DM11

Discrete Mathematics

Goals of course

- Mathematical sophistication

  1. How to read theorems, lemmas, etc.
  2. How to read proofs.
  3. How to state theorems precisely.
  4. How to write rigorous proofs.

- Necessary mathematical background for other courses

  1. Number theory.
  2. Recursive definitions and relations.
  3. More...

Why proofs?

- To better understand some facts.

- Some proofs give us algorithms. Understanding these helps us

  – Use the algorithms correctly to solve other problems.

  – Modify algorithms correctly to solve other problems.

- To ensure that ideas you use in programs are correct.

Logic

- **proposition** - statement which is *true* (**T**) or *false* (**F**)

  – The **truth value** of $1+1=2$ is *true*
  – The **truth value** of $1+1=3$ is *false*

- Operations on propositions:

  – *not, negation*
  $\neg(1+1=3)$ is *true*
  – *and, conjunction*
  $(1+1=2)\wedge(1+1=3)$ is *false*
  – *or, disjunction*
  $(1+1=2)\vee(1+1=3)$ is *true*
  – *xor, exclusive or*
  $(1+1=2)\oplus(1+1=3)$ is *true*
  – *implies, implication*
  $(1+1=2)\rightarrow(1+1=3)$ is *false*

truth tables

$p \wedge q$

p	q	$p \wedge q$
T	T	T
T	F	F
F	T	F
F	F	F

$\neg(p \vee q) \wedge r$

p	q	r	$p \vee q$	$\neg(p \vee q)$	$\neg(p \vee q) \wedge r$
T	T	T	T	F	F
T	T	F	T	F	F
T	F	T	T	F	F
T	F	F	T	F	F
F	T	T	T	F	F
F	T	F	T	F	F
F	F	T	F	T	T
F	F	F	F	T	F

$q \to p$ is the **converse** of $p \to q$

$\neg q \to \neg p$ is the **contrapositive** of $p \to q$

the **biconditional** $p \leftrightarrow q$ means
p if and only if q - *iff*

tautology - a proposition that is always true - $p \vee \neg p$

contradiction - a proposition that is always false - $p \wedge \neg p$

$p \Leftrightarrow q$ means p and q are **logically equivalent**
- $p \leftrightarrow q$ is a tautology

Examples $\begin{cases} p \to q \Leftrightarrow \neg p \vee q \\ p \vee \neg p \Leftrightarrow \mathbf{T} \\ p \wedge \neg p \Leftrightarrow \mathbf{F} \end{cases}$

Example - De Morgan's Laws

$$\neg(p \vee q) \Leftrightarrow \neg p \wedge \neg q$$

$$\neg(p \wedge q) \Leftrightarrow \neg p \vee \neg q$$

p	q	$p \wedge q$	$\neg(p \wedge q)$	$\neg p$	$\neg q$	$\neg p \vee \neg q$
T	T	T	F	F	F	F
T	F	F	T	F	T	T
F	T	F	T	T	F	T
F	F	F	T	T	T	T

Logical Equivalences

Equivalence	**Name for Law**
$p \wedge \mathbf{T} \Leftrightarrow p$ $p \vee \mathbf{F} \Leftrightarrow p$	Identity
$p \vee \mathbf{T} \Leftrightarrow \mathbf{T}$ $p \wedge \mathbf{F} \Leftrightarrow \mathbf{F}$	Domination
$p \wedge p \Leftrightarrow p$ $p \vee p \Leftrightarrow p$	Idempotent
$\neg(\neg p) \Leftrightarrow p$	Double negation
$p \vee q \Leftrightarrow q \vee p$ $p \wedge q \Leftrightarrow q \wedge p$	Commutative
$(p \vee q) \vee r \Leftrightarrow p \vee (q \vee r)$ $(p \wedge q) \wedge r \Leftrightarrow p \wedge (q \wedge r)$	Associative
$p \vee (q \wedge r) \Leftrightarrow (p \vee q) \wedge (p \vee r)$ $p \wedge (q \vee r) \Leftrightarrow (p \wedge q) \vee (p \wedge r)$	Distributive
$\neg(p \vee q) \Leftrightarrow \neg p \wedge \neg q$ $\neg(p \wedge q) \Leftrightarrow \neg p \vee \neg q$	De Morgan's

Think of $\vee$ as plus and $\wedge$ as times.
Then, you get familiar laws for the reals.

Think of **T** as a 1 and **F** as a 0.
Then, you get bit operations
- (integer operations modulo 2).

$a \wedge b$

a	b	$a \wedge b$
0	0	0
0	1	0
1	0	0
1	1	1

$a \vee b$

a	b	$a \vee b$
0	0	0
0	1	1
1	0	1
1	1	1

$a \oplus b$

a	b	$a \oplus b$
0	0	0
0	1	1
1	0	1
1	1	0

These tables can also be written as follows:

$a \wedge b$

$\wedge$	0	1
0	0	0
1	0	1

$a \vee b$

$\vee$	0	1
0	0	1
1	1	1

$a \oplus b$

$\oplus$	0	1
0	0	1
1	1	0

Bit operations can be extended to strings:
bitwise AND - $011010 \wedge 111000 = 011000$
bitwise OR - $011010 \vee 111000 = 111010$
bitwise XOR - $011010 \vee 111000 = 100010$

Proving tautologies
Example $[\neg p \wedge (p \vee q)] \rightarrow q$
By truth tables:

p	q	$\neg p$	$p \vee q$	$r = \neg p \wedge (p \vee q)$	$r \rightarrow q$
T	T	F	T	F	T
T	F	F	T	F	T
F	T	T	T	T	T
F	F	T	F	F	T

Any values for p and q give the result *true* $\Rightarrow$ tautology. □

Using known equivalences:

$$\begin{aligned}
[\neg p \wedge (p \vee q)] \rightarrow q &\Leftrightarrow [(\neg p \wedge p) \vee (\neg p \wedge q)] \rightarrow q \\
&\Leftrightarrow [\mathbf{F} \vee (\neg p \wedge q)] \rightarrow q \\
&\Leftrightarrow (\neg p \wedge q) \rightarrow q \\
&\Leftrightarrow \neg(\neg p \wedge q) \vee q \\
&\Leftrightarrow (p \vee \neg q) \vee q \\
&\Leftrightarrow p \vee (\neg q \vee q) \\
&\Leftrightarrow p \vee \mathbf{T} \\
&\Leftrightarrow \mathbf{T} \qquad \Rightarrow \text{tautology.} \quad \Box
\end{aligned}$$

Proving logical equivalences
Example $p \rightarrow q \;\Leftrightarrow\; \neg q \rightarrow \neg p$. Thus a proposition is equivalent to its contrapositive.
By truth tables:

p	q	$p \rightarrow q$	$\neg p$	$\neg q$	$\neg q \rightarrow \neg p$
T	T	T	F	F	T
T	F	F	F	T	F
F	T	T	T	F	T
F	F	T	T	T	T

Any values for p and q give the save values for the propositions. So the propositions are logically equivalent. □

Using known equivalences:

$$\begin{aligned}
p \rightarrow q &\Leftrightarrow \neg p \vee q \\
&\Leftrightarrow q \vee \neg p \\
&\Leftrightarrow \neg q \rightarrow \neg p
\end{aligned}$$

So they are logically equivalent. □

Predicates and Quantifiers

Example
$R(x, y) = (x + y < 25)$ is a *predicate*; it is also called a *propositional function*.
$R(14, 4)$ is true.

$\forall x P(x)$ is the *universal quantification* of $P(x)$.
Read it as "for all x $P(x)$".

Do we mean all integers x, all real $x \in [0, 1]$,...
What is the *universe of discourse*?
After knowing this, we can determine if the proposition is true.

Example:
$P(x) = (x^2 \geq x)$.
universe of discourse = $\mathbb{Z}$ - $\forall x P(x)$ is true.
universe of discourse = $\mathbb{R}$ - $\forall x P(x)$ is false.
$x = 1/2$ is a counterexample.

$\exists x P(x)$ is the *existential quantification* of $P(x)$.
Read it as "there exists an x such that $P(x)$".

Example:
$P(x) = (x^2 < x)$.
universe of discourse = $\mathbb{Z}$ - $\exists x P(x)$ is false.
universe of discourse = $\mathbb{R}$ - $\exists x P(x)$ is true.
$x = 1/2$ is an instance for which $P(x)$ holds.

Note: $\neg(a \geq b) \;\Leftrightarrow\; a < b$
These examples show how to negate when there are quantifiers:
$\neg \forall x P(x) \;\Leftrightarrow\; \exists x \neg P(x)$
$\neg \exists x P(x) \;\Leftrightarrow\; \forall x \neg P(x)$

Multiple quantifiers

Example: $\forall x \exists y(x+y=0)$
True over the integers; false over the positive integers.
Let $P(x) = \exists y(x+y=0)$.
$\forall x \exists y(x+y=0) \Leftrightarrow \forall x P(x)$.

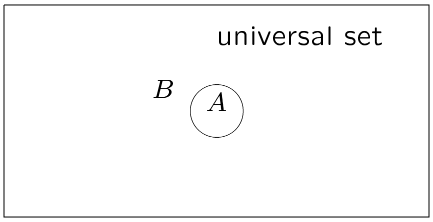
$$\begin{aligned} \neg[\forall x \exists y(x+y=0)] &\Leftrightarrow \neg \forall x P(x). \\ &\Leftrightarrow \exists x \neg P(x). \\ &\Leftrightarrow \exists x \forall y \neg(x+y=0). \\ &\Leftrightarrow \exists x \forall y(x+y \neq 0). \end{aligned}$$

In $P(x)$ the variable y is **bound**,
and the variable x is **free**.

Quantifiers can occur in the middle of propositions:
$\forall x[(x \geq 25) \ \vee \ \exists y((y>0) \wedge (x+y=25))]$

Sets

Venn diagram



$A \subset B$

$A \subseteq B$ means A is a *subset* of B.
$A \subset B$ means A is a *proper subset* of B.

$S = \{1,3,5,7\}$ is a *finite* set, because S has 4 elements, and 4 is a nonnegative integer.
The *cardinality* of S, written $|S|$, is 4.
$\mathbb{Z}$, the set of integers, is *infinite*, i.e. not finite.

The **power set** of S, written $P(S)$, is

$$\begin{aligned} \{ \ &\emptyset, \{1\}, \{3\}, \{5\}, \{7\}, \\ &\{1,3\}, \{1,5\}, \{1,7\}, \{3,5\}, \{3,7\}, \{5,7\}, \\ &\{1,3,5\}, \{1,3,7\}, \{1,5,7\}, \{3,5,7\}, \\ &\{1,3,5,7\}\} \end{aligned}$$

$P(S)$ is the set of all subsets of S.

The **Cartesian product** of the sets A_1, A_2,...,A_n, denoted $A_1 \times A_2 \times ... \times A_n$, is the set of ordered n-tuples $(a_1, a_2, ..., a_n)$ where $a_i \in A_i$ for $1 \leq i \leq n$.

Example: $S \times \{a,b\} =$

$$\begin{aligned} \{ \ &(1,a),(3,a),(5,a),(7,a), \\ &(1,b),(3,b),(5,b),(7,b) \ \} \end{aligned}$$

Operations on sets

Let $S = \{1,3,5,7\}$. Let U, the universal set, be the set of odd positive integers.

- *union*
  $S \cup \{3,9\} = \{1,3,5,7,9\}$
- *intersection*
  $S \cap \{3,9\} = \{3\}$
- *difference*
  $S - \{3,9\} = \{1,5,7\}$
- *complement*
  $\bar{S} = U - S =$ the odd integers > 8

A and B are *disjoint* if $A \cap B = \emptyset$.

Set Identities

Identity	Name for Law
$A \cup \emptyset = A$ $A \cap U = A$	Identity
$A \cup U = U$ $A \cap \emptyset = \emptyset$	Domination
$A \cup A = A$ $A \cap A = A$	Idempotent
$\overline{(\overline{A})} = A$	Complementation
$A \cup B = B \cup A$ $A \cap B = B \cap A$	Commutative
$A \cup (B \cup C) = (A \cup B) \cup C$ $A \cap (B \cap C) = (A \cap B) \cap C$	Associative
$A \cap (B \cup C) = (A \cap B \cup (A \cap C)$ $A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$	Distributive
$\overline{A \cup B} = \overline{A} \cap \overline{B}$ $\overline{A \cap B} = \overline{A} \cup \overline{B}$	De Morgan's

Notice the similarities to the logical equivalences.

Proving set equalities

Example $\overline{A \cup B} = \overline{A} \cap \overline{B}$

By showing that each is a subset of the other:

$$\begin{aligned} x \in \overline{A \cup B} &\Rightarrow x \notin A \cup B \\ &\Rightarrow \neg(x \in A \vee x \in B) \\ &\Rightarrow x \notin A \wedge x \notin B \\ &\Rightarrow x \in \overline{A} \wedge x \in \overline{B} \\ &\Rightarrow x \in \overline{A} \cap \overline{B} \end{aligned}$$

Thus $\overline{A \cup B} \subseteq \overline{A} \cap \overline{B}$.

$$\begin{aligned} x \in \overline{A} \cap \overline{B} &\Rightarrow x \in \overline{A} \wedge x \in \overline{B} \\ &\Rightarrow x \notin A \wedge x \notin B \\ &\Rightarrow \neg(x \in A \vee x \in B) \\ &\Rightarrow x \notin A \cup B \\ &\Rightarrow x \in \overline{A \cup B} \end{aligned}$$

Thus $\overline{A} \cap \overline{B} \subseteq \overline{A \cup B}$. Therefore $\overline{A} \cap \overline{B} = \overline{A \cup B}$.
□

Example: $A - B = A \cap \overline{B}$

Using set builder notation and logical equivalences:

$$\begin{aligned} A - B &= \{x \mid x \in A \wedge x \notin B\} \\ &= \{x \mid x \in A \wedge x \in \overline{B}\} \\ &= \{x \mid x \in A \cap \overline{B}\} \\ &= A \cap \overline{B} \qquad \square \end{aligned}$$

Example: $A \cap \overline{A} = \emptyset$

Using set builder notation and logical equivalences:

$$\begin{aligned} A \cap \overline{A} &= \{x \mid x \in A \wedge x \notin A\} \\ &= \{x \mid x \in A \wedge \neg(x \in A)\} \\ &= \{x \mid \mathbf{F}\} \\ &= \emptyset \qquad \square \end{aligned}$$

Prove $A \cup \overline{A} = U$ similarly.

Example: $\overline{A \cap B \cap C} = \overline{A} \cup \overline{B} \cup \overline{C}$

Using membership tables:

A	B	C	$A \cap B \cap C$	$\overline{A \cap B \cap C}$
0	0	0	0	1
0	0	1	0	1
0	1	0	0	1
0	1	1	0	1
1	0	0	0	1
1	0	1	0	1
1	1	0	0	1
1	1	1	1	0

$\overline{A}$	$\overline{B}$	$\overline{C}$	$\overline{A} \cup \overline{B} \cup \overline{C}$
1	1	1	1
1	1	0	1
1	0	1	1
1	0	0	1
0	1	1	1
0	1	0	1
0	0	1	1
0	0	0	0

No matter which of the three sets any element x is contained in, x is in $\overline{A \cap B \cap C}$ $\Leftrightarrow$ it is contained in $\overline{A} \cup \overline{B} \cup \overline{C}$. □

Example: $A \cap (B - A) = \emptyset$
Using known set identities:

$$\begin{aligned} A \cap (B - A) &= A \cap (B \cap \overline{A}) \\ &= (B \cap \overline{A}) \cap A \\ &= B \cap (\overline{A} \cap A) \\ &= B \cap (A \cap \overline{A}) \\ &= B \cap \emptyset \\ &= \emptyset \qquad \square \end{aligned}$$

Computer representations of sets

Suppose that a universal set U is finite.
We assume without loss of generality (WLOG) that $|U| = n$.

Represent a subset S of U by a binary string of length n.
1 $\Rightarrow$ that the element is present in S.
0 $\Rightarrow$ that the element is not present in S.

Example: $U = \{1, 3, 5, 7, 9\}$.
01010 represents $S_1 = \{3, 7\} \subseteq U$.
11100 represents $S_2 = \{1, 3, 5\} \subseteq U$.
$01010 \vee 11100 = 11110$ represents $S_1 \cup S_2$.
$01010 \wedge 11100 = 01100$ represents $S_1 \cap S_2$.

Functions

Giving keys to elements of array A:

n	1	2	3	4	5	6	7
$f(n)$	2.74	3.91	7.8345	1.6	4.293	8.0	1.234

$S = \{1, 2, 3, 4, 5, 6, 7\}$.
$f : S \to \mathbb{R}$ - a *function*.
f *maps* S to $\mathbb{R}$.
S is the *domain* of f.
$\mathbb{R}$ is the *codomain* of f.
$f(3) = 7.8345$.
7.8345 is the *image* of 3.
3 is *pre-image* of 7.8345.
$\{2.74, 3.91, 7.8345, 1.6, 4.293, 8.0, 1.234\}$ is the *range* of f.
Read the rest of this section in the text.

Sequences and summations

sequence - a function $f : S \subseteq \mathbb{Z} \to T$

strings - finite sequences
01011 $\leftrightarrow$ 0, 1, 0, 1, 1

$\{a_n\}$ where $a_n = 2^n$, $n \geq 0$ - infinite sequence
$\{a_n\} = \{1, 2, 4, 8, ...\}$

$\sum_{i=0}^{k} 2^i = 2^{k+1} - 1$ - **summation**

Why?

Geometric progression - $a,\ ar,\ ar^2,\ ar^3, ..., \ ar^k$

Theorem: $S = \sum_{j=0}^{n} ar^j = \begin{cases} (n+1)a & \text{if } r = 1 \\ \frac{ar^{n+1}-a}{r-1} & \text{if } r \neq 1 \end{cases}$

Proof. Clearly, the result holds for $r = 1$.
Suppose $r \neq 1$.

$$\begin{aligned} rS &= r\sum_{j=0}^{n} ar^j \\ &= \sum_{j=0}^{n} ar^{j+1} \\ &= \sum_{k=1}^{n+1} ar^k \qquad (k = j+1) \\ &= \sum_{k=0}^{n} ar^k + (ar^{n+1} - a) \\ &= S + (ar^{n+1} - a) \end{aligned}$$

$\Rightarrow \ rS - S \ = \ ar^{n+1} - a$
$\Rightarrow \ S \ = \ \frac{ar^{n+1}-a}{r-1}$ if $r \neq 1$. □

Case: $a = 1,\ r = 2 \ \Rightarrow \ \sum_{i=0}^{k} = 2^{k+1} - 1$

Cardinality

Cardinality of a finite set $S = |S|$
= number of elements

Cardinality of infinite sets?

