- Question: How can one figure out p? Flip many times
- Tautology? No: Statistical regularity!

Random Experiment: Flip one Unfair Coin

Flip an unfair (biased, loaded) coin: model

Physical Experiment

Probability Model

Flip Two Fair Coins

- Possible outcomes: $\{HH, HT, TH, TT\} \equiv \{H, T\}^2$.
- Note: $A \times B := \{(a, b) \mid a \in A, b \in B\}$ and $A^2 := A \times A$.
- Likelihoods: $1/4$ each.

Flip Glued Coins

Flips two coins glued together side by side:

Glued coins

50%

50%

- Possible outcomes: $\{HH, TT\}$.
- Likelihoods: $HH : 0.5, TT : 0.5$.
- Note: Coins are glued so that they show the same face.

Flip Glued Coins

Flips two coins glued together side by side:

Glued coins

50%

50%

- Possible outcomes: $\{HT, TH\}$.
- Likelihoods: $HT : 0.5, TH : 0.5$.
- Note: Coins are glued so that they show different faces.

Flip two Attached Coins

Flips two coins attached by a spring:

- Possible outcomes: $\{HH, HT, TH, TT\}$.
- Likelihoods: $HH : 0.4, HT : 0.1, TH : 0.1, TT : 0.4$.
- Note: Coins are attached so that they tend to show the same face, unless the spring twists enough.

Flipping Two Coins

Here is a way to summarize the four random experiments:

- Ω is the set of *possible* outcomes;
- Each outcome has a probability (likelihood);
- The probabilities are ≥ 0 and add up to 1;
- Fair coins: [1]; Glued coins: [3], [4];

  Spring-attached coins: [2];

Flipping Two Coins

Here is a way to summarize the four random experiments:

Important remarks:

- Each outcome describes the two coins.
- E.g., HT is one outcome of the experiment.
- It is wrong to think that the outcomes are $\{H, T\}$ and that one picks twice from that set.
- Indeed, this viewpoint misses the relationship between the two flips.
- Each $\omega \in \Omega$ describes one outcome of the complete experiment.
- Ω and the probabilities specify the random experiment.

Flipping n times

Flip a fair coin n times (some $n \geq 1$):

- Possible outcomes: $\{TT\cdots T, TT\cdots H, \ldots, HH\cdots H\}$.

  Thus, 2^n possible outcomes.
- Note: $\{TT\cdots T, TT\cdots H, \ldots, HH\cdots H\} = \{H, T\}^n$.

  $A^n := \{(a_1, \ldots, a_n) \mid a_1 \in A, \ldots, a_n \in A\}$. $|A^n| = |A|^n$.
- Likelihoods: $1/2^n$ each.

Roll two Dice

Roll a balanced 6-sided die twice:

- Possible outcomes: $\{1,2,3,4,5,6\}^2 = \{(a,b) \mid 1 \leq a,b \leq 6\}$.
- Likelihoods: $1/36$ for each.

Physical Experiment

Probability Model

Probability Space.

1. A "random experiment":

   (a) Flip a biased coin;
   (b) Flip two fair coins;
   (c) Deal a poker hand.

2. A set of possible outcomes: Ω.

   (a) $\Omega = \{H, T\}$;
   (b) $\Omega = \{HH, HT, TH, TT\}$; $|\Omega| = 4$;
   (c) $\Omega = \{\ \underline{A\spadesuit\ A\diamondsuit\ A\clubsuit\ A\heartsuit\ K\spadesuit},\ \underline{A\spadesuit\ A\diamondsuit\ A\clubsuit\ A\heartsuit\ Q\spadesuit}, \ldots\}$
   $|\Omega| = \binom{52}{5}$.

3. Assign a probability to each outcome: $Pr : \Omega \to [0, 1]$.

   (a) $Pr[H] = p, Pr[T] = 1 - p$ for some $p \in [0, 1]$
   (b) $Pr[HH] = Pr[HT] = Pr[TH] = Pr[TT] = \frac{1}{4}$
   (c) $Pr[\ \underline{A\spadesuit\ A\diamondsuit\ A\clubsuit\ A\heartsuit\ K\spadesuit}\ ] = \cdots = 1/\binom{52}{5}$

Probability Space: formalism.

Ω is the **sample space.**
$\omega \in \Omega$ is a **sample point**. (Also called an **outcome**.)
Sample point ω has a probability $Pr[\omega]$ where

- $0 \leq Pr[\omega] \leq 1$;
- $\sum_{\omega \in \Omega} Pr[\omega] = 1$.

$$0 \leq Pr[\omega] \leq 1$$

$$\sum_{\omega} Pr[\omega] = 1$$

Probability Space: Formalism.

In a **uniform probability space** each outcome ω is equally probable: $Pr[\omega] = \frac{1}{|\Omega|}$ for all $\omega \in \Omega$.

Examples:

- Flipping two fair coins, dealing a poker hand are uniform probability spaces.
- Flipping a biased coin is not a uniform probability space.

Probability Space: Formalism

Simplest physical model of a uniform probability space:

Physical experiment

Probability model

A bag of identical balls, except for their color (or a label). If the bag is well shaken, every ball is equally likely to be picked.

$$\Omega = \{\text{white, red, yellow, grey, purple, blue, maroon, green}\}$$

$$Pr[\text{blue}] = \frac{1}{8}.$$

Probability Space: Formalism

Simplest physical model of a non-uniform probability space:

Physical experiment

Probability model

$$\Omega = \{\text{Red, Green, Yellow, Blue}\}$$
$$Pr[\text{Red}] = \frac{3}{10}, Pr[\text{Green}] = \frac{4}{10}, \text{ etc.}$$

Note: Probabilities are restricted to rational numbers: $\frac{N_k}{N}$.

Probability Space: Formalism

Physical model of a general non-uniform probability space:

Physical experiment

Probability model

The roulette wheel stops in sector ω with probability p_ω.

$$\Omega = \{1, 2, 3, \ldots, N\}, Pr[\omega] = p_\omega.$$

An important remark

- The random experiment selects one and only one outcome in Ω.
- For instance, when we flip a fair coin twice
  - $\Omega = \{HH, TH, HT, TT\}$
  - The experiment selects *one* of the elements of Ω.
- In this case, its wrong to think that $\Omega = \{H, T\}$ and that the experiment selects two outcomes.
- Why? Because this would not describe how the two coin flips are related to each other.
- For instance, say we glue the coins side-by-side so that they face up the same way. Then one gets HH or TT with probability 50% each. This is not captured by 'picking two outcomes.'

Example: Monty Hall Problem

:

-
- **one of the other**
-

Controversy

( )

Monty Hall problem

-
- $\Omega =$
  $\{(\quad) (\quad) (\quad)\}$
  $_ \quad :$
-
  $\Omega = \{(\quad) (\quad) (\quad)\}$
  $_$
-

Summary

$\boxed{\text{Modeling Uncertainty: Probability Space}}$

1. Random Experiment
2. Probability Space: $\Omega; Pr[\omega] \in [0,1]; \sum_{\omega} Pr[\omega] = 1$.
3. Uniform Probability Space: $Pr[\omega] = 1/|\Omega|$ for all $\omega \in \Omega$.

Probability problems can be fun

n

( )

BBM 205 Discrete Mathematics
Hacettepe University
http://web.cs.hacettepe.edu.tr/∼bbm205

Lecture 9b: Events, Conditional Probability, Independence, Bayes' Rule

Lecturer: Lale Özkahya

Resources:
Kenneth Rosen, "Discrete Mathematics and App."
http://www.eecs70.org/

CS70: On to Calculation.

Events, Conditional Probability, Independence, Bayes' Rule

1. Probability Basics Review
2. Events
3. Conditional Probability
4. Independence of Events
5. Bayes' Rule

Probability Basics Review

Setup:

- Random Experiment.
  Flip a fair coin twice.
- Probability Space.
  - **Sample Space:** Set of outcomes, Ω.
    $\Omega = \{HH, HT, TH, TT\}$
    (Note: Not $\Omega = \{H, T\}$ with two picks!)
  - **Probability:** $Pr[\omega]$ for all $\omega \in \Omega$.
    $Pr[HH] = \cdots = Pr[TT] = 1/4$
    1. $0 \leq Pr[\omega] \leq 1$.
    2. $\sum_{\omega \in \Omega} Pr[\omega] = 1$.

Set notation review

Figure: Two events

Figure: Union (or)

Figure: Difference (A, not B)

Figure: Complement (not)

Figure: Intersection (and)

Figure: Symmetric difference (only one)

Probability of exactly one 'heads' in two coin flips?

Idea: Sum the probabilities of all the different outcomes that have exactly one 'heads': HT, TH.

This leads to a definition!

Definition:

- An **event,** E**,** is a subset of outcomes: $E \subset \Omega$.
- The **probability of** E is defined as $Pr[E] = \sum_{\omega \in E} Pr[\omega]$.

Event: Example

Physical experiment

Probability model

$$\Omega = \{\text{Red, Green, Yellow, Blue}\}$$

$$Pr[\text{Red}] = \frac{3}{10}, Pr[\text{Green}] = \frac{4}{10}, \text{ etc.}$$

$$E = \{Red, Green\} \Rightarrow Pr[E] = \frac{3+4}{10} = \frac{3}{10} + \frac{4}{10} = Pr[\text{Red}] + Pr[\text{Green}].$$

Probability of exactly one heads in two coin flips?

Sample Space, $\Omega = \{HH, HT, TH, TT\}$.

Uniform probability space:
$Pr[HH] = Pr[HT] = Pr[TH] = Pr[TT] = \frac{1}{4}$.

Event, E, "exactly one heads": $\{TH, HT\}$.

$$Pr[E] = \sum_{\omega \in E} Pr[\omega] = \frac{|E|}{|\Omega|} = \frac{2}{4} = \frac{1}{2}.$$

Example: 20 coin tosses.

20 coin tosses

Sample space: $\Omega =$ set of 20 fair coin tosses.
$\Omega = \{T, H\}^{20} \equiv \{0, 1\}^{20}$; $|\Omega| = 2^{20}$.

- What is more likely?
  - $\omega_1 := (1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1)$, or
  - $\omega_2 := (1,0,1,1,0,0,0,1,0,1,0,1,1,0,1,1,1,0,0,0)$?

  Answer: Both are equally likely: $Pr[\omega_1] = Pr[\omega_2] = \frac{1}{|\Omega|}$.

- What is more likely?

  (E_1) Twenty Hs out of twenty, or
  (E_2) Ten Hs out of twenty?

  Answer: Ten Hs out of twenty.

  Why? There are many sequences of 20 tosses with ten Hs; only one with twenty Hs. $\Rightarrow Pr[E_1] = \frac{1}{|\Omega|} \ll Pr[E_2] = \frac{|E_2|}{|\Omega|}$.

$$|E_2| = \binom{20}{10} = 184,756.$$

Probability of n heads in 100 coin tosses.

$\Omega = \{H, T\}^{100}$; $|\Omega| = 2^{100}$.

Event E_n = 'n heads'; $|E_n| = \binom{100}{n}$

$p_n := Pr[E_n] = \frac{|E_n|}{|\Omega|} = \frac{\binom{100}{n}}{2^{100}}$

Observe:

- Concentration around mean: Law of Large Numbers;
- Bell-shape: Central Limit Theorem.

Roll a red and a blue die.

$$Pr[\text{Sum to } 7] = \frac{6}{36} \qquad Pr[\text{Sum to } 10] = \frac{3}{36}$$

Exactly 50 heads in 100 coin tosses.

Sample space: $\Omega =$ set of 100 coin tosses $= \{H, T\}^{100}$.
$|\Omega| = 2 \times 2 \times \cdots \times 2 = 2^{100}$.

Uniform probability space: $Pr[\omega] = \frac{1}{2^{100}}$.

Event $E =$ "100 coin tosses with exactly 50 heads"

$|E|$?
Choose 50 positions out of 100 to be heads.
$|E| = \binom{100}{50}$.

$$Pr[E] = \frac{\binom{100}{50}}{2^{100}}.$$

Calculation.

Stirling formula (for large n):

$$n! \approx \sqrt{2\pi n}\left(\frac{n}{e}\right)^n.$$

$$\binom{2n}{n} \approx \frac{\sqrt{4\pi n}(2n/e)^{2n}}{[\sqrt{2\pi n}(n/e)^n]^2} \approx \frac{4^n}{\sqrt{\pi n}}.$$

$$Pr[E] = \frac{|E|}{|\Omega|} = \frac{|E|}{2^{2n}} = \frac{1}{\sqrt{\pi n}} = \frac{1}{\sqrt{50\pi}} \approx .08.$$

Exactly 50 heads in 100 coin tosses.

Probability is Additive

Theorem

(a) If events A and B are disjoint, i.e., $A \cap B = \emptyset$, then

$$Pr[A \cup B] = Pr[A] + Pr[B].$$

(b) If events $A_1, \ldots, A_n$ are pairwise disjoint,
i.e., $A_k \cap A_m = \emptyset, \forall k \neq m$, then

$$Pr[A_1 \cup \cdots \cup A_n] = Pr[A_1] + \cdots + Pr[A_n].$$

Proof:

Obvious.

Consequences of Additivity

Theorem

(a) $Pr[A \cup B] = Pr[A] + Pr[B] - Pr[A \cap B]$;

(inclusion-exclusion property)

(b) $Pr[A_1 \cup \cdots \cup A_n] \leq Pr[A_1] + \cdots + Pr[A_n]$;

(union bound)

(c) If $A_1, \ldots A_N$ are a partition of Ω, i.e., pairwise disjoint and $\cup_{m=1}^{N} A_m = \Omega$, then

$$Pr[B] = Pr[B \cap A_1] + \cdots + Pr[B \cap A_N].$$

(law of total probability)

Proof:

(b) is obvious.

Proofs for (a) and (c)? Next...

Inclusion/Exclusion

$$Pr[A \cup B] = Pr[A] + Pr[B] - Pr[A \cap B]$$

$$\begin{aligned} Pr[A] &= x + y \\ Pr[B] &= y + z \\ Pr[A \cap B] &= y \\ Pr[A \cup B] &= x + y + z \end{aligned}$$

Another view. Any $\omega \in A \cup B$ is in $A \cap \overline{B}$, $A \cup B$, or $\overline{A} \cap B$. So, add it up.

Total probability

Assume that Ω is the union of the disjoint sets $A_1, \ldots, A_N$.

Then,

$$Pr[B] = Pr[A_1 \cap B] + \cdots + Pr[A_N \cap B].$$

Indeed, B is the union of the disjoint sets $A_n \cap B$ for $n = 1, \ldots, N$.

In "math": $\omega \in B$ is in exactly one of $A_i \cap B$.

Adding up probability of them, get $Pr[\omega]$ in sum.

..Did I say...

Add it up.

Roll a Red and a Blue Die.

$$|E_1 \cup E_2| = |E_1| + |E_2| - |E_1 \cap E_2|$$

$E_1 =$ 'Red die shows 6'; $E_2 =$ 'Blue die shows 6'

$E_1 \cup E_2 =$ 'At least one die shows 6'

$$Pr[E_1] = \frac{6}{36}, Pr[E_2] = \frac{6}{36}, Pr[E_1 \cup E_2] = \frac{11}{36}.$$

Conditional probability: example.

Two coin flips. First flip is heads. Probability of two heads?
$\Omega = \{HH, HT, TH, TT\}$; Uniform probability space.
Event $A =$ first flip is heads: $A = \{HH, HT\}$.

New sample space: A; uniform still.

Event $B =$ two heads.

The probability of two heads if the first flip is heads.
The probability of B **given** A is $1/2$.

A similar example.

Two coin flips. At least one of the flips is heads.
$\rightarrow$ Probability of two heads?

$\Omega = \{HH, HT, TH, TT\}$; uniform.
Event $A =$ at least one flip is heads. $A = \{HH, HT, TH\}$.

New sample space: A; uniform still.

Event $B =$ two heads.

The probability of two heads if at least one flip is heads.
The probability of B **given** A is $1/3$.

Conditional Probability: A non-uniform example

Physical experiment

Probability model

$$\Omega = \{\text{Red, Green, Yellow, Blue}\}$$

$$Pr[\text{Red}|\text{Red or Green}] = \frac{3}{7} = \frac{Pr[\text{Red} \cap (\text{Red or Green})]}{Pr[\text{Red or Green}]}$$

Another non-uniform example

Consider $\Omega = \{1, 2, \ldots, N\}$ with $Pr[n] = p_n$.
Let $A = \{3, 4\}, B = \{1, 2, 3\}$.

$$Pr[A|B] = \frac{p_3}{p_1 + p_2 + p_3} = \frac{Pr[A \cap B]}{Pr[B]}.$$

Yet another non-uniform example

Consider $\Omega = \{1, 2, \ldots, N\}$ with $Pr[n] = p_n$.
Let $A = \{2, 3, 4\}, B = \{1, 2, 3\}$.

$$Pr[A|B] = \frac{p_2 + p_3}{p_1 + p_2 + p_3} = \frac{Pr[A \cap B]}{Pr[B]}.$$

Conditional Probability.

Definition: The **conditional probability** of B given A is

$$Pr[B|A] = \frac{Pr[A \cap B]}{Pr[A]}$$

In A!

In B?

Must be in $A \cap B$.

$$Pr[B|A] = \frac{Pr[A \cap B]}{Pr[A]}.$$

More fun with conditional probability.

Toss a red and a blue die, sum is 4,
What is probability that red is 1?

$Pr[B|A] = \frac{|B \cap A|}{|A|} = \frac{1}{3}$; versus $Pr[B] = 1/6$.

B is more likely given A.

Yet more fun with conditional probability.

Toss a red and a blue die, sum is 7,
what is probability that red is 1?

Ω : Uniform

$\Omega = \{1, \ldots, 6\}^2$

$A = \{(1, 6), \ldots, (6,1)\}$

$B = \{(1, 1), \ldots, (1, 6)\}$

A = 'sum is 7'

$Pr[B|A] = \frac{|B \cap A|}{|A|} = \frac{1}{6}$; versus $Pr[B] = \frac{1}{6}$.

Observing A does not change your mind about the likelihood of B.

Emptiness..

Suppose I toss 3 balls into 3 bins.
A =“1st bin empty”; B =“2nd bin empty.” What is $Pr[A|B]$?

$$\Omega = \{1,2,3\}^3$$

$$\omega = (\text{bin of red ball}, \text{bin of blue ball}, \text{bin of green ball})$$

$Pr[B] = Pr[\{(a,b,c) \mid a,b,c \in \{1,3\}] = Pr[\{1,3\}^3] = \frac{8}{27}$

$Pr[A \cap B] = Pr[(3,3,3)] = \frac{1}{27}$

$Pr[A|B] = \frac{Pr[A \cap B]}{Pr[B]} = \frac{(1/27)}{(8/27)} = 1/8$; vs. $Pr[A] = \frac{8}{27}$.

A is less likely given B: If second bin is empty the first is more likely to have balls in it.

Three Card Problem

Three cards: Red/Red, Red/Black, Black/Black.
Pick one at random and place on the table. The upturned side is a Red. What is the probability that the other side is Black?
Can't be the BB card, so...prob should be 0.5, right?
R: upturned card is Red; RB: the Red/Black card was selected.
Want $P(RB|R)$.
What's wrong with the reasoning that leads to $\frac{1}{2}$?

$$\begin{aligned} P(RB|R) &= \frac{P(RB \cap R)}{P(R)} \\ &= \frac{\frac{1}{3}\frac{1}{2}}{\frac{1}{3}(1) + \frac{1}{3}\frac{1}{2} + \frac{1}{3}(0)} \\ &= \frac{\frac{1}{6}}{\frac{1}{2}} = \frac{1}{3} \end{aligned}$$

Once you are given R: it is twice as likely that the RR card was picked.

Gambler's fallacy.

Flip a fair coin 51 times.
A = "first 50 flips are heads"
B = "the 51st is heads"
$Pr[B|A]$?

$A = \{HH \cdots HT, HH \cdots HH\}$
$B \cap A = \{HH \cdots HH\}$

Uniform probability space.

$Pr[B|A] = \frac{|B \cap A|}{|A|} = \frac{1}{2}$.

Same as $Pr[B]$.

The likelihood of 51st heads does not depend on the previous flips.

