Propositional Logic

Definitions

- Propositions are sentences that are true or false but not both.
- A contradiction/tautology is a compound proposition which takes the value false/true for all possible truth values of its variables.
- A compound proposition is satisfiable if there is an assignment of truth values to its variables that make it true. Otherwise it is unsatisfiable (contradiction).

Predicates

- A predicate is a sentence that contains finitely many variables, and which becomes a proposition if the variables are given specific values.
    - The domain of a variable in a predicate is the set of all possible values that may be assigned to it e.g. $\mathbb{Z}$.
    - The truth set of a predicate $P(x)$ is the set of all values in the domain that, when assigned to x, make $P(x)$ a true proposition.

Logical Operators/Quantifiers

Operators			**Equivalent to**
Negation	$\neg$	Opposite signs	
Conjunction	$\wedge$	T F F F	
Disjunction	$\vee$	T T T F	
Exclusive Or	$\oplus$	T F F T	
Logical Equivalence	$\equiv$	Same truth table	
Conditional	$\rightarrow$	T F T T	$\neg p \vee q$, $\neg q \rightarrow \neg p$ (contrapositive)
Biconditional	$\iff$	T F F T	$(\neg p \vee q) \wedge (\neg q \vee p)$
Universal Quantifier	$\forall$		$\neg(\forall x \in D, Q(x)) \equiv \exists x \in D : \neg Q(x)$
Existential Quantifier	$\exists$	: "such that"	vice versa ^

Logical Equivalences

Laws		
Commutative Laws	$p \wedge q \equiv q \wedge p$	$p \vee q \equiv q \vee p$
Associative Laws	$(p \wedge q) \wedge r \equiv p \wedge (q \wedge r)$	$(p \vee q) \vee r \equiv p \vee (q \vee r)$
Distributive Laws	$p \wedge (q \vee r) \equiv (p \wedge q) \vee (p \wedge r)$	$p \vee (q \wedge r) \equiv (p \vee q) \wedge (p \vee r)$
Identity Laws	$p \wedge (tautology) \equiv p$	$p \vee (contradiction) \equiv p$
Universal Bound Laws	$p \vee (tautology) \equiv (tautology)$	$p \wedge (contradiction) \equiv (contradiction)$
Negation Laws	$p \vee \neg p \equiv (tautology)$	$p \wedge \neg p \equiv (contradiction)$
Double Negative Law	$\neg(\neg p) \equiv p$	
Idempotent Laws	$p \wedge p \equiv p$	$p \vee p \equiv p$
De Morgan's Laws	$\neg(p \wedge q) \equiv \neg p \vee \neg q$	$\neg(p \vee q) \equiv \neg p \wedge \neg q$
Absorption Laws	$p \vee (p \wedge q) \equiv p$	$p \wedge (p \vee q) \equiv p$
Negations	¬(tautology) is a contradiction	¬(contradiction) is a tautology

Constructions

Constructions	**of** $p \rightarrow q$
Contrapositive (is equivalent)	$\neg q \rightarrow \neg p$
Converse	$q \rightarrow p$
Inverse	$\neg p \rightarrow \neg q$

- The contrapositive of the converse is the inverse.

Arguments

Argument is valid $\iff (p_1 \land \ldots \land p_2) \to c$	
modus ponens (method of affirming)	$(p \land (p \to q)) \to q$
modus tollens (method of denying)	$(\neg q \land (p \to q)) \to \neg p$
Hypothetical syllogism	$((p \to q) \land (q \to r)) \to (p \to r)$
Disjunctive syllogism	$((p \lor q) \land \neg p) \to q$
Addition (generalisation)	$p \to (p \lor q)$
Simplification (specialisation)	$(p \land q) \to p$
Conjunction	$((p) \land (q)) \to (p \land q)$
Resolution	$((p \lor q) \land (\neg p \lor r)) \to (q \lor r)$

TABLE 2 Rules of Inference for Quantified Statements.

Rule of Inference	*Name*
$\dfrac{\forall x\, P(x)}{\therefore\ P(c)}$	Universal instantiation
$\dfrac{P(c) \text{ for an arbitrary } c}{\therefore\ \forall x\, P(x)}$	Universal generalization
$\dfrac{\exists x\, P(x)}{\therefore\ P(c) \text{ for some element } c}$	Existential instantiation
$\dfrac{P(c) \text{ for some element } c}{\therefore\ \exists x\, P(x)}$	Existential generalization

- If the hypothesis is always false, the proposition is true. The proposition is called a vacuous truth. This is an example of the principle of explosion.

Methods of Proof

- Direct proof: $P(x) \to Q(x)$. Choose an arbitrary x from the domain for which $P(x)$ is true, and use logical inference to show that $Q(x)$ is true also.
- Proof by contradiction: To show that p is true, assume that p is false, and use logical inference to prove a contradiction.
- Proof by contraposition: To prove $\forall x, P(x) \to Q(x)$, choose an arbitrary x for which $Q(x)$ is false, and argue by logical inference that $P(x)$ must be false also. Based on $(p \to q) \equiv (\neg q \to \neg p)$.
- Disproof by counterexample: Provide one counterexample, and the statement is disproven.

Incorrect Proof Techniques

- Proof by example: giving several examples for which $P(x)$ and $Q(x)$ are true.
- Begging the question/circular reasoning: Assuming $Q(x)$ within your proof, before you have even proven it.
- Text pages 75, 89 and 90 contain others.

WLOG

- WLOG means that no generality is lost by making a simplifying assumption. If the simple case is true, then trivially all cases must be true.
- E.g. $\forall a, b \in \mathbb{Z}$, if ab and $a + b$ are even, then both a and b are even.

Even and Odd

- The integer n is even if and only if n is twice some integer i.e. n is even if and only if $\exists k \in \mathbb{Z}$ such that $n = 2k$. The set of all integers is $\{n \in \mathbb{Z} | \exists k \in \ Z : n = 2k\}$.
- The integer n is odd if and only if n is twice some integer plus one i.e. n is odd if and only if $\exists k \in \mathbb{Z}$ such that $n = 2k + 1$. The set of all integers is $\{n \in \mathbb{Z} | \exists k \in \ Z : n = 2k + 1\}$.
- A proof for even and odd can be found in lecture 6.

Set Theory

Definitions

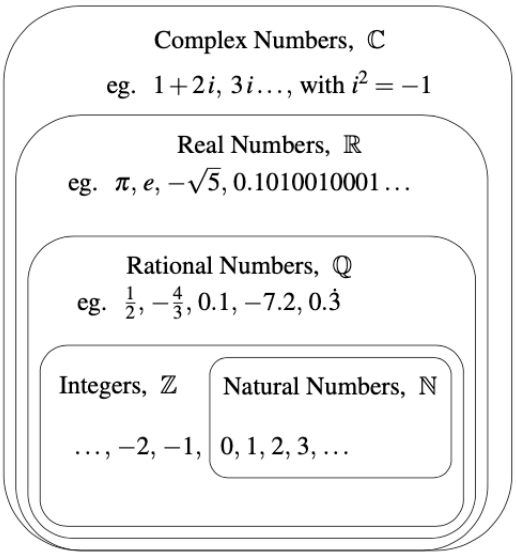


Figure 1.1: Number Sets

- A set S is a collection of objects, which are called the elements of S. If x is in S, we write $x \in S$. If not, we write $x \notin S$.
  - We can list the elements of S with curly braces: $S = \{x_1, x_2, \dots\}$. Finite sets have a countable amount of elements. Infinite sets do not.
  - Order does not matter and repetition is ignored.
  - Two sets are equal if they contain the same elements i.e. $S = T$ means that $\forall x, x \in S \iff x \in T$.
- An empty set, written $\emptyset$, contains no elements at all: $\{\}$. It can also be written as $\forall x, x \notin \emptyset$.