Def. A and B have the same cardinality iff $\exists$ a 1-1 correspondence from A to B.

Example

$$\begin{array}{ccccccc} A = \{ & a, & b, & c, & d & \} \\ & \updownarrow & \updownarrow & \updownarrow & \updownarrow & \\ B = \{ & 1, & 2, & 3, & 4 & \} \end{array}$$

$|A| = |B| = 4$

Example

$I\!N$ - natural numbers - $\{1, 2, 3, ...\}$
$|\mathbb{Z}| = |I\!N|$ Proof?

$$\begin{array}{ccccccc} 1 & 2 & 3 & 4 & 5 & 6 & \ldots \\ \updownarrow & \updownarrow & \updownarrow & \updownarrow & \updownarrow & \updownarrow & \\ 0 & 1 & -1 & 2 & -2 & 3 & \ldots \end{array}$$

$$f(n) = \begin{cases} \frac{n}{2} & \text{if } n \text{ is even} \\ -\frac{n-1}{2} & \text{if } n \text{ is odd} \end{cases}$$

Claim: f is a bijection.
Pf. (1-1) Suppose $f(x) = f(y),\ x, y \in I\!N$.
n even $\Rightarrow \ f(n) \geq 0$
n odd $\Rightarrow \ f(n) < 0$
$\Rightarrow x$ and y both even or both odd.
Both even $\Rightarrow \ \frac{x}{2} = \frac{y}{2} \ \Rightarrow \ x = y$.
Both odd $\Rightarrow \ -\frac{x-1}{2} = -\frac{y-1}{2}$
$\Rightarrow \ x - 1 = y - 1 \ \Rightarrow \ x = y$.
$\Rightarrow f$ is 1-1.

Recall:

$$f(n) = \begin{cases} \frac{n}{2} & \text{if } n \text{ is even} \\ -\frac{n-1}{2} & \text{if } n \text{ is odd} \end{cases}$$

(onto) Suppose $x \in \mathbb{Z}$.
$x = 0 \ \Rightarrow \ f(1) = x$
$x > 0 \ \Rightarrow \ f(2x) = x$
$x < 0 \ \Rightarrow \ f(-(2x+1)) = x$
$\Rightarrow f$ is onto.

Therefore, f is a bijection. □

Def. countable — finite or same cardinality as $I\!N$
uncountable — not countable

Thm. Any subset of a countable set is countable.
Pf. Let S be a countable set and let $T \subseteq S$. Since S is countable, either S is finite or it has the same cardinality as $\mathbb{N}$. If S is finite, then T is also, so T is countable. If S is inifite, there is a function which gives a listing of its elements. List them in order. Remove those which are not in T. You still have a listing. Therefore, T is countable. $\square$

Thm. The positive rational numbers are countable.

Corollary The rational numbers $\mathbb{Q}$ are countable.

Thm. (Cantor) $\mathbb{R}$ - the set of real numbers - is uncountable.
Pf. (by diagonalization - a type of proof by contradiction) Suppose $\mathbb{R}$ is countable.
A subset of a countable set is countable (by Exercise).
$\Rightarrow [0,1)$ is countable.
$\Rightarrow$ The set can be listed - binary expansion

$a_1 = 0.d_{11}d_{12}d_{13}...$
$a_2 = 0.d_{21}d_{22}d_{23}...$
$a_3 = 0.d_{31}d_{32}d_{33}...$
$a_4 = 0.d_{41}d_{42}d_{43}...$
...

Consider $x = 0.b_1b_2b_3...$ where

$$b_i = \begin{cases} 0 & \text{if } d_{ii} = 1 \\ 1 & \text{if } d_{ii} = 0 \end{cases}$$

$x \neq a_i \ \forall i$
$\Rightarrow$ The list is incomplete. $\Rightarrow\!\Leftarrow$
Therefore, $\mathbb{R}$ is uncountable. $\square$

If S' has n elements, how many elements does $P(S')$ have?

We can represent a subset of S' by a binary string of length n. A 1 indicates that the element is present in S'.
Example: $\{0101\}$ represents $\{3,7\} \subseteq S$.
Each binary string of length n represents a different subset. There are 2^n distinct binary strings of length n, so $|P(S')| = 4^n$.

Introduction to Number Theory

Def. Suppose $a, b \in \mathbb{Z}$, $a > 0$.
Suppose $\exists c \in \mathbb{Z}$ s.t. $b = ac$. Then a *divides* b.
$a \mid b$.
a is a *factor* of b.
b is a *multiple* of a.
$e \not| f$ means e does not divide f.

Thm. $a, b, c \in \mathbb{Z}$. Then

1. if $a|b$ and $a|c$, then $a|(b+c)$
2. if $a|b$, then $a|bc \ \forall c \in \mathbb{Z}$
3. if $a|b$ and $b|c$, then $a|c$.

Pf.(of 3) $a|b \ \Rightarrow \ \exists \ d$ s.t. $b = ad$.
$b|c \ \Rightarrow \ \exists \ e$ s.t. $c = be$.
$\Rightarrow \ c = (ad)e$
$\Rightarrow \ c = a(de)$
$\Rightarrow \ a|c$ $\square$

Def. $p \in \mathbb{Z}$, $p > 1$.
p is *prime* if 1 and p are the only positive integers which divide p.
p is *composite* if it is not prime.

Thm. (The Fundamental Theorem of Arithmetic)
Every positive integer can be written uniquely as the product of primes, where the primes are written in nondecreasing order.

Thm. N composite $\Rightarrow$ N has a prime divisor $\leq \sqrt{N}$

Corollary There is an algorithm for factoring N (or testing primality) which does $O(\sqrt{N})$ tests of divisibility.
Problem The length of the input is $n = \lceil \log N \rceil$.
So the running time is $O(2^{n/2})$ - exponential.

Open Problem Does there exist a polynomial time factoring (or primality testing) algorithm?

Thm. $a \in \mathbb{Z}$, $d \in \mathbb{N}$
$\exists$ unique q, r, $0 \leq r < d$ s.t. $a = dq + r$

d − divisor
a − dividend
q − quotient
r − remainder $= a \bmod d$

Def. $gcd(a, b)$ = greatest common divisor of a and b
= largest $d \in \mathbb{Z}$ s.t. $d|a$ and $d|b$

If $gcd(a, b) = 1$, then a and b are *relatively prime*.

Def. $a \equiv b \pmod{m}$
— a is congruent to b modulo m
if $m \mid (a - b)$.

$m \mid (a - b) \;\Rightarrow\; \exists k \in \mathbb{Z}$ s.t. $a = b + km$.

Thm. $a \equiv b \pmod{m}$ $\quad c \equiv d \pmod{m}$
Then $a + c \equiv b + d \pmod{m}$
and $ac \equiv bd \pmod{m}$.
Pf.(of first) $\exists k_1, k_2$ s.t.
$a = b + k_1 m \qquad c = d + k_2 m$

$$\begin{aligned} a + c &= b + k_1 m + d + k_2 m \\ &= b + d + (k_1 + k_2) m \end{aligned} \qquad \square$$

Applications

- Hash functions: $h(k) = k \pmod{m}$
- Pseudorandom number generators
  – the linear congruential method:
  $x_{n+1} = a x_n + c \pmod{m}$
- Cryptology
  - Caesar cipher: $f(c) = c + k \pmod{26}$

a	b	c	...	w	x	y	z
d	e	f	...	z	a	b	c

  - RSA
  - Primality checking:
    Rabin-Miller algorithm

The Extended Euclidean Algorithm

Thm. $a, b \in \mathbb{N}$. $\exists\ s, t \in \mathbb{Z}$ s.t. $sa + tb = gcd(a,b)$.

Pf. Let d be the smallest positive integer in $D = \{xa + yb \mid x, y \in \mathbb{Z}\}$.
$d \in D \;\Rightarrow\; d = x'a + y'b$ for some $x', y' \in \mathbb{Z}$.
$gcd(a,b)|a$ and $gcd(a,b)|b$, so $gcd(a,b)|x'a$, $gcd(a,b)|y'$ and $gcd(a,b)|(x'a+y'b) = d$. We will show that $d|gcd(a,b)$, so $d = gcd(a,b)$. Note $a \in D$.
Suppose $a = dq + r$ with $0 \le r < d$.

$$\begin{aligned} r &= a - dq \\ &= a - q(x'a + y'b) \\ &= (1 - qx')a - (qy')b \end{aligned}$$

$\Rightarrow\; r \in D$
$r < d \;\Rightarrow r = 0 \;\Rightarrow d|a$.
Similarly, one can show that $d|b$.
Therefore, $d|gcd(a,b)$. $\square$

How do you find d, s and t?

Let $d = gcd(a,b)$. Write b as $b = aq + r$ with $0 \le r < d$.
Then, $d|b \;\Rightarrow\; d|(aq + r)$.
Also, $d|a \;\Rightarrow\; d|(aq) \;\Rightarrow\; d|((aq+r) - aq) \;\Rightarrow\; d|r$.

Let $d' = gcd(a, b - aq)$.
Then, $d'|a \;\Rightarrow\; d'|(aq)$
Also, $d'|(b - aq) \;\Rightarrow\; d'|((b - aq) + aq) \;\Rightarrow\; d'|b$.

Thus, $gcd(a,b) = gcd(a, b \pmod{a})$ $= gcd(b \pmod{a}, a)$. This shows how to reduce to a "simpler" problem and gives us the Extended Euclidean Algorithm:

{ Initialize}
$d_0 \leftarrow b$ $\quad s_0 \leftarrow 0$ $\quad t_0 \leftarrow 1$
$d_1 \leftarrow a$ $\quad s_1 \leftarrow 1$ $\quad t_1 \leftarrow 0$
$n \leftarrow 1$
{ Compute next d}
while $d_n > 0$ **do**
begin
$n \leftarrow n + 1$
{ Compute $d_n \leftarrow d_{n-2} \pmod{d_{n-1}}$}
$q_n \leftarrow \lfloor d_{n-2}/d_{n-1} \rfloor$
$d_n \leftarrow d_{n-2} - q_n d_{n-1}$
$s_n \leftarrow q_n s_{n-1} + s_{n-2}$
$t_n \leftarrow q_n t_{n-1} + t_{n-2}$
end
$s \leftarrow (-1)^n s_{n-1}$ $\quad t \leftarrow (-1)^{n-1} t_{n-1}$
$gcd(a,b) \leftarrow d_{n-1}$

The proof of correctness is by induction and the analysis of the running time uses Fibonacci numbers. Postponed.

Finding **multiplicative inverses** modulo m:
Given a and m, find x s.t. $a \cdot x \equiv 1 \pmod{m}$.
Should also find a k, s.t. $ax = 1 + km$.
So solve for an s in an equation $sa + tm = 1$.
This can be done if $gcd(a,m) = 1$.
Just use the Extended Euclidean algorithm.

Solving **linear congruences** modulo m:
Given a congruence $ax \equiv b \pmod{m}$, find x.
Suppose $gcd(a,m) = 1$.

1. Find $s = a^{-1} \pmod{m}$.
2. $x \leftarrow sb \pmod{m}$.

Note: $ax \equiv a(sb) \equiv (as)b \equiv b \pmod{m}$.

Thm (**The Chinese Remainder Theorem**)
Let $m_1, m_2, ..., m_k$ be pairwise relatively prime. For any integers $x_1, x_2, ..., x_k$, there exists $x \in \mathbb{Z}$ s.t. $x \equiv x_i \pmod{m_i}$ for $1 \leq i \leq k$, and this integer is uniquely determined modulo the product $m = m_1 m_2 ... m_k$.

It is also efficiently computable.

CRT Algorithm
For $1 \leq i \leq k$, find u_i such that

$$u_i \equiv 1 \pmod{m_i}$$
$$u_i \equiv 0 \pmod{m_j} \text{ for } j \neq i$$

Compute $x \equiv \sum_{i=1}^{k} x_i u_i \pmod{m}$.

How do you find each u_i?

$u_i \equiv 1 \pmod{m_i}$ $\forall i$
$\Rightarrow$ $\exists$ integers v_i s.t. $u_i + v_i m_i = 1$.
$u_i \equiv 0 \pmod{m_j}$ $\forall j \neq i$
$\Rightarrow$ $\exists$ integers w_i s.t. $u_i = w_i(m/m_i)$.
Thus, $w_i(m/m_i) + v_i m_i = 1$.
Solve for the values v_i and w_i
using the Extended Euclidean Algorithm.

(Note that this is where we need that the m_i are pairwise relatively prime.)
After each w_i is found, the corresponding u_i can be calculated.

The existence of the algorithm proves part of the theorem. What about uniqueness?
Suppose x and y work. Look at $x - y$.

Example: Let $m_1 = 3$, $m_2 = 5$, and $m_3 = 7$.
Suppose

$$\begin{aligned} x_1 &\equiv 2 \pmod{3} \\ x_2 &\equiv 3 \pmod{5} \\ x_3 &\equiv 4 \pmod{7} \end{aligned}$$

To calculate u_1:

$w_1(35) + v_1(3) = 1$
$w_1 = -1$; $v_1 = 12$
$u_1 = (-1)35 \equiv 70 \pmod{105}$

To calculate u_2:

$w_2(21) + v_2(5) = 1$
$w_2 = 1$; $v_2 = -4$
$u_2 = (1)21 \equiv 21 \pmod{105}$

To calculate u_3:

$w_3(15) + v_3(7) = 1$
$w_3 = 1$; $v_3 = -2$
$u_3 = (1)15 \equiv 15 \pmod{105}$

So we can calculate $x \equiv 2 \cdot 70 + 3 \cdot 21 + 4 \cdot 15 \equiv 53 \pmod{105}$.

Applications of the CRT

1. Computer arithmetic with large numbers:
   Suppose you want to perform operations on large numbers.
   Results never larger than m.
   Prefer operations on numbers $\leq s \ll m$.

   Find numbers p_1, $p_2, ..., p_k$, relatively prime.
   If $p_1 \cdot_2 \cdot ... \cdot p_k > m$, $p_i \leq s$ for $1 \leq i \leq k$,

   - Do operations modulo p_i for $1 \leq i \leq k$.
   - Calculate the result using the CRT.

2. Public key cryptography — RSA:

In a *public key cryptosystem*, each user A has

- A public key – P_A
- A secret key – S_A

To send a message m to A,
another user encrypts with P_A.
A decrypts with with S_A.

Example: RSA
Let $N_A = p_A \cdot q_A$, where p_A, q_A prime.
Let $gcd(e_A, (p_A - 1)(q_A - 1)) = 1$.
Let $e_A \cdot d_A \equiv 1 \pmod{(p_A - 1)(q_A - 1)}$.

- $P_A = (N_A, e_A)$
- $S_A = (N_A, d_A)$

To encrypt: $c = E_{P_A}(m) = m^{e_A} \pmod{N_A}$.
To decrypt: $D_{S_A}(c) = c^{d_A} \pmod{N_A}$.

Why does RSA work? CRT +

Fermat's Little Theorem: p is a prime, $p \not| a$.
Then $a^{p-1} \equiv 1 \pmod{p}$ and $a^p \equiv a \pmod{p}$.
Pf. Let $S = \{1, 2, ..., p-1\}$ and $T = \{1{\cdot}a \pmod{p}, 2{\cdot}a \pmod{p}, ..., (p-1){\cdot}a \pmod{p}\}$.
Clearly $T \subseteq S$.
Let $x \in S$, $y = a^{-1} \pmod{p}$, $z = xy \pmod{p}$.
Then $za \equiv x(ya) \equiv x \pmod{p}$, so $x \in T$.
Thus, $S \subseteq T$ and $S = T$.
By Wilson's Theorem, $(p-1)! \equiv -1 \pmod{p}$.
So

$$-1 \equiv \prod_{x \in S} x \equiv \prod_{y \in T} y \equiv (p-1)! a^{p-1} \equiv -a^{p-1} \pmod{p}.$$

$$\Rightarrow \quad a^{p-1} \equiv 1 \pmod{p} \quad \Rightarrow \quad a^p \equiv a \pmod{p}$$

□

Consider $x = D_{S_A}(E_{S_A}(m))$.
Note $\exists k$ s.t. $e_A d_A = 1 + k(p_A - 1)(q_A - 1)$.
$x \equiv (m^{e_A} \pmod{N_A})^{d_A} \pmod{N_A} \equiv m^{e_A d_A} \equiv m^{1+k(p_A-1)(q_A-1)} \pmod{N_A}$.

Consider $x \pmod{p_A}$.
$x \equiv m^{1+k(p_A-1)(q_A-1)} \equiv m{\cdot}(m^{(p_A-1)})^{k(q_A-1)} \equiv m \cdot 1^{k(q_A-1)} \equiv m \pmod{p_A}$.

Consider $x \pmod{q_A}$.
$x \equiv m^{1+k(p_A-1)(q_A-1)} \equiv m{\cdot}(m^{(q_A-1)})^{k(p_A-1)} \equiv m \cdot 1^{k(p_A-1)} \equiv m \pmod{q_A}$.

Apply the Chinese Remainder Theorem:
$gcd(p_A, q_A) = 1$, $\Rightarrow x \equiv m \pmod{N_A}$.
So $D_{S_A}(E_{S_A}(m)) = m$.

Note that anyone who knows p_A and q_A can decrypt, so the security of RSA depends on factoring being hard.

Rules of Inference

Rule	Tautology	Name
p / $\therefore p \vee q$	$p \to (p \vee q)$	Addition
$p \wedge q$ / $\therefore p$	$p \wedge q \to p$	Simplification
p / $p \to q$ / $\therefore q$	$[p \wedge (p \to q)] \to q$	Modus ponens
$\neg q$ / $p \to q$ / $\therefore \neg p$	$[\neg q \wedge (p \to q)] \to \neg p$	Modus tollens
$p \to q$ / $q \to r$ / $\therefore p \to r$	$[(p \to q) \wedge (q \to r)] \to (p \to r)$	Hypothetical syllogism
$p \vee q$ / $\neg p$ / $\therefore q$	$[(p \vee q) \wedge \neg p] \to q$	Disjunctive syllogism

Examples of poor reasoning

Assuming something is a tautology which is not:
If all of the governments predictions were correct, the economy will do well next year. This prediction is incorrect, so the economy will not do well next year.