Product Rule

Recall the definition:

$$Pr[B|A] = \frac{Pr[A \cap B]}{Pr[A]}.$$

Hence,

$$Pr[A \cap B] = Pr[A]Pr[B|A].$$

Consequently,

$$\begin{aligned} Pr[A \cap B \cap C] &= Pr[(A \cap B) \cap C] \\ &= Pr[A \cap B]Pr[C|A \cap B] \\ &= Pr[A]Pr[B|A]Pr[C|A \cap B]. \end{aligned}$$

Product Rule

Theorem Product Rule
Let $A_1, A_2, \ldots, A_n$ be events. Then

$$Pr[A_1 \cap \cdots \cap A_n] = Pr[A_1]Pr[A_2|A_1] \cdots Pr[A_n|A_1 \cap \cdots \cap A_{n-1}].$$

Proof: By induction.
Assume the result is true for n. (It holds for $n = 2$.) Then,

$$\begin{aligned} &Pr[A_1 \cap \cdots \cap A_n \cap A_{n+1}] \\ &\quad = Pr[A_1 \cap \cdots \cap A_n]Pr[A_{n+1}|A_1 \cap \cdots \cap A_n] \\ &\quad = Pr[A_1]Pr[A_2|A_1] \cdots Pr[A_n|A_1 \cap \cdots \cap A_{n-1}]Pr[A_{n+1}|A_1 \cap \cdots \cap A_n], \end{aligned}$$

so that the result holds for $n + 1$. $\square$

Correlation

An example.
Random experiment: Pick a person at random.
Event A: the person has lung cancer.
Event B: the person is a heavy smoker.

Fact:

$$Pr[A|B] = 1.17 \times Pr[A].$$

Conclusion:

- Smoking increases the probability of lung cancer by 17%.
- Smoking causes lung cancer.

Correlation

Event A: the person has lung cancer. Event B: the person is a heavy smoker. $Pr[A|B] = 1.17 \times Pr[A]$.

A second look.

Note that

$$\begin{aligned} Pr[A|B] = 1.17 \times Pr[A] \quad &\Leftrightarrow \quad \frac{Pr[A \cap B]}{Pr[B]} = 1.17 \times Pr[A] \\ &\Leftrightarrow \quad Pr[A \cap B] = 1.17 \times Pr[A]Pr[B] \\ &\Leftrightarrow \quad Pr[B|A] = 1.17 \times Pr[B]. \end{aligned}$$

Conclusion:

- Lung cancer increases the probability of smoking by 17%.
- Lung cancer causes smoking. Really?

Causality vs. Correlation

Events A and B are **positively correlated** if

$$Pr[A \cap B] > Pr[A]Pr[B].$$

(E.g., smoking and lung cancer.)

A and B being positively correlated does not mean that A causes B or that B causes A.

Other examples:

- Tesla owners are more likely to be rich. That does not mean that poor people should buy a Tesla to get rich.
- People who go to the opera are more likely to have a good career. That does not mean that going to the opera will improve your career.
- Rabbits eat more carrots and do not wear glasses. Are carrots good for eyesight?

Total probability

Assume that Ω is the union of the disjoint sets $A_1, \ldots, A_N$.

Then,

$$Pr[B] = Pr[A_1 \cap B] + \cdots + Pr[A_N \cap B].$$

Indeed, B is the union of the disjoint sets $A_n \cap B$ for $n = 1, \ldots, N$. Thus,

$$Pr[B] = Pr[A_1]Pr[B|A_1] + \cdots + Pr[A_N]Pr[B|A_N].$$

Total probability

Assume that Ω is the union of the disjoint sets $A_1, \ldots, A_N$.

$$Pr[B] = Pr[A_1]Pr[B|A_1] + \cdots + Pr[A_N]Pr[B|A_N].$$

Is your coin loaded?

Your coin is fair w.p. $1/2$ or such that $Pr[H] = 0.6$, otherwise.

You flip your coin and it yields heads.

What is the probability that it is fair?

Analysis:

$$A = \text{ 'coin is fair'}, B = \text{ 'outcome is heads'}$$

We want to calculate $P[A|B]$.

We know $P[B|A] = 1/2, P[B|\bar{A}] = 0.6, Pr[A] = 1/2 = Pr[\bar{A}]$

Now,

$$\begin{aligned} Pr[B] &= Pr[A \cap B] + Pr[\bar{A} \cap B] = Pr[A]Pr[B|A] + Pr[\bar{A}]Pr[B|\bar{A}] \\ &= (1/2)(1/2) + (1/2)0.6 = 0.55. \end{aligned}$$

Thus,

$$Pr[A|B] = \frac{Pr[A]Pr[B|A]}{Pr[B]} = \frac{(1/2)(1/2)}{(1/2)(1/2) + (1/2)0.6} \approx 0.45.$$

Is your coin loaded?

A picture:

Imagine 100 situations, among which
$m := 100(1/2)(1/2)$ are such that A and B occur and
$n := 100(1/2)(0.6)$ are such that $\bar{A}$ and B occur.

Thus, among the $m+n$ situations where B occurred, there are m where A occurred.

Hence,

$$Pr[A|B] = \frac{m}{m+n} = \frac{(1/2)(1/2)}{(1/2)(1/2)+(1/2)0.6}.$$

Independence

Definition: Two events A and B are **independent** if

$$Pr[A \cap B] = Pr[A]Pr[B].$$

Examples:

- When rolling two dice, $A =$ sum is 7 and $B =$ red die is 1 are independent;
- When rolling two dice, $A =$ sum is 3 and $B =$ red die is 1 are not independent;
- When flipping coins, $A =$ coin 1 yields heads and $B =$ coin 2 yields tails are independent;
- When throwing 3 balls into 3 bins, $A =$ bin 1 is empty and $B =$ bin 2 is empty are not independent;

Independence and conditional probability

Fact: Two events A and B are **independent** if and only if

$$Pr[A|B] = Pr[A].$$

Indeed: $Pr[A|B] = \frac{Pr[A \cap B]}{Pr[B]}$, so that

$$Pr[A|B] = Pr[A] \Leftrightarrow \frac{Pr[A \cap B]}{Pr[B]} = Pr[A] \Leftrightarrow Pr[A \cap B] = Pr[A]Pr[B].$$

Bayes Rule

Another picture: We imagine that there are N possible causes $A_1, \ldots, A_N$.

$$p_n = Pr[A_n]$$

$$q_n = Pr[B|A_n]$$

$$A_1, \ldots, A_N \text{ disjoint}$$

$$A_1 \cup \cdots \cup A_N = \Omega$$

Imagine 100 situations, among which $100 p_n q_n$ are such that A_n and B occur, for $n = 1, \ldots, N$.
Thus, among the $100 \sum_m p_m q_m$ situations where B occurred, there are $100 p_n q_n$ where A_n occurred.

Hence,

$$Pr[A_n|B] = \frac{p_n q_n}{\sum_m p_m q_m}.$$

Why do you have a fever?

Using Bayes' rule, we find

$$Pr[\text{Flu}|\text{High Fever}] = \frac{0.15 \times 0.80}{0.15 \times 0.80 + 10^{-8} \times 1 + 0.85 \times 0.1} \approx 0.58$$

$$Pr[\text{Ebola}|\text{High Fever}] = \frac{10^{-8} \times 1}{0.15 \times 0.80 + 10^{-8} \times 1 + 0.85 \times 0.1} \approx 5 \times 10^{-8}$$

$$Pr[\text{Other}|\text{High Fever}] = \frac{0.85 \times 0.1}{0.15 \times 0.80 + 10^{-8} \times 1 + 0.85 \times 0.1} \approx 0.42$$

These are the posterior probabilities. One says that 'Flu' is the Most Likely a Posteriori (MAP) cause of the high fever.

Bayes' Rule Operations

Bayes' Rule is the canonical example of how information changes our opinions.

Thomas Bayes

Thomas Bayes

Portrait used of Bayes in a 1936 book,[1] but it is doubtful whether the portrait is actually of him.[2] No earlier portrait or claimed portrait survives.

Born	c. 1701 London, England
Died	7 April 1761 (aged 59) Tunbridge Wells, Kent, England
Residence	Tunbridge Wells, Kent, England
Nationality	English
Known for	Bayes' theorem

Source: Wikipedia.

Thomas Bayes

A

FIG. 3. *Joshua Bayes (1671–1746).*

$$\xrightarrow[\text{Max over } B]{Pr[B|A]}$$

B

Thomas Bayes?

A Bayesian picture of Thomas Bayes.

Testing for disease.

Let's watch TV!!
Random Experiment: Pick a random male.
Outcomes: $(test, disease)$
A - prostate cancer.
B - positive PSA test.

- $Pr[A] = 0.0016$, (.16 % of the male population is affected.)
- $Pr[B|A] = 0.80$ (80% chance of positive test with disease.)
- $Pr[B|\overline{A}] = 0.10$ (10% chance of positive test without disease.)

From http://www.cpcn.org/01_psa_tests.htm and http://seer.cancer.gov/statfacts/html/prost.html (10/12/2011.)

Positive PSA test (B). Do I have disease?

$$Pr[A|B]???$$

Bayes Rule.

Using Bayes' rule, we find

$$P[A|B] = \frac{0.0016 \times 0.80}{0.0016 \times 0.80 + 0.9984 \times 0.10} = .013.$$

A 1.3% chance of prostate cancer with a positive PSA test.

Surgery anyone?

Impotence...

Incontinence..

Death.

Summary

Events, Conditional Probability, Independence, Bayes' Rule

Key Ideas:

- Conditional Probability:

$$Pr[A|B] = \frac{Pr[A \cap B]}{Pr[B]}$$

- Independence: $Pr[A \cap B] = Pr[A]Pr[B]$.
- Bayes' Rule:

$$Pr[A_n|B] = \frac{Pr[A_n]Pr[B|A_n]}{\sum_m Pr[A_m]Pr[B|A_m]}.$$

$Pr[A_n|B] =$ posterior probability; $Pr[A_n] =$ prior probability .

- All these are possible:

$$Pr[A|B] < Pr[A]; Pr[A|B] > Pr[A]; Pr[A|B] = Pr[A].$$

BBM 205 Discrete Mathematics
Hacettepe University
http://web.cs.hacettepe.edu.tr/∼bbm205

Lecture 10a: Random Variables

Lecturer: Lale Özkahya

Resources:
Kenneth Rosen, "Discrete Mathematics and App."
http://www.eecs70.org/

Balls in bins

One throws m balls into $n > m$ bins.

Balls in bins

One throws m balls into $n > m$ bins.

Theorem:
$Pr[\text{no collision}] \approx \exp\{-\frac{m^2}{2n}\}$, for large enough n.

Balls in bins

Theorem:

$Pr[\text{no collision}] \approx \exp\{-\frac{m^2}{2n}\}$, for large enough n.

Balls in bins

Theorem:
$Pr[\text{no collision}] \approx \exp\{-\frac{m^2}{2n}\}$, for large enough n.

In particular, $Pr[\text{no collision}] \approx 1/2$ for $m^2/(2n) \approx \ln(2)$, i.e.,

$$m \approx \sqrt{2\ln(2)n} \approx 1.2\sqrt{n}.$$

E.g., $1.2\sqrt{20} \approx 5.4$.

Roughly, $Pr[\text{collision}] \approx 1/2$ for $m = \sqrt{n}$. ($e^{-0.5} \approx 0.6$.)

The Calculation.

A_i = no collision when ith ball is placed in a bin.

$Pr[A_i | A_{i-1} \cap \cdots \cap A_1] = (1 - \frac{i-1}{n})$.

no collision = $A_1 \cap \cdots \cap A_m$.

Product rule:

$Pr[A_1 \cap \cdots \cap A_m] = Pr[A_1] Pr[A_2 | A_1] \cdots Pr[A_m | A_1 \cap \cdots \cap A_{m-1}]$

$$\Rightarrow Pr[\text{no collision}] = \left(1 - \frac{1}{n}\right) \cdots \left(1 - \frac{m-1}{n}\right).$$

Hence,

$$\begin{aligned} \ln(Pr[\text{no collision}]) &= \sum_{k=1}^{m-1} \ln(1 - \frac{k}{n}) \approx \sum_{k=1}^{m-1} (-\frac{k}{n})^{(*)} \\ &= -\frac{1}{n} \frac{m(m-1)}{2}^{(\dagger)} \approx -\frac{m^2}{2n} \end{aligned}$$

$^{(*)}$ We used $\ln(1 - \varepsilon) \approx -\varepsilon$ for $|\varepsilon| \ll 1$.

$^{(\dagger)}$ $1 + 2 + \cdots + m - 1 = (m-1)m/2$.

Approximation

$$\exp\{-x\} = 1 - x + \frac{1}{2!}x^2 + \cdots \approx 1 - x, \text{ for } |x| \ll 1.$$

Hence, $-x \approx \ln(1-x)$ for $|x| \ll 1$.

Today's your birthday, it's my birthday too..

Probability that m people all have different birthdays?
With $n = 365$, one finds

$Pr[\text{collision}] \approx 1/2$ if $m \approx 1.2\sqrt{365} \approx 23$.

If $m = 60$, we find that

$$Pr[\text{no collision}] \approx \exp\{-\frac{m^2}{2n}\} = \exp\{-\frac{60^2}{2 \times 365}\} \approx 0.007.$$

If $m = 366$, then $Pr[\text{no collision}] = 0$. (No approximation here!)

Checksums!

Consider a set of m files.
Each file has a checksum of b bits.
How large should b be for $Pr[\text{share a checksum}] \leq 10^{-3}$?

Claim: $b \geq 2.9\ln(m) + 9$.

Proof:

Let $n = 2^b$ be the number of checksums.
We know $Pr[\text{no collision}] \approx \exp\{-m^2/(2n)\} \approx 1 - m^2/(2n)$.
Hence,

$$\begin{aligned} &Pr[\text{no collision}] \approx 1 - 10^{-3} \Leftrightarrow m^2/(2n) \approx 10^{-3} \\ &\Leftrightarrow 2n \approx m^2 10^3 \Leftrightarrow 2^{b+1} \approx m^2 2^{10} \\ &\Leftrightarrow b + 1 \approx 10 + 2\log_2(m) \approx 10 + 2.9\ln(m). \end{aligned}$$

Note: $\log_2(x) = \log_2(e)\ln(x) \approx 1.44\ln(x)$.

Coupon Collector Problem.

There are n different baseball cards.
(Brian Wilson, Jackie Robinson, Roger Hornsby, ...)

One random baseball card in each cereal box.

Theorem: If you buy m boxes,

(a) $Pr[\text{miss one specific item}] \approx e^{-\frac{m}{n}}$

(b) $Pr[\text{miss any one of the items}] \leq ne^{-\frac{m}{n}}$.

Coupon Collector Problem: Analysis.

Event A_m = 'fail to get Brian Wilson in m cereal boxes'

Fail the first time: $(1 - \frac{1}{n})$
Fail the second time: $(1 - \frac{1}{n})$
And so on ... for m times. Hence,

$$\begin{aligned} Pr[A_m] &= (1 - \frac{1}{n}) \times \cdots \times (1 - \frac{1}{n}) \\ &= (1 - \frac{1}{n})^m \\ ln(Pr[A_m]) &= m \ln(1 - \frac{1}{n}) \approx m \times (-\frac{1}{n}) \\ Pr[A_m] &\approx \exp\{-\frac{m}{n}\}. \end{aligned}$$

For $p_m = \frac{1}{2}$, we need around $n \ln 2 \approx 0.69n$ boxes.

Collect all cards?

Experiment: Choose m cards at random with replacement.

Events: E_k = 'fail to get player k' , for k = 1, . . . , n

Probability of failing to get at least one of these n players:

$$p := Pr[E_1 \cup E_2 \cdots \cup E_n]$$

How does one estimate p? Union Bound:

$$p = Pr[E_1 \cup E_2 \cdots \cup E_n] \leq Pr[E_1] + Pr[E_2] \cdots Pr[E_n].$$

$$Pr[E_k] \approx e^{-\frac{m}{n}}, k = 1, \ldots, n.$$

Plug in and get

$$p \leq ne^{-\frac{m}{n}}.$$

Collect all cards?

Thus,

$$Pr[\text{missing at least one card}] \leq ne^{-\frac{m}{n}}.$$

Hence,

$$Pr[\text{missing at least one card}] \leq p \text{ when } m \geq n\ln(\frac{n}{p}).$$

To get $p = 1/2$, set $m = n\ln(2n)$.

$(p \leq ne^{-\frac{m}{n}} \leq ne^{-\ln(n/p)} \leq n(\frac{p}{n}) \leq p.)$

E.g., $n = 10^2 \Rightarrow m = 530; n = 10^3 \Rightarrow m = 7600.$

Quick Review.

Bayes' Rule, Mutual Independence, Collisions and Collecting

Main results:

- Bayes' Rule: $Pr[A_m|B] = p_m q_m/(p_1 q_1 + \cdots + p_M q_M)$.
- Product Rule:
  $Pr[A_1 \cap \cdots \cap A_n] = Pr[A_1]Pr[A_2|A_1]\cdots Pr[A_n|A_1 \cap \cdots \cap A_{n-1}]$.
- Balls in bins: m balls into $n > m$ bins.

$$Pr[\text{no collisions}] \approx \exp\{-\frac{m^2}{2n}\}$$

- Coupon Collection: n items. Buy m cereal boxes.

  $Pr[\text{miss one specific item}] \approx e^{-\frac{m}{n}}$; $Pr[\text{miss any one of the items}] \leq ne^{-\frac{m}{n}}$.

Key Mathematical Fact: $\ln(1-\varepsilon) \approx -\varepsilon$.

Random Variables

Random Variables

1. Random Variables.
2. Expectation
3. Distributions.

Questions about outcomes ...

Experiment: roll two dice.
Sample Space: $\{(1,1),(1,2),\ldots,(6,6)\} = \{1,\ldots,6\}^2$
How many pips?

Experiment: flip 100 coins.
Sample Space: $\{HHH\cdots H, THH\cdots H, \ldots, TTT\cdots T\}$
How many heads in 100 coin tosses?

Experiment: choose a random student in cs70.
Sample Space: $\{Adam, Jin, Bing, \ldots, Angeline\}$
What midterm score?

Experiment: hand back assignments to 3 students at random.
Sample Space: $\{123, 132, 213, 231, 312, 321\}$
How many students get back their own assignment?

In each scenario, each outcome gives a number.

The number is a (known) function of the outcome.

Random Variables.

A **random variable,** X**,** for an experiment with sample space Ω is a function $X : \Omega \to \Re$.

Thus, $X(\cdot)$ assigns a real number $X(\omega)$ to each $\omega \in \Omega$.

$X(\omega_1) = 2.1; X(\omega_2) = 7.0; X(\omega_3) = X(\omega_4) = -3.9$

The function $X(\cdot)$ is defined on the outcomes Ω.

The function $X(\cdot)$ is not random, not a variable!

What varies at random (from experiment to experiment)? The outcome!

Example 1 of Random Variable

Experiment: roll two dice.
Sample Space: $\{(1,1),(1,2),\ldots,(6,6)\} = \{1,\ldots,6\}^2$
Random Variable X: number of pips.
$X(1,1) = 2$
$X(1,2) = 3,$
$\vdots$
$X(6,6) = 12,$
$X(a,b) = a+b, (a,b) \in \Omega.$

Example 2 of Random Variable

Experiment: flip three coins

Sample Space: $\{HHH, THH, HTH, TTH, HHT, THT, HTT, TTT\}$

Winnings: if win 1 on heads, lose 1 on tails: X

$X(HHH) = 3$ $\quad X(THH) = 1$ $\quad X(HTH) = 1$ $\quad X(TTH) = -1$

$X(HHT) = 1$ $\quad X(THT) = -1$ $\quad X(HTT) = -1$ $\quad X(TTT) = -3$

Number of pips in two dice.

"What is the likelihood of getting *n* pips?"

$$Pr[X = 10] = 3/36 = Pr[X^{-1}(10)]; Pr[X = 8] = 5/36 = Pr[X^{-1}(8)].$$

Distribution

The probability of X taking on a value a.

Definition: The **distribution** of a random variable X, is $\{(a, Pr[X = a]) : a \in \mathscr{A}\}$, where $\mathscr{A}$ is the range of X.

$$Pr[X = a] := Pr[X^{-1}(a)] \text{ where } X^{-1}(a) := \{\omega \mid X(\omega) = a\}.$$

Handing back assignments

Experiment: hand back assignments to 3 students at random.
Sample Space: $\Omega = \{123, 132, 213, 231, 312, 321\}$
How many students get back their own assignment?
Random Variable: values of $X(\omega) : \{3, 1, 1, 0, 0, 1\}$

Distribution:

$$X = \begin{cases} 0, & \text{w.p. } 1/3 \\ 1, & \text{w.p. } 1/2 \\ 3, & \text{w.p. } 1/6 \end{cases}$$

Flip three coins

Experiment: flip three coins
Sample Space: $\{HHH, THH, HTH, TTH, HHT, THT, HTT, TTT\}$
Winnings: if win 1 on heads, lose 1 on tails. X
Random Variable: $\{3, 1, 1, -1, 1, -1, -1, -3\}$

Distribution:

$$X = \begin{cases} -3, & \text{w. p. } 1/8 \\ -1, & \text{w. p. } 3/8 \\ 1, & \text{w. p. } 3/8 \\ 3 & \text{w. p. } 1/8 \end{cases}$$

Number of pips.