- A one-element set $\{x\}$ is not the same as x.
- $A = \{x \in S | P(x)\}$ means that the elements of A are precisely those elements $x \in S$ for which the predicate $P(x)$ is true i.e. A is the truth set of P and S is the domain of P.

Russell's Paradox

- Given the set: $S = \{x | x \notin x\}$.
  - If $S \in S$, then by definition of S, we have $S \notin S$.
  - If $S \notin S$, then by definition of S, we have $S \in S$.
- To avoid problems such as Russell's paradox, we can attempt to define all sets recursively.
  - Base: $\emptyset$ is a set.
  - Recursion: We define several operations that build new sets from old.
- Using the empty set as the starting point:
  - $\{\emptyset\} = \{\{\}\}$
  - $\{\emptyset, \{\emptyset\}\} = \{\{\}, \{\{\}\}\}$
  - $\{\emptyset, \{\emptyset\}, \{\emptyset, \{\emptyset\}\}\}$ = $\{\{\}, \{\{\}\}, \{\{\{\}\}\}\}$
- Giving these set names:
  - 0 = $\emptyset$
  - 1 = $\{\emptyset\}$ = $\{0\}$
  - 2 = $\{\emptyset, \{\emptyset\}\}$ = $\{0, 1\}$
  - 3 = $\{\emptyset, \{\emptyset\}, \{\emptyset, \{\emptyset\}\}\}$ = $\{0, 1, 2\}$

Relationships and Operations

Operators		
Union	$S \cup T$	$S \cup T = (x \in S \vee x \in T)$
Intersection	$S \cap T$	$S \cap T = \{x \in S \wedge x \in T\}$
Difference (S not in T)	$S \setminus T$	$S \setminus T = \{x \in S \wedge x \in T\}$
Complement (U not in S)	$\overline{S}$	$\overline{S} = \{x \in U \| x \notin S\}$

- U Is the universal set in which an individual is working.
- $[a, b]$ denotes the set of all real numbers x for which $a \leq x \leq b$: $[a, b] = \{x \in \mathbb{R} | a \leq x \leq b\}$. Circle brackets to exclude an endpoint.
- $\bigcup_{i=1}^{5} \{i, 2i\} = \{1, 2, 3, 4, 5, 6, 8, 10\}$ and $\bigcap_{i=1}^{\infty} [i, -i] = [-1, 1]$.

Subsets

- For sets S and T, we say that S is a subset of T if every element of S belongs to T also.
  - We write this as $S \subseteq T$. Formally, $\forall x, x \in S \rightarrow x \in T$.
  - S is a proper subset of T if $S \subseteq T$ and $S \neq T$.

Proving Relationships

- To prove $S \subseteq T$, we need to show $\forall x, x \in S \rightarrow x \in T$. Method: Choose an arbitrary $x \in S$, and prove that $x \in T$ also.
- To prove S = T, we need to show $\forall x, x \in S \iff x \in T$. Using the equivalency, $p \iff q \equiv (p \rightarrow q) \wedge (q \rightarrow p)$, choose two arbitrary $x \in S$, $x \in T$, and prove that $x \in T$, $x \in S$.
  - We can also do this using $S \subseteq T \wedge T \subseteq S$.

Cardinality

- If S is a finite set, then the cardinality of S is the number of distinct elements thaat S contains. We write this as $|S|$ e.g. if $S = \{3, \{3\}, 3\}$ then $|S| = 2$.
- If S is an infinite set, we write that $|S| = \infty$, although there are many different infinities; two infinite sets may not have the same cardinality.

Venn Diagrams Exist.

Power Sets

- For any set S, the power of set of S is the set of all subsets of S. We write this as: $P(S) = \{X | X \subseteq S\}$.
  - If $S = \{3, 5\}$, then $P(S) = \{\emptyset, \{3\}, \{5\}, \{3, 5\}\}$.
- Cardinality of power sets is $|P(S)| = 2^n$ where n is the number of elements in the set.

Cartesian Product

- The Cartesian product $A \times B$ of sets A and B is: $A \times B = \{(a, b) | a \in A, b \in B\}$. The order and repetition of the ordered pair matters.
  - If $A = \emptyset$ and $B = \emptyset$, then $A \times B = \emptyset$. We write $A^2 = A \times A$, where $|A^k| = |A|^k$. From geometry, the plane is thus: $\mathbb{R}^2 = \mathbb{R} \times \mathbb{R}$.
- Suppose X and Y are finite sets with $|X| = n$ and $|Y| = m$. Therefore, $|X \times Y| = n \cdot m$.
- We can generalise from pairs to ordered triples/quadruples/n-tuples, where $a_k \in A_k$ for sets $A_1, A_2, A_3, \ldots$
  - $(a_1, a_2, \ldots, a_n) = (b_1, b_2, \ldots, b_n) \iff a_k = b_k$ for each $k \in \{1, \ldots, n\}$.
  - The set of all ordered n-tuples is denoted $A_1 \times A_2 \times \ldots \times A_n$.

Functions

- Let X and Y be sets. The function f from X to Y, written $f : X \rightarrow Y$ assigns to each $x \in X$ a unique element $y \in Y$.
  - We can abbreviate $f(x)$ to $x \mapsto y$. For example, $g : \mathbb{N} \rightarrow \mathbb{N}$ defined by $n \mapsto 2^n$.
- Formally, we can think of a function as a subset of a Cartesian product: A function $f : X \rightarrow Y$ is a subset $\Gamma \subseteq X \times Y$ such that, for each $x \in X$, there is a unique $y \in Y$ such that $(x, y) \in \Gamma$.
  - This can be expressed as $\Gamma = \{(x, f(x)) | x \in X\} \subseteq X \times Y$.

More Definitions

- If $f : X \to Y$ is a function, then:
  - X is called the domain of f.
  - Y is called the co-domain of f.
  - If $x \in X$, then $f(x)$ is called the image of x.
  - If $A \subseteq X$, then $f(A) = \{f(x)|x \in A\} = \{y|\exists x \in A : f(x) = y\} \subseteq Y$ is called the image of A. The entire set $f(X)$ is called the range of f.
  - If $y \in Y$, then $f^{-1}(y) = \{x \in X|f(x) = y\} \subseteq X$ is called the preimage of y.
  - If $B \subset Y$, then $f^{-1}(B) = \{x \in X|f(x) \in B\} \subseteq X$ is called the preimage of B.

More Functions

- The assignment $f(x) = x$ for each $x \in X$ defines the identity function $\iota x : X \to X$.
- If $f : X \to Y$ is a function, then a function $P(X) \to P(Y)$ is defined by $A \mapsto f(A)$.
- If $f : X \to Y$ is a function, then a function $P(Y) \to P(X)$ is defined by $B \mapsto f^{-1}(B)$.