Confusing $\exists$ with $\forall$:
We have given lots of money to programs to help poor people. See, this person collects money from two of these programs and she wears a mink coat. Therefore, all of this money is wasted. (It should be cut off.)

Circular reasoning:
To prove that there are only a finite number of primes. Let n be the largest prime. Then all numbers greater than n are composite. There are only n positive integers less than or equal to n, so any subset of them is finite. Thus, there are only a finite number of primes.

Read other examples in the textbook!

Proof types

To prove: If $10|n$, then $2|n$.

- **Direct proof:**
  $10|n \;\Rightarrow\; \exists a \in \mathbb{Z}$ s.t. $n = 10a$.
  $\Rightarrow \; n = (2 \cdot 5)a \;\Rightarrow\; n = 2(5a) \;\Rightarrow\; 2|n$
  $\square$
- **Indirect proof (uses the contrapositive):**
  Suppose $2 \not| n$. Then $\not\exists a \in \mathbb{Z}$ s.t. $n = 2a$. $\Rightarrow \; \forall a \in \mathbb{Z}\; n \neq 2a \;\Rightarrow\; \forall b \in \mathbb{Z}\; n \neq 2(5b) \;\Rightarrow\; \forall b \in \mathbb{Z}\; n \neq 10b \;\Rightarrow\; 10 \not| b$.
  Thus, if $10|n$, then $2|n$. $\square$
- **Proof by contradiction:**
  <u>Suppose $10|n$, but $2 \not| n$</u>. Then $\exists a \in \mathbb{Z}$ s.t. $n = 10a$ and $\forall b \in \mathbb{Z}\; n \neq 2b$. Thus, $\forall b \in \mathbb{Z}\; 10a \neq 2b$. Consider $b = 5a$. Then, $10a \neq 2(5a) = 10a$. Contradiction. Therefore, the assumption underlined must be false. Thus, if $10|n$, then $2|n$. $\square$

More proof types

- Constructive existence proof:
  Thm $\forall n \geq 1 \;\exists\; n$ consecutive composite integers.
  Pf. Let $x = (n+1)! + 1$ and $1 \leq i \leq n$. $(i+1) \mid (n+1)! + (i+1) = x+i$. Since $1 < i+1 < x+i$, $x+i$ is composite. These are n consecutive composite integers. $\square$
- Nonconstructive existence proof:
  Thm $\forall n \;\exists\; p > n$ such that p is prime.
  Pf. Let $n \in \mathbb{Z}$ and $m = n! + 1$.
  Suppose $\exists q$ s.t. $1 < q \leq n$ and $q|m$. Then $q|(n!+1) \;\Rightarrow\; q|(n!+1-n!) \;\Rightarrow\; q|1$. This is a contradiction, so any prime dividing m must be larger than n. So $\exists$ a prime larger than n. $\square$
- Proof by counterexample:
  Thm Not all odd numbers are prime.
  Pf. $9 = 3 \cdot 3$, so 9 is an odd number which is not prime. $\square$

Mathematical Induction

The Well-Ordering Property:
Every nonempty set of nonnegative integers has a least element.

Why important? Every set of counter-examples has a least element.
To show that something (about the positive integers) is true, prove that there is no counterexample.

Mathematcial Induction: To prove $P(n)$ holds $\forall$ positive integers n:

- **Basis step:** Prove that $P(1)$ holds.

- **Inductive step:** Show that if $P(n)$ holds, then $P(n+1)$ also holds $\forall n \geq 1$.

$P(n)$ is the inductive hypothesis.
Assuming this is not circular reasoning!
This shows that there is no least counter-example, so there is no counterexample.

Def. The *harmonic numbers* is $\{H_n \mid n \geq 1\}$ where

$$H_n = \sum_{i=1}^{n} \frac{1}{i}$$

Example: $H(4) = 1 + \frac{1}{2} + \frac{1}{3} + \frac{1}{4} = \frac{25}{12}$
Recall:

$$\sum_{i=1}^{n} \frac{1}{2^{i-1}} = 2 - \frac{1}{2^{n-1}},$$

so it is bounded by 2.

Thm. $H_{2^n} \geq 1 + \frac{n}{2} \quad \forall n \geq 0$.
Pf. (By induction on n
— $P(n)$ is "$H_{2^n} \geq 1 + \frac{n}{2}$.")
Basis step: $n = 0$. $H_{2^0} = H_1 = 1 \geq 1 + \frac{0}{2}$. $\surd$
Inductive step: Assume $H_{2^n} \geq 1 + \frac{n}{2}$.

$$\begin{aligned}
H_{2^{n+1}} &= \textstyle\sum_{i=1}^{2^{n+1}} \frac{1}{i} \\
&= H_{2^n} + \textstyle\sum_{i=2^n+1}^{2^{n+1}} \frac{1}{i} \\
&\geq (1 + \tfrac{n}{2}) + \textstyle\sum_{i=2^n+1}^{2^{n+1}} \frac{1}{i} && \text{by IH} \\
&\geq (1 + \tfrac{n}{2}) + \textstyle\sum_{i=2^n+1}^{2^{n+1}} \frac{1}{2^{n+1}} \\
&= (1 + \tfrac{n}{2}) + 2^n \cdot \tfrac{1}{2^{n+1}} \\
&= (1 + \tfrac{n}{2}) + \tfrac{1}{2} \\
&= (1 + \tfrac{n+1}{2}) && \surd
\end{aligned}$$

By induction, $H_{2^n} \geq 1 + \frac{n}{2} \quad \forall n \geq 0$. $\square$

Thm. Let S contain n elements. Then S has 2^n subsets.
Pf. (By induction on n – $P(n)$ is "Every set with n elements has 2^n subsets.")
Basis step: $n = 0$. $S = \emptyset$. The only subset of S is $\emptyset$. $2^0 = 1$. $\surd$
Inductive step: Assume that every set with $n \geq 0$ elements has 2^n subsets. Suppose S has $n+1$ elements. S has at least one element. Choose one x and consider $T = S - \{x\}$. T has n elements, so by the induction hypothesis, T has 2^n subsets. For each subset R of T there are 2 subsets of S, R and $R \cup \{x\}$. All subsets of S can be expressed like this. Thus S has $2 \cdot 2^n = 2^{n+1}$ subsets. $\surd$
By induction, if S has n elements, it has 2^n subsets. $\square$

The second principle of mathematical induction.

To prove $P(n)$ holds $\forall$ positive integers n:

- **Basis step:** Prove that $P(1)$ holds.
- **Inductive step:** Show that if $P(1)$, $P(2)$,..., $P(n)$ hold, then $P(n+1)$ also holds $\forall n \geq 1$.

Note: this is no stronger than ordinary mathematical induction, but it can be easier to use.

Recall the Extended Euclidean Algorithm:

{ Initialize}
$d_0 \leftarrow b \qquad s_0 \leftarrow 0 \qquad t_0 \leftarrow 1$
$d_1 \leftarrow a \qquad s_1 \leftarrow 1 \qquad t_1 \leftarrow 0$
$n \leftarrow 1$
{ Compute next d}
while $d_n > 0$ **do**
begin
$n \leftarrow n+1$
{ Compute $d_n \leftarrow d_{n-2} \pmod{d_{n-1}}$}
$q_n \leftarrow \lfloor d_{n-2}/d_{n-1} \rfloor$
$d_n \leftarrow d_{n-2} - q_n d_{n-1}$
$s_n \leftarrow q_n s_{n-1} + s_{n-2}$
$t_n \leftarrow q_n t_{n-1} + t_{n-2}$
end
$s \leftarrow (-1)^n s_{n-1} \qquad t \leftarrow (-1)^{n-1} t_{n-1}$
$gcd(a,b) \leftarrow d_{n-1}$

Prove correctness by induction.

Proof of correctness of the Extended Euclidean Algorithm

Claim: $gcd(a,b) = gcd(d_k, d_{k-1})$ for $1 \leq k \leq n$.
Pf. (by induction on k)
Basis step: $k = 1$. $d_1 = a$, $d_0 = b$, so $gcd(a,b) = gcd(d_1, d_{1-1})$. $\surd$
Inductive step: Assume that $gcd(a,b) = gcd(d_k, d_{k-1})$ and $1 \leq k \leq n-1$. Then $d_{k+1} = d_{k-1} - q_{k+1} d_k$, so, as argued before, $gcd(d_k, d_{k-1}) = gcd(d_{k+1}, d_k)$. By the inductive hypothesis, this is equal to $gcd(a,b)$. $\surd$
Thus, $gcd(a,b) = gcd(d_k, d_{k-1})$ for $1 \leq k \leq n$. $\square$
Since $d_n = 0$, $gcd(d_n, d_{n-1}) = d_{n-1}$, which is the result produced by the algorithm. By the previous claim, this is the correct greatest common divisor.

Claim: $(-1)^{k-1} s_k a + (-1)^k t_k b = d_k$ for $0 \leq k \leq n-1$.
Pf. (by induction on k)
Basis step: $k = 0$. $d_0 = b$, $s_0 = 0$, and $t_0 = 1$, so $(-1) \cdot 0 \cdot a + (1) \cdot 1 \cdot b = d_0$. $\surd$
$k = 1$. $d_1 = a$, $s_1 = 1$, and $t_1 = 0$, so $(1) \cdot 1 \cdot a + (-1) \cdot 0 \cdot b = d_1$. $\surd$
Inductive step: Assume $(-1)^{k-1} s_k a + (-1)^k t_k b = d_k$ for $0 \leq k < k'$, where $2 \leq k' \leq n-1$.

$$\begin{aligned} d_{k'} &= d_{k'-2} - q_{k'} d_{k'-1} \\ &= ((-1)^{k'-3} s_{k'-2} a + (-1)^{k'-2} t_{k'-2} b) \\ &\quad - q_{k'}((-1)^{k'-2} s_{k'-1} a + (-1)^{k'-1} t_{k'-1} b) \\ &= (-1)^{k'-1}(s_{k'-2} + q_{k'} s_{k'-1}) a \\ &\quad + (-1)^{k'}(t_{k'-2} + q_{k'} t_{k'-1}) b \\ &= (-1)^{k'-1} s_{k'} a + (-1)^{k'} t_{k'} b \qquad \surd \end{aligned}$$

Thus, the claim holds. $\square$
Since the algorithm sets $s = (-1)^n s_{n-1}$ and $t = (-1)^{n-1} t_{n-1}$, at the end $sa + tb = d_{n-1} = gcd(a,b)$. Thus, the algorithm is correct.

Recursive definitions – inductive definitions

1. Recursively defined functions (on $I\!N$)

- Define the function on the first k integers.
- Give some rule for determining the value at n in terms of the previous k integers.

Example: Define $F(n) = n! = n \cdot (n-1) \cdots 2 \cdot 1$.
$F(0) = 1$
$F(n+1) = (n+1)F(n)$

Example: Define the *Fibonacci numbers.*
$f_0 = 0$
$f_1 = 1$
$f_n = f_{n-1} + f_{n-2}$

$\{0, 1, 1, 2, 3, 5, 8, 13, 21, 34, 55, 89, ...\}$

Thm. For $n \geq 3$, $f_n > \alpha^{n-2}$,
where $\alpha = (1 + \sqrt{5})/2$ (the Golden Ratio).
Pf. (by induction on n)
Basis step: $n = 3$. $f_3 = 2 > \alpha = \alpha^{3-2}$. $\surd$
$n = 4$. $f_4 = 3 > (3 + \sqrt{5})/2 =$
$(1 + 2\sqrt{5} + 5)/4 = \alpha^{4-2}$. $\surd$
Inductive step: Suppose $f_k > \alpha^{k-2}$, for $3 \leq k \leq n$, where $n \geq 4$.
Note α is a solution of $x^2 - x - 1 = 0$, so $\alpha^2 = \alpha + 1$.

$$\begin{aligned} f_{n+1} &= f_n + f_{n-1} \\ &> \alpha^{n-2} + \alpha^{n-3} \quad \text{by IH} \\ &= (\alpha+1)\alpha^{n-3} \\ &= \alpha^2\alpha^{n-3} \\ &= \alpha^{n-1} \qquad \surd \end{aligned}$$

By induction, $f_n > \alpha^{n-2}$, for all $n \geq 3$. $\square$

Thm. [Lamé's Theorem] $a, b \in Z\!\!\!Z$. $a \geq b > 0$.
The number of divisions used by the Extended Euclidean Algorithm with a and b as input $\leq 5\lfloor \log_{10} a \rfloor + 6$.

Pf. The alg. computes $d_k = d_{k-2} \pmod{d_{k-1}}$ for $2 \leq k \leq n$, so there are $n-1$ divisions.
The quotients $q_i \geq 1$ for $i \geq 3$.
We prove by induction that $d_{n-i} \geq f_{i+1}$ for $1 \leq i \leq n-1$.
Basis step:
$d_{n-1} > 0 \;\Rightarrow\; d_{n-1} \geq 1 = f_2$
$d_{n-2} > d_{n-1} \;\Rightarrow\; d_{n-2} \geq 2 = f_3$ $\surd$
Inductive step: Assume true for $1 \leq i < k'$ where $3 \leq k' \leq n-1$.
$d_{n-k'+2} = d_{n-k'} - q_{n-k'+2} \cdot d_{n-k'+1} \Rightarrow$

$$\begin{aligned} d_{n-k'} &\geq d_{n-(k'-2)} + d_{n-(k'-1)} \\ &\geq f_{k'-1} + f_{k'} \\ &= f_{k'+1}. \qquad \surd \end{aligned}$$

Thus, $a = d_1 \geq f_n \geq \alpha^{n-2}$.

$$\begin{aligned} \lfloor \log_{10} a \rfloor + 1 &\geq \log_{10} a \\ &\geq \log_{10}(\alpha^{n-2}) \\ &= (n-2)\log_{10}\alpha \\ &> (n-2)/5. \end{aligned}$$

Thus, the number of division
$= n - 1 \leq 5(\lfloor \log_{10} a \rfloor + 1) + 1$. $\square$

Recursively defined sets

To define a set X recursively:

- Define a set of basis elements S_0.
- Define a finite set of operators.
- Define an infinite sequence of sets $\{S_i \mid i \geq 0\}$.
  $S_{i+1} = S_i \bigcup$ all of the objects which can be obtained from objects in S_i by applying one of the operators.
  $X = \bigcup_{i \geq 0} S_i$.

Example: The set of strings Σ^* over an alphabet Σ:
$S_0 = \{\lambda\}$ — λ is the empty string.
The operations are concatenation on the right by elements of Σ.

$x \in \Sigma^*$, $a \in \Sigma \Rightarrow xa \in X$.
S_n is the set of all strings of length $\leq n$.
Σ^* contains all strings of finite length.

Defining concatenation of strings:
$x, y \in \Sigma^*$, $a \in \Sigma$.
$x \odot \lambda = x$, and $x \odot ya = (x \odot y)a$.

Structural induction - for proving properties of recursively defined sets.
Same as mathematical induction.

Structural induction. Suppose X is a recursively defined set with basis elements S_0. Let $\{S_i \mid i \geq 0\}$ be the sequence of sets defined by the basis elements and operators. Let P be a propositional function defined on the elements of X. To prove $P(x)$ holds for all $x \in X$:

- **Basis step:** Prove that $P(y)$ holds for all $y \in S_0$.
- **Inductive step:** Show that if $P(y)$ holds for all $y \in S_n$, then $P(z)$ also holds for all $z \in S_{n+1}$.

Inductive step – prove for each operator $\oplus$ in the recursive definition that if P holds for some elements x_1, x_2,..., x_k, then P holds for the element $\oplus(x_1, x_2, ..., x_k)$.

Structural induction for strings:
Basis step: Prove that P holds for the empty string.
Then, show that if P holds for $y \in \Sigma^*$, it holds for every string ya where $a \in \Sigma$.

Define reverse of a string:
$\lambda^{\mathrm{R}} = \lambda$, $(xa)^{\mathrm{R}} = [a] \odot (x^{\mathrm{R}})$, if $x \in \Sigma^*$ and $a \in \Sigma$.
Want to prove that $(x^{\mathrm{R}})^{\mathrm{R}} = x$.

Claim: $\forall x \in \Sigma^*$, $\forall a \in \Sigma$, $([a] \odot x)^{\mathrm{R}} = x^{\mathrm{R}}a$.
Pf. (By structural induction on x.)
Basis step: $x = \lambda$. $([a] \odot x)^{\mathrm{R}} = ([a] \odot \lambda)^{\mathrm{R}} = (\lambda a)^{\mathrm{R}} = [a] \odot (\lambda)^{\mathrm{R}} = [a] \odot \lambda = \lambda a = \lambda^{\mathrm{R}} a = x^{\mathrm{R}} a$. $\surd$
Inductive step: Assume that for some $y \in \Sigma^*$, for all $a \in \Sigma$, $([a] \odot y)^{\mathrm{R}} = y^{\mathrm{R}}a$.
Let $x = yb$ where $b \in \Sigma$.
$([a] \odot x)^{\mathrm{R}} = ([a] \odot yb)^{\mathrm{R}} = (([a] \odot y)b)^{\mathrm{R}} = [b] \odot ([a] \odot y)^{\mathrm{R}} = [b] \odot (y^{\mathrm{R}}a)$ (by the IH). By the definitions of concatenation and reverse, $[b] \odot (y^{\mathrm{R}}a) = ([b] \odot y^{\mathrm{R}})a = (yb)^{\mathrm{R}}a = x^{\mathrm{R}}a$. $\surd$
The claim follows by structural induction. $\square$

Thm. $\forall x \in \Sigma^*$, $(x^{\mathrm{R}})^{\mathrm{R}} = x$.
Pf. (By structural induction on x.)
Basis step: $x = \lambda$. $(x^{\mathrm{R}})^{\mathrm{R}} = (\lambda^{\mathrm{R}})^{\mathrm{R}} = \lambda^{\mathrm{R}} = \lambda$. $\surd$
Inductive step: Assume that $(y^{\mathrm{R}})^{\mathrm{R}} = y$. Let $x = ya$ where $a \in \Sigma$. Then $(x^{\mathrm{R}})^{\mathrm{R}} = ((ya)^{\mathrm{R}})^{\mathrm{R}} = ([a] \odot (y^{\mathrm{R}}))^{\mathrm{R}} = (y^{\mathrm{R}})^{\mathrm{R}}a = ya = x$. $\surd$
The theorem follows by structural induction. $\square$

Def. A **binary tree** is either empty or it consists of a node called the *root* together with two binary trees called the *left subtree* and the *right subtree* of the root. A node is a *leaf* if its left and right subtree are empty; otherwise it is an *internal node*.