Experiment: roll two dice.

Expectation.

How did people do on the midterm?

Distribution.

Summary of distribution?

Average!

Expectation - Definition

Definition: The **expected value** of a random variable X is

$$E[X] = \sum_a a \times Pr[X = a].$$

The expected value is also called the mean.

According to our intuition, we expect that if we repeat an experiment a large number N of times and if $X_1, \ldots, X_N$ are the successive values of the random variable, then

$$\frac{X_1 + \cdots + X_N}{N} \approx E[X].$$

That is indeed the case, in the same way that the fraction of times that $X = x$ approaches $Pr[X = x]$.

This (nontrivial) result is called the Law of Large Numbers.

The subjectivist(bayesian) interpretation of $E[X]$ is less obvious.

Expectation: A Useful Fact

Theorem:

$$E[X] = \sum_{\omega} X(\omega) \times Pr[\omega].$$

Proof:

$$\begin{aligned} E[X] &= \sum_{a} a \times Pr[X = a] \\ &= \sum_{a} a \times \sum_{\omega : X(\omega) = a} Pr[\omega] \\ &= \sum_{a} \sum_{\omega : X(\omega) = a} X(\omega) Pr[\omega] \\ &= \sum_{\omega} X(\omega) Pr[\omega] \end{aligned}$$

$\square$

Distributive property of multiplication over addition.

An Example

Flip a fair coin three times.

$\Omega = \{HHH, HHT, HTH, THH, HTT, THT, TTH, TTT\}$.

$X =$ number of H's: $\{3, 2, 2, 2, 1, 1, 1, 0\}$.

Thus,

$$\sum_{\omega} X(\omega) Pr[\omega] = \{3+2+2+2+1+1+1+0\} \times \frac{1}{8}.$$

Also,

$$\sum_{a} a \times Pr[X = a] = 3 \times \frac{1}{8} + 2 \times \frac{3}{8} + 1 \times \frac{3}{8} + 0 \times \frac{1}{8}.$$

What's the answer? Uh.... $\frac{3}{2}$

Expectation and Average.

There are n students in the class;

$X(m)$ = score of student m, for $m = 1, 2, \ldots, n$.

"Average score" of the n students: add scores and divide by n:

$$\text{Average} = \frac{X(1) + X(1) + \cdots + X(n)}{n}.$$

Experiment: choose a student uniformly at random.
Uniform sample space: $\Omega = \{1, 2, \cdots, n\}, Pr[\omega] = 1/n$, for all ω.
Random Variable: midterm score: $X(\omega)$.
Expectation:

$$E(X) = \sum_{\omega} X(\omega) Pr[\omega] = \sum_{\omega} X(\omega) \frac{1}{n}.$$

Hence,

$$\text{Average } = E(X).$$

This holds for a uniform probability space.

Named Distributions.

Some distributions come up over and over again.

...like “choose” or “stars and bars”....

Let’s cover some.

The binomial distribution.

Flip n coins with heads probability p.

Random variable: number of heads.

Binomial Distribution: $Pr[X = i]$, for each i.

How many sample points in event "$X = i$"?
i heads out of n coin flips $\implies \binom{n}{i}$

What is the probability of ω if ω has i heads?
Probability of heads in any position is p.
Probability of tails in any position is $(1 - p)$.
So, we get

$$Pr[\omega] = p^i (1 - p)^{n-i}.$$

Probability of "$X = i$" is sum of $Pr[\omega]$, $\omega \in$ "$X = i$".

$$Pr[X = i] = \binom{n}{i} p^i (1 - p)^{n-i}, i = 0, 1, \ldots, n : B(n, p) \text{ distribution}$$

The binomial distribution.

$\binom{n}{m}$ outcomes with m Hs and $n-m$ Ts

$$\Longrightarrow \boxed{Pr[X=m] = \binom{n}{m} p^m (1-p)^{n-m}}$$

Error channel and...

A packet is corrupted with probability p.

Send $n+2k$ packets.

Probability of at most k corruptions.

$$\sum_{i \leq k} \binom{n+2k}{i} p^i (1-p)^{n+2k-i}.$$

Also distribution in polling, experiments, etc.

Expectation of Binomial Distibution

Parameter p and n. What is expectation?

$$Pr[X = i] = \binom{n}{i} p^i (1-p)^{n-i}, i = 0, 1, \ldots, n : B(n, p) \text{ distribution}$$

$$E[X] = \sum_i i \times Pr[X = i].$$

Uh oh? Well... It is pn.

Proof? After linearity of expectation this is easy.

Waiting is good.

Uniform Distribution

Roll a six-sided balanced die. Let X be the number of pips (dots). Then X is equally likely to take any of the values $\{1,2,\ldots,6\}$. We say that X is *uniformly distributed* in $\{1,2,\ldots,6\}$.

More generally, we say that X is uniformly distributed in $\{1,2,\ldots,n\}$ if $Pr[X=m]=1/n$ for $m=1,2,\ldots,n$.
In that case,

$$E[X]=\sum_{m=1}^{n} mPr[X=m]=\sum_{m=1}^{n} m\times\frac{1}{n}=\frac{1}{n}\frac{n(n+1)}{2}=\frac{n+1}{2}.$$

Geometric Distribution

Let's flip a coin with $Pr[H] = p$ until we get H.

For instance:

$$\begin{aligned}
\omega_1 &= H, \text{ or} \\
\omega_2 &= T\ H, \text{ or} \\
\omega_3 &= T\ T\ H, \text{ or} \\
\omega_n &= T\ T\ T\ T\ \cdots\ T\ H.
\end{aligned}$$

Note that $\Omega = \{\omega_n, n = 1, 2, \ldots\}$.

Let X be the number of flips until the first H. Then, $X(\omega_n) = n$.

Also,

$$Pr[X = n] = (1-p)^{n-1}p,\ n \geq 1.$$

Geometric Distribution

$$Pr[X = n] = (1 - p)^{n-1} p, n \geq 1.$$

Geometric Distribution

$$Pr[X = n] = (1-p)^{n-1}p, n \geq 1.$$

Note that

$$\sum_{n=1}^{\infty} Pr[X_n] = \sum_{n=1}^{\infty} (1-p)^{n-1}p = p\sum_{n=1}^{\infty}(1-p)^{n-1} = p\sum_{n=0}^{\infty}(1-p)^{n}.$$

Now, if $|a| < 1$, then $S := \sum_{n=0}^{\infty} a^n = \frac{1}{1-a}$. Indeed,

$$\begin{aligned} S &= 1 + a + a^2 + a^3 + \cdots \\ aS &= \quad a + a^2 + a^3 + a^4 + \cdots \\ (1-a)S &= 1 + a - a + a^2 - a^2 + \cdots = 1. \end{aligned}$$

Hence,

$$\sum_{n=1}^{\infty} Pr[X_n] = p\,\frac{1}{1-(1-p)} = 1.$$

Geometric Distribution: Expectation

$$X =_D G(p), \text{ i.e., } Pr[X = n] = (1-p)^{n-1}p, n \geq 1.$$

One has

$$E[X] = \sum_{n=1}^{\infty} nPr[X = n] = \sum_{n=1}^{\infty} n(1-p)^{n-1}p.$$

Thus,

$$\begin{aligned}
E[X] &= p + 2(1-p)p + 3(1-p)^2p + 4(1-p)^3p + \cdots \\
(1-p)E[X] &= (1-p)p + 2(1-p)^2p + 3(1-p)^3p + \cdots \\
pE[X] &= p + (1-p)p + (1-p)^2p + (1-p)^3p + \cdots \\
&\qquad \text{by subtracting the previous two identities} \\
&= \sum_{n=1}^{\infty} Pr[X = n] = 1.
\end{aligned}$$

Hence,

$$E[X] = \frac{1}{p}.$$

Poisson

Experiment: flip a coin n times. The coin is such that $Pr[H] = \lambda/n$.
Random Variable: X - number of heads. Thus, $X = B(n, \lambda/n)$.

Poisson Distribution is distribution of X "for large n."

Poisson

Experiment: flip a coin n times. The coin is such that $Pr[H] = \lambda/n$.
Random Variable: X - number of heads. Thus, $X = B(n, \lambda/n)$.
Poisson Distribution is distribution of X "for large n."
We expect $X \ll n$. For $m \ll n$ one has

$$\begin{aligned}
Pr[X = m] &= \binom{n}{m} p^m (1-p)^{n-m}, \text{ with } p = \lambda/n \\
&= \frac{n(n-1)\cdots(n-m+1)}{m!} \left(\frac{\lambda}{n}\right)^m \left(1 - \frac{\lambda}{n}\right)^{n-m} \\
&= \frac{n(n-1)\cdots(n-m+1)}{n^m} \frac{\lambda^m}{m!} \left(1 - \frac{\lambda}{n}\right)^{n-m} \\
&\approx^{(1)} \frac{\lambda^m}{m!} \left(1 - \frac{\lambda}{n}\right)^{n-m} \approx^{(2)} \frac{\lambda^m}{m!} \left(1 - \frac{\lambda}{n}\right)^{n} \approx \frac{\lambda^m}{m!} e^{-\lambda}.
\end{aligned}$$

For (1) we used $m \ll n$; for (2) we used $(1 - a/n)^n \approx e^{-a}$.

Poisson Distribution: Definition and Mean

Definition Poisson Distribution with parameter $\lambda > 0$

$$X = P(\lambda) \Leftrightarrow Pr[X = m] = \frac{\lambda^m}{m!} e^{-\lambda}, m \geq 0.$$

Fact: $E[X] = \lambda$.

Proof:

$$\begin{aligned} E[X] &= \sum_{m=1}^{\infty} m \times \frac{\lambda^m}{m!} e^{-\lambda} = e^{-\lambda} \sum_{m=1}^{\infty} \frac{\lambda^m}{(m-1)!} \\ &= e^{-\lambda} \sum_{m=0}^{\infty} \frac{\lambda^{m+1}}{m!} = e^{-\lambda} \lambda \sum_{m=0}^{\infty} \frac{\lambda^m}{m!} \\ &= e^{-\lambda} \lambda e^{\lambda} = \lambda. \end{aligned}$$

$\square$

Simeon Poisson

The Poisson distribution is named after:

Equal Time: B. Geometric

The geometric distribution is named after:

I could not find a picture of D. Binomial, sorry.

Summary

Random Variables

- A random variable X is a function $X : \Omega \to \Re$.
- $Pr[X = a] := Pr[X^{-1}(a)] = Pr[\{\omega \mid X(\omega) = a\}]$.
- $Pr[X \in A] := Pr[X^{-1}(A)]$.
- The distribution of X is the list of possible values and their probability: $\{(a, Pr[X = a]), a \in \mathscr{A}\}$.
- $E[X] := \sum_a aPr[X = a]$.
- Expectation is Linear.
- $B(n, p), U[1 : n], G(p), P(\lambda)$.

BBM 205 Discrete Mathematics
Hacettepe University
http://web.cs.hacettepe.edu.tr/∼bbm205

Lecture 10b: Expectation and Important Distributions

Lecturer: Lale Özkahya

Resources:
Kenneth Rosen, "Discrete Mathematics and App."
http://www.eecs70.org/

Random Variables: Definitions

Definition

A random variable, X, for a random experiment with sample space Ω is a function $X : \Omega \to \Re$.

Thus, $X(\cdot)$ assigns a real number $X(\omega)$ to each $\omega \in \Omega$.

Definitions

(a) For $a \in \Re$, one defines

$$X^{-1}(a) := \{\omega \in \Omega \mid X(\omega) = a\}.$$

(b) For $A \subset \Re$, one defines

$$X^{-1}(A) := \{\omega \in \Omega \mid X(\omega) \in A\}.$$

(c) The probability that $X = a$ is defined as

$$Pr[X = a] = Pr[X^{-1}(a)].$$

(d) The probability that $X \in A$ is defined as

$$Pr[X \in A] = Pr[X^{-1}(A)].$$

(e) The distribution of a random variable X, is

$$\{(a, Pr[X = a]) : a \in \mathscr{A}\},$$

where $\mathscr{A}$ is the *range* of X. That is, $\mathscr{A} = \{X(\omega), \omega \in \Omega\}$.

Expectation - Definition

Definition: The **expected value** (or mean, or expectation) of a random variable X is

$$E[X] = \sum_a a \times Pr[X = a].$$

Theorem:

$$E[X] = \sum_\omega X(\omega) \times Pr[\omega].$$

An Example

Flip a fair coin three times.

$\Omega = \{HHH, HHT, HTH, THH, HTT, THT, TTH, TTT\}$.

$X =$ number of H's: $\{3, 2, 2, 2, 1, 1, 1, 0\}$.

Thus,

$$\sum_{\omega} X(\omega) Pr[\omega] = \{3+2+2+2+1+1+1+0\} \times \frac{1}{8}.$$

Also,

$$\sum_{a} a \times Pr[X = a] = 3 \times \frac{1}{8} + 2 \times \frac{3}{8} + 1 \times \frac{3}{8} + 0 \times \frac{1}{8}.$$

Win or Lose.

Expected winnings for heads/tails games, with 3 flips?
Recall the definition of the random variable X:
$\{HHH, HHT, HTH, HTT, THH, THT, TTH, TTT\} \to \{3, 1, 1, -1, 1, -1, -1, -3\}$.

$$E[X] = 3 \times \frac{1}{8} + 1 \times \frac{3}{8} - 1 \times \frac{3}{8} - 3 \times \frac{1}{8} = 0.$$

Can you ever win 0?

Apparently: expected value is not a common value, by any means.

The expected value of X is not the value that you expect!
It is the average value per experiment, if you perform the experiment many times:

$$\frac{X_1 + \cdots + X_n}{n}, \text{ when } n \gg 1.$$

The fact that this average converges to $E[X]$ is a theorem:
the Law of Large Numbers. (See later.)

Law of Large Numbers

An Illustration: Rolling Dice

Indicators

Definition

Let A be an event. The random variable X defined by

$$X(\omega) = \begin{cases} 1, & \text{if } \omega \in A \\ 0, & \text{if } \omega \notin A \end{cases}$$

is called the indicator of the event A.

Note that $Pr[X = 1] = Pr[A]$ and $Pr[X = 0] = 1 - Pr[A]$.

Hence,

$$E[X] = 1 \times Pr[X = 1] + 0 \times Pr[X = 0] = Pr[A].$$

This random variable $X(\omega)$ is sometimes written as

$$1\{\omega \in A\} \text{ or } 1_A(\omega).$$

Thus, we will write $X = 1_A$.

Linearity of Expectation

Theorem: Expectation is linear

$$E[a_1 X_1 + \cdots + a_n X_n] = a_1 E[X_1] + \cdots + a_n E[X_n].$$

Proof:

$$\begin{aligned}
&E[a_1 X_1 + \cdots + a_n X_n] \\
&\quad = \sum_{\omega} (a_1 X_1 + \cdots + a_n X_n)(\omega) Pr[\omega] \\
&\quad = \sum_{\omega} (a_1 X_1(\omega) + \cdots + a_n X_n(\omega)) Pr[\omega] \\
&\quad = a_1 \sum_{\omega} X_1(\omega) Pr[\omega] + \cdots + a_n \sum_{\omega} X_n(\omega) Pr[\omega] \\
&\quad = a_1 E[X_1] + \cdots + a_n E[X_n].
\end{aligned}$$

$\square$

Note: If we had defined $Y = a_1 X_1 + \cdots + a_n X_n$ has had tried to compute $E[Y] = \sum_y y Pr[Y = y]$, we would have been in trouble!

Using Linearity - 1: Pips (dots) on dice

Roll a die n times.

X_m = number of pips on roll m.

$X = X_1 + \cdots + X_n$ = total number of pips in n rolls.

$$\begin{aligned} E[X] &= E[X_1 + \cdots + X_n] \\ &= E[X_1] + \cdots + E[X_n], \text{ by linearity} \\ &= nE[X_1], \text{ because the } X_m \text{ have the same distribution} \end{aligned}$$

Now,

$$E[X_1] = 1 \times \frac{1}{6} + \cdots + 6 \times \frac{1}{6} = \frac{6 \times 7}{2} \times \frac{1}{6} = \frac{7}{2}.$$

Hence,

$$E[X] = \frac{7n}{2}.$$

Note: Computing $\sum_x xPr[X = x]$ directly is not easy!

Using Linearity - 2: Random assignments Example

Hand out assignments at random to n students.

$X =$ number of students that get their own assignment back.

$X = X_1 + \cdots + X_n$ where
$X_m = 1\{\text{student } m \text{ gets his/her own assignment back}\}$.

One has

$$\begin{aligned}
E[X] &= E[X_1 + \cdots + X_n] \\
&= E[X_1] + \cdots + E[X_n], \text{ by linearity} \\
&= nE[X_1], \text{ because all the } X_m \text{ have the same distribution} \\
&= nPr[X_1 = 1], \text{ because } X_1 \text{ is an indicator} \\
&= n(1/n), \text{ because student 1 is equally likely} \\
&\qquad\qquad \text{to get any one of the } n \text{ assignments} \\
&= 1.
\end{aligned}$$

Note that linearity holds even though the X_m are not independent.

Note: What is $Pr[X = m]$? Tricky ....

Using Linearity - 3: Binomial Distribution.

Flip n coins with heads probability p. X - number of heads

Binomial Distibution: $Pr[X = i]$, for each i.

$$Pr[X = i] = \binom{n}{i} p^i (1-p)^{n-i}.$$

$$E[X] = \sum_i i \times Pr[X = i] = \sum_i i \times \binom{n}{i} p^i (1-p)^{n-i}.$$

Uh oh. ... Or... a better approach: Let

$$X_i = \begin{cases} 1 & \text{if } i\text{th flip is heads} \\ 0 & \text{otherwise} \end{cases}$$

$E[X_i] = 1 \times Pr[\text{“}heads\text{''}] + 0 \times Pr[\text{“}tails\text{''}] = p.$

Moreover $X = X_1 + \cdots X_n$ and

$E[X] = E[X_1] + E[X_2] + \cdots E[X_n] = n \times E[X_i] = np.$

Calculating $E[g(X)]$

Let $Y = g(X)$. Assume that we know the distribution of X.

We want to calculate $E[Y]$.

Method 1: We calculate the distribution of Y:

$$Pr[Y = y] = Pr[X \in g^{-1}(y)] \text{ where } g^{-1}(x) = \{x \in \Re : g(x) = y\}.$$

This is typically rather tedious!

Method 2: We use the following result.

Theorem:

$$E[g(X)] = \sum_x g(x) Pr[X = x].$$

Proof:

$$\begin{aligned} E[g(X)] &= \sum_\omega g(X(\omega)) Pr[\omega] = \sum_x \sum_{\omega \in X^{-1}(x)} g(X(\omega)) Pr[\omega] \\ &= \sum_x \sum_{\omega \in X^{-1}(x)} g(x) Pr[\omega] = \sum_x g(x) \sum_{\omega \in X^{-1}(x)} Pr[\omega] \\ &= \sum_x g(x) Pr[X = x]. \end{aligned}$$

$\square$

An Example

Let X be uniform in $\{-2,-1,0,1,2,3\}$.

Let also $g(X) = X^2$. Then (method 2)

$$\begin{aligned} E[g(X)] &= \sum_{x=-2}^{3} x^2 \frac{1}{6} \\ &= \{4+1+0+1+4+9\}\frac{1}{6} = \frac{19}{6}. \end{aligned}$$

Method 1 - We find the distribution of $Y = X^2$:

$$Y = \begin{cases} 4, & \text{w.p. } \frac{2}{6} \\ 1, & \text{w.p. } \frac{2}{6} \\ 0, & \text{w.p. } \frac{1}{6} \\ 9, & \text{w.p. } \frac{1}{6}. \end{cases}$$

Thus,

$$E[Y] = 4\frac{2}{6} + 1\frac{2}{6} + 0\frac{1}{6} + 9\frac{1}{6} = \frac{19}{6}.$$

Center of Mass

The expected value has a *center of mass* interpretation:

$$\sum_n {}_n\big({}_n - \quad \big) = 0$$

$$\Leftrightarrow \quad = \sum_n {}_n \; {}_n = \quad [\quad]$$

Monotonicity

Definition

Let X, Y be two random variables on Ω. We write $X \leq Y$ if $X(\omega) \leq Y(\omega)$ for all $\omega \in \Omega$, and similarly for $X \geq Y$ and $X \geq a$ for some constant a.