Basic Properties of Functions

- Functions $f, g : X \to Y$ are equal, written $f = g$ if and only if $f(x) = g(x)$ for all $x \in X$.
- Let $x \in \mathbb{R}$ be a real number. The floor of x, denoted $\lfloor x \rfloor$, is the unique integer n such that $n \leq x < n + 1$.
- Let $x \in \mathbb{R}$ be a real number. The ceiling of x, denoted $\lceil x \rceil$, is the unique integer n such that $n - 1 < x \leq n$.

Properties of Functions

- Let $f : X \to Y$. Then:
  - f is onto, or surjective, or a surjection, if: $\forall y \in Y, \exists x \in X$ such that $f(x) = y$. Every element of Y is the image of something.
  - f is one-to-one, or injective, or an injection, if: $\forall x_1, x_2 \in X, f(x_1) = f(x_2) \to x_1 = x_2$. Different elements of X have different images.
  - f is a one-to-one correspondence, or bijective, or a bijection, if f is both one-to-one and onto. The elements of X are "paired off" with the elements of Y.

Finite Sets

- Suppose X and Y are finite sets.
  - If $|X| > |Y|$, then there is no injective function $X \to Y$.
  - If $|X| < |Y|$, then there is no surjective function $X \to Y$.
  - There is a bijective function $X \to Y$ if and only if $|X| = |Y|$.
- Put differently:
  - If $f : X \to Y$ is injective, then $|X| \leq |Y|$.
  - If $f : X \to Y$ is surjective, then $|X| \geq |Y|$.
  - If $f : X \to Y$ is bijective, then $|X| = |Y|$.
- For finite sets with $|X| = |Y|$, the following statements are equivalent.
  - $f : X \to Y$ is injective/bijective/surjective.

Composition of Functions

- If $f : X \to Y$ and $g : Y \to Z$, then the composition, $g \circ f : X \to Z$, if defined by $(g \circ f)(x) = g(f(x))$ for all $x \in X$. We sometimes write this as gf.

Sequences

- A sequence can be written: $T_0, T_1, T_2, \ldots$. We write this sequence as $(T_n)_{n\geq 0}$ or $(T_n)_{n=0}^{\infty}$ or $\{T_n\}$.
- Each number T_i in the sequence is called a term. We conjectured that: $T_0 = 0$ and $T_n = 2T_{n-1} + 1$ for $n \geq 1$. This is called a recursive definition for (T_n).
- We also conjectured that: $T_n = 2^n - 1$. This is an explicit formula, or closed formula, for T_n.
- Sequences can be finite: $(a_n)_{n=1}^{6}$ or infinite: $(g_n)_{n\geq 0}$. The index doesn't have to start at 0, be written as n.
    - Alternating sequences alternate between positive and negative: $((-\frac{1}{2})^n)_{n\geq 0}$.
    - Other famous sequences include the Fibonacci sequence, factorials, Conway's look-and-say.

Induction

- To define a sequence recursively, we need:
    - Initial conditions, which directly specify one or more terms that begin the sequence: $F_0 = 0, F_1 = 1$
    - A recurrence relation, which defines every other term using earlier terms: $F_n = F_{n-1} + F_{n-2}$ for $n \geq 2$.

Notation for Sums

- For a sequence $(a_i)_{i=m}^{n}$, we can add some or all of its terms: $\sum_{i=m}^{n} a_i = a_m + a_{m+1} + a_{m+2} + \ldots + a_{n-1} + a_n$
- If $m = n$, the sum only has only term, if $n < m$ then the sum is empty and we define it to be zero.
- The dummy variable (in the above case i) is only relevant inside the sum, which means you can reuse it outside the sum. You can also perform a change of variable.

Arithmetic with Finite Sums

- There are many ways to manipulate finite sums.
    - Adding/subtracting over the same range: $\sum_{i=m}^{n} a_i \pm \sum_{i=m}^{n} b_i = \sum_{i=m}^{n} (a_i \pm b_i)$.
    - Taking out a common factor: $\sum_{i=m}^{n} ca_i = c \sum_{i=m}^{n} a_i$.
    - Combining consecutive indices: $\sum_{i=p}^{q} a_i + \sum_{i=q+1}^{r} a_i = \sum_{i=p}^{r} a_i$ if $p \leq q \leq r$.
    - Index shift: $\sum_{i=m}^{n} a_i = \sum_{i=m+p}^{n+p} a_{i-p} = \sum_{i=m-q}^{n-q} a_{i+q}$.

- Telescoping sums: $\sum_{i=m}^{n}(a_i - a_{i+1}) = a_m - a_{n+1}$ if $m \leq n$.

Division

- If $n, d \in \mathbb{Z}$, then n is divisible by d if and only if there exists some $k \in \mathbb{Z}$ such that $n = kd$. We write $d|n$, and say "d divides n" or "d is a divisor of n".
  - If n is not divisible by d, we write that $d \nmid n$.
  - For all $a, b, m \in Z$, if a and b are divisible by m, then $a + b$ is divisible by m. Symbolically: $\forall a, b, m \in \mathbb{Z}, (m|a) \wedge (m|b) \rightarrow m|(a + b), m|(a - b), m|(a \times b)$.
  - $\forall a, b, m \in \mathbb{Z}$, if ab is divisible by m, then both a and b dont have to be divisible by m.
- Let $n, d \in \mathbb{Z}$. If $|n| \geq 1$ and $d \mid n$, then $0 < |d| \leq |n|$.

The Quotient-Remainder Theorem

- Given any integer n and positive integer d, there exists unique integers q and r such that $n = qd + r$ and $0 \leq r < d$. We call q the quotient, and r the remainder.
- Using the QRT, we can prove that all squared numbers end in one of $\{0, 1, 4, 5, 6, 9\}$, and given any three digit number written twice (i.e. $123 \rightarrow 123123$), this number is divisible by 7 without remainder.

Modular Arithmetic

- Modular arithmetic is a fancy way of writing down arguments like this in a more elegant, concise form.
  - We can group integers into equivalence classes of numbers that are equivalent mod $n \in \mathbb{Z}$.
- If n and m leave the same remainder after division by d, we say that they are congruent modulo d i.e. $n \equiv m$ (mod d) . The relationship is symmetric.
  - If $n = qd + r$, then $n \equiv r$ (mod d).
  - $n \equiv m$ (mod d) if and only if $d \mid (m - n)$.
  - $n \equiv 0$ (mod d) if and only if $d \mid n$.
- If $a \equiv b$ (mod d) and $n \equiv m$ (mod d), then:
  - $an \equiv bm$ (mod d)
  - $a + n \equiv b + m$ (mod d)
  - $a - n \equiv b - m$ (mod d)
- $10^k \equiv 1$ (mod 9) for each $k \in \mathbb{N}$. So: $n = a_m 10^m + \ldots + a_1 10 + a_0$ where $n \in \mathbb{N}$, and its digits are $n = a_m a_{m-1} \ldots a_1 a_0$.
  - $9|n$ if and only if the sum of the digits of n is divisible by 9.

Prime Factorisation

- The natural number $n \in \mathbb{N}$ is said to be written as a product of primes if there is a natural number $m \in \mathbb{N}$ and prime numbers $p_1, \ldots, p_m$ such that: $n = p_1 \cdot p_2 \cdot \ldots \cdot p_m = \prod_{k=1}^{m} p_k$.
- Every natural number $n > 1$ can be written as a product of primes.
- There are infinitely many prime numbers (Euclid).