Structural induction for binary trees:
Basis step: Prove that P holds for the empty tree.
Inductive step: Show that if P holds for 2 subtrees L and R, it holds for the tree with root r, left subtree L, and right subtree R.

Thm. A binary tree with n internal nodes has at most $n+1$ leaves.
Pf. (By structural induction.)
Basis step: The empty tree has 0 internal nodes and 0 leaves. $\surd$
Inductive step: Suppose L is a binary tree; n_1 internal nodes; $\leq n_1 + 1$ leaves.
Suppose R is a binary tree; n_2 internal nodes; $\leq n_2 + 1$ leaves.
Suppose T has root r, left subtree L and right subtree R.
If both L and R are empty, then r is a leaf, so there are zero internal nodes and one leaf. $\surd$
Otherwise, r is an internal node of T.
The internal nodes of L and R are also internal nodes of T.
The leaves of L and R are the leaves of T.
Thus T has $1 + n_1 + n_2$ internal nodes and $\leq (n_1+1)+(n_2+1) = 2 + n_1 + n_2$ leaves. $\surd$
The theorem follows by structural induction.

Counting

- How many elements are there in $A_1 \times A_2 \times ... \times A_n$?

  Answer: $|A_1| \cdot |A_2| \cdots |A_n|$.

- How many bits strings are of length n?

  Answer: 2 possibilities for each bit, so 2^n.

- How many subsets are there of a set of size n?

  Answer: There is a 1-1 correspondence between these subsets and the bit strings of length n, so 2^n.

- How many functions $f : A \to B$, where $|A| = m$, $|B| = n$?

  Answer: There are n possible values for $f(a)$ $\forall a \in A$, so n^m.

- How many 1-1 functions $f : A \to B$, where $|A| = m$, $|B| = n$?
  Note: $m \le n$.

  Answer: n possibilities for the 1st
  $n - 1$ possibilities for the 2nd
  $n - 2$ possibilities for the 3rd
  ... $n - m + 1$ possibilities for last

  So $\prod_{i=0}^{m-1}(n-i) = \frac{n!}{(n-m)!}$.

- The Inclusion-Exclusion Principle. (2 sets)
  $|A \cup B| = |A| + |B| - |A \cap B|$

- Tree diagrams.

The Pigeonhole Principle

Thm. [The Pigeonhole Principle] If $\ge k+1$ pigeons go into k holes, then ≥ 1 hole has ≥ 2 pigeons.

Thm.
[The Generalized Pigeonhole Principle] If N pigeons go into k holes, then ≥ 1 hole has $\ge \lceil N/k \rceil$ pigeons.
Pf. Suppose no hole contains $> \lceil N/k \rceil - 1$ pigeons. Then, the number of pigeons is $\le k(\lceil N/k \rceil - 1) < k((N/k + 1) - 1) = N$. Contradiction. So ≥ 1 hole has $\ge \lceil N/k \rceil$ pigeons. $\Box$

Examples:

- Let $A = \{1, 2, 3, 4, 5, 6, 7, 8\}$.
  Suppose $B \subset A$, $|B| = 5$. Must there be 2 integers in B which sum to 9?
  Yes. Consider $\{\{1, 8\}, \{2, 7\}, \{3, 6\}, \{4, 5\}\}$. Since there are 5 elements from these 4 subsets, some subset must have 2.

- Suppose you have 10 red balls and 10 blue balls. How many do I have to take before I am sure I have ≥ 3 of the same color?
  5. There are 2 sets. 5 is the least N s.t. $\lceil N/2 \rceil = 3$.

- $3, 19, 13, 7, 12, 8, 4, 9, 29, 31, 2, 16, 1, 15, 21, 5, 6$.
  Does this sequence have a subsequence of length 5 which is strictly increasing or strictly decreasing?
  Yes. 3,7,8,29,31. With $n^2 + 1$ elements, must have $n + 1$ increasing or decreasing.

Ramsey Theory
Suppose there are 6 people.
Each pair is either friends or enemies.
There is a subset of 3 — all friends or all enemies.

Why? Consider one person P. There are 5 other people. By the Pigeonhole Principle, either at least 3 are friends or at least 3 are enemies of P.

Suppose A, B, and C are friends of P. If any pair are friends, we have a subset of 3 friends. If they are all enemies, they form a subset of 3 enemies.

Suppose A, B, and C are enemies of P. If any pair are enemies, we have a subset of 3 enemies. If they are all friends, they form a subset of 3 friends.

Permutations

Permutation – an ordering of some objects.
r-**permutation** – an ordering of r objects from a set.

Example: $5, 2, 3$ — 3-permutation of $\{1, 2, 3, 4, 5\}$.

Thm. The number of r-permutations of n elements is

$$P(n,r) = n(n-1)(n-2)\cdots(n-r+1) = n!/(n-r)!$$

Pf. Same argument as for the number of 1-1 functions $f : A \to B$, where $|A| = r$, $|B| = n$, since A can be thought of as the position number. $\square$

Combinations

r-**combination** – an unordered subset of r objects from a set.

Example: $\{2, 3, 5\}$ — 3-combination of $\{1, 2, 3, 4, 5\}$.

Thm. The number of r-combinations of n elements is

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

Pf. Any r-permutation is an ordering of an r-combination. There are $P(r,r) = r!$ ways to order an r-permutation. Thus,

$$P(n,r) = C(n,r) \cdot P(r,r)$$

and

$$C(n,r) = \frac{n!}{(n-r)!}/r! = \frac{n!}{r!(n-r)!} \qquad \square$$

Example 1: A *traveling salesman* starts in city 1, has to visit 5 other cities (6 in all), each exactly once, and then return home. How many tours are there?

Answer: $P(5,5) = 5!/0! = 5 \cdot 4 \cdot 3 \cdot 2 \cdot 1 = 120$.

Example 2: Suppose there are 36 numbers in a lottery and you have to guess 7 correct to win. How many different possibilities are you choosing from?

Answer: $\frac{36!}{7!29!} = 8,347,680$. So you have less than 1 chance in 8,000,000 of winning big.

Notation: $\binom{n}{r} = C(n,r)$
is a *binomial coefficient*.

Binomial Coefficients

Thm. [Pascal's Identity] $n \geq k \geq 1$.
$C(n+1,k) = C(n,k-1) + C(n,k)$

Pf.

$$\begin{aligned} C(n+1,k) &= \frac{(n+1)!}{k!(n+1-k)!} \\ &= \frac{(n+1)n!}{k!(n-k)!(n+1-k)} \\ &= \frac{(n+1-k)n!}{k!(n-k)!(n+1-k)} \\ &\quad + \frac{k \cdot n!}{k!(n-k)!(n+1-k)} \\ &= C(n,k) + \frac{n!}{(k-1)!(n+1-k)!} \\ &= C(n,k) + C(n,k-1) \end{aligned}$$ $\square$

Thm. $n \geq 1$. $\sum_{k=0}^{n} C(n,k) = 2^n$.
Pf. Consider a set with n elements.
There are $C(n,k)$ different subsets of size k.
Thus, the total number of subsets is $\sum_{k=0}^{n} C(n,k)$.
The number of subsets of a set of size n is 2^n. $\square$

Thm. [Vandermonde's Identity]
$n \geq r \geq 0$. $m \geq r$.
$C(m+n,r) = \sum_{k=0}^{r} C(m,r-k)C(n,k)$.
Pf. Suppose $|A| = n$, $|B| = m$.
Let $C = A \cup B$. $|C| = m+n$.
$C(m+n,r)$ = number of ways to choose r elements from C.
Same as choosing k from A and $r-k$ from B.
$C(m,r-k)C(n,k)$ ways — for fixed k.
$\sum_{k=0}^{r} C(m,r-k)C(n,k)$ ways in all. Thus,
$C(m+n,r) = \sum_{k=0}^{r} C(m,r-k)C(n,k)$. $\square$

Pascal's Triangle

$$\begin{array}{c}
\binom{0}{0} \\
\binom{1}{0} \quad \binom{1}{1} \\
\binom{2}{0} \quad \binom{2}{1} \quad \binom{2}{2} \\
\binom{3}{0} \quad \binom{3}{1} \quad \binom{3}{2} \quad \binom{3}{3} \\
\binom{4}{0} \quad \binom{4}{1} \quad \binom{4}{2} \quad \binom{4}{3} \quad \binom{4}{4} \\
\binom{5}{0} \quad \binom{5}{1} \quad \binom{5}{2} \quad \binom{5}{3} \quad \binom{5}{4} \quad \binom{5}{5} \\
\binom{6}{0} \quad \binom{6}{1} \quad \binom{6}{2} \quad \binom{6}{3} \quad \binom{6}{4} \quad \binom{6}{5} \quad \binom{6}{6} \\
\binom{7}{0} \quad \binom{7}{1} \quad \binom{7}{2} \quad \binom{7}{3} \quad \binom{7}{4} \quad \binom{7}{5} \quad \binom{7}{6} \quad \binom{7}{7} \\
\ldots
\end{array}$$

$$\begin{array}{c}
1 \\
1 \quad 1 \\
1 \quad 2 \quad 1 \\
1 \quad 3 \quad 3 \quad 1 \\
1 \quad 4 \quad 6 \quad 4 \quad 1 \\
1 \quad 5 \quad 10 \quad 10 \quad 5 \quad 1 \\
1 \quad 6 \quad 15 \quad 20 \quad 15 \quad 6 \quad 1 \\
1 \quad 7 \quad 21 \quad 35 \quad 35 \quad 21 \quad 7 \quad 1 \\
\ldots
\end{array}$$

$$\begin{aligned}
(x+y)^1 &= x + y \\
(x+y)^2 &= x^2 + 2xy + y^2 \\
(x+y)^3 &= x^3 + 3x^2y + 3xy^2 + y^3 \\
(x+y)^4 &= x^4 + 4x^3y + 6x^2y^2 + 4xy^3 + y^4 \\
(x+y)^5 &= x^5 + 5x^4y + 10x^3y^2 + 10x^2y^3 + 5xy^4 + y^5
\end{aligned}$$

Thm. [The Binomial Theorem]
$n \geq 1$. x, y variables.
$(x+y)^n = \sum_{j=0}^{n} C(n,j) x^{n-j} y^j$

$$= \binom{n}{0} x^n + \binom{n}{1} x^{n-1} y + \binom{n}{2} x^{n-2} y^2 + \ldots + \binom{n}{n-1} x y^{n-1} + \binom{n}{n} y^n .$$

Pf. (By induction on n)
Basis step: $n = 1$.
$(x+y)^1 = C(1,0)x^1y^0 \; + \; C(1,1)x^0y^1$. $\surd$
Inductive step: Suppose $n \geq 1$
and $(x+y)^n = \sum_{j=0}^{n} C(n,j) x^{n-j} y^j$.

$$\begin{aligned}(x+y)^{n+1} &= (x+y)(x+y)^n \\ &= x(\textstyle\sum_{j=0}^{n} C(n,j) x^{n-j} y^j) \\ &\quad + y(\textstyle\sum_{j=0}^{n} C(n,j) x^{n-j} y^j) \\ &= (\textstyle\sum_{j=0}^{n} C(n,j) x^{n+1-j} y^j) \\ &\quad + (\textstyle\sum_{j=0}^{n} C(n,j) x^{n-j} y^{j+1}) \\ &= (\textstyle\sum_{j=0}^{n} C(n,j) x^{n+1-j} y^j) \\ &\quad + (\textstyle\sum_{k=1}^{n+1} C(n,k-1) x^{n+1-k} y^k)\end{aligned}$$

For the last step $j = k-1$.

$$\begin{aligned}&= (\textstyle\sum_{j=0}^{n} C(n,j) x^{n+1-j} y^j) \\ &\quad + (\textstyle\sum_{k=1}^{n+1} C(n,k-1) x^{n+1-k} y^k) \\ &= C(n,0)x^{n+1}y^0 \; + \; C(n,n)x^0y^{n+1} \\ &\quad + \; \textstyle\sum_{k=1}^{n} (C(n,k) + C(n,k-1)) x^{n+1-k} y^k \\ &= C(n,0)x^{n+1}y^0 \; + \; C(n+1,n+1)x^0y^{n+1} \\ &\quad + \; \textstyle\sum_{k=1}^{n} C(n+1,k) x^{n+1-k} y^k \\ &= \textstyle\sum_{k=0}^{n+1} C(n+1,k) x^{n+1-k} y^k \qquad \surd\end{aligned}$$

So by induction, $(x+y)^n = \sum_{j=0}^{n} C(n,j) x^{n-j} y^j$ for all $n \geq 1$. $\square$

Alternate proof:
$(x+y)^n \; = \; (x+y)(x+y)\cdots(x+y)$
n terms.
Result — choose an x or a y from each term.
How many ways to get $x^{n-j}y^j$?
Choose which terms give a y — $C(n,j)$.
Thus, $(x+y)^n = \sum_{j=0}^{n} C(n,j) x^{n-j} y^j$ for $n \geq 1$.
$\square$

Using the binomial theorem

Thm. $n \geq 1$. $\sum_{k=0}^{n} C(n,k) = 2^n$.
Pf. $2^n = (1+1)^n = \sum_{k=0}^{n} C(n,k) 1^{n-k} 1^k = \sum_{k=0}^{n} C(n,k)$ $\square$

Thm. $n \geq 1$. $\sum_{k=0}^{n} (-1)^k C(n,k) = 0$.
Pf. $0 = (1+(-1))^n = \sum_{k=0}^{n} C(n,k) 1^{n-k} (-1)^k = \sum_{k=0}^{n} (-1)^k C(n,k)$ $\square$

Discrete Probability

Example:
6-sided dice — 2
experiment — throwing the dice, getting 2 numbers
sample space — all pairs of numbers (i,j), where $1 \leq i \leq j \leq 6$.
event — pairs of numbers that sum to 7
probability that the sum is 7 — $\frac{6}{6 \cdot 6} = \frac{1}{6}$.

Example:
Deck of 52 playing cards:
experiment — taking 13 different cards at random
sample space — all hands containing 13 cards
event — all hands with no face cards or aces

probability — $\frac{\frac{36!}{13!23!}}{\frac{52!}{13!39!}} = \frac{36!39!}{23!52!} = .00363896...$

Example: 36 numbered balls
experiment — randomly choosing 7 balls and then 4 more
sample space — all pairs of sets of numbers (A, B), where all numbers are distinct, $|A| = 7$, $|B| = 4$.
event — a specific set of pairs $(A_1, B_1), ..., (A_k, B_k)$, such that each A_i contains 6 of the 7 distinct numbers $x_1, x_2, x_3, x_4, x_5, x_6, x_7$ and each B_i contains the number missing from A_i.
probability is $\frac{k}{\frac{36!}{7!29!} \cdot \frac{29!}{4! \cdot 25!}} = \frac{k}{\frac{36!}{4!7!25!}}$.

So what is k? There are 7 possibilities for which x_j is not in A_i, 29 possibilities for that last number in A_i, and $28!/(3!25!)$ possibilities for the extra elements in B_i. So $k = 7 \cdot 29 \cdot 28!/(3!25!)$, and the entire probability is
$\frac{7 \cdot 29!/(3!25!)}{\frac{36!}{4!7!25!}} = \frac{4 \cdot 7 \cdot 29! 7!}{36!} < 3.3542 * 10^{-6}$.

Fact. If all outcomes of a finite sample space S are equally likely, the probability of an event E is $p(E) = |E|/|S|$. This distribution of probabilities is called the *uniform distribution*.

Def. Events $A_1, A_2, ...$ are *pairwise mutually exclusive* if $A_i \cap A_j = \emptyset$ for $i \neq j$.

Suppose sample space $S = \{x_1, x_2, ..., x_n\}$, and probability of x_i is $p(x_i)$.
Must have

- $0 \leq p(x_i) \leq 1 \quad \forall i$
- $\sum_{i=1}^{n} p(x_i) = 1$.
- For any events $A_1, A_2, ...A_k$ that are pairwise mutually exclusive, $p(\bigcup_i A_i) = \sum_i p(A_i)$.

Suppose sample space $S = \{x_i \mid i \geq 1\}$. Then,

- $0 \leq p(x_i) \leq 1 \quad \forall i$
- $\sum_{i=1}^{\infty} p(x_i) = 1$.
- For any events $A_1, A_2, ...$ that are pairwise mutually exclusive, $p(\bigcup_i A_i) = \sum_i p(A_i)$.

Example:
experiment — a fair coin is flipped until "heads".
sample space — $S = \{H, TH, TTH, ..., T^nH, ...\}$.
event — 1 sequence of flips
probability — $p(T^nH) = (1/2)^{n+1}$ for $n \geq 0$.

Suppose the coin is biased — $p(\text{heads}) = p$; $p(\text{tails}) = q = 1 - p$.
$p(T^nH) = q^n p$ — the *geometric distribution*

The probability of event E is

$$p(E) = \sum_{x_i \in E} p(x_i)$$

.

For fair 6-sided dice: $p(\{5, 6\}) = 1/3$.

Suppose a die is loaded so

$$\begin{aligned} p(1) &= 1/3 \\ p(2) &= 2/15 \\ p(3) &= 2/15 \\ p(4) &= 2/15 \\ p(5) &= 2/15 \\ p(6) &= 2/15 \end{aligned}$$

Then $\sum_{i=1}^{6} p(i) = 1/3 + 5(2/15) = 1$.

For this die, $p(\{5, 6\}) = 4/15 < 1/3$.

With 2 of these dice $p(\text{sum} = 7)$ is

$$2 \cdot \frac{1}{3} \cdot \frac{2}{15} + 4 \cdot \frac{2}{15} \cdot \frac{2}{15} = 4/25 < 1/6$$

Thm. $p(\overline{E}) = 1 - p(E)$.
Pf. $\sum_{i=1}^{n} p(x_i) = 1 = p(E) + p(\overline{E})$.
Thus, $P(\overline{E}) = 1 - p(E)$. □

Substituting ∞ for n changes nothing.