Facts

(a) If $X \geq 0$, then $E[X] \geq 0$.

(b) If $X \leq Y$, then $E[X] \leq E[Y]$.

Proof

(a) If $X \geq 0$, every value a of X is nonnegative. Hence,

$$E[X] = \sum_a aPr[X = a] \geq 0.$$

(b) $X \leq Y \Rightarrow Y - X \geq 0 \Rightarrow E[Y] - E[X] = E[Y - X] \geq 0.$

$\square$

Example:

$$B = \cup_m A_m \Rightarrow 1_B(\omega) \leq \sum_m 1_{A_m}(\omega) \Rightarrow Pr[\cup_m A_m] \leq \sum_m Pr[A_m].$$

Summary

Random Variables

- A random variable X is a function $X : \Omega \to \Re$.
- $Pr[X = a] := Pr[X^{-1}(a)] = Pr[\{\omega \mid X(\omega) = a\}]$.
- $Pr[X \in A] := Pr[X^{-1}(A)]$.
- The distribution of X is the list of possible values and their probability: $\{(a, Pr[X = a]), a \in \mathscr{A}\}$.
- $E[X] := \sum_a aPr[X = a]$.
- Expectation is Linear.

Indicator Random Variable

Definition

Let A be an event. The random variable X defined by

$$X(\omega) = \begin{cases} 1, & \text{if } \omega \in A \\ 0, & \text{if } \omega \notin A \end{cases}$$

is called the indicator of the event A.

Note that $Pr[X = 1] = Pr[A]$ and $Pr[X = 0] = 1 - Pr[A]$.

Hence,

$$E[X] = 1 \times Pr[X = 1] + 0 \times Pr[X = 0] = Pr[A].$$

This random variable $X(\omega)$ is sometimes written as

$$1\{\omega \in A\} \text{ or } 1_A(\omega).$$

Thus, we will write $X = 1_A$.

Uniform Distribution

Roll a six-sided balanced die. Let X be the number of pips (dots). Then X is equally likely to take any of the values $\{1,2,\ldots,6\}$. We say that X is *uniformly distributed* in $\{1,2,\ldots,6\}$.

More generally, we say that X is uniformly distributed in $\{1,2,\ldots,n\}$ if $Pr[X=m]=1/n$ for $m=1,2,\ldots,n$.
In that case,

$$E[X]=\sum_{m=1}^{n} mPr[X=m]=\sum_{m=1}^{n} m\times\frac{1}{n}=\frac{1}{n}\frac{n(n+1)}{2}=\frac{n+1}{2}.$$

Geometric Distribution

Let's flip a coin with $Pr[H] = p$ until we get H.

For instance:

$$\begin{aligned}
\omega_1 &= H, \text{ or} \\
\omega_2 &= T\ H, \text{ or} \\
\omega_3 &= T\ T\ H, \text{ or} \\
\omega_n &= T\ T\ T\ T\ \cdots\ T\ H.
\end{aligned}$$

Note that $\Omega = \{\omega_n, n = 1, 2, \ldots\}$.

Let X be the number of flips until the first H. Then, $X(\omega_n) = n$.

Also,

$$Pr[X = n] = (1-p)^{n-1}p,\ n \geq 1.$$

Geometric Distribution

$$Pr[X = n] = (1 - p)^{n-1} p, n \geq 1.$$

Geometric Distribution

$$Pr[X = n] = (1-p)^{n-1}p, n \geq 1.$$

Note that

$$\sum_{n=1}^{\infty} Pr[X_n] = \sum_{n=1}^{\infty} (1-p)^{n-1}p = p\sum_{n=1}^{\infty}(1-p)^{n-1} = p\sum_{n=0}^{\infty}(1-p)^{n}.$$

Now, if $|a| < 1$, then $S := \sum_{n=0}^{\infty} a^n = \frac{1}{1-a}$. Indeed,

$$\begin{aligned} S &= 1 + a + a^2 + a^3 + \cdots \\ aS &= \quad a + a^2 + a^3 + a^4 + \cdots \\ (1-a)S &= 1 + a - a + a^2 - a^2 + \cdots = 1. \end{aligned}$$

Hence,

$$\sum_{n=1}^{\infty} Pr[X_n] = p\,\frac{1}{1-(1-p)} = 1.$$

Geometric Distribution: Expectation

$$X =_D G(p), \text{ i.e., } Pr[X = n] = (1-p)^{n-1}p, n \geq 1.$$

One has

$$E[X] = \sum_{n=1}^{\infty} nPr[X = n] = \sum_{n=1}^{\infty} n(1-p)^{n-1}p.$$

Thus,

$$\begin{aligned}
E[X] &= p + 2(1-p)p + 3(1-p)^2p + 4(1-p)^3p + \cdots \\
(1-p)E[X] &= (1-p)p + 2(1-p)^2p + 3(1-p)^3p + \cdots \\
pE[X] &= p + (1-p)p + (1-p)^2p + (1-p)^3p + \cdots \\
&\qquad \text{by subtracting the previous two identities} \\
&= \sum_{n=1}^{\infty} Pr[X = n] = 1.
\end{aligned}$$

Hence,

$$E[X] = \frac{1}{p}.$$

Coupon Collectors Problem.

Experiment: Get coupons at random from n until collect all n coupons.
Outcomes: $\{123145..., 56765...\}$
Random Variable: X - length of outcome.

$E[X]$=?

Time to collect coupons

X-time to get n coupons.

X_1 - time to get first coupon. Note: $X_1 = 1$. $E(X_1) = 1$.

X_2 - time to get second coupon after getting first.

$Pr[\text{“get second coupon”}|\text{“got } \cancel{\text{milk}} \text{ first coupon”}] = \frac{n-1}{n}$

$E[X_2]$? Geometric ! ! ! $\implies E[X_2] = \frac{1}{p} = \frac{1}{\frac{n-1}{n}} = \frac{n}{n-1}$.

$Pr[\text{“getting } i\text{th coupon}|\text{“got } i-1\text{rst coupons”}] = \frac{n-(i-1)}{n} = \frac{n-i+1}{n}$

$E[X_i] = \frac{1}{p} = \frac{n}{n-i+1}, i = 1, 2, \ldots, n.$

$$\begin{aligned} E[X] &= E[X_1] + \cdots + E[X_n] = \frac{n}{n} + \frac{n}{n-1} + \frac{n}{n-2} + \cdots + \frac{n}{1} \\ &= n(1 + \frac{1}{2} + \cdots + \frac{1}{n}) =: nH(n) \approx n(\ln n + \gamma) \end{aligned}$$

Review: Harmonic sum

$$H(n) = 1 + \frac{1}{2} + \cdots + \frac{1}{n} \approx \int_1^n \frac{1}{x} dx = \ln(n).$$

.

A good approximation is

$H(n) \approx \ln(n) + \gamma$ where $\gamma \approx 0.58$ (Euler-Mascheroni constant).

Geometric Distribution: Memoryless

Let X be $G(p)$. Then, for $n \geq 0$,

$$Pr[X > n] = Pr[\text{ first } n \text{ flips are } T] = (1-p)^n.$$

Theorem

$$Pr[X > n+m | X > n] = Pr[X > m], m, n \geq 0.$$

Proof:

$$\begin{aligned} Pr[X > n+m | X > n] &= \frac{Pr[X > n+m \text{ and } X > n]}{Pr[X > n]} \\ &= \frac{Pr[X > n+m]}{Pr[X > n]} \\ &= \frac{(1-p)^{n+m}}{(1-p)^n} = (1-p)^m \\ &= Pr[X > m]. \end{aligned}$$

Geometric Distribution: Memoryless - Interpretation

$$Pr[X > n+m | X > n] = Pr[X > m], m, n \geq 0.$$

$$Pr[X > n+m | X > n] = Pr[A|B] = Pr[A] = Pr[X > m].$$

The coin is memoryless, therefore, so is X.

Geometric Distribution: Yet another look

Theorem: For a r.v. X that takes the values $\{0,1,2,\ldots\}$, one has

$$E[X] = \sum_{i=1}^{\infty} Pr[X \geq i].$$

[See later for a proof.]

If $X = G(p)$, then $Pr[X \geq i] = Pr[X > i-1] = (1-p)^{i-1}$.

Hence,

$$E[X] = \sum_{i=1}^{\infty} (1-p)^{i-1} = \sum_{i=0}^{\infty} (1-p)^{i} = \frac{1}{1-(1-p)} = \frac{1}{p}.$$

Expected Value of Integer RV

Theorem: For a r.v. X that takes values in $\{0,1,2,\ldots\}$, one has

$$E[X] = \sum_{i=1}^{\infty} Pr[X \geq i].$$

Proof: One has

$$\begin{aligned}
E[X] &= \sum_{i=1}^{\infty} i \times Pr[X = i] \\
&= \sum_{i=1}^{\infty} i\{Pr[X \geq i] - Pr[X \geq i+1]\} \\
&= \sum_{i=1}^{\infty} \{i \times Pr[X \geq i] - i \times Pr[X \geq i+1]\} \\
&= \sum_{i=1}^{\infty} \{i \times Pr[X \geq i] - (i-1) \times Pr[X \geq i]\} \\
&= \sum_{i=1}^{\infty} Pr[X \geq i].
\end{aligned}$$

$\square$

Poisson

Experiment: flip a coin n times. The coin is such that $Pr[H] = \lambda/n$.
Random Variable: X - number of heads. Thus, $X = B(n, \lambda/n)$.

Poisson Distribution is distribution of X "for large n."

Poisson

Experiment: flip a coin n times. The coin is such that $Pr[H] = \lambda/n$.
Random Variable: X - number of heads. Thus, $X = B(n, \lambda/n)$.
Poisson Distribution is distribution of X "for large n."
We expect $X \ll n$. For $m \ll n$ one has

$$\begin{aligned}
Pr[X = m] &= \binom{n}{m} p^m (1-p)^{n-m}, \text{ with } p = \lambda/n \\
&= \frac{n(n-1)\cdots(n-m+1)}{m!} \left(\frac{\lambda}{n}\right)^m \left(1 - \frac{\lambda}{n}\right)^{n-m} \\
&= \frac{n(n-1)\cdots(n-m+1)}{n^m} \frac{\lambda^m}{m!} \left(1 - \frac{\lambda}{n}\right)^{n-m} \\
&\approx^{(1)} \frac{\lambda^m}{m!} \left(1 - \frac{\lambda}{n}\right)^{n-m} \approx \frac{\lambda^m}{m!} \left(1 - \frac{\lambda}{n}\right)^{n} \approx^{(2)} \frac{\lambda^m}{m!} e^{-\lambda}.
\end{aligned}$$

For (1) we used $m \ll n$; for (2) we used $(1 - a/n) \approx e^{-a/n}$ for $a/n \ll 1$.

Poisson Distribution: Definition and Mean

Definition Poisson Distribution with parameter $\lambda > 0$

$$X = P(\lambda) \Leftrightarrow Pr[X = m] = \frac{\lambda^m}{m!} e^{-\lambda}, m \geq 0.$$

Fact: $E[X] = \lambda$.

Proof:

$$\begin{aligned} E[X] &= \sum_{m=1}^{\infty} m \times \frac{\lambda^m}{m!} e^{-\lambda} = e^{-\lambda} \sum_{m=1}^{\infty} \frac{\lambda^m}{(m-1)!} \\ &= e^{-\lambda} \sum_{m=0}^{\infty} \frac{\lambda^{m+1}}{m!} = e^{-\lambda} \lambda \sum_{m=0}^{\infty} \frac{\lambda^m}{m!} \\ &= e^{-\lambda} \lambda e^{\lambda} = \lambda. \end{aligned}$$

$\square$

Summary.

$\boxed{\text{Distributions}}$

- $U[1,\ldots,n] : Pr[X=m] = \frac{1}{n}, m = 1,\ldots,n;$
  $E[X] = \frac{n+1}{2};$
- $B(n,p) : Pr[X=m] = \binom{n}{m} p^m (1-p)^{n-m}, m = 0,\ldots,n;$
  $E[X] = np;$
- $G(p) : Pr[X=n] = (1-p)^{n-1} p, n = 1,2,\ldots;$
  $E[X] = \frac{1}{p};$
- $P(\lambda) : Pr[X=n] = \frac{\lambda^n}{n!} e^{-\lambda}, n \geq 0;$
  $E[X] = \lambda.$

BBM 205 Discrete Mathematics
Hacettepe University

Lecture 11a: Maximum Matchings

Lale Özkahya

Resources:
Kenneth Rosen, "Discrete Mathematics and App."
http://www.cs.cmu.edu/ ./15251/schedule.html

Some motivating real-world examples

matching machines and jobs

Job 1

Job 2

⋮ ⋮

Job n

Some motivating real-world examples

matching professors and courses

⋮

15-110

15-112

15-122

15-150

15-251

⋮

Some motivating real-world examples

matching rooms and courses

GHC 4401	15-110
DH 2210	15-112
GHC 5222	15-122
WEH 7500	15-150
DH 2315	15-251
⋮	⋮

Some motivating real-world examples

matching students and internships

Macrosoft

Google

UMBRELLA CORPORATION

"Our business is life itself..."

Some motivating real-world examples

matching kidney donors and patients

How do you solve a problem like this?

1. Formulate the problem

2. **Ask**: Is there a trivial algorithm?

3. **Ask**: Is there a better algorithm?

4. Find and analyze

Remember the CS life lesson

First step: Formulate the problem

Purpose:

- Get rid of all the distractions
- Identify the crux of the problem
- Get a clean mathematical model that is easy to reason about.

Bipartite Graphs

$G = (V, E)$ is **bipartite** if:

- there exists a bipartition of V into X and Y
- each edge connects a vertex in X to a vertex in Y

Given a graph $G = (V, E)$, we could ask, is it bipartite?

Bipartite Graphs

Given a graph $G = (V, E)$, we could ask, is it bipartite?

Poll

Is this graph bipartite?

- Yes
- No
- Beats me

Poll Answer

Is this graph bipartite?

bipartite = 2-colorable

Color the vertices with **2** **colors** so that no edge's endpoints get the same color.

Important Characterization

<u>An obstruction for being bipartite:</u>

Contains a cycle of odd length.

Is this the only type of obstruction?

<u>Theorem:</u>

A graph G=(V,E) is bipartite **<u>if and only if</u>** it contains no cycles of odd length.

Bipartite Graphs

Often we write the bipartition explicitly:

$$G = (X, Y, E)$$

Bipartite Graphs

Great at modeling relations between two classes of objects.

Examples:

X = machines, Y = jobs

An edge $\{x, y\}$ means x is capable of doing y.

X = professors, Y = courses

An edge $\{x, y\}$ means x can teach y.

X = students, Y = internship jobs

An edge $\{x, y\}$ means x and y are interested in each other.

$\vdots$

Matchings in bipartite graphs

Often, we are interested in finding a **matching** in a bipartite graph

A **matching** :

A <u>subset of the edges</u> that do not share an endpoint.

Matchings in bipartite graphs

Often, we are interested in finding a **matching** in a bipartite graph

matching

A **matching** :

A <u>subset of the edges</u> that do not share an endpoint.

Matchings in bipartite graphs

Often, we are interested in finding a **matching** in a bipartite graph

matching

A **matching** :

A <u>subset of the edges</u> that do not share an endpoint.

Matchings in bipartite graphs

Often, we are interested in finding a **matching** in a bipartite graph

not a matching

A **matching** :

A subset of the edges that do not share an endpoint.

Matchings in bipartite graphs

Often, we are interested in finding a **matching** in a bipartite graph

maximum matching

Maximum matching: a matching with largest number of edges (among all possible matchings).

Matchings in bipartite graphs

Often, we are interested in finding a **matching** in a bipartite graph

maximal matching

Cannot add more edges.

"*local optimum*"

Maximal matching: a matching which cannot contain any more edges.

Matchings in bipartite graphs

Often, we are interested in finding a **matching** in a bipartite graph

perfect matching

a necessary condition for perfect matching: $|X| = |Y|$

Perfect matching: a matching that covers all vertices.

Poll

How many different perfect matchings does the graph have (in terms of n)?

$$|X| = |Y| = n$$

Important Note

We can define matchings for <u>non-bipartite</u> graphs as well.

Important Note

We can define matchings for <u>non-bipartite</u> graphs as well.

Maximum matching problem

The problem we want to solve is:

Maximum matching problem

Input: A graph $G = (V, E)$.

Output: A maximum matching in G.

Bipartite maximum matching problem

Actually, we want to solve the following restriction:

Bipartite maximum matching problem

Input: A *bipartite* graph $G = (X, Y, E)$.

Output: A maximum matching in G.

How do you solve a problem like this?

1. Formulate the problem

2. **Ask**: Is there a trivial algorithm?

3. **Ask**: Is there a better algorithm?

4. Find and analyze

Bipartite maximum matching problem

Bipartite maximum matching problem

Input: A *bipartite* graph $G = (X, Y, E)$.

Output: A maximum matching in G.

Is there a (trivial) algorithm to solve this problem?

- Try all possible subsets of the edges.

Running time: $\Omega(2^m)$

How do you solve a problem like this?

1. Formulate the problem

2. **Ask**: Is there a trivial algorithm?

3. **Ask**: Is there a better algorithm?

4. Find and analyze

Bipartite maximum matching problem

A good first attempt:

What if we picked edges "greedily"?

Bipartite maximum matching problem

A good first attempt:

What if we picked edges "greedily"?

Bipartite maximum matching problem

A good first attempt:

What if we picked edges "greedily"?

Bipartite maximum matching problem

A good first attempt:

What if we picked edges "greedily"?

maximal matching

but not maximum

Is there a way to get out of this *local optimum*?

What is interesting about the path 4 – 8 – 2 – 5 – 1 – 7 ?

Bipartite maximum matching problem

A good first attempt:

What if we picked edges "greedily"?

maximal matching

but not maximum

Bipartite maximum matching problem

A good first attempt:

What if we picked edges "greedily"?

now maximum

Important Definition: Augmenting paths

Let **M** be some matching.

An ***alternating path*** with respect to **M** is a path in **G** such that:

- the edges in the path alternate between being in **M** and not being in **M**

An ***augmenting path*** with respect to **M** is an alternating path such that:

- the first and last vertices are **not** matched by **M**

Important Definition: Augmenting paths

matching = red edges

Augmenting path:

4-8-2-5-1-7

augmenting path $\Longrightarrow$ can obtain a bigger matching.

Important Definition: Augmenting paths

matching = red edges

Augmenting path:

2-5-1-7

An augmenting path
need **not** contain
all the edges of the matching.

augmenting path $\Longrightarrow$ can obtain a bigger matching.

Important Definition: Augmenting paths

matching = red edges

Augmenting path:

4-8

An augmenting path
need **not** contain
any of the edges of the matching.

augmenting path $\Longrightarrow$ can obtain a bigger matching.

Augmenting paths and maximum matchings

augmenting path $\implies$ can obtain a bigger matching.

In fact, it turns out:

no augmenting path $\implies$ maximum matching.

Theorem:

A matching **M** is maximum **if and only if** there is **no** augmenting path with respect to **M**.

Augmenting paths and maximum matchings

Proof:

If there is an augmenting path with respect to **M**, we saw that **M** is not maximum.

Want to show:

If **M** not maximum, there is an augmenting path w.r.t. **M**.

Let **M*** be a maximum matching. |**M***| > |**M**|.

Let **S** be the set of edges contained in **M*** or **M** but not both.

S = (**M*** ∪ **M**) - (**M** ∩ **M***)

Augmenting paths and maximum matchings

Proof (continued):

Let **S** be the set of edges contained in **M*** or **M** but not both.

S = (**M*** ∪ **M**) - (**M** ∩ **M***)

(will find an augmenting path in **S**)

What does **S** look like?

Each vertex has degree 1 or 2. (why?)

So **S** is a collection of disjoint **cycles** and **paths**.

(exercise)

The edges alternate **red** and **blue**.

Augmenting paths and maximum matchings

Proof (continued):

Let **S** be the set of edges contained in **M*** or **M** but not both.

S = (**M*** ∪ **M**) - (**M** ∩ **M***)

So **S** is a collection of disjoint **cycles** and **paths**.

The edges alternate **red** and **blue**.

red > # blue in **S**

red = # blue in **cycles**

So ∃ a **path** with # red > # blue.

This is an ***augmenting path*** with respect to **M**. □

Augmenting paths and maximum matchings

Theorem:

A matching **M** is maximum **if and only if** there is **no** augmenting path with respect to **M**.

Summary of proof:

$\Longrightarrow$

If there is an augmenting path, not a max matching.

$\Longleftarrow$

If the matching **M** is not maximum, $\exists$ **M*** s.t. |**M***| > |**M**|.