The Fundamental Theorem of Arithmetic

- Given any integer $n > 1$, there exists a natural number k, pairwise distinct prime numbers $p_1, p_2, \ldots, p_k$, and natural numbers $e_1, e_2, \ldots, e_k$ such that: $n = p_1^{e_1} \cdot p_2^{e_2} \cdot \ldots \cdot p_k^{e_k} = \prod_{i=1}^{k} p_i^{e_i}$, and any other expression for n as a product of prime numbers is identical to this except possibly for the order in which the factors are written.

Unique Prime Factorisation

- Suppose $d|576$ for some $d \in \mathbb{N}$. Then $d \cdot k = 576$ for some $k \in \mathbb{N}$. Expressing d as a product of primes: $d = p_1 \ldots p_r$ and expressing k as a product of primes $k = q_1 \ldots q_s$, then $p_1 \ldots p_r \cdot q_1 \ldots q_s = 2^6 \cdot 3^2$.
- Therefore $d = 2^i \cdot 3^j$ with $0 \leq i \leq 6$ and $0 \leq j \leq 2$.

Applications of Unique Factorisation

- For natural numbers a, b:
  - The greatest common divisor of integers $a, b \neq 0$ is the largest $d \in \mathbb{N}$ for which $d|a$ and $d|b$. We write this as $gcd(a, b)$.
    - If you have the prime factorisations, just take the smallest power of each prime that appears in both integers, and ignore any negative signs.
  - The least common multiple of the positive integers a and b is the smallest $n \in \mathbb{N}$ for which $a|n$ and $b|n$. We write this as $lcm(a, b)$.
    - If you have the prime factorisations, just take the largest power of each prime regardless of if they appear in both.
- Two integers $a, b \in \mathbb{Z}$ are called coprime if $gcd(a, b) = 1$. If $a, b \in \mathbb{Z}$ are coprime, and $ab = c^3$ for some $c \in \mathbb{Z}$, then $a = d^3$ and $b = e^3$ for some $d, e \in \mathbb{Z}$.
  - If the product of two coprime integers is a cube, then each of the integers is a cube also.
- For all $a, b \in \mathbb{Z}$, $gcd(a, b) = gcd(b, a - b)$. The Euclidian algorithm is as follows:
  - From this, we can see that: For all $a, b \in \mathbb{Z}$, if $a = bq + r$, then $gcd(a, b) = gcd(b, r)$ i.e. $gcd(a, b) = gcd(b, a\%b)$. The gcd is the last non-zero remainder.
  - If one or both of a, b are negative, then just ignore the negative signs.
  - If $a = 0$ and $b = 0$, then $gcd(a, b)$ is not defined.

Representation of Integers

- Let b be an integer greater than 1. Every positive integer n can be expressed uniquely in the form: $n = a_k b^k + a_{k-1} b^{k-1} + \ldots + a_1 b + a_0$, where k is a non-negative integer, $a_0, \ldots, a_k$ are non-negative integers less than b and $a_k \neq 0$.
- In order to write a number in base b:
  - Repeatedly divide by b, and write down the remainders.
  - Stop when you reach zero.
  - The remainders will give the digits in reverse order.

Algorithm Complexity

- Running time of an algorithm is the number of steps required for it to finish, depending on how large the input is: $f : \mathbb{N} \to \mathbb{N}$; Input size $\mapsto$ Number of steps required.

O-Notation

- O-notation is a mathematical notation that describes how quickly a function f grows compared to some function g when both their arguments tend towards infinity.
  - Let f and g be functions from a subset of $\mathbb{R}$ to $\mathbb{R}$. Then $f(x)$ is in $O(g(x))$ if there exists constants C and k such that for all $x \in A, x \geq k : |f(x)| \leq C|g(x)|$.
  - It is better to write $f(x) \in O(g(x))$ since $f(x) = O(g(x))$ ignores that it is a one-way equality.
- For a given g it assigns a (large) set of functions which all grow at roughly the same rate, or slower. The comparsion is done by checking whether f is an element of this set.
- C And k are the witnesses of the statement "$f(x)$ is in $O(g(x))$".
- For this to make sense, f must have the following properties:
  - $f, g : A \subset \mathbb{R} \to \mathbb{R}$.
  - A is unbounded (for all $k \in A$ there exists infintely many $x \in A$ such that $x \geq k$).

Example

- $f(n) = \sum_{k=1}^{n} k^2 = 1^2 + 2^2 + \ldots + n^2 = \frac{1}{3}n(n + \frac{1}{2})(n + 1) = \frac{1}{3}n^3 + \frac{1}{2}n^2 + \frac{1}{6}n$.
- Claim: $f(n) - \frac{1}{3}n^3 \in O(n^2)$:
  - $f(x) - \frac{1}{3}n^3 = \frac{1}{2}n^2 + \frac{1}{6}n$.
  - $\frac{1}{2}n^2 \leq \frac{1}{2}n^2$; $\frac{1}{6}n \leq \frac{1}{6}n^2$ when $n \geq 1$.
  - $f(x) - \frac{1}{3}n^3 = \frac{1}{2}n^2 + \frac{1}{6}n \leq \frac{2}{3}n^2 = \frac{2}{3}g(n)$.
  - $C = \frac{2}{3}, k = 1$ are witnesses and $f(n) \in \frac{1}{3}n^3 + O(n^2)$.

More Operations

- $f(n) \in O(f(n))$
- $O(c \cdot f(n)) = O(f(n))$ if c is a constant.
- $O(f(n) + f(n)) = O(f(n))$
- $O(f(n)g(n)) = f(n) \cdot O(g(n))$.
  - Note that we use "=" to mean set equality.
- In order to prove that a function is not in $O(n)$, we can prove this by contradiction.

Important Examples

- Exponentials always grow faster than polynomials.
- $n! \in O(n^n)$
- $log(n!) \in O(nlog(n))$
- $C^n \in O(n!), c > 0$

Lower Bounds and Exact Order

- Definition of Ω-notation: Let f and g be functions from a subset of $\mathbb{R}$ to $\mathbb{R}$. Then $f(x)$ is in $\Omega(g(x))$ if there exists constants C and k such that for all $x \in A, x \geq k : |f(x)| \geq C|g(x)|$.
  - Matrix multiplication is in $\Omega(n^2)$.
- Definition of Θ-notation: Let f and g be functions from a subset of $\mathbb{R}$ to $\mathbb{R}$. Then $f(x)$ is in $\Theta(g(x))$ if $f(x) \in O(g(x))$ and $f(x) \in \Omega(g(x))$. It means that $f(x)$ and $g(x)$ are of the same order i.e. they are within a constant of each other.

P vs NP

- A decision problem is a yes/no question, for which we wish to find an algorithm.
  - Input: $k \in \mathbb{N}$. Question: Is k prime?
- P vs NP is about which decision problems you can solve quickly, and which problems are inherently difficult.
  - A decision problem is in the class P if you can solve it quickly.
    - Input: $k, l \in \mathbb{N}$. Question: Are k and l coprime. Solution: Euclidian algorithm.
  - A decision problem is in the class NP if, when the answer is "yes", I can give you information that lets you verify my solution quickly.
    - Input: $k \in \mathbb{N}$. Question: Is k composite? Information: I give you the prime factorisation of k. We know how to "quickly" test whether k is composite.
- An algorithm is considered fast if its running time is bounded by a polynomial.
- The question P vs NP asks: are P and NP the same? That is, if a "yes" solution is fast to verify, doe sthat mean the problem must be fast to solve?
  - The hardest problems in NP are called NP-complete: a fast solution to any of these would give a fast solution to every problem in NP.
  - Example of NP-Complete: Input: A set $S \subseteq \mathbb{Z}$. Does S have a subset whose sum is 0?