Thm. $p(E_1 \cup E_2) = p(E_1) + p(E_2) - p(E_1 \cap E_2)$.
Pf.

$$\begin{aligned} p(E_1 \cup E_2) &= \sum_{x_i \in E_1 \cup E_2} p(x_i) \\ &= \sum_{x_i \in E_1} p(x_i) + \sum_{x \in E_2} p(x_i) \\ &\quad - \sum_{x \in E_1 \cap E_2} p(x_i) \\ &= p(E_1) + p(E_2) - p(E_1 \cap E_2). \end{aligned}$$ □

Def. The **conditional probability** of E given F is $p(E|F) = \frac{p(E \cap F)}{p(F)}$.

Example: With 2 fair dice, what is the probability that the sum is 7, given that both dice are ≥ 3?
Answer:
p(sum is 7 and both $\geq$ 3)/p(both $\geq$ 3) = $2 \cdot \frac{1}{6} \cdot \frac{1}{6} / (\frac{2}{3} \cdot \frac{2}{3}) = 1/8$.

Thm. [Baye's Theorem] For events A, B, with $p(A) > 0$, $p(B) > 0$, $p(A|B) = \frac{p(A)p(B|A)}{p(B)}$.
Pf. By the def of conditional probability,
$p(A \cap B) = p(B)p(A|B)$.
$p(A \cap B) = p(A)p(B|A)$.
So $p(B)p(A|B) = p(A)p(B|A)$.
Divide both sides by $p(B)$. □

Example: With 2 loaded dice.
(with $p(1) = 1/3$ and the others 2/15),
what is the probability that the sum is 7, given that both dice are ≥ 3?
Answer:
p(sum is 7 and both ≥ 3)/p(both ≥ 3)
$= 2 \cdot \frac{2}{15} \cdot \frac{2}{15} / (\frac{8}{15} \cdot \frac{8}{15}) = 1/8$.

Example: Suppose you are on a game show. The host asks you to choose 1 of 3 doors for a prize. You choose door A.
The host opens another door B.
No big prize there.
You are told you can switch your choice.
Should you switch?

Answer: Yes.
p(prize behind A | not behind B)
$= p(A \wedge \neg B)/p(\neg B) = 1/2$ is not the answer.
B was chosen after A.
p(prize behind A | the host chose B) is the correct probability.
You could have chosen 3 doors. If you chose the prize, the host has 2 choices; otherwise only 1.

Your choice	Location of Prize	Prob of B
A	A	(1/3)(1/2)
A	B	(1/3)(0)
A	C	(1/3)(1)

So p(prize behind A and host chose B) = (1/3)(1/2) = 1/6.
p(host chose B) = (1/3)(1/2)+(1/3) = 1/2.
Thus, p(prize behind A | the host chose B) = (1/6)/(1/2)=1/3,
and p(prize behind A | the host chose B) = 2/3.

Another tricky example: Suppose you know that A has two children and you are told that one is a girl. What is the probability that the other is also a girl?

Wrong answer: 1/2 since there is always a 50-50 chance for a boy or a girl. This is wrong even assuming that the probabilities are exactly 50-50 and that the events are independent.
This is only correct if you said the oldest or the youngest, etc.

Correct answer: 1/3.
There are 4 possibilities: (G,G), (G,B), (B,G), (B,B).
"One is a girl" only rules out the last!.

Def. E and F are **independent** iff $p(E \cap F) = p(E)p(F)$.

Fact: If E and F are independent, then $p(E|F) = p(E \cap F)/p(F) = p(E)$.

Example: With 2 fair dice, the probability that the sum is 7 is *not* independent of both dice being ≥ 3.
$p(\text{sum is 7 and both} \geq 3)/p(\text{both} \geq 3) = 2 \cdot \frac{1}{6} \cdot \frac{1}{6}/(\frac{2}{3} \cdot \frac{2}{3}) = 1/8 \neq 1/6 = p(\text{sum is 7})$.

Example: The probability that the sum is 7 given that both dice are ≥ 3 is independent of the probabilities of the dice being 1 or 2. (Assuming that all probabilities are nonzero.)

Bernoulli trial: – an experiment with 2 outcomes; success with probability p, failure with probability $q = 1 - p$.

Bernoulli trials: — k independent repetitions of a Bernoulli trial.

Example: Consider 2 fair dice. Throw k times. What is the probability that the first time you get a sum of 7 is on throw j, where $j \leq k$.
Answer: $(\frac{5}{6})^{j-1}(\frac{1}{6})$.
This is the same as the biased coin with probabilities p of "heads" and $q = 1 - p$ of "tails". The answer is from the geometric distribution: $q^{j-1}p$.
Try tree diagrams.

Thm. The probability of k successes in n independent Bernoulli trials is $\binom{n}{k} p^k q^{n-k}$.
Pf. The outcome is $(x_1, x_2, ..., x_n)$, where
$x_i = \begin{cases} S, & \text{if the } i\text{th trial is success} \\ F, & \text{if the } i\text{th trial is failure} \end{cases}$
$p((x_1, x_2, ..., x_n))$, with k successes and $n - k$ failures is $p^k q^{n-k}$. There are $\binom{n}{k}$ outcomes with k successes and $n - k$ failures. So the probability is $\binom{n}{k} p^k q^{n-k}$. $\square$

The **binomial distribution** is
$b(k; n, p) = \binom{n}{k} p^k q^{n-k}$.

Note $\sum_{k=0}^{n} \binom{n}{k} p^k q^{n-k} = (p + q)^n = 1$.

Def. A *random variable* is a function $f : S \to I\!R$.

Def. For a finite sample space $S = \{s_1, s_2, ..., s_n\}$, the *expected value* of the random variable $X(s)$ is $E(X) = \sum_{i=1}^{n} p(s_i)X(s_i)$.

Def. For a countably infinite sample space $S = \{s_1 \mid i \geq 1\}$, the *expected value* of the random variable $X(s)$ is $E(X) = \sum_{i=1}^{\infty} p(s_i)X(s_i)$.

Example: What is the expected number of successes in n Bernoulli trials? Probability of success = p. Probability of failure = $q = 1-p$.

Answer:

$$\begin{aligned}
E[X] &= \sum_{k=1}^{n} kp(X=k) \\
&= \sum_{k=1}^{n} k \binom{n}{k} p^k q^{n-k} \\
&= \sum_{k=1}^{n} k \left(\frac{n!}{k!(n-k)!}\right) p^k q^{n-k} \\
&= \sum_{k=1}^{n} k \left(\frac{n(n-1)!}{k(k-1)!(n-1-k+1)!}\right) p^k q^{n-k} \\
&= n \sum_{k=1}^{n} \left(\frac{(n-1)!}{(k-1)!(n-1-(k-1))!}\right) p^k q^{n-k} \\
&= n \sum_{k=1}^{n} \binom{n-1}{k-1} p^k q^{n-k} \\
&= np \sum_{k=1}^{n} \binom{n-1}{k-1} p^{k-1} q^{n-k} \\
&= np \sum_{j=0}^{n-1} \binom{n-1}{j} p^j q^{n-1-j} \\
&= np(p+q)^{n-1} \\
&= np. \qquad \square
\end{aligned}$$

Example: What is the expected value of the first successful Bernoulli trial?

Answer:

$$\begin{aligned}
\sum_{i=1}^{\infty} iq^{i-1}p &= \sum_{i=1}^{\infty} iq^{i-1} - \sum_{i=1}^{\infty} iq^i \\
&= \sum_{j=0}^{\infty} (j+1)q^j - \sum_{i=1}^{\infty} iq^i \quad (j = i-1) \\
&= 1 + \sum_{j=1}^{\infty} (j+1-j)q^j \\
&= 1 + \sum_{j=1}^{\infty} q^j \\
&= 1 + (\sum_{j=0}^{\infty} q^j) - 1 \\
&= \lim_{k\to\infty} \frac{q^{k+1}-1}{q-1}
\end{aligned}$$

(Since this is a geometric progression.)

Since $q < 1$, $\lim_{k\to\infty} q^{k+1} = 0$.

Thus, $\displaystyle\sum_{i=1}^{\infty} iq^{i-1}p = \frac{-1}{q-1} = \frac{1}{p}$.

With a fair die, the expected number of throws before a 1 is 6.

Linearity of Expectation

A *linear funtion* has the form

$f(X_1, X_2, ..., X_n) = a_0 + a_1X_1 + a_2X_2 + ... + a_nX_n$

where $a_i \in I\!R$ for $0 \leq i \leq n$.

Thm. Let f be a linear function, S be a sample space, and $X_1, X_2, ..., X_n$ be random variables defined on S. Then, $E[f(X_1, X_2, ..., X_n)] = f(E[X_1], E[X_2], ..., E[X_n])$.

Pf. Let $f(X_1, ..., X_n) = a_0 + a_1X_1 + ... + a_nX_n$ where $a_i \in I\!R$ for $0 \leq i \leq n$. Then,

$$\begin{aligned}
&E[f(X_1, ..., X_n)] \\
&= \sum_{s\in S} p(s)f(X_1(s), ..., X_n(s)) \\
&= \sum_{s\in S} p(s)(a_0 + a_1X_1(s) + ... + a_nX_n(s)) \\
&= \sum_{s\in S} (p(s)a_0 + \sum_{i=1}^{n} p(s)a_iX_i(s)) \\
&= a_0 + \sum_{i=1}^{n} (\sum_{s\in S} p(s)a_iX_i(s)) \\
&= a_0 + \sum_{i=1}^{n} (a_i \sum_{s\in S} p(s)X_i(s)) \\
&= f(E[X_1], E[X_2], ..., E[X_n]). \qquad \square
\end{aligned}$$

If two random variables X and Y are independent, then $E[XY] = E[X] \cdot E[Y]$.

The **variance** of a random variable is *Var*$[X]$ $= E[(X - E[X])^2]$.

Var$[X] = E[X^2 - 2XE[X] + E^2[X]]$.
By the linearity of expectations, this is
$E[X^2] - 2E[XE[X]] + E[E^2[X]]$.
Since $E[X]$ is a real number, this is
$E[X^2] - 2E^2[X] + E^2[X]$.
Thus *Var*$[X]$ is also $E[X^2] - E^2[X]$.

If X and Y are independent random variables, then *Var*$[X+Y] =$ *Var*$[X]$ $+$*Var*$[Y]$. If $X_1, X_2, ..., X_n$ are pairwise independent random variables, then *Var*$[\sum_{i=1}^{n} X_i] = \sum_{i=1}^{n}$ *Var*$[X_i]$.

The **standard deviation** of a random variable is the positive square root of the variance.

The variance of a geometric distribution can be shown to q/p (recall that the expectation is $1/p$).

Consider Bernoulli trials.
X_i — how many successes in the ith trial.
Must be either 0 or 1,
so $E[X_i^2] = E[X_i] = p$.
The variance of the binomial distribution —
Var$[X]$ = *Var*$[\sum_{i=1}^{n} X_i] = \sum_{i=1}^{n}$ *Var*$[X_i]$
since the X_i are pairwise independent.
Var$[X_i] = E[X_i^2] - E^2[X_i] = p - p^2 = pq$.
Thus for the binomial distribution, *Var*$[X]$ = $\sum_{i=1}^{n} pq = npq$.

Permutations with repetition — sampling with replacement

Example

A bag with 3 red balls,
4 green balls, and
5 blue balls.

- Take a ball out.
- Put it back and take another ball out.
- Put it back and take another ball out.

Probability all 3 are blue $= \frac{5^3}{12^3} = 125/1728$.

Combinations with repetition

Example
Same experiment. 3 red. 4 green. 5 blue.
How many ways to get 1 red ball and 2 green balls? $3 \cdot 4^2$
Note: not interested in order.
Represent result as: $n_1 \mid n_2 \mid n_3$, where
n_1 — number of red balls
n_2 — number of green balls
n_3 — number of blue balls
We wanted: $1 \mid 2 \mid 0$.
Could also represent as $*|**|$.
How many distinct results are there? $(n_1 \mid n_2 \mid n_3)$
The number of ways to write down 3 *s and 2 |s.
There are 5 places. Choose 3 places for stars.
Answer: $\binom{5}{3} = \frac{5!}{3!2!} = 10$.

Example
Same experiment with n types of balls and r choices.
(An r-combination from n elements, with repetition.)
There are r stars and $n-1$ bars.
Number of distinct results is $\binom{n+r-1}{r}$.

Example
How many ways can you choose k integers ≥ 0, $(n_1, n_2, ..., n_k)$, such that $\sum_{i=1}^{k} n_i = n$?
Answer: $\binom{k+n-1}{n}$.
How many nonnegative solutions are there to

$$n_1 + n_2 + n_3 + n_4 = 7?$$

Answer: $\binom{4+7-1}{7} = \frac{10!}{7!3!} = 120$.

Example
A bag with 3 red balls, 4 green balls, and 5 blue balls.
Consider all red balls the same, etc.
How many orderings are there of the balls?
There are (3+4+5)! orderings.
Given one ordering, $(x_1, x_2, ..., x_{12})$, there are 3!4!5! which are identical to it.
The number of distinct orderings is $\frac{12!}{3!4!5!} = 27,720$.

Example
How many ways to put n pigeons in k holes, if the ith hole should have n_i pigeons?
Answer: $\binom{n}{n_1}$ for the first hole.
$\binom{n-n_1}{n_2}$ for the second hole.
$\binom{n-n_1-n_2}{n_3}$ for the third hole.
Etc. Multiply these together to get: $\frac{n!}{n_1!n_2!\cdots n_k!}$

Generating permutations

Suppose you want to test your sorting algorithm.
Try it on all permutations of $\{1, 2, ..., n\}$.
How large should n be?
If n is 10, this is 10!=3,628,800, so not more than 10.

Lexicographic order would be nice.
$(2, 4, 7, 3, 8, 1, 9, 6, 5, 10) <$
$(2, 4, 7, 5, 6, 10, 9, 1, 3, 8)$ because $3 < 5$.

How to get the next larger permutation?
(after $(a_1, a_2, ..., a_n)$)
(1,2,3),(1,3,2),(2,1,3),(2,3,1),(3,1,2),(3,2,1).

(1,2,3),(1,3,2),(2,1,3),(2,3,1),(3,1,2),(3,2,1).

If $a_{n-1} < a_n$, switch them.

Otherwise, find largest j with $a_j < a_{j+1}$.
$a_{j+1} > a_{j+2} > ... > a_n$.
Find the least value $a_k > a_j$ from $\{a_{j+1}, a_{j+2}, ..., a_n\}$.
Put a_k where a_j was.
Put the remaining elements from $\{a_j, a_{j+1}, ..., a_n\}$ in increasing order.

$(1, 2, 5, 10, 7, \mathbf{4}, 9, 8, 6, 3) \rightarrow$
$(1, 2, 5, 10, 7, \mathbf{6}, 3, 4, 8, 9)$.

procedure next_permutation($a_1, a_2, ..., a_n$)
{ a permutation of $(1, 2, ..., n)$, $\neq (n, n-1, ..., 1)$}
$j \leftarrow n-1$
while $a_j > a_{j+1}$
    $j \leftarrow j-1$
{ j is largest subscript with $a_j < a_{j+1}$}

$k \leftarrow n$
while $a_j > a_k$
    $k \leftarrow k-1$
{ a_k is smallest value $> a_j$ to right of a_j}

switch a_j and a_k
$r \leftarrow n$
$s \leftarrow j+1$
while $r > s$
    switch a_r and a_s
    $r \leftarrow r-1$
    $s \leftarrow s+1$
{ this reverses the order after a_j}

Generating combinations

Combinations are subsets:
So use binary strings to represent them.
Lexicographic order of strings is increasing order of integers.
$(000), (001), (010), (011),$
$(100), (101), (110), (111)$.
To get next integer, find rightmost 0.
Change it to 1. Change all 1's to right to 0's.

For an r-combination, have r ones in the string.
$(00111), (01011), (01101), (01110), (10011),$
$(10101), (10110), (11001), (11010), (11100)$.
Find the rightmost 01.
Change it to 10.
Move all the 1's to the right as far right as possible.

To find the next r-combination of $(1, 2, ..., n)$:
$(11100) \leftrightarrow (1, 2, 3)$ — the smallest.
Lexicographic order is not lexicographic order of strings.
To get the next, find the rightmost 10.
Change it to 01.
Move all the 1's to the right as far left as possible.

To do this directly:
Suppose have $(a_1, a_2, ..., a_r)$.
If this looks like
$(a_1, ..., a_k, n-r+k+1, n-r+k+2, ..., n-1, n)$,
should change a_k to $a_k + 1$.
$a_{k+1} \leftarrow a_k + 2$.
$a_{k+2} \leftarrow a_k + 3$.
Generally, $a_j = a_k + j - k + 1$, for $k+1 \leq j \leq r$.
Suppose $n = 6$. Consider $(2, 5, 6) \rightarrow (3, 4, 5)$.

procedure next_combination($a_1, a_2, ..., a_r$)
{ a r-combination of $(1, 2, ..., n)$,
$\neq (n-r+1, ..., n)$}
$i \leftarrow r$
while $a_i = n - r + i$
    $i \leftarrow i-1$
$a_i \leftarrow a_i + 1$
for $j \leftarrow i+1$ **to** r
    $a_j \leftarrow a_i + j - i + 1$ (typo in text)

The Principle of Inclusion-Exclusion

$A_1, A_2, ..., A_n$ finite sets.

$|A_1 \cup A_2| = |A_1| + |A_2| - |A_1 \cap A_2|$.

$|A_1 \cup A_2 \cup A_3| = |A_1| + |A_2| + |A_3| - |A_1 \cap A_2|$
$\quad -|A_1 \cap A_2| - |A_1 \cap A_2| + |A_1 \cap A_2 \cap A_3|$.

Thm.[The Principle of Inclusion-Exclusion]

$$|A_1 \cup A_2 \cup ... \cup A_n| = \sum_{1 \leq i \leq n} |A_i| - \sum_{1 \leq i < j \leq n} |A_i \cap A_j| + \sum_{1 \leq i < j < k \leq n} |A_i \cap A_j \cap A_k| - ... + (-1)^{n+1} |A_1 \cap A_2 \cap ... \cap A_n|.$$

Pf. (by induction on n)
Basis step: $n = 2$. Proven earlier.
Inductive step:

Applications
Counting number of elements with or without the properties $P_1, P_2, ..., P_n$:

$N(P_{i_1} P_{i_2} ... P_{i_k})$ — number of elements with properties $P_{i_1}, P_{i_2}, ..., P_{i_k}$.
$N(P'_{i_1} P'_{i_2} ... P'_{i_k})$ — number of elements with none of the properties $P_{i_1}, P_{i_2}, ..., P_{i_k}$.