Can find an augmenting path w.r.t. **M** in

the "symmetric difference" of **M*** and **M**.

Next time:

- Algorithm to find a maximum matching in bipartite graphs.

- Stable matchings.

BBM 205 Discrete Mathematics
Hacettepe University

Lecture 11b: Stable Matchings

Lale Özkahya

Resources:
http://www.cs.cmu.edu/ ./15251/schedule.html
http://www.cs.princeton.edu/courses/archive/spring13/cos423

From Last Time

Bipartite maximum matching problem

Bipartite maximum matching problem

Input: A *bipartite* graph $G = (X, Y, E)$.

Output: A maximum matching in G.

Important Definition: Augmenting paths

Let **M** be some matching.

An ***augmenting path*** with respect to **M** is an alternating path such that:

- the first and last vertices are **not** matched by **M**

Algorithm to find maximum matching

Theorem:

A matching **M** is maximum **if and only if** there is **no** augmenting path with respect to **M**.

Algorithm:

- Start with a single edge as your matching **M**.
- Repeat until there is no augmenting path w.r.t. **M**:
  - Find an augmenting path with respect to **M**.
  - Update **M** according to the augmenting path.

OK, but how do you find an augmenting path?

Algorithm to find augmenting path

Algorithm to find augmenting path

- direct edges **<u>not</u>** in **M** from left to right (X to Y).

Algorithm to find augmenting path

- direct edges <u>**not**</u> in **M** from left to right (X to Y).
- direct edges in **M** from right to left (Y to X).

Algorithm to find augmenting path

- direct edges **not** in **M** from left to right (X to Y).
- direct edges in **M** from right to left (Y to X).

Observation:

There is an augmenting path **iff**
there is a directed path from an *unmatched* $x \in X$
to an *unmatched* $y \in Y$.

Algorithm to find augmenting path

<u>Algorithm</u>:

- for each *unmatched* $x \in X$:
  - do DFS(x), stop when you find *unmatched* $y \in Y$.

<u>Running time</u>: $O(n+m)$

Important Note

Theorem:

A matching **M** is maximum **if and only if** there is **no** augmenting path with respect to **M**.

This theorem holds for all graphs.

The algorithm works for bipartite graphs.

How do you solve a problem like this?

1. Formulate the problem

2. **Ask**: Is there a trivial algorithm?

3. **Ask**: Is there a better algorithm?

4. Find and analyze

Hall's Theorem

Characterization for perfect matchings

Often we are interested in perfect matchings.

<u>An obstruction:</u>

$$|X| \neq |Y|$$

Characterization for perfect matchings

Often we are interested in perfect matchings.

<u>An obstruction:</u>

If $|X| > |Y|$, we cannot "cover" all the nodes in X.

If $|X| > |N(X)|$, we cannot "cover" all the nodes in X.

Characterization for perfect matchings

Often we are interested in perfect matchings.

$$S = \{1, 3, 4\}$$

$$N(S) = \{5, 7\}$$

<u>An obstruction:</u>

For $S \subseteq X$:

if $|S| > |N(S)|$, we cannot "cover" all the nodes in S.

Characterization for perfect matchings

Is this the only type of obstruction?

Theorem [Hall's Theorem]:

Let $G = (X, Y, E)$ be a bipartite graph.

There is a matching covering all vertices in X **iff**

$$\forall S \subseteq X : \quad |S| \leq |N(S)| \, .$$

Corollary:

$G = (X, Y, E)$ has a perfect matching **iff**

$|X| = |Y|$ and $\forall S \subseteq X, \quad |S| \leq |N(S)| \, .$

An application of Hall's Theorem

Rank: 1 2 3 4 5 6 7 8 9 10 J Q K

Suppose a deck of cards is dealt into 13 piles of 4 cards each.

<u>Claim</u>: there is a way to select one card from each pile so that you have one card from each rank.

An application of Hall's Theorem

$$|X| = |Y|$$

<u>Want to show:</u>

For any $S \subseteq X$, $|S| \leq |N(S)|$.

An application of Hall's Theorem

For any $S \subseteq X$,
total weight coming out $= 4|S|$.

All this weight is absorbed by $N(S)$.

Each $y \in N(S)$ absorbs ≤ 4 units of this weight.

$\implies N(S)$ absorbs ≤ $4|N(S)|$ units. $\implies \cancel{4}|S| \leq \cancel{4}|N(S)|$

Stable matching problem

2-Sided Markets

A market with 2 distinct groups of participants each with their own preferences.

2-Sided Markets

1.

2.

3.

4.

1. Alice
2. Bob
3. Charlie
4. David

-
-
-

Other examples:

medical residents - hospitals

students - colleges

professors - colleges

⋮

1. Bob
2. David
3. Alice
4. Charlie

Aspiration: A Good Centeralized System

What can go wrong?

Suppose Alice gets "matched" with Macrosoft.
Charlie gets "matched" with Umbrella.

But, say, Alice prefers Umbrella over Macrosoft and Umbrella prefers Alice over Charlie.

Formalizing the problem

An instance of the problem can be represented as a *complete bipartite graph* + *preference list of each node.*

$$|X| = |Y| = n$$

Goal: Find a stable matching.

Formalizing the problem

What is a stable matching?

1. It has to be a perfect matching.

2. Cannot contain an unstable pair:

   A pair (x, y) <u>unmatched</u>
   but they prefer each other over their current partners.

Formalizing the problem

What is a stable matching?

(a, e) is an unstable pair.

1. It has to be a perfect matching.

2. Cannot contain an unstable pair:

   A pair (x, y) <u>unmatched</u>
   but they prefer each other over their current partners.

Formalizing the problem

An instance of the problem can be represented as a *complete bipartite graph* + *preference list of each node.*

$$|X| = |Y| = n$$

Goal: Find a stable matching.

(Is it guaranteed to always exist?)

A variant: Roommate problem

A non-bipartite version

(c,b,d) a

c (b,a,d)

(a,c,d) b

d (a,c,b)

Does this have a stable matching?

Stable matching: Is there a trivial algorithm?

Trivial algorithm:

Try all possible perfect matchings,

and check if it is stable.

perfect matchings in terms $n = |X|$:

Stable matching: Is there a trivial algorithm?

Trivial algorithm:

Try all possible perfect matchings,

and check if it is stable.

perfect matchings in terms $n = |X|$: $n!$

Gale–Shapley demo

hospitals' preference lists

	1st	2nd	3rd	4th	5th
Atlanta	Wayne	Val	Yolanda	Zeus	Xavier
Boston	Yolanda	Wayne	Val	Xavier	Zeus
Chicago	Wayne	Zeus	Xavier	Yolanda	Val
Dallas	Val	Yolanda	Xavier	Wayne	Zeus
Eugene	Wayne	Yolanda	Val	Zeus	Xavier

students' preference lists

	1st	2nd	3rd	4th	5th
Val	Eugene	Atlanta	Boston	Dallas	Chicago
Wayne	Chicago	Boston	Dallas	Atlanta	Eugene
Xavier	Boston	Chicago	Dallas	Eugene	Atlanta
Yolanda	Atlanta	Eugene	Dallas	Chicago	Boston
Zeus	Dallas	Boston	Eugene	Chicago	Atlanta

Gale–Shapley demo

hospitals' preference lists

	1st	2nd	3rd	4th	5th
Atlanta	Wayne	Val	Yolanda	Zeus	Xavier
Boston	Yolanda	Wayne	Val	Xavier	Zeus
Chicago	Wayne	Zeus	Xavier	Yolanda	Val
Dallas	Val	Yolanda	Xavier	Wayne	Zeus
Eugene	Wayne	Yolanda	Val	Zeus	Xavier

students' preference lists

	1st	2nd	3rd	4th	5th
Val	Eugene	Atlanta	Boston	Dallas	Chicago
Wayne	Chicago	Boston	Dallas	Atlanta	Eugene
Xavier	Boston	Chicago	Dallas	Eugene	Atlanta
Yolanda	Atlanta	Eugene	Dallas	Chicago	Boston
Zeus	Dallas	Boston	Eugene	Chicago	Atlanta

Atlanta proposes to Wayne

Gale–Shapley demo

hospitals' preference lists

	1st	2nd	3rd	4th	5th
Atlanta	Wayne	Val	Yolanda	Zeus	Xavier
Boston	Yolanda	Wayne	Val	Xavier	Zeus
Chicago	Wayne	Zeus	Xavier	Yolanda	Val
Dallas	Val	Yolanda	Xavier	Wayne	Zeus
Eugene	Wayne	Yolanda	Val	Zeus	Xavier

students' preference lists

	1st	2nd	3rd	4th	5th
Val	Eugene	Atlanta	Boston	Dallas	Chicago
Wayne	Chicago	Boston	Dallas	Atlanta	Eugene
Xavier	Boston	Chicago	Dallas	Eugene	Atlanta
Yolanda	Atlanta	Eugene	Dallas	Chicago	Boston
Zeus	Dallas	Boston	Eugene	Chicago	Atlanta

Atlanta proposes to Wayne

Wayne accepts
(since previously unmatched)

Gale–Shapley demo

hospitals' preference lists

	1st	2nd	3rd	4th	5th
Atlanta	Wayne	Val	Yolanda	Zeus	Xavier
Boston	Yolanda	Wayne	Val	Xavier	Zeus
Chicago	Wayne	Zeus	Xavier	Yolanda	Val
Dallas	Val	Yolanda	Xavier	Wayne	Zeus
Eugene	Wayne	Yolanda	Val	Zeus	Xavier

students' preference lists

	1st	2nd	3rd	4th	5th
Val	Eugene	Atlanta	Boston	Dallas	Chicago
Wayne	Chicago	Boston	Dallas	Atlanta	Eugene
Xavier	Boston	Chicago	Dallas	Eugene	Atlanta
Yolanda	Atlanta	Eugene	Dallas	Chicago	Boston
Zeus	Dallas	Boston	Eugene	Chicago	Atlanta

Boston proposes to Yolanda

Gale–Shapley demo

hospitals' preference lists

	1st	2nd	3rd	4th	5th
Atlanta	Wayne	Val	Yolanda	Zeus	Xavier
Boston	Yolanda	Wayne	Val	Xavier	Zeus
Chicago	Wayne	Zeus	Xavier	Yolanda	Val
Dallas	Val	Yolanda	Xavier	Wayne	Zeus
Eugene	Wayne	Yolanda	Val	Zeus	Xavier

students' preference lists

	1st	2nd	3rd	4th	5th
Val	Eugene	Atlanta	Boston	Dallas	Chicago
Wayne	Chicago	Boston	Dallas	Atlanta	Eugene
Xavier	Boston	Chicago	Dallas	Eugene	Atlanta
Yolanda	Atlanta	Eugene	Dallas	Chicago	Boston
Zeus	Dallas	Boston	Eugene	Chicago	Atlanta

Boston proposes to Yolanda

Yolanda accepts
(since previously unmatched)

Gale–Shapley demo

hospitals' preference lists

	1st	2nd	3rd	4th	5th
Atlanta	Wayne	Val	Yolanda	Zeus	Xavier
Boston	Yolanda	Wayne	Val	Xavier	Zeus
Chicago	Wayne	Zeus	Xavier	Yolanda	Val
Dallas	Val	Yolanda	Xavier	Wayne	Zeus
Eugene	Wayne	Yolanda	Val	Zeus	Xavier

students' preference lists

	1st	2nd	3rd	4th	5th
Val	Eugene	Atlanta	Boston	Dallas	Chicago
Wayne	Chicago	Boston	Dallas	Atlanta	Eugene
Xavier	Boston	Chicago	Dallas	Eugene	Atlanta
Yolanda	Atlanta	Eugene	Dallas	Chicago	Boston
Zeus	Dallas	Boston	Eugene	Chicago	Atlanta

Chicago proposes to Wayne

Gale–Shapley demo

hospitals' preference lists

	1st	2nd	3rd	4th	5th
Atlanta	Wayne	Val	Yolanda	Zeus	Xavier
Boston	Yolanda	Wayne	Val	Xavier	Zeus
Chicago	Wayne	Zeus	Xavier	Yolanda	Val
Dallas	Val	Yolanda	Xavier	Wayne	Zeus
Eugene	Wayne	Yolanda	Val	Zeus	Xavier

students' preference lists

	1st	2nd	3rd	4th	5th
Val	Eugene	Atlanta	Boston	Dallas	Chicago
Wayne	Chicago	Boston	Dallas	Atlanta	Eugene
Xavier	Boston	Chicago	Dallas	Eugene	Atlanta
Yolanda	Atlanta	Eugene	Dallas	Chicago	Boston
Zeus	Dallas	Boston	Eugene	Chicago	Atlanta

Chicago proposes to Wayne

Wayne accepts
(and renounces Atlanta)

Gale–Shapley demo

hospitals' preference lists

	1st	2nd	3rd	4th	5th
Atlanta	Wayne	Val	Yolanda	Zeus	Xavier
Boston	Yolanda	Wayne	Val	Xavier	Zeus
Chicago	Wayne	Zeus	Xavier	Yolanda	Val
Dallas	Val	Yolanda	Xavier	Wayne	Zeus
Eugene	Wayne	Yolanda	Val	Zeus	Xavier

students' preference lists

	1st	2nd	3rd	4th	5th
Val	Eugene	Atlanta	Boston	Dallas	Chicago
Wayne	Chicago	Boston	Dallas	Atlanta	Eugene
Xavier	Boston	Chicago	Dallas	Eugene	Atlanta
Yolanda	Atlanta	Eugene	Dallas	Chicago	Boston
Zeus	Dallas	Boston	Eugene	Chicago	Atlanta

Atlanta proposes to Val

Gale–Shapley demo

hospitals' preference lists

	1st	2nd	3rd	4th	5th
Atlanta	Wayne	Val	Yolanda	Zeus	Xavier
Boston	Yolanda	Wayne	Val	Xavier	Zeus
Chicago	Wayne	Zeus	Xavier	Yolanda	Val
Dallas	Val	Yolanda	Xavier	Wayne	Zeus
Eugene	Wayne	Yolanda	Val	Zeus	Xavier

students' preference lists

	1st	2nd	3rd	4th	5th
Val	Eugene	Atlanta	Boston	Dallas	Chicago
Wayne	Chicago	Boston	Dallas	Atlanta	Eugene
Xavier	Boston	Chicago	Dallas	Eugene	Atlanta
Yolanda	Atlanta	Eugene	Dallas	Chicago	Boston
Zeus	Dallas	Boston	Eugene	Chicago	Atlanta

Atlanta proposes to Val

Val accepts
(since previously unmatched)

Gale–Shapley demo

hospitals' preference lists

	1st	2nd	3rd	4th	5th
Atlanta	Wayne	Val	Yolanda	Zeus	Xavier
Boston	Yolanda	Wayne	Val	Xavier	Zeus
Chicago	Wayne	Zeus	Xavier	Yolanda	Val
Dallas	Val	Yolanda	Xavier	Wayne	Zeus
Eugene	Wayne	Yolanda	Val	Zeus	Xavier

students' preference lists

	1st	2nd	3rd	4th	5th
Val	Eugene	Atlanta	Boston	Dallas	Chicago
Wayne	Chicago	Boston	Dallas	Atlanta	Eugene
Xavier	Boston	Chicago	Dallas	Eugene	Atlanta
Yolanda	Atlanta	Eugene	Dallas	Chicago	Boston
Zeus	Dallas	Boston	Eugene	Chicago	Atlanta

Dallas proposes to Val

Gale–Shapley demo

hospitals' preference lists

	1st	2nd	3rd	4th	5th
Atlanta	Wayne	Val	Yolanda	Zeus	Xavier
Boston	Yolanda	Wayne	Val	Xavier	Zeus
Chicago	Wayne	Zeus	Xavier	Yolanda	Val
Dallas	Val	Yolanda	Xavier	Wayne	Zeus
Eugene	Wayne	Yolanda	Val	Zeus	Xavier

students' preference lists

	1st	2nd	3rd	4th	5th
Val	Eugene	Atlanta	Boston	Dallas	Chicago
Wayne	Chicago	Boston	Dallas	Atlanta	Eugene
Xavier	Boston	Chicago	Dallas	Eugene	Atlanta
Yolanda	Atlanta	Eugene	Dallas	Chicago	Boston
Zeus	Dallas	Boston	Eugene	Chicago	Atlanta

Dallas proposes to Val

Val rejects
(since she prefers Atlanta)

Gale–Shapley demo

hospitals' preference lists

	1st	2nd	3rd	4th	5th
Atlanta	Wayne	Val	Yolanda	Zeus	Xavier
Boston	Yolanda	Wayne	Val	Xavier	Zeus
Chicago	Wayne	Zeus	Xavier	Yolanda	Val
Dallas	Val	Yolanda	Xavier	Wayne	Zeus
Eugene	Wayne	Yolanda	Val	Zeus	Xavier

students' preference lists

	1st	2nd	3rd	4th	5th
Val	Eugene	Atlanta	Boston	Dallas	Chicago
Wayne	Chicago	Boston	Dallas	Atlanta	Eugene
Xavier	Boston	Chicago	Dallas	Eugene	Atlanta
Yolanda	Atlanta	Eugene	Dallas	Chicago	Boston
Zeus	Dallas	Boston	Eugene	Chicago	Atlanta

Dallas proposes to Yolanda

Gale–Shapley demo

hospitals' preference lists

	1st	2nd	3rd	4th	5th
Atlanta	Wayne	Val	Yolanda	Zeus	Xavier
Boston	Yolanda	Wayne	Val	Xavier	Zeus
Chicago	Wayne	Zeus	Xavier	Yolanda	Val
Dallas	Val	Yolanda	Xavier	Wayne	Zeus
Eugene	Wayne	Yolanda	Val	Zeus	Xavier

students' preference lists

	1st	2nd	3rd	4th	5th
Val	Eugene	Atlanta	Boston	Dallas	Chicago
Wayne	Chicago	Boston	Dallas	Atlanta	Eugene
Xavier	Boston	Chicago	Dallas	Eugene	Atlanta
Yolanda	Atlanta	Eugene	Dallas	Chicago	Boston
Zeus	Dallas	Boston	Eugene	Chicago	Atlanta

Dallas proposes to Yolanda

Dianne accepts
(and renounces Boston)

Gale–Shapley demo

hospitals' preference lists

	1st	2nd	3rd	4th	5th
Atlanta	Wayne	Val	Yolanda	Zeus	Xavier
Boston	Yolanda	Wayne	Val	Xavier	Zeus
Chicago	Wayne	Zeus	Xavier	Yolanda	Val
Dallas	Val	Yolanda	Xavier	Wayne	Zeus
Eugene	Wayne	Yolanda	Val	Zeus	Xavier

students' preference lists

	1st	2nd	3rd	4th	5th
Val	Eugene	Atlanta	Boston	Dallas	Chicago
Wayne	Chicago	Boston	Dallas	Atlanta	Eugene
Xavier	Boston	Chicago	Dallas	Eugene	Atlanta
Yolanda	Atlanta	Eugene	Dallas	Chicago	Boston
Zeus	Dallas	Boston	Eugene	Chicago	Atlanta

Boston proposes to Wayne

Gale–Shapley demo

hospitals' preference lists

	1st	2nd	3rd	4th	5th
Atlanta	Wayne	Val	Yolanda	Zeus	Xavier
Boston	Yolanda	Wayne	Val	Xavier	Zeus
Chicago	Wayne	Zeus	Xavier	Yolanda	Val
Dallas	Val	Yolanda	Xavier	Wayne	Zeus
Eugene	Wayne	Yolanda	Val	Zeus	Xavier

students' preference lists

	1st	2nd	3rd	4th	5th
Val	Eugene	Atlanta	Boston	Dallas	Chicago
Wayne	Chicago	Boston	Dallas	Atlanta	Eugene
Xavier	Boston	Chicago	Dallas	Eugene	Atlanta
Yolanda	Atlanta	Eugene	Dallas	Chicago	Boston
Zeus	Dallas	Boston	Eugene	Chicago	Atlanta

Boston proposes to Wayne

Wayne rejects
(since he prefers Chicago)

Gale–Shapley demo

hospitals' preference lists

	1st	2nd	3rd	4th	5th
Atlanta	Wayne	Val	Yolanda	Zeus	Xavier
Boston	Yolanda	Wayne	Val	Xavier	Zeus
Chicago	Wayne	Zeus	Xavier	Yolanda	Val
Dallas	Val	Yolanda	Xavier	Wayne	Zeus
Eugene	Wayne	Yolanda	Val	Zeus	Xavier

students' preference lists

	1st	2nd	3rd	4th	5th
Val	Eugene	Atlanta	Boston	Dallas	Chicago
Wayne	Chicago	Boston	Dallas	Atlanta	Eugene
Xavier	Boston	Chicago	Dallas	Eugene	Atlanta
Yolanda	Atlanta	Eugene	Dallas	Chicago	Boston
Zeus	Dallas	Boston	Eugene	Chicago	Atlanta