Induction

- The Principle of Mathematical Induction. Let $P(n)$ be a predicate that is defined for all integers $n \geq a, a \in N$. Suppose:
  - $P(a)$ is true;
  - For all integers $n \geq a, P(n) \rightarrow P(n+1)$. Note that we never directly prove that $P(n)$ is true $(n \neq a)$.
  - Then $P(n)$ is true for all integers $n \geq a$.
- To prove $P(n)$ for all integers $n \geq a$, you need to:
  - Prove $P(a)$. This is called the basis step.
  - Prove that $P(n) \rightarrow P(n+1)$ for all integers $n \geq a$. Here you must:
    - Assume $P(n)$ is true for some particular but arbitrary $n \geq a$. This is called the inductive hypothesis.
    - Using this, show that $P(n+1)$ is also true. This is called the inductive step.

Other Ways to Think About Induction

- Essentially, the principle of mathematical induction states that the following argument is valid (similar to modus ponens):
  - $P(0)$
  - $\forall n \in \mathbb{N}, P(n) \to P(n+1)$
  - $\therefore \forall n \in \mathbb{N}, P(n)$
- Let X be a set of natural numbers with the following properties:
  - The number 0 is in X.
  - For all $n \in \mathbb{N}$, if n is in X, then $n+1$ is also in X.
  - Then X is the set of all natural numbers, i.e. $X = \mathbb{N}$. Here, X is the truth set of $P(n)$.
- The Principle of Strong Mathematical Induction. It can be shown that this is equivalent to the ordinary principle of mathematical induction. Let $P(n)$ be a predicate that is defined for all integers $n \geq a$, and let $b \geq a$. Suppose:
  - Basis step: $P(a), P(a+1), \ldots, P(b-1), P(b)$ are all true.
  - Inductive step: For all integers $n \geq b$, if $P(a), P(a+1), \ldots, P(n-1), P(n)$ are all true, then $P(n+1)$ is also true.
  - Therefore, $P(n)$ is true for all integers $n \geq a$.

Recursion

- Defining a mathematical expression using itself. Easiest case: define $f : \mathbb{N} \to \mathbb{R}$ by:
  - $f(0) = a$. This is the initial condition.
  - $f(n+1) = g(n, f(n))$ where g is known. This is the recurrence relation.
- Converting a recursive definition into an explicit formula is usually difficult, and you cannot always do it.
  - If you can guess the formula however, it is easy to prove. Just check that it satisfies the recursive definition, by substituting in your formula, and checking that the equality holds.

Guessing the Recursive Formula

- Expand how each term is constructed.

  Consider $(a_n)_{n\geq 0}$, with $a_0 = 7$ and $a_n = a_{n-1} + 3$ for $n \geq 1$.

  $a_0 = 7$
  $a_1 = 7 + 3$
  $a_2 = 7 + 3 + 3$
  $a_3 = 7 + 3 + 3 + 3$

  Conjecture: $a_n = 7 + 3n$

  Proof:

  - $a_0 = 7 + 3 \cdot 0 = 7$, which satisfies the initial conditions
  - $a_n = 7 + 3n = 7 + 3(n-1) + 3 = a_{n-1} + 3$, which satisfies the recurrence relation

- Compute several terms and look for a pattern.

  Consider $(a_n)_{n\geq 0}$, with:

  $$a_0 = 0, \quad a_1 = 1, \quad a_n = \frac{1}{2}(a_{n-1} + a_{n-2} + 1) + 3(n-1) \text{ for } n \geq 2$$

  Some initial terms:
  $a_0 = 0, \quad a_1 = 1, \quad a_2 = 4, \quad a_3 = 9, \quad a_4 = 16, \quad a_5 = 25$

  Conjecture: $a_n = n^2$

  Proof:

  - $a_0 = 0^2 = 0$ and $a_1 = 1^2 = 1$, satisfying the initial conditions
  - For the recurrence relation, we start on the right hand side:
    $\frac{1}{2}(a_{n-1} + a_{n-2} + 1) + 3(n-1) =$
    $\frac{1}{2}[(n-1)^2 + (n-2)^2 + 1] + 3(n-1) =$
    $\frac{1}{2}(2n^2 - 6n + 5 + 1) + 3n - 3 = n^2 - 3n + 3 + 3n - 3 = n^2$
    $= a_n$, which satisfies the recurrence relation

- Other methods include: The Online Encyclopaedia of Integer Sequences (oeis.org), generating functions, transform the sequence, examine differences etc.
- Note that guessing can be dangerous, and even though it may hold for the first few terms, it may not hold for all n e.g. Fermat primes.

Counting

- Suppose $m, n \in \mathbb{Z}$ and $m \leq n$, there are $n - m - 1$ integers between m and n.
- Sample space S: All possible outcomes for a random process or experiment.
  - $S = \{\{H_1, H_2\}, \{H_1, T_2\}, \{T_1, H_2\}, \{T_1, T_2\}\}$ for two coins.
- Event E: Subset/partition of the sample space e.g. $E_1 = \{\{H_1, H_2\}\}$ for two heads.
- Equally likely probability formula for a finite sample space: If all outcomes in S are equally likely, then: $P(E) = \frac{\text{number of outcomes in E}}{\text{number of outcomes in S}} = \frac{|E|}{|S|}$/
  - If you have a finite sample space, and you can count, then you can determine probabilities.
- We multiply if we must make decision A and then decision B. We add if we make either A or B.
- It's okay to overcount, as long as you subtract off the unwanted solutions later!

For Finite Sets

- Multiplication: $|S \times T| = |S| \times |T|$
- Addition: If S and T are disjoint, then $|S \cup T| = |S| + |T|$
- Subtraction: If $T \subseteq S$ then $|S \setminus T| = |S| - |T|$
- For any finite set S, the number of subsets of S with an even number of elements is equal to the number of subsets of S with an odd number of elements.

Common Counting Problems

Case 1: Order Matters, Repetition Allowed

- Given a set $S = \{0, 1, 2, 3, 4, 5, 6, 7, 8, 9\}$, picking four digits from this set would yield $S \times S \times S \times S = S^4$.
    - Questions e.g. 3 entrees, 7 mains, 5 desserts, how many combinations.

The Multiplication Rule

- If a decision process can be broken down into a finite number of independent steps, and there are n_i possible outcomes for the ith step, then the entire process can be carried out in $\prod n_i$ possible ways.
- In particular, for finite sets $S_1, S_2, \ldots, S_k$, $|S_1 \times S_2 \times S_k| = \prod_{i=1}^{k} |S_i|$.

Case 2: Order Matters, Repetition Not Allowed

- If you choose k elements from a set S with n elements, and order matters and repetition is not allowed, then there are:
    - n possibilities for the first choice
    - $n - 1$ for the second
    - $n - 2$ for the third
    - $n - k - 1$ for the kth.
- The total number of choices is: $n \cdot (n-1) \cdot (n-2) \ldots (n-k+1) = \frac{n!}{(n-k)!}$. Examples include: allocate 5 breadcrumbs to 13 ducklings.
- We call this number $P(n, k)$. The related sample space is: $\{(a_1, \ldots, a_k) \in S^k | a_i \neq a_j \text{if } i \neq j\} \subseteq S^k$.