A_i — subset of elements with property P_i.
N — total number of elements
$N(P_{i_1} P_{i_2} ... P_{i_k}) = |A_{i_1} \cap A_{i_2} \cap ... \cap A_{i_k}|$
$N(P'_{i_1} P'_{i_2} ... P'_{i_k}) = N - |A_{i_1} \cup A_{i_2} \cup ... \cup A_{i_k}|$

$$N(P'_{i_1} P'_{i_2} ... P'_{i_k}) = N - \sum_{1 \leq i \leq n} N(P_i) + \sum_{1 \leq i < j \leq n} N(P_i P_j) - ... + (-1)^n N(P_1, P_2, ..., P_n).$$

How many solutions are there to $n_1 + n_2 + n_3 + n_4 = 7$, if you must have $n_1, n_2, n_3 \leq 3$ and $n_4 \leq 4$.

Property P_1 is $n_1 > 3$.
Property P_2 is $n_2 > 3$.
Property P_3 is $n_3 > 3$.
Property P_4 is $n_4 > 4$.

$$\begin{aligned} N(P'_1 P'_2 P'_3 P'_4) = N & - N(P_1) - N(P_2) - N(P_3) - N(P_4) \\ & + N(P_1 P_2) + N(P_1 P_3) + N(P_1 P_4) \\ & + N(P_2 P_3) + N(P_2 P_4) + N(P_3 P_4) \\ & - N(P_1 P_2 P_3) - N(P_1 P_2 P_4) \\ & - N(P_1 P_3 P_4) - N(P_2 P_3 P_4) \\ & + N(P_1 P_2 P_3 P_4) \end{aligned}$$

- N = total number = $\binom{4+7-1}{7} = 120$.
- $N(P_1)$ = number with $n_1 \geq 4$ $= \binom{4+3-1}{3} = 20$
- $N(P_2) = N(P_3) = N(P_1) = 20$
- $N(P_4)$ = number with $n_4 \geq 5$ $= \binom{4+2-1}{2} = 10$
- $N(P_1 P_2)$ = number with $n_1 \geq 4$ and $n_2 \geq 4$. This is impossible, so 0.
- All intersections of at least 2 are impossible.

Thus, $N(P'_1 P'_2 P'_3 P'_4) = 120 - 3(20) - 10 + 0 = 50$.

The Sieve of Eratosthenes
To get a list of primes $\leq B$:

Create an empty list P.
Create a list L of the integers $2..B$.
while L is not empty **do**
  Remove p — smallest element in L.
  Insert p in P.
  Delete all multiples of p from L.

How many primes are there ≤ 19?
How many numbers are in P if $B = 19$?

Property $P_i(x)$ is: ith prime$< x \leq 19$
and ith prime divides x.
Every composite ≤ 19 is divisible by a prime $\leq \sqrt{19} \leq 5$.

Answer: $N(P_1'P_2'P_3') = N$
$- N(P_1) - N(P_2) - N(P_3)$
$+ N(P_1P_2) + N(P_1P_3) + N(P_2P_3) -$
$N(P_1P_2P_3)$

- $N = 18$.
- $N(P_1) = 8$ (2 divides x)
- $N(P_2) = 5$ (3 divides x)
- $N(P_3) = 2$ (5 divides x)
- $N(P_1P_2) = 3$ (6 divides x)
- $N(P_1P_3) = 1$ (10 divides x)
- $N(P_2P_3) = 1$ (15 divides x)
- $N(P_1P_2P_3) = 0$ (30 divides x)

$N(P_1'P_2'P_3') = 18-8-5-2+3+1+1-0 = 8.$

Derangements
Example: Peter likes betting. He hears:

- Team A is expected to win.
- Team B is expected to be 2nd.
- Team C is expected to be 3rd.
- Team D is expected to be 4th.
- Team E is expected to be 5th.

Peters bets on these 5 events.
In how many ways can he lose all bets?

A *derangement* is a permutation with no object in its original position.

Property P_i — place i was correct.
Want $D = N(P_1'P_2'P_3'P_4'P_5')$.

$$\begin{aligned} D &= N - \sum_{i=1}^{5} N(P_i) + \sum_{1\leq i<j\leq 5} N(P_iP_j) \\ &- \sum_{1\leq i<j<k\leq 5} N(P_iP_jP_k) \\ &+ \sum_{1\leq i<j<k<l\leq 5} N(P_iP_jP_kP_l) \\ &- N(P_1P_2P_3P_4P_5). \end{aligned}$$

$N = 5!$.
$N(P_i) = (5-1)! \quad \forall i$.
$N(P_iP_j) = (5-2)! \quad \forall i,j$.
$N(P_iP_jP_k) = (5-3)! \quad \forall i,j,k$.
$N(P_iP_jP_kP_l) = (5-4)! \quad \forall i,j,k,l$.
$N(P_1P_2P_3P_4P_5) = (5-5)!$.

How many terms are there in each sum?

$\sum_{i=1}^{5} N(P_i)$ — $\binom{5}{1}$.

$\sum_{1\le i<j\le 5} N(P_iP_j)$ — $\binom{5}{2}$.

$\sum_{1\le i<j<k\le 5} N(P_iP_jP_k)$ — $\binom{5}{3}$.

$\sum_{1\le i<j<k<l\le 5} N(P_iP_jP_kP_l)$ — $\binom{5}{4}$.

$N(P_1P_2P_3P_4P_5)$ — $\binom{5}{5}$.

$$D = 5! - (5-1)!\tfrac{5!}{1!(5-1)!} + (5-2)!\tfrac{5!}{2!(5-2)!} - (5-3)!\tfrac{5!}{3!(5-3)!} + (5-4)!\tfrac{5!}{4!(5-4)!} - (5-5)!\tfrac{5!}{5!(5-5)!}$$

$$D = 5!\left(1 - \tfrac{1}{1!} + \tfrac{1}{2!} - \tfrac{1}{3!} + \tfrac{1}{4!} - \tfrac{1}{5!}\right) = 44.$$

Thm. The number of derangements of a set with n elements is

$$D_n = n!\left(1 - \frac{1}{1!} + \frac{1}{2!} - ... + (-1)^n\frac{1}{n!}\right)$$

Suppose all permutations were equally likely. What is the probability of a derangement?

$D_n/n! = 1 - \frac{1}{1!} + \frac{1}{2!} - ... + (-1)^n\frac{1}{n!}$

The infinte sum

$$1 - \frac{1}{1!} + \frac{1}{2!} - ... + (-1)^n\frac{1}{n!} + ...$$

gives $1/e \approx 0.368$.

$\frac{1}{e} - \frac{1}{(n+1)!} \le \frac{D_n}{n!} \le \frac{1}{e} + \frac{1}{(n+1)!}$.

The probability of at least one object being fixed is approximately $1 - 1/e \approx 0.632$.

Recurrence relations

Example: Tower of Hanoi.
Move n disks from peg 1 to peg 2.
H_n - number of moves used for n disks.
Solution: Move $n-1$ pegs to peg 3, using H_{n-1} moves. Move the bottom disk to peg 2. Move $n-1$ pegs from peg 3 to peg 2, using H_{n-1} moves.
Initial condition: $H_1 = 1$.
Recurrence relation: $H_n = 2H_{n-1} + 1$.

Example: Count comparisons in *Mergesort*.
Assume $n = 2^k$.
Initial condition: $M_1 = 0$.
Recurrence relation: $M_n \le 2M_{n/2} + (n-1)$.

Example: Fibonacci numbers.
Initial conditions: $f_0 = 0$, $f_1 = 1$.
Recurrence relation: $f_n = f_{n-1} + f_{n-2}$.

Solving recurrence relations
— iterative approach

Tower of Hanoi: $H_1 = 1$. $H_n = 2H_{n-1} + 1$.

$$\begin{aligned} H_n &= 2H_{n-1} + 1 \\ &= 2(2H_{n-2} + 1) + 1 \\ &= 2^2H_{n-2} + 2 + 1 \\ &= 2^2(2H_{n-3} + 1) + 2 + 1 \\ &= 2^3H_{n-3} + 2^2 + 2 + 1 \\ &. \\ &. \\ &. \\ &= 2^{n-1}H_1 + 2^{n-2} + ... + 2 + 1 \\ &= 2^{n-1} + 2^{n-2} + ... + 2 + 1 \\ &= 2^n - 1 \end{aligned}$$

Prove this correct using induction!

Solving recurrence relations
— some forms

Def. *Linear homogeneous recurrence relation of degree k with constant coefficients* —

$$a_n = c_1 a_{n-1} + c_2 a_{n-2} + \dots + c_k a_{n-k}$$

$c_i \in \mathbb{R}$ for $1 \leq i \leq k$, $c_k \neq 0$.

Examples: Fibonacci numbers.
$f_n = f_{n-1} + f_{n-2}$ — degree 2.

Number of strings with n binary digits.
$a_n = 2a_{n-1}$ — degree 1.

Examples which are not:
$H_n = 2H_{n-1} + 1$ — not homogeneous
$a_n = a_{n-1} + a_{n-2}^2$ — not linear
$a_n = na_{n-1} + a_{n-2}$ — nonconstant coefficient

Idea: Look for solutions $a_n = r^n$, $r \in \mathbb{R}$.

Def. The *characteristic equation* of

$$a_n = c_1 a_{n-1} + c_2 a_{n-2} + \dots + c_k a_{n-k}$$

is

$$r^k - c_1 r^{k-1} - c_2 r^{k-2} - \dots - c_{k-1} r - c_k = 0.$$

The solutions of this are the *characteristic roots*.

Degree 2

Thm. 1. Suppose $r^2 - c_1 r - c_2 = 0$ has roots r_1, r_2, $r_1 \neq r_2$. Then, $\{a_n\}$ is a solution to $a_n = c_1 a_{n-1} + c_2 a_{n-2}$ iff there are constants α_1, α_2 such that $a_n = \alpha_1 r_1^n + \alpha_2 r_2^n$, $\forall n$.

Example: Fibonacci numbers.
Initial conditions: $f_0 = 0$, $f_1 = 1$.
Recurrence relation: $f_n = f_{n-1} + f_{n-2}$.
Characteristic equation: $r^2 - r - 1 = 0$.
Characteristic roots: $r_1 = (1 + \sqrt{5})/2$,
$r_2 = (1 - \sqrt{5})/2$.
Theorem implies: $f_n = \alpha_1 \left(\frac{1+\sqrt{5}}{2}\right)^n + \alpha_2 \left(\frac{1-\sqrt{5}}{2}\right)^n$.

Initial conditions:
$f_0 = 0 = \alpha_1 + \alpha_2$.
$f_1 = 1 = \alpha_1 \left(\frac{1+\sqrt{5}}{2}\right) + \alpha_2 \left(\frac{1-\sqrt{5}}{2}\right)$

Solving for α_1, α_2:
$\alpha_2 = -\alpha_1$.
$1 = \alpha_1 \left(\frac{1+\sqrt{5}}{2}\right) - \alpha_1 \left(\frac{1-\sqrt{5}}{2}\right)$
$2 = 2\alpha_1\sqrt{5}$
$\alpha_1 = 1/\sqrt{5}$. $\alpha_2 = -1/\sqrt{5}$.

Solution: $f_n = \frac{1}{\sqrt{5}} \left(\frac{1+\sqrt{5}}{2}\right)^n - \frac{1}{\sqrt{5}} \left(\frac{1-\sqrt{5}}{2}\right)^n$.

Recall the following:
Thm. For $n \geq 3$, $f_n > \alpha^{n-2}$,
where $\alpha = (1 + \sqrt{5})/2$ (the Golden Ratio).

The previous argument gives:

$f_n = \frac{1}{\sqrt{5}} \left(\frac{1+\sqrt{5}}{2}\right)^n - \frac{1}{\sqrt{5}} \left(\frac{1-\sqrt{5}}{2}\right)^n$.

$\left(\frac{1-\sqrt{5}}{2}\right) \approx -0.618$.

So this term vanishes quickly.

Proof of Theorem 1:

Thm. 2. Suppose $r^2 - c_1 r - c_2 = 0$ has only one root r_0. Then, $\{a_n\}$ is a solution to $a_n = c_1 a_{n-1} + c_2 a_{n-2}$ iff there are constants α_1, α_2 such that $a_n = \alpha_1 r_0^n + \alpha_2 n r_0^n$, $\forall n$.

Example: Initial conditions: $a_0 = 0$, $a_1 = 2$.
Recurrence relation: $a_n = 4a_{n-1} - 4a_{n-2}$.
Characteristic equation: $r^2 - 4r + 4 = 0 = (r-2)(r-2)$.
Characteristic root: $r_0 = 2$.
Theorem implies: $a_n = \alpha_1 2^n + \alpha_2 n 2^n$.
Initial conditions:
$a_0 = 0 = \alpha_1$.
$a_1 = 2 = \alpha_1(2) + \alpha_2(1)(2) = 2\alpha_2$.
Solution: $a_n = n2^n$. Sequence: 0,2,8,24,64,...

Degree $n \geq 2$

Thm. Suppose $r^n - c_1 r^{n-1} - c_2 r^{n-2} - ... - c_{k-1} r - c_k$ has roots $r_1, r_2, ..., r_s$, where root r_i has multiplicity m_i. Then, $\{a_n\}$ is a solution to $a_n = c_1 a_{n-1} + c_2 a_{n-2} + ... + c_k a_{n-k}$ iff there are constants α_{i_j} such that $a_n = \sum_{i=1}^{s} \sum_{j=1}^{m_i} \alpha_{i_j} n^{j-1} r_i^n$, $\forall n$.

Generating functions

The *generating function* for the sequence $a_0, a_1, a_2, ..., a_k, ...$ is

$$G(x) = a_0 + a_1 x + a_2 x^2 + ... + a_k x^k + ... = \sum_{k=0}^{\infty} a_k x^k$$

Example: The generating function for $1, 3, 3, 1$ is $G(x) = 1 + 3x + 3x^2 + x^3 = (1+x)^3$.

Example: The generating function for $1, 1, 1, 1, 1$ is $1 + x + x^2 + x^3 + x^4 = (1 - x^5)/(1 - x)$.

Example: The generating function for $1, 1, 1, ...$ is $1 + x + x^2 + ... + x^k + ... = 1/(1-x)$ for $|x| < 1$.

Example: The generating function for $1, a, a^2, ...$ is $1 + ax + a^2x^2 + ... + a^k x^k + ... = 1/(1 - ax)$ for $|ax| < 1$.

Example: More generally, $1/(1 - ax)^{m+1} = \sum_{k=0}^{\infty} \binom{m+k}{m} a^k x^k$.

Facts. Let $f(x) = \sum_{k=0}^{\infty} a_k x^k$, $g(x) = \sum_{k=0}^{\infty} b_k x^k$.

- $f(x) + g(x) = \sum_{k=0}^{\infty} (a_k + b_k) x^k$
- $f(x)g(x) = \sum_{k=0}^{\infty} (\sum_{j=0}^{k} a_j b_{k-j}) x^k$

Example: How many solutions are there to $n_1 + n_2 + n_3 + n_4 = 7$, if you must have $n_1, n_2, n_3 \leq 3$ and $n_4 \leq 4$.
Answer: The coefficient of x^7 in $(1+x+x^2+x^3)\cdot(1+x+x^2+x^3)\cdot(1+x+x^2+x^3)\cdot(1 + x + x^2 + x^3 + x^4)$ — 50. Choose term with exponent corresponding to value of n_i. The exponents then must sum to 7.

Example: Initial condition: $a_0 = 1$.
Recurrence relation: $a_n = 5a_{n-1}$.
Generating function: $G(x) = \sum_{k=0}^{\infty} a_k x^k$.
Find a nice expression for $G(x)$:

$$\begin{aligned} & G(x) - 5xG(x) \\ & = \sum_{k=0}^{\infty} a_k x^k - 5 \sum_{k=1}^{\infty} a_{k-1} x^k \\ & = a_0 + \sum_{k=1}^{\infty} (a_k - 5a_{k-1}) x^k \\ & = 1 \end{aligned}$$

Thus, $(1 - 5x)G(x) = 1$.

$$\begin{aligned} G(x) &= \frac{1}{1-5x} \\ &= \sum_{n=0}^{\infty} 5^n x^n \end{aligned}$$

So, $a_n = 5^n$ for $n \geq 0$.

Example: Fibonacci numbers.
Initial conditions: $f_0 = 0$, $f_1 = 1$.
Recurrence relation: $f_n = f_{n-1} + f_{n-2}$.
Generating function: $G(x) = \sum_{k=0}^{\infty} f_k x^k$.
Find a nice expression for $G(x)$:

$$\begin{aligned}
&G(x) - xG(x) - x^2G(x) \\
&\quad = \sum_{k=0}^{\infty} f_k x^k - \sum_{k=1}^{\infty} f_{k-1}x^k - \sum_{k=2}^{\infty} f_{k-2}x^k \\
&\quad = f_0 + f_1 x - f_0 x + \sum_{k=2}^{\infty} (f_k - f_{k-1} - f_{k-2})x^k \\
&\quad = 0 + 1x - 0x \\
&\quad = x
\end{aligned}$$

Thus, $(1 - x - x^2)G(x) = x$.

$$\begin{aligned}
\text{Let} \quad r_1 &= \left(\tfrac{1+\sqrt{5}}{2}\right) \\
r_2 &= \left(\tfrac{1-\sqrt{5}}{2}\right)
\end{aligned}$$

$$\begin{aligned}
G(x) &= \tfrac{x}{1-x-x^2} \\
&= \tfrac{x}{(1-r_1x)(1-r_2x)} \\
&= \tfrac{1/\sqrt{5}}{(1-r_1x)} + \tfrac{-1/\sqrt{5}}{(1-r_2x)} \\
&= \tfrac{1}{\sqrt{5}} \sum_{n=0}^{\infty} r_1^n x^n - \frac{1}{\sqrt{5}} \sum_{n=0}^{\infty} r_2^n x^n. \\
&= \sum_{n=0}^{\infty} (\frac{1}{\sqrt{5}} r_1^n - \frac{1}{\sqrt{5}} r_2^n) x^n.
\end{aligned}$$

So $a_n = \frac{1}{\sqrt{5}} r_1^n - \frac{1}{\sqrt{5}} r_2^n$ for $n \geq 0$.