Boston proposes to Val

Gale–Shapley demo

hospitals' preference lists

	1st	2nd	3rd	4th	5th
Atlanta	Wayne	Val	Yolanda	Zeus	Xavier
Boston	Yolanda	Wayne	Val	Xavier	Zeus
Chicago	Wayne	Zeus	Xavier	Yolanda	Val
Dallas	Val	Yolanda	Xavier	Wayne	Zeus
Eugene	Wayne	Yolanda	Val	Zeus	Xavier

students' preference lists

	1st	2nd	3rd	4th	5th
Val	Eugene	Atlanta	Boston	Dallas	Chicago
Wayne	Chicago	Boston	Dallas	Atlanta	Eugene
Xavier	Boston	Chicago	Dallas	Eugene	Atlanta
Yolanda	Atlanta	Eugene	Dallas	Chicago	Boston
Zeus	Dallas	Boston	Eugene	Chicago	Atlanta

Boston proposes to Val

Val rejects
(since she prefers Atlanta)

Gale–Shapley demo

hospitals' preference lists

	1st	2nd	3rd	4th	5th
Atlanta	Wayne	Val	Yolanda	Zeus	Xavier
Boston	Yolanda	Wayne	Val	Xavier	Zeus
Chicago	Wayne	Zeus	Xavier	Yolanda	Val
Dallas	Val	Yolanda	Xavier	Wayne	Zeus
Eugene	Wayne	Yolanda	Val	Zeus	Xavier

students' preference lists

	1st	2nd	3rd	4th	5th
Val	Eugene	Atlanta	Boston	Dallas	Chicago
Wayne	Chicago	Boston	Dallas	Atlanta	Eugene
Xavier	Boston	Chicago	Dallas	Eugene	Atlanta
Yolanda	Atlanta	Eugene	Dallas	Chicago	Boston
Zeus	Dallas	Boston	Eugene	Chicago	Atlanta

Boston proposes to Xavier

Gale–Shapley demo

hospitals' preference lists

	1st	2nd	3rd	4th	5th
Atlanta	Wayne	Val	Yolanda	Zeus	Xavier
Boston	Yolanda	Wayne	Val	Xavier	Zeus
Chicago	Wayne	Zeus	Xavier	Yolanda	Val
Dallas	Val	Yolanda	Xavier	Wayne	Zeus
Eugene	Wayne	Yolanda	Val	Zeus	Xavier

students' preference lists

	1st	2nd	3rd	4th	5th
Val	Eugene	Atlanta	Boston	Dallas	Chicago
Wayne	Chicago	Boston	Dallas	Atlanta	Eugene
Xavier	Boston	Chicago	Dallas	Eugene	Atlanta
Yolanda	Atlanta	Eugene	Dallas	Chicago	Boston
Zeus	Dallas	Boston	Eugene	Chicago	Atlanta

Boston proposes to Xavier

Xavier accepts
(since previously unmatched)

Gale–Shapley demo

hospitals' preference lists

	1st	2nd	3rd	4th	5th
Atlanta	Wayne	Val	Yolanda	Zeus	Xavier
Boston	Yolanda	Wayne	Val	Xavier	Zeus
Chicago	Wayne	Zeus	Xavier	Yolanda	Val
Dallas	Val	Yolanda	Xavier	Wayne	Zeus
Eugene	Wayne	Yolanda	Val	Zeus	Xavier

students' preference lists

	1st	2nd	3rd	4th	5th
Val	Eugene	Atlanta	Boston	Dallas	Chicago
Wayne	Chicago	Boston	Dallas	Atlanta	Eugene
Xavier	Boston	Chicago	Dallas	Eugene	Atlanta
Yolanda	Atlanta	Eugene	Dallas	Chicago	Boston
Zeus	Dallas	Boston	Eugene	Chicago	Atlanta

Eugene proposes to Wayne

Gale–Shapley demo

hospitals' preference lists

	1st	2nd	3rd	4th	5th
Atlanta	Wayne	Val	Yolanda	Zeus	Xavier
Boston	Yolanda	Wayne	Val	Xavier	Zeus
Chicago	Wayne	Zeus	Xavier	Yolanda	Val
Dallas	Val	Yolanda	Xavier	Wayne	Zeus
Eugene	Wayne	Yolanda	Val	Zeus	Xavier

students' preference lists

	1st	2nd	3rd	4th	5th
Val	Eugene	Atlanta	Boston	Dallas	Chicago
Wayne	Chicago	Boston	Dallas	Atlanta	Eugene
Xavier	Boston	Chicago	Dallas	Eugene	Atlanta
Yolanda	Atlanta	Eugene	Dallas	Chicago	Boston
Zeus	Dallas	Boston	Eugene	Chicago	Atlanta

Eugene proposes to Wayne

Wayne rejects
(since he prefers Chicago)

Gale–Shapley demo

hospitals' preference lists

	1st	2nd	3rd	4th	5th
Atlanta	Wayne	Val	Yolanda	Zeus	Xavier
Boston	Yolanda	Wayne	Val	Xavier	Zeus
Chicago	Wayne	Zeus	Xavier	Yolanda	Val
Dallas	Val	Yolanda	Xavier	Wayne	Zeus
Eugene	Wayne	Yolanda	Val	Zeus	Xavier

students' preference lists

	1st	2nd	3rd	4th	5th
Val	Eugene	Atlanta	Boston	Dallas	Chicago
Wayne	Chicago	Boston	Dallas	Atlanta	Eugene
Xavier	Boston	Chicago	Dallas	Eugene	Atlanta
Yolanda	Atlanta	Eugene	Dallas	Chicago	Boston
Zeus	Dallas	Boston	Eugene	Chicago	Atlanta

Eugene proposes to Yolanda

Gale–Shapley demo

hospitals' preference lists

	1st	2nd	3rd	4th	5th
Atlanta	Wayne	Val	Yolanda	Zeus	Xavier
Boston	Yolanda	Wayne	Val	Xavier	Zeus
Chicago	Wayne	Zeus	Xavier	Yolanda	Val
Dallas	Val	Yolanda	Xavier	Wayne	Zeus
Eugene	Wayne	Yolanda	Val	Zeus	Xavier

students' preference lists

	1st	2nd	3rd	4th	5th
Val	Eugene	Atlanta	Boston	Dallas	Chicago
Wayne	Chicago	Boston	Dallas	Atlanta	Eugene
Xavier	Boston	Chicago	Dallas	Eugene	Atlanta
Yolanda	Atlanta	Eugene	Dallas	Chicago	Boston
Zeus	Dallas	Boston	Eugene	Chicago	Atlanta

Eugene proposes to Yolanda

Yolanda accepts (and renounces Dallas)

Gale–Shapley demo

hospitals' preference lists

	1st	2nd	3rd	4th	5th
Atlanta	Wayne	Val	Yolanda	Zeus	Xavier
Boston	Yolanda	Wayne	Val	Xavier	Zeus
Chicago	Wayne	Zeus	Xavier	Yolanda	Val
Dallas	Val	Yolanda	Xavier	Wayne	Zeus
Eugene	Wayne	Yolanda	Val	Zeus	Xavier

students' preference lists

	1st	2nd	3rd	4th	5th
Val	Eugene	Atlanta	Boston	Dallas	Chicago
Wayne	Chicago	Boston	Dallas	Atlanta	Eugene
Xavier	Boston	Chicago	Dallas	Eugene	Atlanta
Yolanda	Atlanta	Eugene	Dallas	Chicago	Boston
Zeus	Dallas	Boston	Eugene	Chicago	Atlanta

Dallas proposes to Xavier

Gale–Shapley demo

hospitals' preference lists

	1st	2nd	3rd	4th	5th
Atlanta	Wayne	Val	Yolanda	Zeus	Xavier
Boston	Yolanda	Wayne	Val	Xavier	Zeus
Chicago	Wayne	Zeus	Xavier	Yolanda	Val
Dallas	Val	Yolanda	Xavier	Wayne	Zeus
Eugene	Wayne	Yolanda	Val	Zeus	Xavier

students' preference lists

	1st	2nd	3rd	4th	5th
Val	Eugene	Atlanta	Boston	Dallas	Chicago
Wayne	Chicago	Boston	Dallas	Atlanta	Eugene
Xavier	Boston	Chicago	Dallas	Eugene	Atlanta
Yolanda	Atlanta	Eugene	Dallas	Chicago	Boston
Zeus	Dallas	Boston	Eugene	Chicago	Atlanta

Dallas proposes to Xavier

Xavier rejects
(since he prefers Boston)

Gale–Shapley demo

hospitals' preference lists

	1st	2nd	3rd	4th	5th
Atlanta	Wayne	Val	Yolanda	Zeus	Xavier
Boston	Yolanda	Wayne	Val	Xavier	Zeus
Chicago	Wayne	Zeus	Xavier	Yolanda	Val
Dallas	Val	Yolanda	Xavier	Wayne	Zeus
Eugene	Wayne	Yolanda	Val	Zeus	Xavier

students' preference lists

	1st	2nd	3rd	4th	5th
Val	Eugene	Atlanta	Boston	Dallas	Chicago
Wayne	Chicago	Boston	Dallas	Atlanta	Eugene
Xavier	Boston	Chicago	Dallas	Eugene	Atlanta
Yolanda	Atlanta	Eugene	Dallas	Chicago	Boston
Zeus	Dallas	Boston	Eugene	Chicago	Atlanta

Dallas proposes to Wayne

Gale–Shapley demo

hospitals' preference lists

	1st	2nd	3rd	4th	5th
Atlanta	Wayne	Val	Yolanda	Zeus	Xavier
Boston	Yolanda	Wayne	Val	Xavier	Zeus
Chicago	Wayne	Zeus	Xavier	Yolanda	Val
Dallas	Val	Yolanda	Xavier	Wayne	Zeus
Eugene	Wayne	Yolanda	Val	Zeus	Xavier

students' preference lists

	1st	2nd	3rd	4th	5th
Val	Eugene	Atlanta	Boston	Dallas	Chicago
Wayne	Chicago	Boston	Dallas	Atlanta	Eugene
Xavier	Boston	Chicago	Dallas	Eugene	Atlanta
Yolanda	Atlanta	Eugene	Dallas	Chicago	Boston
Zeus	Dallas	Boston	Eugene	Chicago	Atlanta

Dallas proposes to Wayne

Wayne rejects
(since he prefers Chicago)

Gale–Shapley demo

hospitals' preference lists

	1st	2nd	3rd	4th	5th
Atlanta	Wayne	Val	Yolanda	Zeus	Xavier
Boston	Yolanda	Wayne	Val	Xavier	Zeus
Chicago	Wayne	Zeus	Xavier	Yolanda	Val
Dallas	Val	Yolanda	Xavier	Wayne	Zeus
Eugene	Wayne	Yolanda	Val	Zeus	Xavier

students' preference lists

	1st	2nd	3rd	4th	5th
Val	Eugene	Atlanta	Boston	Dallas	Chicago
Wayne	Chicago	Boston	Dallas	Atlanta	Eugene
Xavier	Boston	Chicago	Dallas	Eugene	Atlanta
Yolanda	Atlanta	Eugene	Dallas	Chicago	Boston
Zeus	Dallas	Boston	Eugene	Chicago	Atlanta

Dallas proposes to Zeus

Gale–Shapley demo

hospitals' preference lists

	1st	2nd	3rd	4th	5th
Atlanta	Wayne	Val	Yolanda	Zeus	Xavier
Boston	Yolanda	Wayne	Val	Xavier	Zeus
Chicago	Wayne	Zeus	Xavier	Yolanda	Val
Dallas	Val	Yolanda	Xavier	Wayne	Zeus
Eugene	Wayne	Yolanda	Val	Zeus	Xavier

students' preference lists

	1st	2nd	3rd	4th	5th
Val	Eugene	Atlanta	Boston	Dallas	Chicago
Wayne	Chicago	Boston	Dallas	Atlanta	Eugene
Xavier	Boston	Chicago	Dallas	Eugene	Atlanta
Yolanda	Atlanta	Eugene	Dallas	Chicago	Boston
Zeus	Dallas	Boston	Eugene	Chicago	Atlanta

Dallas proposes to Zeus

Zeus accepts
(since previously unmatched)

Gale–Shapley demo

hospitals' preference lists

	1st	2nd	3rd	4th	5th
Atlanta	Wayne	Val	Yolanda	Zeus	Xavier
Boston	Yolanda	Wayne	Val	Xavier	Zeus
Chicago	Wayne	Zeus	Xavier	Yolanda	Val
Dallas	Val	Yolanda	Xavier	Wayne	Zeus
Eugene	Wayne	Yolanda	Val	Zeus	Xavier

students' preference lists

	1st	2nd	3rd	4th	5th
Val	Eugene	Atlanta	Boston	Dallas	Chicago
Wayne	Chicago	Boston	Dallas	Atlanta	Eugene
Xavier	Boston	Chicago	Dallas	Eugene	Atlanta
Yolanda	Atlanta	Eugene	Dallas	Chicago	Boston
Zeus	Dallas	Boston	Eugene	Chicago	Atlanta

STOP
(stable matching)

The Gale-Shapley proposal algorithm

While there is a man m who is not matched:

- Let w be the highest ranked woman in m's list to whom m has not proposed yet.
- If w is unmatched, or w prefers m over her current match:
  - Match m and w.
    (The previous match of w is now unmatched.)

Cool, but does it work correctly?

- Does it always terminate?
- Does it always find a stable matching?

(Does a stable matching always exist?)

Gale-Shapley algorithm analysis

Theorem:

The *Gale-Shapley proposal algorithm* always terminates with a stable matching after at most n^2 iterations.

A *constructive* proof that a stable matching always exists.

3 things to show:

1. Number of iterations is at most n^2.
2. The algorithm terminates with a perfect matching.
3. The matching has no unstable pairs.

Gale-Shapley algorithm analysis

1. Number of iterations is at most n^2.

iterations = # proposals

No man proposes to a woman more than once.

So each man makes at most n proposals.

There are n men in total.

$$\implies \quad \text{\# proposals} \quad \leq n^2.$$

$$\implies \quad \text{\# iterations} \quad \leq n^2.$$

Gale-Shapley algorithm analysis

2. The algorithm terminates with a perfect matching.

If we don't have a perfect matching:

A man is not matched

$\Longrightarrow$ All women must be matched

$\Longrightarrow$ All men must be matched.

Contradiction

<u>Second implication:</u>

There are an equal number of men and women.

Gale-Shapley algorithm analysis

2. The algorithm terminates with a perfect matching.

If we don't have a perfect matching:

A man is not matched

$\Longrightarrow$ All women must be matched

$\Longrightarrow$ All men must be matched.

Contradiction

<u>First implication:</u>

Observe: once a woman is matched, she stays matched.

A man got rejected by every woman:

case1: she was already matched, or

case2: she got a better offer

Either way, she was matched at some point.

Gale-Shapley algorithm analysis

3. The matching has <u>no</u> unstable pairs.

"<u>Improvement" Lemma</u>:
(i) A man can only go down in his preference list.
(ii) A woman can only go up in her preference list.

Unstable pair:
(m,w) <u>unmatched</u>
but they prefer each other.

Consider any <u>unmatched</u> (m,w). **WTS**: it cannot be unstable.

Case 1: m never proposed to w

by (i), m prefers w' over w

Case 2: m proposed to w

w rejected m $\implies$ by (ii), w prefers m' over m

$\square$

Further questions

Theorem:

The *Gale-Shapley proposal algorithm* always terminates with a stable matching after at most n^2 iterations.

Does the order of how we pick men matter?

Would it lead to different matchings?

Is the algorithm "fair"?

Does this algorithm favor men or women or neither?

Further questions

m and w are *valid partners* if there is a stable matching in which they are matched.

best(m) = highest ranked valid partner of m

Theorem:
Gale-Shapley algorithm returns $\{(m, \text{best}(m)) : m \in X\}$.

Not at all obvious this would be a matching,
let alone a stable matching!

Further questions

worst(w) = lowest ranked valid partner of w

Theorem:

Gale-Shapley algorithm returns $\{(\text{worst}(w), w) : w \in Y\}$.

Real-world applications

Variants of the Gale-Shapley algorithm
is used for:

- matching medical students and hospitals
- matching students to high schools (e.g. in New York)
- matching students to universities (e.g. in Hungary)
- matching users to servers

⋮

The Gale-Shapley Proposal Algorithm (1962)

Nobel Prize in Economics 2012

"for the theory of stable allocations and the practice of market design."

BBM 205 Discrete Mathematics
Hacettepe University
http://web.cs.hacettepe.edu.tr/∼bbm205

Lecture 2: Proof Techniques

Lecturer: Lale Özkahya

Resources:
Kenneth Rosen, "Discrete Mathematics and App."
cs.colostate.edu/ cs122/.Spring15/home_resources.php
http://www.cs.nthu.edu.tw/ wkhon/math16.html

Proof Terminology

Theorem: statement that can be shown to be true

Proof: a valid argument that establishes the truth of a theorem

Axioms: statements we assume to be true

Lemma: a less important theorem that is helpful in the proof of other results

Corollary: theorem that can be established directly from a theorem that has been proved

Conjecture: statement that is being *proposed* to be a true statement

Learning objectives

- Direct proofs
- Proof by contrapositive
- Proof by contradiction
- Proof by cases

Technique #1: Direct Proof

- Direct Proof:
  - First step is a premise
  - Subsequent steps use rules of inference or other premises
  - Last step proves the conclusion

Methods of Proving

- A direct proof of a conditional statement

$$p \rightarrow q$$

  first assumes that p is true, and uses axioms, definitions, previously proved theorems, with rules of inference, to show that q is also true

- The above targets to show that the case where p is true and q is false never occurs
  - Thus, $p \rightarrow q$ is always true

Direct Proof (Example 1)

- Show that

  if n is an odd integer, then n^2 is odd.

- Proof :

  Assume that n is an odd integer. This implies that there is some integer k such that

$$n = 2k + 1.$$

Then $n^2 = (2k+1)^2 = 4k^2 + 4k + 1 = 2(2k^2 + 2k) + 1.$

Thus, n^2 is odd.

Direct Proof (Example 2)

- Show that

  if m and n are both square numbers,

  then m n is also a square number.

- Proof :

  Assume that m and n are both squares. This implies that there are integers u and v such that

$$m = u^2 \quad \text{and} \quad n = v^2.$$

  Then $m\,n = u^2 v^2 = (uv)^2$. Thus, m n is a square.

Class Exercise

- Prove: If n is an even integer, then n^2 is even.
  - If $\boldsymbol{n}$ is even, then $\boldsymbol{n = 2k}$ for some integer $\boldsymbol{k}$.
  - $\boldsymbol{n^2 = (2k)^2 = 4k^2}$
  - Therefore, $\boldsymbol{n = 2(2k^2)}$, which is even.

Can you do the formal version?

	Step	Reason
1.	n is even	Premise
2.	$\exists k \in Z \;\; n = 2k$	Def of even integer in (1)
3.	$n^2 = (2k)^2$	Squaring (2)
4.	$= 4k^2$	Algebra on (3)
5.	$= 2(2k^2)$	Algebra on (4)
6.	$\therefore\ n^2$ is even	Def even int, from (5)

Technique #2: Proof by Contrapositive

- A direct proof, but starting with the contrapositive equivalence:
  - $p \rightarrow q \equiv \neg q \rightarrow \neg p$
- If you are asked to prove $p \rightarrow q$
- you instead prove $\neg q \rightarrow \neg p$
- Why? Sometimes, it may be easier to directly prove $\neg q \rightarrow \neg p$ than $p \rightarrow q$

Methods of Proving

- The proof by contraposition method makes use of the equivalence

$$p \rightarrow q \quad \equiv \quad \neg q \rightarrow \neg p$$

- To show that the conditional statement $p \rightarrow q$ is true, we first assume $\neg q$ is true, and use axioms, definitions, proved theorems, with rules of inference, to show $\neg p$ is also true

Proof by Contraposition (Example 1)

- Show that

  if $3n + 2$ is an odd integer, then n is odd.

- Proof :

  Assume that n is even. This implies that

  $$n = 2k \text{ for some integer } k.$$

  Then, $3n + 2 = 3(2k) + 2 = 6k + 2 = 2(3k + 1)$, so that $3n + 2$ is even. Since the negation of conclusion implies the negation of hypothesis, the original conditional statement is true

Proof by Contraposition (Example 2)

- Show that

  if $n = a\,b$, where a and b are positive,

  then $a \le \sqrt{n}$ or $b \le \sqrt{n}$.