Order Does Not Matter, Repetition Not Allowed

- To choose k items from n:
    - If order does matter, there are $P(n, k) = \frac{n!}{(n-k)!}$ possibilities.
    - We have counted each unordered set $k!$ times, so when order does not matter, there are $\frac{n!}{k!(n-k)!}$ possibilities.
    - Examples include: How many poker hands are there (choose 5 from 52)?
- We write this as $\binom{n}{k}$, pronounced as "n choose k":
    - $\binom{n}{k} = \frac{n!}{k!(n-k)!} = \frac{n \cdot (n-1) \cdot (n-2) \cdot \ldots \cdot (n-k+1)}{k!}$.
- For all $n, k \in Z$ with $0 \leq k \leq n$: $\binom{n}{k} = \binom{n}{n-k}$. This is because choosing which k objects to take is the same as choosing which $n - k$ objects to leave behind.

Order Does Not Matter, Repetition Allowed

- Choosing k items from a set of n items such that order does not matter and repetition is allowed is the same as counting all possible arrangements of k crosses and $n-1$ bars, and this number is: $\frac{(k+n-1)!}{k!(n-1)!} = \binom{n+k-1}{n-1}$.
- Examples include: How many ways are there to put 2 not distinguished balls into 3 distinguished boxes?

Binomial Theorem

- If S is a set with $|S| = n$, then the number of subsets of S with exactly k elements is: $\binom{n}{k} = \frac{n!}{k!(n-k)!}$.
  - This is called the binomial coefficient.
- The Binomial Theorem states that: For all $a, b \in \mathbb{R}$, and all $n \in \mathbb{N}$:
  - $(a+b)^n = \sum_{k=0}^{n} \binom{n}{k} a^k b^{n-k}$

Algebraic vs Combinatorial Proof of $\binom{n}{k} = \binom{n-1}{k} + \binom{n-1}{k-1}$

Algebraic

$$
\begin{aligned}
& \binom{n-1}{k} + \binom{n-1}{k-1} \\
= \; & \frac{(n-1)!}{k!(n-k-1)!} + \frac{(n-1)!}{(k-1)!(n-k)!} \\
= \; & \frac{(n-1)!}{(k-1)!(n-k-1)!} \cdot \left[\frac{1}{k} + \frac{1}{n-k}\right] \\
= \; & \frac{(n-1)!}{(k-1)!(n-k-1)!} \cdot \left[\frac{n-k}{k(n-k)} + \frac{k}{k(n-k)}\right] \\
= \; & \frac{(n-1)!}{(k-1)!(n-k-1)!} \cdot \frac{n}{k(n-k)} \\
= \; & \frac{n!}{k!(n-k)!}
\end{aligned}
$$

Combinatorial

Let $|S| = n$. Then $\binom{n}{k}$ is the number of ways of choosing a subset $A \subseteq S$ with $|A| = k$.

Let $S = \{s_1, s_2, \ldots, s_n\}$.
We take cases according to whether or not $s_1 \in A$:

- The number of subsets $A \subseteq S$ with $|A| = k$ and $s_1 \in A$ is $\binom{n-1}{k-1}$.
- The number of subsets $A \subseteq S$ with $|A| = k$ and $s_1 \notin A$ is $\binom{n-1}{k}$.

Therefore $\binom{n}{k} = \binom{n-1}{k} + \binom{n-1}{k-1}$.

Inclusion-Exclusion Principle

- For any sets A and B: $|A \cup B| = |A| + |B| - |A \cap B|$. Based on this rule, we find that:
- For sets $A_1, A_2, \ldots, A_n$:

$$\begin{aligned}|A_1 \cup \ldots \cup A_n| = \sum_i |A_i| - \sum_{i<j} |A_i \cap A_j| \\ + \sum_{i<j<k} |A_i \cap A_j \cap A_k| \\ - \sum_{i<j<k<l} |A_i \cap A_j \cap A_k \cap A_l| \\ + \ldots \pm |A_1 \cap A_2 \cap \ldots \cap A_n|\end{aligned}$$

- To measure the size (cardinality) of the set union, we measure each set individually, subtract all two-way intersections, add all three-way intersection, subtract all four-way intersections and so on.
- This can be applied to questions such as: There are ten students and ten marked quizzes. If quizzes are given at random, what is the probability that nobody receives their own quiz?

Pigeonhole Principle

- If you have n pigeons sitting in k pigeonholes, and if $n > k$, then at least one of the pigeonholes contains at least two pigeons.
- Generalised: If you have n pigeons sitting in k pigeonholes, and if $n > k \cdot m$, then at least one of the pigeonholes contains at least $m + 1$ pigeons.
    - The original pigeonhole principle corresponds to the case $m = 1$.

Catalan Numbers

- $b_n = p_n = t_n = \frac{1}{n+1}\binom{2n}{n} = \frac{(2n)!}{(n+1)!n!}$. The first few terms are: 1, 2, 5, 14.
- A recurrence relation that follows this is: $b_{n+1} = b_0 b_n + b_1 b_{n-1} + \ldots + b_n b_0 = \sum_{k=0}^{n} b_k b_{n-k}$

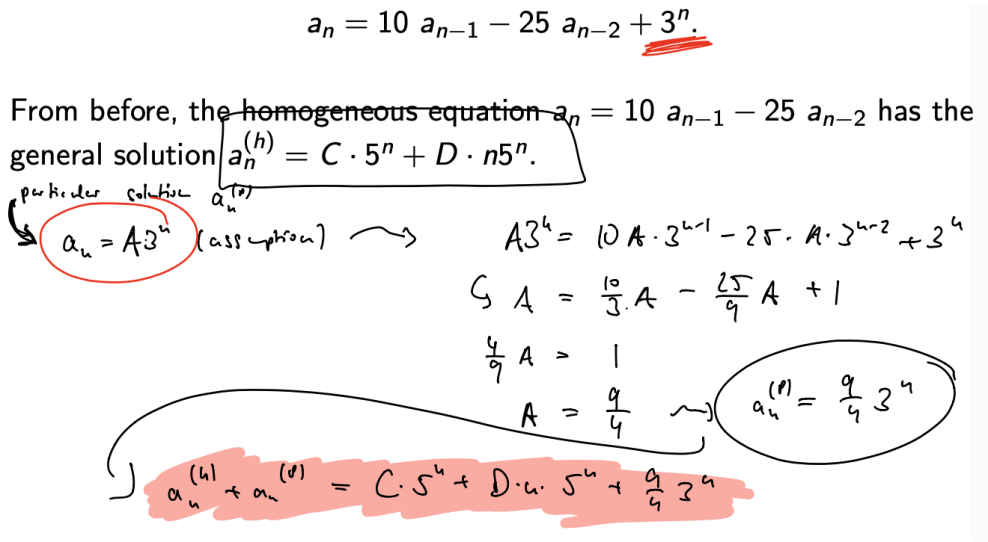
Examples

- How many balanced strings can be made with n left-brackets and n right-brackets?
- How many specials paths exist on an $n \times n$ grid going from bottom left to bottom right, never crossing the diagonal?
- How many rooted binary trees exist with exactly n vertices.