Nonhomogeneous linear recurrences

Tower of Hanoi: $H_1 = 1$. $H_n = 2H_{n-1} + 1$.
Note: Can define $H_0 = 0$.
Generating function: $G(x) = \sum_{k=0}^{\infty} H_k x^k$.
Find a nice expression for $G(x)$.

$$\begin{aligned}
&(1 - 2x)(1 - x)G(x) \\
&\quad = G(x) - 3xG(x) + 2x^2G(x) \\
&\quad = \sum_{k=0}^{\infty} H_k x^k - 3 \sum_{k=1}^{\infty} H_{k-1}x^k \\
&\qquad + 2 \sum_{k=2}^{\infty} H_{k-2}x^k \\
&\quad = H_0 + H_1 x - 3H_0 x \\
&\qquad + \sum_{k=2}^{\infty} (H_k - 3H_{k-1} + 2H_{k-2})x^k \\
&\quad = 0 + x - 0x \\
&\qquad + \sum_{k=2}^{\infty} ((2H_{k-1} + 1) - 3H_{k-1} + 2H_{k-2})x^k \\
&\quad = x + \sum_{k=2}^{\infty} (1 - (2H_{k-2} + 1) + 2H_{k-2})x^k \\
&\quad = x
\end{aligned}$$

$$\begin{aligned}
G(x) &= \tfrac{x}{(1-2x)(1-x)} \\
&= \tfrac{1}{(1-2x)} + \tfrac{-1}{(1-x)} \\
&= \sum_{n=0}^{\infty} 2^n x^n - \sum_{n=0}^{\infty} 1^n x^n. \\
&= \sum_{n=0}^{\infty} (2^n - 1)x^n.
\end{aligned}$$

More generally, if the recurrence relation is

$$a_n = c_1 a_{n-1} + c_2 a_{n-2} + \ldots + c_k a_{n-k} + b^n p(n),$$

where b is a constant and $p(n)$ is a polynomial of degree d, then the characteristic equation is

$$(r^k - c_1 r^{k-1} - c_2 r^{k-2} - \ldots - c_{k-1} r - c_k)(r - b)^{d+1} = 0.$$

Example: Initial conditions: $a_0 = 0$, $a_1 = 2$.
Recurrence relation: $a_n = 2a_{n-1} + 2^n$.
Characteristic equation: $(r - 2)(r - 2) = 0$.
Characteristic root: $r_0 = 2$.
Solution: $a_n = n2^n$.

Divide-and-conquer

Divide-and-conquer recurrence relations often have the form

$$T(n) = aT(n/b) + g(n).$$

To solve them, do a change of variable.
Replace n by b^k.
Use previous techniques.

Example: Mergesort. Assume $n = 2^k$.
Initial condition: $M_1 = 0$.
Recurrence relation: $M_n \le 2M_{n/2} + (n-1)$.

Thm. Let T be an increasing function s.t.

$$T(n) = aT(n/b) + cn^d$$

if $n = b^k$, $a \ge 1$, $b \ge 1$ an integer, $c, d \ge 0$. Then

$$T(n) = \begin{cases} O(n^d) & \text{if } a < b^d \\ O(n^d \log n) & \text{if } a = b^d \\ O(n^{\log_b a}) & \text{if } a > b^d \end{cases}$$

In the special case where $d = 0$:

$$T(n) = aT(n/b) + c$$

$$T(n) = \begin{cases} O(\log n) & \text{if } a = 1 \\ O(n^{\log_b a}) & \text{if } a > 1 \end{cases}$$

Relations

Def. A *binary relation* from a set A to a set B is a subset of $A \times B$.

Representations:

- List
- Bipartite graph
- Table

Example:
A={ Peter, Marianne, Niels} and
B={ vanilla, strawberry, chocolate, mint}.
R={ (Peter,vannilla), (Peter,chocolate), (Marianne,vanilla), (Marianne,strawberry), (Niels,vanilla), (Niels,chocolate), (Niels,mint)}

Example: Functions.
$f : A \to B$ defines R.
$(a, b) \in R$ iff $f(a) = b$.

Example: Relations on a set. $A = B$.
Let $A = B = \mathbb{Z}$. Relation — $\le$.
$(13, 25) \in R$.
$(25, 25) \in R$.
$(25, 13) \notin R$.

Possible properties of relations:

- *Reflexive:* $\forall a \in A,\ (a,a) \in R$.
  $\leq$ is reflexive.
  $<$ is not reflexive.

- *Symmetric:* $\forall a, b \in A$,
  $(a,b) \in R \Rightarrow (b,a) \in R$.

- *Antisymmetric:* $\forall a, b \in A$,
  $(a,b) \in R$ and $(b,a) \in R \Rightarrow a = b$.
  $\leq$ and $<$ are antisymmetric.

- *Transitive:* $\forall a, b, c \in A$,
  $(a,b) \in R,\ (b,c) \in R\ \Rightarrow\ (a,c) \in R$.
  $\leq$ and $<$ are transitive.

Operations on relations:

Def. The *composite* of relations R and S, $R \circ S$, where R is from A to B and S is from B to C is

$$\{(a,c) \mid \exists b \in B \text{ s.t. } (a,b) \in R \text{ and } (b,c) \in S\}$$

Example:
$R = \{(1,1),(1,3),(1,5),(2,4),(3,5)\}$
$S = \{((1,2),(1,3),(2,3),(2,4),(3,4),(3,5),(4,5)\}$
$R \circ S = \{(1,2),(1,3),(1,4),(1,5),(2,5)\}$

Def. R a relation on a set A.
$R^1 = R$
$R^{n+1} = R^n \circ R$.

Thm. R is transitive iff $R^n \subseteq R\ \ \forall n \geq 1$.

Pf. ($\Rightarrow$) Suppose R is transitive.
Prove by induction on n that $R^n \subseteq R\ \ \forall n \geq 1$.
Basis step: $n = 1$. $R^1 = R \subseteq R$. $\surd$
Inductive step: Suppose that $R^n \subseteq R$.
$R^{n+1} = R^n \circ R =$
$\{(a,c) \mid \exists b \in A \text{ s.t. } (a,b) \in R^n\ \wedge\ (b,c) \in R\}$.
Since $R^n \subseteq R$, $R^{n+1} \subseteq$
$\{(a,c) \mid \exists b \in A \text{ s.t. } (a,b) \in R\ \wedge\ (b,c) \in R\}$.
But since R is transitive,
if $\exists b \in A$ s.t. $(a,b),(b,c) \in R$, then $(a,c) \in R$.
Therefore, $R^{n+1} \subseteq R$.
By induction, $R^n \subseteq R\ \ \forall n \geq 1$. $\surd$

($\Leftarrow$) Suppose $R^n \subseteq R\ \ \forall n \geq 1$.
Suppose $(a,b),(b,c) \in R$.
Then $(a,c) \in R \circ R = R^2$.
Since $R^2 \subseteq R$, $(a,c) \in R$, so R is transitive.
$\square$

Def. An *n-ary relation* on the sets $A_1, A_2, ..., A_n$ is a subset of $A_1 \times A_2 \times ... \times A_n$. $A_1, A_2, ..., A_n$ are the *domains* and n is the *degree*.

Example: Database of patients.
A_1 = CPR number
A_2 = Last name
A_3 = First names
A_4 = Date admitted to hospital
A_5 = primary diagnosis

(110393-1750, Larsen, Marianne Boyar, 110393, newborn)
(190944-1951, Hansen, Hans Jørgen, 110393, lung cancer)
etc.

These are 2 *records*.
Each record has 5 *fields*.

A relation can be represented by a *table*.

Operations in databases (and elsewhere)

Def. The *projection* $P_{i_1 i_2 \ldots i_m}$ maps $(a_1, a_2, ..., a_n)$ to $(a_{i_1}, a_{i_2}, ..., a_{i_m})$.
$P_{13}(1, 5, 25, 625) = (1, 25)$

Def.
R — relation of degree m
S — relation of degree n
join$_p(R, S)$, $p \leq \min(m, n)$
— relation of degree $m + n - p$

$$\{(a_1, a_2, ..., a_{m-p}, c_1, c_2, ..., c_p, b_1, b_2, ..., b_{n-p}) \mid (a_1, a_2, ..., a_{m-p}, c_1, c_2, ..., c_p) \in R \text{ and } (c_1, c_2, ..., c_p, b_1, b_2, ..., b_{n-p}) \in S\}.$$

Example:
$R \subseteq$ { (CPR #, Last Name, First, Date admitted, Diagnosis)}
$S \subseteq$ { (Date admitted, Diagnosis, Treatment)}

join$_2(R, S)$
adds treatment onto everyone in R.

If > 1 treatment for some date and diagnosis — each person admitted that day with that diagnosis will have one record inserted for every treatment.

Representing binary relations by matrices

R – a relation from $A = \{a_1, a_2, ..., a_m\}$ to $B = \{b_1, b_2, ..., b_n\}$.
Let $M_R = [m_{ij}]$ be an $m \times n$ matrix, where

$$M_{ij} = \begin{cases} 1 & \text{if } (a_i, b_j) \in R \\ 0 & \text{if } (a_i, b_j) \notin R \end{cases}$$

Example:
A={ Peter, Marianne, Niels} and
B={ vanilla, strawberry, chocolate, mint}.
R={ (Peter,vannilla), (Peter,chocolate), (Marianne,vanilla), (Marianne,strawberry), (Niels,vanilla), (Niels,chocolate), (Niels,mint)}

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

If R is a relation on a set A, M_R is square.
R is reflexive iff $m_{ii} = 1 \quad \forall i$.
R is symmetric iff $M_R = M_R^t$
R is antisymmetric iff $m_{ij} = 1$
$\Rightarrow m_{ji} = 0$ for $i \neq j$.

Union: $R \cup S$
Def. $A = [a_{ij}]$, $B = [b_{ij}]$, $m \times n$, 0-1 matrices.
The *join* of A and B is $A \vee B = [a_{ij} \vee b_{ij}]$.
$M_{R \cup S} = M_R \vee M_S$.

Intersection: $R \cap S$
Def. $A = [a_{ij}]$, $B = [b_{ij}]$, $m \times n$, 0-1 matrices.
The *meet* of A and B is $A \wedge B = [a_{ij} \wedge b_{ij}]$.
$M_{R \cap S} = M_R \wedge M_S$.

Composite: $R \circ S$
Def. $A = [a_{ij}]$, $B = [b_{ij}]$, $m \times k$, 0-1 matrix.
$B = [b_{ij}]$, $k \times n$, 0-1 matrix.
The *Boolean product* of A and B is $A \odot B = [c_{ij} = (a_{i1} \wedge b_{1j}) \vee (a_{i2} \wedge b_{2j}) \vee ... \vee (a_{ik} \wedge b_{kj})]$.
$M_{R \circ S} = M_R \odot M_S$.

Powers of relations:

Def. A — $m \times m$, 0-1 matrix.
The r^{th} Boolean power of A is

$$\begin{aligned} A^{[0]} &= I_n \\ A^{[n+1]} &= A^{[n]} \odot A \end{aligned}$$

$M_{R^n} = M_R^{[n]}$.

Example:
$R = \{(1,1),(1,3),(1,5),(2,4),(3,4)\}$

1	0	1	0	1
0	0	0	1	0
0	0	0	1	0
0	0	0	0	0
0	0	0	0	0

Directed graphs of relations

Reflexive — loops at every vertex

Symmetric — all edges come in pairs one going in each direction

Antisymmetric — only 1 edge between any 2 vertices

Transitive — if ∃ a path from u to v, then ∃ an edge from u to v

Closure

P – property –transitivity
R – relation
S – *closure* of R w.r.t. P
The following must hold:

- $R \subseteq S$
- S has property P
- If $R \subseteq T$, T has P, then $S \subseteq T$

Example: Train network.
Cities = $\{A, B, C, D, E\}$
∃ train from A to B and B to A
∃ train from A to C and C to A
∃ train from B to E and E to B
∃ train from C to E and E to C
∃ train from D to E and E to D

Thm. R a relation. G – digraph for R.
∃ a path of length n from a to b in G
iff $(a,b) \in R^n$.
Pf. (by ind on n)
Basis step: $n = 1$.
$(a,b) \in R$ iff path of length 1. $\surd$
Inductive step: Assume true for some $n \geq 1$:
$\forall a, b \in V(G)$ ∃ a path of length n from a to b
iff $(a,b) \in R^n$.

∃ a path of length $n+1$ from a to b

iff ∃ c s.t. ∃ a path of length n from a to c and an edge (c,b)

iff ∃ c s.t. $(a,c) \in R^n$ and $(c,b) \in R$ (IH)

iff $(a,b) \in R^{n+1}$. $\surd$

By induction, ∃ a path of length n from a to b in G iff $(a,b) \in R^n$. $\square$

Warshall's Algorithm

$W_0 = M_R$
$W_k = [w_{ij}^{(k)}]$ where

$$w_{ij}^{(k)} = \begin{cases} 1 & \text{if } \exists\, P = \{v_i, v_{s_1}, v_{s_2}, ..., v_{s_t}, v_j\} \\ & \text{s.t. } s_r \in \{1, ..., k\}\ \forall r \\ 0 & \text{otherwise} \end{cases}$$

procedure Warshall (M_R: $n \times n$, 0-1 matrix)
$W \leftarrow M_R$
for $k \leftarrow 1$ **to** n
  for $i \leftarrow 1$ **to** n
    for $j \leftarrow 1$ **to** n
      $w_{ij} \leftarrow w_{ij} \vee (w_{ik} \wedge w_{kj})$
return $(W = M_{R^*})$

$2n^3$ bit operations.

Equivalence relation —
reflexive, symmetric, transitive

Example: $R = \{(a, b) \mid a \equiv b \pmod{m}\}$

Equivalence class of x —
all elements related to x
$[a]_R = \{s \mid (a, s) \in R\}$
$b \in [a]_R$ is a *representative* of the class.

An equivalence relation *partitions* a set.

Thm. R — equivalence relation on A
The following are equivalent:

1. aRb
2. $[a] = [b]$
3. $[a] \cap [b] \neq \emptyset$

Pf. ($1 \Rightarrow 2$) Suppose $x \in [a]$.
Then, $(a, x) \in R$.
R reflexive and $(a, b) \in R \Rightarrow\ (b, a) \in R$.
R transitive, so $(b, a) \in R$, $(a, x) \in R$
$\Rightarrow\ (b, x) \in R$. Thus, $x \in [b]$, so $[a] \subseteq [b]$.
Similarly $[b] \subseteq [a]$, so $[a] = [b]$.

($2 \Rightarrow 3$) $a \in [a]$.
$[a] = [b]\ \Rightarrow\ a \in [b]\ \Rightarrow\ a \in [a] \cap [b]$.
Thus, $[a] \cap [b] \neq \emptyset$.

($3 \Rightarrow 1$) Suppose $x \in [a] \cap [b]$.
$(a, x) \in R\ \wedge\ (b, x) \in R$.
R symmetric, so $(x, b) \in R$.
R transitive, so $(a, b) \in R$. □

Def. R a relation on S.
R is a *partial order* if it is
reflexive, antisymmetric, transitive.
(S, R) — *partially ordered set, poset*.

Example: Containment of line segments on the real line.

a, b are *comparable* in (S, R) if $(a, b) \in R$ or $(b, a) \in R$. Otherwise they are *incomparable*.

Def. If every pair of elements in S is comparable, S is a *totally ordered set* (*linearly ordered set, chain*), and R is a *total order* (*linear order*).

$\mathbb{Z}$ and $\mathbb{R}$ are totally ordered under $\leq$.

Def. A totally ordered set is *well-ordered* if every nonempty set has a least element.

$\mathbb{Z}$ is well ordered.

Hasse diagrams —

- start with the digraph for the partial order
- remove loops
- remove edges implied by transitivity
- remove arrows (they all go upwards)

An element $a \in (S, R)$ is *maximal* if $(a, b) \in R \;\Rightarrow\; a = b$. It is *minimal* if $(b, a) \in R \;\Rightarrow\; a = b$.

An element $a \in (S, R)$ is the *greatest element* if $\forall\ b \in S$, $(b, a) \in R$. It is the *least element* if $\forall\ b \in S$, $(a, b) \in R$.

An element $a \in S$ is an *upper bound* of $A \subseteq S$ if $\forall u \in A$, $(u, a) \in R$. The *least upper bound* is less than or equal to any other upper bound of A.

An element $a \in S$ is an *lower bound* of $A \subseteq S$ if $\forall u \in A$, $(a, u) \in R$. The *greatest lower bound* is greater than or equal to any other lower bound of A.

Lattice — a poset in which every pair of elements has both a least upper bound and a greatest lower bound.

Example: $(P(S), \subseteq)$, where S is a set.
The least upper bound of A and B is $A \cup B$.
The greatest lower bound of A and B is $A \cap B$.

Operations in databases (and elsewhere)

Def. The *projection* $P_{i_1 i_2 \ldots i_m}$ maps $(a_1, a_2, \ldots, a_n)$ to $(a_{i_1}, a_{i_2}, \ldots, a_{i_m})$.
$P_{13}(1, 5, 25, 625) = (1, 25)$

Def.
R — relation of degree m
S — relation of degree n
join$_p(R, S)$, $p \leq \min(m, n)$
— relation of degree $m + n - p$

$$\{(a_1, a_2, \ldots, a_{m-p}, c_1, c_2, \ldots, c_p, b_1, b_2, \ldots, b_{n-p}) \mid (a_1, a_2, \ldots, a_{m-p}, c_1, c_2, \ldots, c_p) \in R \text{ and } (c_1, c_2, \ldots, c_p, b_1, b_2, \ldots, b_{n-p}) \in S\}.$$

Example:
$R \subseteq$ { (CPR #, Last Name, First, Date admitted, Diagnosis)}
$S \subseteq$ { (Date admitted, Diagnosis, Treatment)}

join$_2(R, S)$
adds treatment onto everyone in R.

If > 1 treatment for some date and diagnosis — each person admitted that day with that diagnosis will have one record inserted for every treatment.

Representing binary relations by matrices

R – a relation from $A = \{a_1, a_2, ..., a_m\}$ to $B = \{b_1, b_2, ..., b_n\}$.