- Proof :

  Assume that both a and b are larger than $\sqrt{n}$. Thus, $a\,b > n$ so that $n \ne a\,b$. Since the negation of conclusion implies the negation of hypothesis, the original conditional statement is true

Proof by contrapositive

Prove: If n^2 is an even integer, then n is even.

$(n^2 \text{ even}) \rightarrow (n \text{ even})$

By the contrapositive: This is the same as showing that

- $\neg(n \text{ even}) \rightarrow \neg(n^2 \text{ even})$
- If $\boldsymbol{n}$ is odd, then n^2 is odd.
- We already proved this on slides 4 and 5.

Since we have proved the contrapositive:

$\neg(n \text{ even}) \rightarrow \neg(n^2 \text{ even})$

We have also proved the original hypothesis:

$(n^2 \text{ even}) \rightarrow (n \text{ even})$

Technique #3: Proof by contradiction

Prove: If p then q.

Proof strategy:

- Assume the negation of q.
- In other words, assume that $p \wedge \neg q$ is true.
- Then arrive at a contradiction $p \wedge \neg p$ (or something that contradicts a known fact).
- Since this cannot happen, our assumption must be wrong.
- Thus, $\neg q$ is false. q is true.

Proof by contradiction example

Prove: If *(3n+2) is odd,* then *n is odd*.

Proof:

- Given: (3n+2) is odd.
- Assume that n is not odd, that is n is even.
- If n is even, there is some integer k such that n=2k.
- (3n+2) = (3(2k)+2)=6k+2 = 2(3k+1), which is 2 times a number.
- Thus 3n+2 turned out to be even, but we know it's odd.
- This is a contradiction. Our assumption was wrong.
- Thus, n must be odd.

Proof by Contradiction Example

Prove that the $\sqrt{2}$ is irrational.

Assume that "$\sqrt{2}$ is irrational" is false, that is, $\sqrt{2}$ is rational.

Hence, $\sqrt{2} = \frac{a}{b}$ and a and b have no common factors. The fraction is in its lowest terms.

So $a^2 = 2b^2$, which means a must be even,

Hence, $a = 2c$

Therefore, $b^2 = 2c^2$ then b must be even, which means a and b must have common factors.

Contradiction.

Technique #4: Proof by cases

- Given a problem of the form:
  - $(p_1 \vee p_2 \vee \ \ldots \vee p_n) \rightarrow q$
  - where $p_1, p_2, \ldots p_n$ are the cases

- This is equivalent to the following:
  - $[(p_1 \rightarrow q) \wedge (p_2 \rightarrow q) \wedge \ldots \wedge (p_n \rightarrow q)]$
- So prove all the clauses are true.

Proof by cases (example)

- Prove: If n is an integer, then $n^2 \geq n$
  - $(n = 0 \vee n \geq 1 \vee n \leq -1) \rightarrow n^2 \geq n$
- Show for all the three cases, i.e.,
  - $(n = 0 \rightarrow n^2 \geq n) \wedge (n \geq 1 \rightarrow n^2 \geq n)$
    $\wedge$ $(n \leq -1 \rightarrow n^2 \geq n)$
- Case 1: Show that $n = 0 \rightarrow n^2 \geq n$
  - When $n=0$, $n^2 = 0$.
  - $0=0$ ☺

Proof by cases (example contd)

- Case 2: Show that $n \geq 1 \rightarrow n^2 \geq n$
  - Multiply both sides of the inequality $n \geq 1$ by n
  - We get $n^2 \geq n$

Proof by cases (example contd)

- Case 3: Show that $n \leq -1 \rightarrow n^2 \geq n$
  - Given $n \leq -1$,
  - We know that n^2 cannot be negative, i.e., $n^2 > 0$
  - We know that $0 > -1$
  - Thus, $n^2 > -1$. We also know that $-1 \geq n$ (given)
  - Therefore, $n^2 \geq n$

Proof by Cases Example

Theorem: Given two real numbers x and y, $abs(x*y)=abs(x)*abs(y)$

Exhaustively determine the premises

Case p1: x>=0, y>=0, so x*y>=0 so abs(x*y)=x*y and abs(x)=x and abs(y)=y so abs(x)*abs(y)=x*y

Case p2: x<0, y>=0

Case p3: x>=0, y<0

Case p4: x<0, y<0

Methods of Proving

- When proving bi-conditional statement, we may make use of the equivalence

$$p \leftrightarrow q \equiv (p \rightarrow q) \wedge (q \rightarrow p)$$

- In general, when proving several propositions are equivalent, we can use the equivalence

$$p_1 \leftrightarrow p_2 \leftrightarrow \ldots \leftrightarrow p_k$$
$$\equiv (p_1 \rightarrow p_2) \wedge (p_2 \rightarrow p_3) \wedge \ldots \wedge (p_k \rightarrow p_1)$$

Proofs of Equivalence (Example)

- Show that the following statements about the integer n are equivalent :

  p := "n is even"

  q := "n – 1 is odd"

  r := "n^2 is even"

- To do so, we can show the three propositions

$$p \rightarrow q, \quad q \rightarrow r, \quad r \rightarrow p$$

are all true. Can you do so ?

Methods of Proving

- A proof of the proposition of the form $\exists x$ P(x) is called an existence proof
- Sometimes, we can find an element s, called a witness, such that P(s) is true

  This type of existence proof is constructive
- Sometimes, we may have non-constructive existence proof, where we do not find the witness

Existence Proof (Examples)

- Show that there is a positive integer that can be written as the sum of cubes of positive integers in two different ways.
- Proof: $1729 = 1^3 + 12^3 = 9^3 + 10^3$

- Show that there are irrational numbers r and s such that r^s is rational.
- Hint: Consider $(\sqrt{2}^{\sqrt{2}})^{\sqrt{2}}$

Common Mistakes in Proofs

- Show that $1 = 2$.
- Proof : Let a be a positive integer, and $b = a$.

Step	**Reason**
1. $a = b$	Given
2. $a^2 = a\,b$	Multiply by a in (1)
3. $a^2 - b^2 = a\,b - b^2$	Subtract by b^2 in (2)
4. $(a - b)(a + b) = b(a - b)$	Factor in (3)
5. $a + b = b$	Divide by $(a - b)$ in (4)
6. $2b = b$	By (1) and (5)
7. $2 = 1$	Divide by b in (6)

Common Mistakes in Proofs

- Show that

  if n^2 is an even integer, then n is even.

- Proof :

  Suppose that n^2 is even.

  Then $n^2 = 2k$ for some integer k.

  Let $n = 2m$ for some integer m.

  Thus, n is even.

Common Mistakes in Proofs

- Show that

  if x is real number, then x^2 is positive.

- Proof : There are two cases.

  Case 1: x is positive

  Case 2: x is negative

In Case 1, x^2 is positive.

In Case 2, x^2 is also positive

Thus, we obtain the same conclusion in all cases, so that the original statement is true.

Proof Strategies

- Adapting Existing Proof
- Show that

  $\sqrt{3}$ is irrational.

- Instead of searching for a proof from nowhere, we may recall some similar theorem, and see if we can slightly modify (adapt) its proof to obtain what we want

Proof Strategies

- Sometimes, it may be difficult to prove a statement q directly
- Instead, we may find a statement p with the property that $p \rightarrow q$, and then prove p

  Note: If this can be done, by Modus Ponens, q is true
- This strategy is called backward reasoning

Backward Reasoning (Example)

- Show that for distinct positive real numbers x and y,

$$0.5\,(x + y) > (x\,y)^{0.5}$$

- Proof: By backward reasoning strategy, we find that

1. $0.25\,(x + y)^2 > x\,y \quad \rightarrow \quad 0.5\,(x + y) > (x\,y)^{0.5}$
2. $(x + y)^2 > 4\,x\,y \quad \rightarrow \quad 0.25\,(x + y)^2 > x\,y$
3. $x^2 + 2\,x\,y + y^2 > 4\,x\,y \quad \rightarrow \quad (x + y)^2 > 4\,x\,y$
4. $x^2 - 2\,x\,y + y^2 > 0 \quad \rightarrow \quad x^2 + 2\,x\,y + y^2 > 4\,x\,y$
5. $(x - y)^2 > 0 \quad \rightarrow \quad x^2 - 2\,x\,y + y^2 > 0$
6. $(x - y)^2 > 0$ is true, since x and y are distinct.

Thus, the original statement is true.

Interesting Examples

Can a checkerboard be tiled by 1×2 dominoes?

Interesting Examples

What if the top left corner is removed ?

Interesting Examples

What if the lower right corner is also removed ?

BBM 205 Discrete Mathematics
Hacettepe University
http://web.cs.hacettepe.edu.tr/∼bbm205

Lecture 3: Method of Induction

Lecturer: Lale Özkahya

Resources:
http://www.eecs70.org
cs.colostate.edu/ cs122/.Spring15/home_resources.php
http://www.cs.nthu.edu.tw/ wkhon/math16.html

Introduction

- What is the formula of the sum of the first n positive odd integers?

$$1 = 1$$

$$1 + 3 = 4$$

$$1 + 3 + 5 = 9$$

$$1 + 3 + 5 + 7 = 16$$

$$1 + 3 + 5 + 7 + 9 = 25$$

Introduction

- It is reasonable to guess that the sum is n^2
- To do so, we may use a method, called mathematical induction, to prove that the guess is correct
- How to do that ?

Mathematical Induction

- Firstly, to simplify the discussion, we define a propositional function P(n), where

P(n) := "The sum of first n positive odd integers is n^2"

so that our target is to show $\forall n\ P(n)$ is true

Mathematical Induction

- Next, we are going to show the following two statements to be true:

  1. $P(1)$, called basic step
  2. $\forall n\ (P(n) \rightarrow P(n+1))$, called inductive step, where domain of n is all positive integers

- If both can be shown true, then we can conclude that $\forall n\ P(n)$ is true [why?]

Back to the Example

- We let

  P(n) := "The sum of first n positive odd integers is n^2"

  and we hope to use mathematical induction to show $\forall n\ P(n)$ is true

- Can we show the basic step to be true?
- Can we show the inductive step to be true?

Back to the Example

- Can we show the basic step to be true?

- The basic step is P(1), which is :

  P(1) := "The sum of first 1 positive odd integers is 1^2"

  This is obviously true.

Back to the Example

- Can we show the inductive step to be true?
- The inductive step is $\forall n\ (P(n) \rightarrow P(n+1))$
- To show it is true, we focus on an arbitrary chosen k, and see if $P(k) \rightarrow P(k+1)$ is true
  - If so, by universal generalization,

$$\forall n\ (P(n) \rightarrow P(n+1)) \text{ is true}$$

Back to the Example

- Suppose that P(k) is true. That is,

  P(k) := "The sum of first k positive odd integers is k^2"

  This implies $1 + 3 + \ldots + (2k - 1) = k^2$.

  Then, we have

$$1 + 3 + \ldots + (2k - 1) + (2k + 1) = k^2 + (2k + 1)$$
$$= (k + 1)^2,$$

  so that P(k+1) is true if P(k) is true.

Remark

- Note : When we show that the inductive step is true, we do not show P(k+1) is true.

  Instead, we show the conditional statement

  $$P(k) \rightarrow P(k+1) \text{ is true.}$$

  This allows us to use P(k) as the premise, and gives us an easier way to show P(k+1)

- Once basic step and inductive step are proven, by mathematical induction, $\forall n\ P(n)$ is true

Remark

- Mathematical induction is a very powerful technique, because we show just two statements, but this can imply infinite number of cases to be correct
- However, the technique does not help us find new theorems. In fact, we have to obtain the theorem (by guessing) in the first place, and induction is then used to formally confirm the theorem is correct

More Examples

- Ex 1: Show that for all positive integer n,

$$n < 2^n$$

- Ex 2: Show that for all positive integer n,

$$n^3 - n \text{ is divisible by } 3$$

- Ex 3: Show that for all positive integer n,

$$1^2 + 2^2 + 3^2 + \ldots + n^2 = n(n+1)(2n+1) / 6$$

Using a Different Basic Step

- When we apply the induction technique, it is not necessary to have P(1) as the basic step

- We may replace the basic step by P(k) for some fixed k. If both basic step and inductive step are true, this will imply that

$$\forall n \geq k\ (P(n))$$

More Examples

- Ex 4: Show that for all positive integer $n \geq 4$,

$$2^n < n!$$

- Ex 5: Show that for all non-negative integer n,

$$1 + 2 + \ldots + 2^n = 2^{n+1} - 1$$

- Ex 6: Show that for non-negative integer n,

$7^{n+2} + 8^{2n+1}$ is divisible by 57

Interesting Examples

Snowball Fight

- There are 2n + 1 people
- Each must throw to the nearest
- All with distinct distance apart
- Show that at least one is not hit by any snowball

Interesting Examples

Tiling (Again!)

- A big square of size $2^n \times 2^n$
- Somewhere inside, a 1×1 small square is removed
- Show that the remaining board can always be tiled by L-shaped dominoes :

each consists of three 1×1 squares

Strong Induction

- An alternative form of induction, called strong induction, uses a different inductive step:

$$\forall n\ (\ (P(1) \wedge P(2) \wedge \ldots \wedge P(n)) \rightarrow P(n+1)\ )$$

- The basic step is still to prove P(1) to be true
- Again, if both the basic and inductive steps are true, then we can conclude that $\forall n\ P(n)$ is true [how?]

Examples

- Ex 1:

Define the nth Fibonacci number, F_n , as follows:

$$F_0 = 1, \quad F_1 = 1,$$

$$F_n = F_{n-1} + F_{n-2}, \quad \text{when } n \geq 2$$

By the above recursive definition, we get the first few Fibonacci numbers :

1, 1, 2, 3, 5, 8, 13, 21, 34, 55, …

Examples

- Ex 1 (continued):

  Show that F_n can be computed by the formula

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

Examples

- Ex 2: Quicksort is a recursive algorithm for sorting a collection of distinct numbers :

  1. If there is at most 1 number to sort, done
  2. Else, pick any number x from the collection, and use x to divide the remaining numbers into two groups:

     those smaller than x, those larger than x.

     Next, apply Quicksort to sort each group

     (putting x in-between afterwards)

Examples

- For instance, suppose the input collection of numbers contains 1, 4, 3, 10, 7, 2
- First round, say we pick x = 3
- Then we will form two groups S and L:

$$S = \{ 1, 2 \} \text{ and } L = \{ 4, 10, 7 \}$$

- After that, we apply Quicksort on each group, and in the end, we report

$$\text{Quicksort}(S),\ x,\ \text{Quicksort}(L)$$

Examples

- Ex 2 (continued):

  Show that Quicksort can correctly sort any collection of n distinct numbers

Interesting Example

Peg Solitaire

- There are pegs on a board
- A peg can jump over another one into an adjacent empty square, so that the jumped-over peg is eliminated
- Target: Can we eliminate all but one peg ?

Interesting Example

- Show that if we start with $n \times n$ pegs (arranged as a square) on a board with infinite size, and n is not divisible by 3, then we can eliminate all but one peg

- Hint: Let P(n) denote the above proposition. Show that P(1) and P(2) are true, and for all n, $P(3n+1) \rightarrow P(3n+5)$, $P(3n+2) \rightarrow P(3n+4)$ are true

Common Mistakes

- Show that

  P(n) = “any n cats will have the same color”

  is true for all positive integer n.

- Proof: The basic step P(1) is obviously true.

  Next, assume P(k) is true. Then, when we have k + 1 cats, we can remove one of them, say y, so that by P(k), they will have the same color

Common Mistakes

- Proof (continued):

  Now, we exchange the removed cat with one of the other k cats :

Common Mistakes

- Proof (continued):

  Then, by P(k) again, y must have the same color as the other k – 1 cats.

  This implies all the cats are of the same color!

- What's wrong with the proof ?

$\forall k \in . \ \sum_{i=1}^{k} i = \frac{k(k+1)}{2}$

$\forall k \in . \ \sum_{i=1}^{k} i = \frac{k(k+1)}{2}$

(By induction) Let $P(k)$ be the predicate "$\sum_{i=1}^{k} i = \frac{k(k+1)}{2}$"

Base Case: $\sum_{i=1}^{0} i = 0 = \frac{0(0+1)}{2}$, thus $P(0)$ is true

Inductive Hypothesis: Let $k \geq 0$. We assume that $P(k)$ is true, *i.e.* $\sum_{i=1}^{k} i = \frac{k(k+1)}{2}$

Inductive Step:

$$\begin{aligned} \sum_{i=1}^{k+1} i &= \left[\sum_{i=1}^{k} i\right] + (k+1) \\ &= \frac{k(k+1)}{2} + (k+1) \qquad \text{(by I.H.)} \\ &= \frac{k(k+1)+2(k+1)}{2} \\ &= \frac{(k+1)(k+2)}{2} \end{aligned}$$

Thus $P(k+1)$ is true $\square$

$\forall k \geq 4.\ 2^k < k!$

$\forall k \geq 4.\ 2^k < k!$

(By induction) Let $P(k)$ be the predicate "$2^k < k!$"

Base Case: $2^4 = 16 < 24 = 4!$, thus $P(4)$ is true

Inductive Hypothesis: Let $k \geq 4$. We assume that $P(k)$ is true, *i.e.* $2^k < k!$

Inductive Step:

$$\begin{aligned} 2^{k+1} &= 2 \cdot 2^k \\ &< 2 \cdot k! && \text{(by I.H.)} \\ &< (k+1) \cdot k! && (k \geq 4) \\ &= (k+1)! \end{aligned}$$

Thus $P(k+1)$ is true $\square$

Another Induction Proof.

Theorem: For every $n \in N$, $n^3 - n$ is divisible by 3. ($3|(n^3 - n)$).

Another Induction Proof.

Theorem: For every $n \in N$, $n^3 - n$ is divisible by 3. ($3|(n^3 - n)$).

Proof:

Another Induction Proof.

Theorem: For every $n \in N$, $n^3 - n$ is divisible by 3. ($3|(n^3 - n)$).

Proof: By induction.

Another Induction Proof.

Theorem: For every $n \in N$, $n^3 - n$ is divisible by 3. $(3|(n^3 - n)$).

Proof: By induction.
Base Case: $P(0)$ is "$(0^3) - 0$" is divisible by 3.

Another Induction Proof.

Theorem: For every $n \in N$, $n^3 - n$ is divisible by 3. $(3|(n^3 - n)\ )$.

Proof: By induction.
Base Case: $P(0)$ is "$(0^3) - 0$" is divisible by 3. Yes!

Another Induction Proof.

Theorem: For every $n \in N$, $n^3 - n$ is divisible by 3. $(3|(n^3 - n)$).

Proof: By induction.
Base Case: $P(0)$ is "$(0^3) - 0$" is divisible by 3. Yes!
Induction Step: $(\forall k \in N), P(k) \implies P(k+1)$

Another Induction Proof.

Theorem: For every $n \in N$, $n^3 - n$ is divisible by 3. $(3|(n^3 - n)$).

Proof: By induction.
Base Case: $P(0)$ is "$(0^3) - 0$" is divisible by 3. Yes!
Induction Step: $(\forall k \in N), P(k) \implies P(k+1)$
Induction Hypothesis: $k^3 - k$ is divisible by 3.

Another Induction Proof.

Theorem: For every $n \in N$, $n^3 - n$ is divisible by 3. $(3|(n^3 - n)$).

Proof: By induction.
Base Case: $P(0)$ is "$(0^3) - 0$" is divisible by 3. Yes!
Induction Step: $(\forall k \in N), P(k) \implies P(k+1)$
Induction Hypothesis: $k^3 - k$ is divisible by 3.
or $k^3 - k = 3q$ for some integer q.

Another Induction Proof.

Theorem: For every $n \in N$, $n^3 - n$ is divisible by 3. $(3|(n^3 - n)$).

Proof: By induction.
Base Case: $P(0)$ is "$(0^3) - 0$" is divisible by 3. Yes!
Induction Step: $(\forall k \in N), P(k) \implies P(k+1)$
Induction Hypothesis: $k^3 - k$ is divisible by 3.
or $k^3 - k = 3q$ for some integer q.

Another Induction Proof.

Theorem: For every $n \in N$, $n^3 - n$ is divisible by 3. ($3|(n^3 - n)$).

Proof: By induction.
Base Case: $P(0)$ is "$(0^3) - 0$" is divisible by 3. Yes!
Induction Step: $(\forall k \in N), P(k) \implies P(k+1)$
Induction Hypothesis: $k^3 - k$ is divisible by 3.
or $k^3 - k = 3q$ for some integer q.

$(k+1)^3 - (k+1)$

Another Induction Proof.

Theorem: For every $n \in N$, $n^3 - n$ is divisible by 3. ($3|(n^3 - n)$).