Recurrences Revisited

- Consider a linear homoegenous recurrence relation of order 2: $a_n = \alpha a_{n-1} + \beta a_{n-2}$.
    - Factor $x^2 - \alpha x - \beta = (x - \lambda_1)(x - \lambda_2)$
    - If $\lambda_1 \neq \lambda_2$, then $a_n = A\lambda_1^n + B\lambda_2^n$ for some constants A and B.
    - If $\lambda_1 = \lambda_2$, then $a_n = C\lambda^n + D\lambda^n$ where $\lambda = \lambda_1 = \lambda_2$ for some constants C and D.
    - The solutions for a_n are the general solution of the reccurence relation.
    - Initial conditions (values of a_0 and a_1) determine constants A, B or C, D.

- Consider a linear non-homoegenous recurrence relation of order 2:
  $a_n = \alpha a_{n-1} + \beta a_{n-2} + F(n)$.
  - Find one particular solution $a_n^{(p)}$ by poking around.
  - Determine the general solution $a_n^{(h)}$ to the homogeneous equation:
    $a_n = \alpha a_{n-1} + \beta a_{n-2}$.
  - The general solution of the non-homogeneous recurrence relation is then given by $a_n^{(p)}$ and $a_n^{(h)}$.
  - When we are given initial conditions, i.e. values of a_0 and a_1, then these determine the constants A, B or C, D.



- How many ways can we write an integer as a sum of two squares? i.e.
  $n = a^2 + b^2$ where $a, b \in \mathbb{Z}$. As n grows, the average of $f(1), \ldots, f(n)$ approaches π.

Probability Basics

- Set-Up
  - Sample space S (also called probability space). Here $|S| < \infty$.
  - $x \in S$ is called an outcome.
  - $E \subset S$ is called an event.
  - Set of events is a subset of the powerset of S, so $\mathcal{E} \subset \mathcal{P}(S)$.
  - For $x \in S$, $\{x\}$ is called an elementary event.
- Probability distribution: Assign probabilities to outcomes. $p : S \to [0, 1]$ s.t. $\sum_{x \in S} p(x) = 1$.
  Extend to events $E \subset S$ by adding up the probabilities of outcomes of E.

Random Variable

- We may want to assign values to events. A random variable is a function $X : S \to \mathbb{R}$ defined on the outcomes of a sample space.

Example: S outcomes of the roll of two fair dice, so $|S| = 36$.
$X : S \to \mathbb{R}$, $X(s) =$ "sum of the values of the two dice in outcome $s \in S$"
Then $X(S) = \{2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12\}$.
We have:

$2 = X(⚀, ⚀)$
$3 = X(⚀, ⚁) = X(⚁, ⚀)$
$4 = X(⚀, ⚂) = X(⚁, ⚁) = X(⚂, ⚀)$
$5 = X(⚀, ⚃) = X(⚁, ⚂) = X(⚂, ⚁) = X(⚃, ⚀)$
$6 = X(⚀, ⚄) = X(⚁, ⚃) = X(⚂, ⚂) = X(⚃, ⚁) = X(⚄, ⚀)$
$7 = X(⚀, ⚅) = X(⚁, ⚄) = X(⚂, ⚃) = X(⚃, ⚂) = X(⚄, ⚁) = X(⚅, ⚀)$
$8 = X(⚁, ⚅) = X(⚂, ⚄) = X(⚃, ⚃) = X(⚄, ⚂) = X(⚅, ⚁)$
$9 = X(⚂, ⚅) = X(⚃, ⚄) = X(⚄, ⚃) = X(⚅, ⚂)$
$10 = X(⚃, ⚅) = X(⚄, ⚄) = X(⚅, ⚃)$
$11 = X(⚄, ⚅) = X(⚅, ⚄)$
$12 = X(⚅, ⚅)$

- Distribution of a random variable: Set of all pairs $(r, p(X = r))$. You can think of the image $X(S) = \{r | \exists s \in S(r = X(s)\}$ of the random variable as a new sample space with probability distribution $p(r) = p(X = r)$ because:

$$\sum_{r \in X(s)} p(X = r) = 1$$

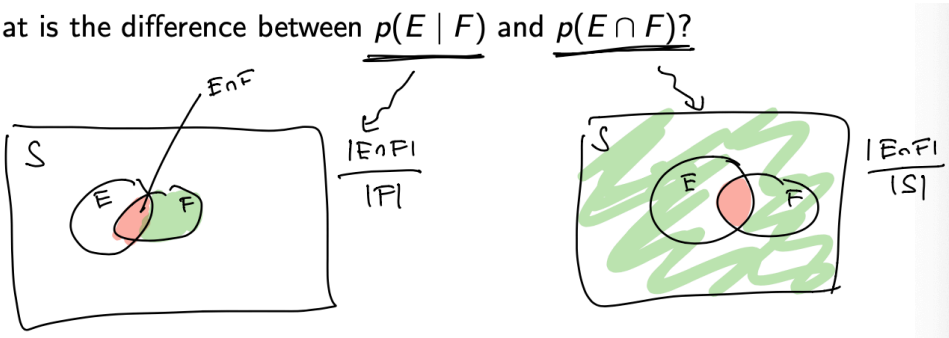
Example from above:
The sample space is $S = \{2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12\}$
The new probability distribution is $p : S \to [0, 1]$ given by

$$p(2) = p(12) = \frac{1}{36}; \quad p(3) = p(11) = \frac{1}{18};$$
$$p(4) = p(10) = \frac{1}{12}; \quad p(5) = p(9) = \frac{1}{9};$$
$$p(6) = p(8) = \frac{5}{36}; \quad p(7) = \frac{1}{6}$$

Conditional Probability

- Let E and F be events with $p(F) > 0$. The conditional probability of E given F is $p(E|F) = \frac{p(E \cap F)}{p(F)}$.

What is the difference between $p(E \mid F)$ and $p(E \cap F)$?



Independence

- Assume that $p(E) > 0$ and $p(F) > 0$. The conditional probability of E given F is $p(E|F) = \frac{p(E \cap F)}{p(F)}$.
- The follow are equivalent:
  - $p(E|F) = p(E)$ or equivalently, $p(F|E) = p(F)$.
  - $p(E \cap F) = p(E)p(F)$.
- If any of the above hold, E and F are independent.

Bayes' Theorem

Let E and F be two events with $p(E) > 0$ and $p(F) > 0$.
The complement $\overline{F} = S \setminus F$ gives a partition

$$E = (E \cap F) \cup (E \cap \overline{F})$$

and hence:

$$p(E) = p(E \cap F) + p(E \cap \overline{F}).$$

Then

$$\begin{aligned} p(E) &= p(E \cap F) + p(E \cap \overline{F}) \\ &= p(E \mid F) \cdot p(F) + p(E \mid \overline{F}) \cdot p(\overline{F}) \end{aligned}$$