Let $M_R = [m_{ij}]$ be an $m \times n$ matrix, where

$$M_{ij} = \begin{cases} 1 & \text{if } (a_i, b_j) \in R \\ 0 & \text{if } (a_i, b_j) \notin R \end{cases}$$

Example:

A={ Peter, Marianne, Niels} and
B={ vanilla, strawberry, chocolate, mint}.
R={ (Peter,vannilla), (Peter,chocolate), (Marianne,vanilla), (Marianne,strawberry), (Niels,vanilla), (Niels,chocolate), (Niels,mint)}

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

If R is a relation on a set A, M_R is square.
R is reflexive iff $m_{ii} = 1 \quad \forall i$.
R is symmetric iff $M_R = M_R^t$
R is antisymmetric iff $m_{ij} = 1$
$\Rightarrow m_{ji} = 0$ for $i \neq j$.

Union: $R \cup S$
Def. $A = [a_{ij}]$, $B = [b_{ij}]$, $m \times n$, 0-1 matrices.
The *join* of A and B is $A \vee B = [a_{ij} \vee b_{ij}]$.
$M_{R \cup S} = M_R \vee M_S$.

Intersection: $R \cap S$
Def. $A = [a_{ij}]$, $B = [b_{ij}]$, $m \times n$, 0-1 matrices.
The *meet* of A and B is $A \wedge B = [a_{ij} \wedge b_{ij}]$.
$M_{R \cap S} = M_R \wedge M_S$.

Composite: $R \circ S$
Def. $A = [a_{ij}]$, $B = [b_{ij}]$, $m \times k$, 0-1 matrix.
$B = [b_{ij}]$, $k \times n$, 0-1 matrix.
The *Boolean product* of A and B is $A \odot B = [c_{ij} = (a_{i1} \wedge b_{1j}) \vee (a_{i2} \wedge b_{2j}) \vee ... \vee (a_{ik} \wedge b_{kj})]$.
$M_{R \circ S} = M_R \odot M_S$.

Powers of relations:
Def. A — $m \times m$, 0-1 matrix.
The r^{th} Boolean power of A is

$$\begin{aligned} A^{[0]} &= I_n \\ A^{[n+1]} &= A^{[n]} \odot A \end{aligned}$$

$M_{R^n} = M_R^{[n]}$.

Example:
$R = \{(1,1), (1,3), (1,5), (2,4), (3,4)\}$

1	0	1	0	1
0	0	0	1	0
0	0	0	1	0
0	0	0	0	0
0	0	0	0	0

Directed graphs of relations

Reflexive — loops at every vertex

Symmetric — all edges come in pairs one going in each direction

Antisymmetric — only 1 edge between any 2 vertices

Transitive — if $\exists$ a path from u to v, then $\exists$ an edge from u to v

Closure

P – property –transitivity
R – relation
S – *closure* of R w.r.t. P
The following must hold:

- $R \subseteq S$
- S has property P
- If $R \subseteq T$, T has P, then $S \subseteq T$

Example: Train network.
Cities $= \{A, B, C, D, E\}$
$\exists$ train from A to B and B to A
$\exists$ train from A to C and C to A
$\exists$ train from B to E and E to B
$\exists$ train from C to E and E to C
$\exists$ train from D to E and E to D

Thm. R a relation. G – digraph for R.
$\exists$ a path of length n from a to b in G
iff $(a,b) \in R^n$.
Pf. (by ind on n)
Basis step: $n = 1$.
$(a,b) \in R$ iff path of length 1. $\surd$
Inductive step: Assume true for some $n \geq 1$:
$\forall a, b \in V(G)$ $\exists$ a path of length n from a to b
iff $(a,b) \in R^n$.

$\exists$ a path of length $n+1$ from a to b

iff $\exists$ c s.t. $\exists$ a path of length n from a to c
and an edge (c,b)

iff $\exists$ c s.t. $(a,c) \in R^n$ and $(c,b) \in R$ (IH)

iff $(a,b) \in R^{n+1}$. $\surd$

By induction, $\exists$ a path of length n from a to b in G iff $(a,b) \in R^n$. $\square$

Warshall's Algorithm

$W_0 = M_R$
$W_k = [w_{ij}^{(k)}]$ where

$$w_{ij}^{(k)} = \begin{cases} 1 & \text{if } \exists\ P = \{v_i, v_{s_1}, v_{s_2}, ..., v_{s_t}, v_j\} \\ & \quad \text{s.t. } s_r \in \{1, ..., k\}\ \forall r \\ 0 & \text{otherwise} \end{cases}$$

procedure Warshall (M_R: $n \times n$, 0-1 matrix)
$W \leftarrow M_R$
for $k \leftarrow 1$ **to** n
    for $i \leftarrow 1$ **to** n
        for $j \leftarrow 1$ **to** n
            $w_{ij} \leftarrow w_{ij} \vee (w_{ik} \wedge w_{kj})$
return $(W = M_{R^*})$

$2n^3$ bit operations.

Equivalence relation —
reflexive, symmetric, transitive

Example: $R = \{(a,b) \mid a \equiv b \pmod{m}\}$

Equivalence class of x —
all elements related to x
$[a]_R = \{s \mid (a,s) \in R\}$
$b \in [a]_R$ is a *representative* of the class.

An equivalence relation *partitions* a set.

Thm. R — equivalence relation on A
The following are equivalent:

1. aRb
2. $[a] = [b]$
3. $[a] \cap [b] \neq \emptyset$

Pf. $(1 \Rightarrow 2)$ Suppose $x \in [a]$.
Then, $(a, x) \in R$.
R reflexive and $(a, b) \in R \Rightarrow (b, a) \in R$.
R transitive, so $(b, a) \in R$, $(a, x) \in R$
$\Rightarrow (b, x) \in R$. Thus, $x \in [b]$, so $[a] \subseteq [b]$.
Similarly $[b] \subseteq [a]$, so $[a] = [b]$.

$(2 \Rightarrow 3)$ $a \in [a]$.
$[a] = [b] \Rightarrow a \in [b] \Rightarrow a \in [a] \cap [b]$.
Thus, $[a] \cap [b] \neq \emptyset$.

$(3 \Rightarrow 1)$ Suppose $x \in [a] \cap [b]$.
$(a, x) \in R \wedge (b, x) \in R$.
R symmetric, so $(x, b) \in R$.
R transitive, so $(a, b) \in R$. □

Def. R a relation on S.
R is a *partial order* if it is
reflexive, antisymmetric, transitive.
(S, R) — *partially ordered set, poset.*

Example: Containment of line segments on the real line.

a, b are *comparable* in (S, R) if $(a, b) \in R$ or $(b, a) \in R$. Otherwise they are *incomparable*.

Def. If every pair of elements in S is comparable, S is a *totally ordered set* (*linearly ordered set, chain*), and R is a *total order* (*linear order*).

$\mathbb{Z}$ and $\mathbb{R}$ are totally ordered under $\leq$.

Def. A totally ordered set is *well-ordered* if every nonempty set has a least element.

$\mathbb{Z}$ is well ordered.

Hasse diagrams —

- start with the digraph for the partial order
- remove loops
- remove edges implied by transitivity
- remove arrows (they all go upwards)

An element $a \in (S, R)$ is *maximal* if $(a, b) \in R \Rightarrow a = b$. It is *minimal* if $(b, a) \in R \Rightarrow a = b$.

An element $a \in (S, R)$ is the *greatest element* if $\forall\ b \in S$, $(b, a) \in R$. It is the *least element* if $\forall\ b \in S$, $(a, b) \in R$.

An element $a \in S$ is an *upper bound* of $A \subseteq S$ if $\forall u \in A$, $(u, a) \in R$. The *least upper bound* is less than or equal to any other upper bound of A.

An element $a \in S$ is an *lower bound* of $A \subseteq S$ if $\forall u \in A$, $(a, u) \in R$. The *greatest lower bound* is greater than or equal to any other lower bound of A.

Lattice — a poset in which every pair of elements has both a least upper bound and a greatest lower bound.

Example: $(P(S), \subseteq)$, where S is a set.
The least upper bound of A and B is $A \cup B$.
The greatest lower bound of A and B is $A \cap B$.

Huffman Encoding

Text files in ASCII waste space.
ASCII uses 8 bits for every character.
Redundancy in all natural languages.

Use short bit strings for the most frequent n-tuples and longer strings for others.
One can represent English text using $\approx$ 1.5 bits/character on average.

Huffman codes are prefix codes.
Also called *prefix-free codes*.
No code representing a character is a prefix of another code.

Example: "0110" and "011010" could not both be codes in a prefix-free encoding;
the first is a prefix of the second.

Prefix-free encoding makes decoding easy.

Example: Suppose $f(a) = 01$, $f(b) = 0010$, $f(c) = 0011$, $f(d) = 10$, $f(e) = 11$.
010011111000110111 decodes as
01 0011 11 10 0011 01 11: *acedcae*.
Need just one pass through the string.

To do Huffman encoding, build a binary tree.
k possible characters.
Probability of character c_i is p_i.
$0 < p_i \leq 1 \ \forall\ i$, and $\sum_{i=1}^{k} p_i = 1$.
One leaf in the binary tree for each character.
Each node in the tree contains a key, a probability.

procedure Huffman($p_1, p_2, ..., p_k$)

Create 1 tree for each character.
Put the probabilities for the characters in roots.

Sort the roots according to their probabilities.

while $\exists > 1$ tree **do**
- Remove the 2 trees whose roots have the smallest probabilities, p_i and p_j.
- Make these two trees the left and right subtrees of a new root.
- Give the new root probability $p_i + p_j$.
- Insert the new tree, maintaining sorted order for roots.

end do

return pointer to root of remaining tree

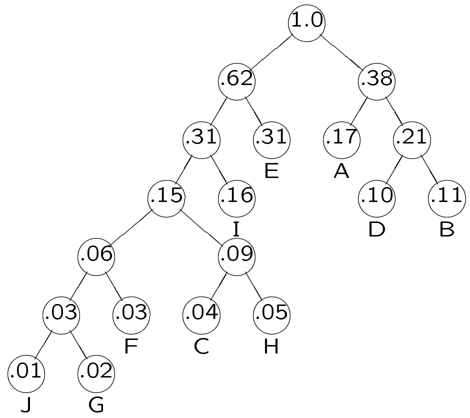
Each pass through the loop decreases the number of existing trees by one.
So there are $k - 1$ passes through the loop.

Get short paths from the root to the leaves with frequently occurring characters.

The path from the root to a leaf will define the encoding.

Edge going to a left child — 0
Edge going to a right child — 1

A	B	C	D	E	F	G	H	I	J
.17	.11	.04	.10	.31	.03	.02	.05	.16	.01



The encoding of the different characters is as follows:

A	B	C	D	E
10	111	00010	110	01
F	**G**	**H**	**I**	**J**
00001	000001	00011	001	000000

Codes are paths in the tree, so no code is a prefix of another.

Algebra

The definition of a group

Def. A *group* is a set G closed under a binary operation $\odot$ such that

- *Associative Law:*
  $\forall x, y, z \in G,\ x \odot (y \odot z) = (x \odot y) \odot z.$
- *Identity:* $\exists e \in G$, the *identity*:
  $\forall x \in G,\ e \odot x = x \odot e = x.$
- *Inverse:*
  $\forall x \in G,\ \exists y \in G$ s.t. $x \odot y = y \odot x = e.$
  $y = x^{-1}$ is the *inverse* of x.

Examples: The integers $\mathbb{Z}$, the reals $\mathbb{R}$, and the rationals $\mathbb{Q}$ are groups under addition.

Example: $\mathbb{R} - \{0\}$ under multiplication.

Def. For finite groups, the number of elements in a group G, written $|G|$, is the *order* of the group.

Example: $\mathbb{Z}_n$, the integers modulo n, under addition. The order of $\mathbb{Z}_n$ is n.

Example: $\mathbb{Z}_n^*$, the positive integers less than n which are relatively prime to n, under multiplication. The order of $\mathbb{Z}_n^*$ is $\phi(n)$, where ϕ is the Euler ϕ-function.

The above examples are all *abelian groups* — the operation is commutative: $x \odot y = y \odot x$ for all x, y in the group.
Not all groups are abelian.

Example: S_n, the symmetric group on n letters, is the set of permutations of the set $\{1, 2, ..., n\}$.
S_n is not an abelian group.

Any permutation can be written as a product of cycles.

A *transposition* is a cycle of length 2, (ij). A permutation is *even* iff it can be expressed as a product of an even number of transpositions.

The *symmetric group* of a set X is Sym(X).

Subgroups

Def. Let G be a group, and $H \subseteq G$.
H is a *subgroup* of G if H itself is a group w.r.t. the operation in G. The *order* of a subgroup is its cardinality.

Suppose G is a group and $H \subseteq G$. Then H is a subgroup of G iff the following hold:

- $\forall x, y \in H$, $x \odot y \in H$.
  (H is closed under the group operation.)
- The identity is in H.
- $\forall x \in H$, $x^{-1} \in H$.

Example: Any group is a subgroup of itself.

Example: If e is the identity in G, $\{e\}$ is a subgroup of G.

Example: The even integers are a subgroup of the integers under addition.

Example: $\left\{ \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 1 & 2 & 3 & 4 & 5 \end{pmatrix}, \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 2 & 1 & 4 & 3 & 5 \end{pmatrix} \right\}$ is a subgroup of S_5.

Example: A_n, the set of all even permutations on n letters, is a subgroup of S_n. It is called the *alternating group* on n letters.

Example: The set $\{0, 5, 10\}$ is a subgroup of $\mathbb{Z}_{15}$ under addition.

Example: The set $\{1, 2, 4\}$ is a subgroup of $\mathbb{Z}_7^*$.

Thm. Suppose S is a nonempty collection of subgroups of a group G. Let $H' = \bigcap_{H \in S} H$. Then H' is a subgroup of G.

Generators

Def. Let H be a subset of a group G, and let S be the collection of all subgroups of G which contain H. Then, $<H> = \bigcap_{G' \in S} G'$ is the *subgroup generated by* H.

$H \subseteq G$.
$<H> = \{h_1 \odot h_2 \odot \ldots \odot h_n \mid h_i \text{ or } h_i^{-1} \in H \quad \forall i\}$.

Def. A group or subgroup is said to be *cyclic* if it is generated by a single element.
Such an element is a *generator* or *primitive element*.

Def. The *order* of an element of a group G is the order of the subgroup that element generates.

Thm. Suppose G is a group with identity e, and $g \in G$ has finite order m. Then m is the least positive integer such that $g^m = e$.

Example: The set $\{2\}$ generates the subgroup $\{1, 2, 4\}$ of $\mathbb{Z}_7^*$.
Thus, it is a cyclic subgroup.
The order of the element 2 is 3.
The set $\{3\}$ generates all of $\mathbb{Z}_7^*$, so it is a cyclic group.

Fact. $\mathbb{Z}_p^*$ is cyclic whenever p is prime.

Lagrange's Theorem

Def. Let $H < G$, $x \in G$. $Hx = \{h \odot x \mid h \in H\}$ is a *right coset* of H in G.

Lemma Let $H < G$. All right cosets of H contain $|H|$ elements.

The relation $R =$
$\{(a, b) \mid a$ and b are in the same right coset of $H\}$ is an equivalence relation.

Lemma Let H be a subgroup of G, $x, y \in G$. Then either $Hx = Hy$ or $Hx \cap Hy = \emptyset$.

These two lemmas tell us that the group G must be a disjoint union of right cosets of any subgroup H, all of which have the same size.

Thm. [Lagrange] If G is a finite group and $H < G$, then the order of H divides the order of G.

Corollary Let G be a finite group and $g \in G$. The order of the element g divides the order of the group G.

Corollary Suppose $|G| = n$ and $g \in G$. Then $g^n = e$, where e is the identity in G.
Pf. Let s be the order of the subgroup generated by g. By a previous theorem, $g^s = e$, where e is the identity. By Lagrange's Theorem, s divides n, so there is an integer c such that $n = sc$. Note that $g^n = g^{sc} = (g^s)^c = e^c = e$, so the corollary follows. □

Thm. [Fermat's Little Theorem] If p is prime and a is an integer not divisible by p, then $a^{p-1} \equiv 1 \pmod{p}$.

Using this same group theory, we can show that encryption with RSA, followed by decryption with RSA yields the original message, if the message is less than n and relatively prime to n.

$n = pq$ where p and q are large primes.
$|\mathbb{Z}_n^*| = \phi(n) = (p-1)(q-1)$.
If the message M is less than n and relatively prime to n, $M \in \mathbb{Z}_n^*$.
$ed \equiv 1 \pmod{(p-1)(q-1)}$,
so $\exists k \in \mathbb{Z}$ s.t. $ed = 1 + k(p-1)(q-1)$.

If $C \equiv M^e \pmod{n}$, then $C^d \equiv (M^e)^d \equiv M^{ed} \equiv M^{1+k(p-1)(q-1)} \equiv M \cdot M^{k(p-1)(q-1)} \equiv M \cdot (M^{\phi(n)})^k \equiv M \cdot 1^k \equiv M \pmod{n}$.

Rings and fields

Def. A *ring* is a set R closed under two binary operations $+$ and $\bullet$ s.t.

- R1. R is an abelian group with respect to the operation $+$.
- R2. The operation $\bullet$ is associative.
- R3. [Distributive Laws] $\forall x, y, z \in R$, the following hold:

$$\begin{aligned} x \bullet (y+z) &= x \bullet y + x \bullet z \\ (y+z) \bullet x &= y \bullet x + z \bullet x \end{aligned}$$

The first operation $+$ is called addition and the second operation $\bullet$ is called multiplication.

Example: $\{0\}$ is the *trivial ring*.
$0 + 0 = 0$ and $0 \bullet 0 = 0$.
A *nontrivial ring* is a ring with more than one element.

The identity element with respect to addition is called *zero*, and all other elements are called *nonzero elements*.

If the ring R has an identity element i with respect to multiplication, then for all $x \in G$, $i \bullet x = x \bullet i = x$, and R is said to be a *ring with identity*. This identity is denoted by 1.

The ring R is *commutative* if for all $x, y \in R$, $x \bullet y = y \bullet x$.

Examples: $\mathbb{Z}$ and $\mathbb{R}$ are both commutative rings with identity.

Def. A *field* is a nontrivial commutative ring with identity in which every nonzero element has a multiplicative inverse.

Examples: $\mathbb{R}$ is a field and $\mathbb{Q}$ is a field, but $\mathbb{Z}$ is not.

Example: $\mathbb{Z}_n$ is a field when n is prime. It is a ring when n is composite, but not a field.