Proof: By induction.
Base Case: $P(0)$ is "$(0^3) - 0$" is divisible by 3. Yes!
Induction Step: $(\forall k \in N), P(k) \implies P(k+1)$
Induction Hypothesis: $k^3 - k$ is divisible by 3.
or $k^3 - k = 3q$ for some integer q.

$(k+1)^3 - (k+1) = k^3 + 3k^2 + 3k + 1 - (k+1)$

Another Induction Proof.

Theorem: For every $n \in N$, $n^3 - n$ is divisible by 3. ($3|(n^3 - n)$).

Proof: By induction.
Base Case: $P(0)$ is "$(0^3) - 0$" is divisible by 3. Yes!
Induction Step: $(\forall k \in N), P(k) \implies P(k+1)$
Induction Hypothesis: $k^3 - k$ is divisible by 3.
or $k^3 - k = 3q$ for some integer q.

$$\begin{aligned}(k+1)^3 - (k+1) &= k^3 + 3k^2 + 3k + 1 - (k+1) \\ &= k^3 + 3k^2 + 2k\end{aligned}$$

Another Induction Proof.

Theorem: For every $n \in N$, $n^3 - n$ is divisible by 3. $(3|(n^3 - n)$).

Proof: By induction.
Base Case: $P(0)$ is "$(0^3) - 0$" is divisible by 3. Yes!
Induction Step: $(\forall k \in N), P(k) \implies P(k+1)$
Induction Hypothesis: $k^3 - k$ is divisible by 3.
or $k^3 - k = 3q$ for some integer q.

$$\begin{aligned}(k+1)^3 - (k+1) &= k^3 + 3k^2 + 3k + 1 - (k+1) \\ &= k^3 + 3k^2 + 2k \\ &= (k^3 - k) + 3k^2 + 3k\end{aligned}$$

Another Induction Proof.

Theorem: For every $n \in N$, $n^3 - n$ is divisible by 3. ($3|(n^3 - n)$).

Proof: By induction.
Base Case: $P(0)$ is "$(0^3) - 0$" is divisible by 3. Yes!
Induction Step: $(\forall k \in N), P(k) \implies P(k+1)$
Induction Hypothesis: $k^3 - k$ is divisible by 3.
or $k^3 - k = 3q$ for some integer q.

$$\begin{aligned}(k+1)^3 - (k+1) &= k^3 + 3k^2 + 3k + 1 - (k+1) \\ &= k^3 + 3k^2 + 2k \\ &= (k^3 - k) + 3k^2 + 3k \quad \text{Subtract/add } k\end{aligned}$$

Another Induction Proof.

Theorem: For every $n \in N$, $n^3 - n$ is divisible by 3. ($3|(n^3 - n)$).

Proof: By induction.
Base Case: $P(0)$ is "$(0^3) - 0$" is divisible by 3. Yes!
Induction Step: $(\forall k \in N), P(k) \implies P(k+1)$
Induction Hypothesis: $k^3 - k$ is divisible by 3.
or $k^3 - k = 3q$ for some integer q.

$$\begin{aligned}(k+1)^3 - (k+1) &= k^3 + 3k^2 + 3k + 1 - (k+1) \\ &= k^3 + 3k^2 + 2k \\ &= (k^3 - k) + 3k^2 + 3k \quad \text{Subtract/add } k \\ &= 3q + 3(k^2 + k)\end{aligned}$$

Another Induction Proof.

Theorem: For every $n \in N$, $n^3 - n$ is divisible by 3. ($3|(n^3 - n)$).

Proof: By induction.
Base Case: $P(0)$ is "$(0^3) - 0$" is divisible by 3. Yes!
Induction Step: $(\forall k \in N), P(k) \implies P(k+1)$
Induction Hypothesis: $k^3 - k$ is divisible by 3.
or $k^3 - k = 3q$ for some integer q.

$$\begin{aligned}
(k+1)^3 - (k+1) &= k^3 + 3k^2 + 3k + 1 - (k+1) \\
&= k^3 + 3k^2 + 2k \\
&= (k^3 - k) + 3k^2 + 3k \quad \text{Subtract/add } k \\
&= 3q + 3(k^2 + k) \qquad \text{Induction Hyp.}
\end{aligned}$$

Another Induction Proof.

Theorem: For every $n \in N$, $n^3 - n$ is divisible by 3. ($3|(n^3 - n)$).

Proof: By induction.
Base Case: $P(0)$ is "$(0^3) - 0$" is divisible by 3. Yes!
Induction Step: $(\forall k \in N), P(k) \implies P(k+1)$
Induction Hypothesis: $k^3 - k$ is divisible by 3.
or $k^3 - k = 3q$ for some integer q.

$$\begin{aligned}
(k+1)^3 - (k+1) &= k^3 + 3k^2 + 3k + 1 - (k+1) \\
&= k^3 + 3k^2 + 2k \\
&= (k^3 - k) + 3k^2 + 3k \quad \text{Subtract/add } k \\
&= 3q + 3(k^2 + k) \qquad \text{Induction Hyp.} \quad \text{Factor.}
\end{aligned}$$

Another Induction Proof.

Theorem: For every $n \in N$, $n^3 - n$ is divisible by 3. ($3|(n^3 - n)$).

Proof: By induction.
Base Case: $P(0)$ is "$(0^3) - 0$" is divisible by 3. Yes!
Induction Step: $(\forall k \in N), P(k) \implies P(k+1)$
Induction Hypothesis: $k^3 - k$ is divisible by 3.
or $k^3 - k = 3q$ for some integer q.

$$\begin{aligned}
(k+1)^3 - (k+1) &= k^3 + 3k^2 + 3k + 1 - (k+1) \\
&= k^3 + 3k^2 + 2k \\
&= (k^3 - k) + 3k^2 + 3k \quad \text{Subtract/add } k \\
&= 3q + 3(k^2 + k) \qquad \text{Induction Hyp.} \quad \text{Factor.} \\
&= 3(q + k^2 + k)
\end{aligned}$$

Another Induction Proof.

Theorem: For every $n \in N$, $n^3 - n$ is divisible by 3. ($3|(n^3 - n)$).

Proof: By induction.
Base Case: $P(0)$ is "$(0^3) - 0$" is divisible by 3. Yes!
Induction Step: $(\forall k \in N), P(k) \implies P(k+1)$
Induction Hypothesis: $k^3 - k$ is divisible by 3.
or $k^3 - k = 3q$ for some integer q.

$$\begin{aligned}
(k+1)^3 - (k+1) &= k^3 + 3k^2 + 3k + 1 - (k+1) \\
&= k^3 + 3k^2 + 2k \\
&= (k^3 - k) + 3k^2 + 3k \quad \text{Subtract/add } k \\
&= 3q + 3(k^2 + k) \qquad \text{Induction Hyp.} \quad \text{Factor.} \\
&= 3(q + k^2 + k) \qquad \text{(Un)Distributive} + \text{ over } \times
\end{aligned}$$

Another Induction Proof.

Theorem: For every $n \in N$, $n^3 - n$ is divisible by 3. ($3|(n^3 - n)$).

Proof: By induction.
Base Case: $P(0)$ is "$(0^3) - 0$" is divisible by 3. Yes!
Induction Step: $(\forall k \in N), P(k) \implies P(k+1)$
Induction Hypothesis: $k^3 - k$ is divisible by 3.
or $k^3 - k = 3q$ for some integer q.

$$\begin{aligned}
(k+1)^3 - (k+1) &= k^3 + 3k^2 + 3k + 1 - (k+1) \\
&= k^3 + 3k^2 + 2k \\
&= (k^3 - k) + 3k^2 + 3k \quad \text{Subtract/add } k \\
&= 3q + 3(k^2 + k) \qquad \text{Induction Hyp.} \quad \text{Factor.} \\
&= 3(q + k^2 + k) \qquad \text{(Un)Distributive } + \text{ over } \times
\end{aligned}$$

Or $(k+1)^3 - (k+1) = 3(q + k^2 + k)$.

Another Induction Proof.

Theorem: For every $n \in N$, $n^3 - n$ is divisible by 3. ($3|(n^3 - n)$).

Proof: By induction.
Base Case: $P(0)$ is "$(0^3) - 0$" is divisible by 3. Yes!
Induction Step: $(\forall k \in N), P(k) \implies P(k+1)$
Induction Hypothesis: $k^3 - k$ is divisible by 3.
or $k^3 - k = 3q$ for some integer q.

$$\begin{aligned}
(k+1)^3 - (k+1) &= k^3 + 3k^2 + 3k + 1 - (k+1) \\
&= k^3 + 3k^2 + 2k \\
&= (k^3 - k) + 3k^2 + 3k \quad \text{Subtract/add } k \\
&= 3q + 3(k^2 + k) \qquad \text{Induction Hyp.} \quad \text{Factor.} \\
&= 3(q + k^2 + k) \qquad \text{(Un)Distributive} + \text{ over } \times
\end{aligned}$$

Or $(k+1)^3 - (k+1) = 3(q + k^2 + k)$.

$(q + k^2 + k)$ is integer (closed under addition and multiplication).

Another Induction Proof.

Theorem: For every $n \in N$, $n^3 - n$ is divisible by 3. ($3|(n^3 - n)$).

Proof: By induction.
Base Case: $P(0)$ is "$(0^3) - 0$" is divisible by 3. Yes!
Induction Step: $(\forall k \in N), P(k) \implies P(k+1)$
Induction Hypothesis: $k^3 - k$ is divisible by 3.
or $k^3 - k = 3q$ for some integer q.

$$\begin{aligned}
(k+1)^3 - (k+1) &= k^3 + 3k^2 + 3k + 1 - (k+1) \\
&= k^3 + 3k^2 + 2k \\
&= (k^3 - k) + 3k^2 + 3k \quad \text{Subtract/add } k \\
&= 3q + 3(k^2 + k) \qquad \text{Induction Hyp.} \quad \text{Factor.} \\
&= 3(q + k^2 + k) \qquad \text{(Un)Distributive} + \text{ over } \times
\end{aligned}$$

Or $(k+1)^3 - (k+1) = 3(q + k^2 + k)$.

$(q + k^2 + k)$ is integer (closed under addition and multiplication).
$\implies (k+1)^3 - (k+1)$ is divisible by 3.

Another Induction Proof.

Theorem: For every $n \in N$, $n^3 - n$ is divisible by 3. ($3|(n^3 - n)$).

Proof: By induction.
Base Case: $P(0)$ is "$(0^3) - 0$" is divisible by 3. Yes!
Induction Step: $(\forall k \in N), P(k) \implies P(k+1)$
Induction Hypothesis: $k^3 - k$ is divisible by 3.
or $k^3 - k = 3q$ for some integer q.

$$\begin{aligned}
(k+1)^3 - (k+1) &= k^3 + 3k^2 + 3k + 1 - (k+1) \\
&= k^3 + 3k^2 + 2k \\
&= (k^3 - k) + 3k^2 + 3k \quad \text{Subtract/add } k \\
&= 3q + 3(k^2 + k) \qquad \text{Induction Hyp.} \quad \text{Factor.} \\
&= 3(q + k^2 + k) \qquad \text{(Un)Distributive } + \text{ over } \times
\end{aligned}$$

Or $(k+1)^3 - (k+1) = 3(q + k^2 + k)$.

$(q + k^2 + k)$ is integer (closed under addition and multiplication).
$\implies (k+1)^3 - (k+1)$ is divisible by 3.
Thus, $(\forall k \in N) P(k) \implies P(k+1)$

Another Induction Proof.

Theorem: For every $n \in N$, $n^3 - n$ is divisible by 3. ($3|(n^3 - n)$).

Proof: By induction.
Base Case: $P(0)$ is "$(0^3) - 0$" is divisible by 3. Yes!
Induction Step: $(\forall k \in N), P(k) \implies P(k+1)$
Induction Hypothesis: $k^3 - k$ is divisible by 3.
or $k^3 - k = 3q$ for some integer q.

$$\begin{aligned}
(k+1)^3 - (k+1) &= k^3 + 3k^2 + 3k + 1 - (k+1) \\
&= k^3 + 3k^2 + 2k \\
&= (k^3 - k) + 3k^2 + 3k \quad \text{Subtract/add } k \\
&= 3q + 3(k^2 + k) \qquad \text{Induction Hyp.} \quad \text{Factor.} \\
&= 3(q + k^2 + k) \qquad \text{(Un)Distributive } + \text{ over } \times
\end{aligned}$$

Or $(k+1)^3 - (k+1) = 3(q + k^2 + k)$.

$(q + k^2 + k)$ is integer (closed under addition and multiplication).
$\implies (k+1)^3 - (k+1)$ is divisible by 3.
Thus, $(\forall k \in N) P(k) \implies P(k+1)$
Thus, theorem holds by induction.

Another Induction Proof.

Theorem: For every $n \in N$, $n^3 - n$ is divisible by 3. $(3|(n^3 - n)$).

Proof: By induction.
Base Case: $P(0)$ is "$(0^3) - 0$" is divisible by 3. Yes!
Induction Step: $(\forall k \in N), P(k) \implies P(k+1)$
Induction Hypothesis: $k^3 - k$ is divisible by 3.
or $k^3 - k = 3q$ for some integer q.

$$\begin{aligned}
(k+1)^3 - (k+1) &= k^3 + 3k^2 + 3k + 1 - (k+1) \\
&= k^3 + 3k^2 + 2k \\
&= (k^3 - k) + 3k^2 + 3k \quad \text{Subtract/add } k \\
&= 3q + 3(k^2 + k) \qquad \text{Induction Hyp.} \quad \text{Factor.} \\
&= 3(q + k^2 + k) \qquad \text{(Un)Distributive } + \text{ over } \times
\end{aligned}$$

Or $(k+1)^3 - (k+1) = 3(q + k^2 + k)$.

$(q + k^2 + k)$ is integer (closed under addition and multiplication).
$\implies (k+1)^3 - (k+1)$ is divisible by 3.
Thus, $(\forall k \in N) P(k) \implies P(k+1)$
Thus, theorem holds by induction. $\square$

Four Color Theorem.

Theorem: Any map can be colored so that those regions that share an edge have different colors.

Four Color Theorem.

Theorem: Any map can be colored so that those regions that share an edge have different colors.

Check Out: “Four corners”.

Four Color Theorem.

Theorem: Any map can be colored so that those regions that share an edge have different colors.

Check Out: “Four corners”.
States connected at a point, can have same color.

Four Color Theorem.

Theorem: Any map can be colored so that those regions that share an edge have different colors.

Check Out: “Four corners”.
States connected at a point, can have same color.
(Couldn’t find a map where they did though.)

Four Color Theorem.

Theorem: Any map can be colored so that those regions that share an edge have different colors.

Check Out: “Four corners”.
States connected at a point, can have same color.
(Couldn’t find a map where they did though.)

Quick Test: Which states?

Four Color Theorem.

Theorem: Any map can be colored so that those regions that share an edge have different colors.

Check Out: “Four corners”.
States connected at a point, can have same color.
(Couldn’t find a map where they did though.)

Quick Test: Which states? Utah.

Four Color Theorem.

Theorem: Any map can be colored so that those regions that share an edge have different colors.

Check Out: “Four corners”.
States connected at a point, can have same color.
(Couldn’t find a map where they did though.)

Quick Test: Which states? Utah. Colorado.

Four Color Theorem.

Theorem: Any map can be colored so that those regions that share an edge have different colors.

Check Out: “Four corners”.
States connected at a point, can have same color.
(Couldn’t find a map where they did though.)

Quick Test: Which states? Utah. Colorado. New Mexico.

Four Color Theorem.

Theorem: Any map can be colored so that those regions that share an edge have different colors.

Check Out: “Four corners”.
States connected at a point, can have same color.
(Couldn’t find a map where they did though.)

Quick Test: Which states? Utah. Colorado. New Mexico. Arizona.

Two color theorem: example.

Any map formed by dividing the plane into regions by drawing straight lines can be properly colored with two colors.

Two color theorem: example.

Any map formed by dividing the plane into regions by drawing straight lines can be properly colored with two colors.

Two color theorem: example.

Any map formed by dividing the plane into regions by drawing straight lines can be properly colored with two colors.

Two color theorem: example.

Any map formed by dividing the plane into regions by drawing straight lines can be properly colored with two colors.

Two color theorem: example.

Any map formed by dividing the plane into regions by drawing straight lines can be properly colored with two colors.

Two color theorem: example.

Any map formed by dividing the plane into regions by drawing straight lines can be properly colored with two colors.

Fact: Swapping red and blue gives another valid colors.

Two color theorem: example.

Any map formed by dividing the plane into regions by drawing straight lines can be properly colored with two colors.

Fact: Swapping red and blue gives another valid colors.

Two color theorem: proof illustration.

Base Case.

Two color theorem: proof illustration.

R

B

Base Case.

Two color theorem: proof illustration.

1. Add line.

Two color theorem: proof illustration.

1. Add line.
2. Get inherited color for split regions

Two color theorem: proof illustration.

1. Add line.
2. Get inherited color for split regions
3. Switch on one side of new line.

(Fixes conflicts along line, and makes no new ones.)

Two color theorem: proof illustration.

1. Add line.
2. Get inherited color for split regions
3. Switch on one side of new line.

(Fixes conflicts along line, and makes no new ones.)

Two color theorem: proof illustration.

1. Add line.
2. Get inherited color for split regions
3. Switch on one side of new line.

(Fixes conflicts along line, and makes no new ones.)

Two color theorem: proof illustration.

1. Add line.
2. Get inherited color for split regions
3. Switch on one side of new line.

(Fixes conflicts along line, and makes no new ones.)

Two color theorem: proof illustration.

1. Add line.
2. Get inherited color for split regions
3. Switch on one side of new line.

(Fixes conflicts along line, and makes no new ones.)

Two color theorem: proof illustration.

1. Add line.
2. Get inherited color for split regions
3. Switch on one side of new line.

(Fixes conflicts along line, and makes no new ones.)

Two color theorem: proof illustration.

1. Add line.
2. Get inherited color for split regions
3. Switch on one side of new line.

(Fixes conflicts along line, and makes no new ones.)

Two color theorem: proof illustration.

1. Add line.
2. Get inherited color for split regions
3. Switch on one side of new line.

(Fixes conflicts along line, and makes no new ones.)

Two color theorem: proof illustration.

1. Add line.
2. Get inherited color for split regions
3. Switch on one side of new line.

(Fixes conflicts along line, and makes no new ones.)

Two color theorem: proof illustration.

1. Add line.
2. Get inherited color for split regions
3. Switch on one side of new line.

(Fixes conflicts along line, and makes no new ones.)

Two color theorem: proof illustration.

1. Add line.
2. Get inherited color for split regions
3. Switch on one side of new line.

(Fixes conflicts along line, and makes no new ones.)

Algorithm gives $P(k) \implies P(k+1)$.

Two color theorem: proof illustration.

1. Add line.
2. Get inherited color for split regions
3. Switch on one side of new line.

(Fixes conflicts along line, and makes no new ones.)

Algorithm gives $P(k) \implies P(k+1)$. $\square$

Every natural number $k > 1$ can be written as a product of primes

(By induction) Let $P(k)$ be the predicate "k can be written as a product of primes"

Base Case: Since 2 is a prime number, $P(2)$ is true

Inductive Hypothesis: Let $k \geq 1$. We assume that $P(k)$ is true, *i.e.* "k can be written as a product of primes"

Inductive Step: We distinguish two cases: *(i)* Case $k+1$ is a prime, then $P(k+1)$ is true; *(ii)* Case $k+1$ is not a prime. Then by definition of primality, there must exist $1 < n, m < k+1$ such that $k+1 = n \cdot m$. But then we know by I.H. that n and m can be written as a product of primes (since $n, m \leq k$). Therefore, $k+1$ can also be written as a product of primes. Thus, $P(k+1)$ is true $\square$

Every natural number $k > 1$ can be written as a product of primes

(By induction) Let $P(k)$ be the predicate "k can be written as a product of primes"

Base Case: Since 2 is a prime number, $P(2)$ is true

Inductive Hypothesis: Let $k \geq 1$. We assume that $P(k)$ is true, *i.e.* "k can be written as a product of primes"

Inductive Step: We distinguish two cases: *(i)* Case $k+1$ is a prime, then $P(k+1)$ is true; *(ii)* Case $k+1$ is not a prime. Then by definition of primality, there must exist $1 < n, m < k+1$ such that $k+1 = n \cdot m$. But then we know by I.H. that n and m can be written as a product of primes (since $n, m \leq k$). Therefore, $k+1$ can also be written as a product of primes. Thus, $P(k+1)$ is true $\square$

$\longrightarrow$ If we had only assumed $P(k)$ to be true, then we could not apply our I.H. to n and m