We also have

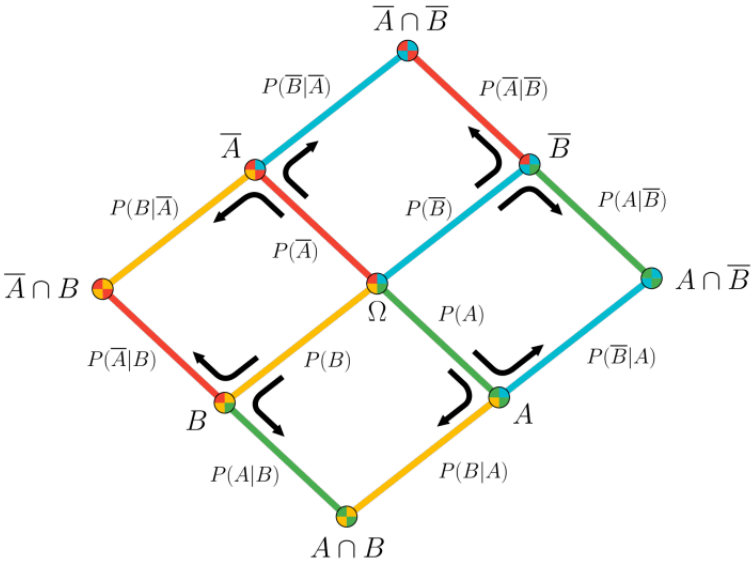
$$p(F \mid E) \cdot p(E) = p(F \cap E) = p(E \cap F) = p(E \mid F) \cdot p(F)$$

and so

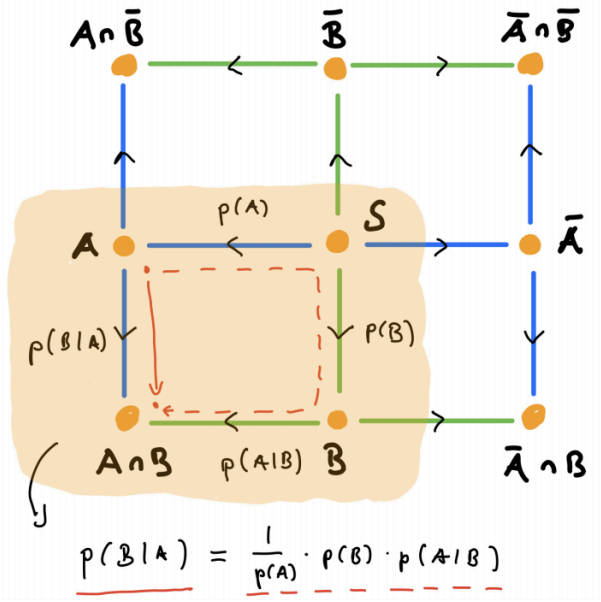
$$p(F \mid E) = \frac{p(F)}{p(E)} \cdot p(E \mid F)$$

- Suppose E and F are events from a sample space S with $p(E) > 0$ and $p(F) > 0$. Then:

$$p(F|E) = \frac{p(F)}{p(E|F) \cdot p(F) + p(E|\overline{F}) \cdot p(\overline{F})} \cdot p(E|F)$$

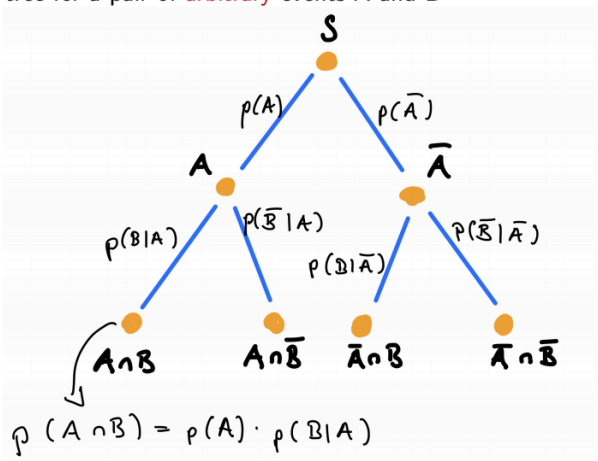


$$P(A|B) \cdot P(B) = P(A \cap B) = P(B|A) \cdot P(A)$$

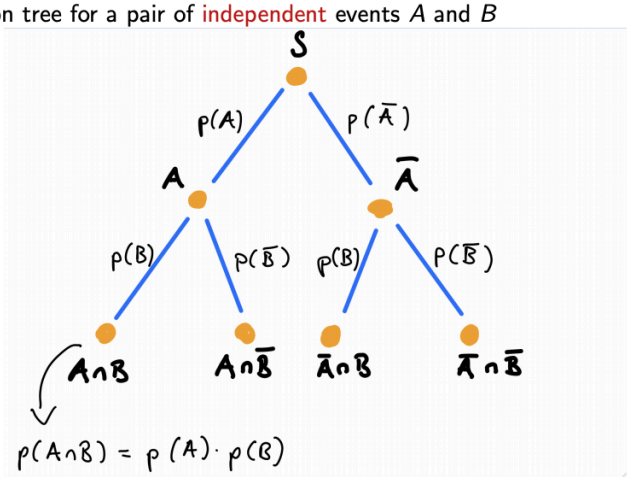


Decision Trees

Decision tree for a pair of arbitrary events A and B



Decision tree for a pair of independent events A and B



Expected Value

- The expected value, also called the expectation or mean, of a random variable X on a sample space S is equal to:

$$E(X) = \sum_{s \in S} p(s)X(s)$$

- Consider the example of throwing two die:

$S = \{(\boxdot, \boxdot), \ldots, (\boxplus, \boxplus)\}$,
$p : S \to [0,1]$, $p(s) = \frac{1}{|S|} = \frac{1}{36}$,
$X : S \to \mathbb{R}$, $X(s) =$ "sum of the values of the two dice in outcome $s \in S$"

expected no. of points on dice

$$E(X) = \sum_{s \in S} p(s)X(s) = \frac{1}{36} \sum_{s \in S} X(s)$$
$$= p((\boxdot, \boxdot)) \cdot 2 + p((\boxdot, \boxdot)) \cdot 3 + p((\boxdot, \boxdot)) \cdot 3 + \ldots$$
$$= \frac{1}{36} \cdot 2 + \frac{1}{36} \cdot 3 + \frac{1}{36} \cdot 3 + \frac{1}{36} \cdot 4 \ldots$$
$$= \frac{1}{36}(2 + 3 + 3 + 4 + 4 + 4 + 5 + 5 + 5 + 5 + \ldots)$$
$$= 7$$

Variance

- The variance of the random variable X on the sample space S is equal to:

$$Var(X) = \sum_{s \in S} (X(s) - E(X))^2 p(s)$$

- Useful Identities:
  - Linearity of Expectation: X, Y random variables, $a, b \in \mathbb{R}$, then:

$$E(X_1 + X_2 + \ldots + X_n) = E(X_1) + E(X_2) + \ldots + E(X_n)$$
$$E(aX + b) = aE(X) + b$$

  - Variance:

$$Var(X) = E(X^2) - E(X)^2 = E((X - E(X))^2)$$

- If X and Y are independent, then:

$$E(XY) = E(X) \cdot E(Y) \text{ and } Var(X + Y) = Var(X) + Var(Y)$$

Relations

- There is a relation between two things if there is a connection between them.
- Recall: A function $f : X \to Y$ is a subset $\Gamma \subseteq X \times Y$ such that for each $x \in X$, there is a unique $y \in Y$ such that $(x, y) \in \Gamma$.
- Let X and Y be sets. A relation R from X to Y is a subset of $X \times Y$. This is a generalisation of a function.
  - We write $(x, y) \in R$ also as $x \, \mathrm{R} \, y$ or $x \sim y$, and say that x is related to y.
  - The complementary relation to R is $\overline{R} = (X \times Y) \setminus R$.
  - If $X = Y$, we say that R is a relation on X.

Three Properties

- Let R be a relation on the non-empty set X. Then:
  - R is reflexive provided that $(x, x) \in R$ for $x \in X$.
  - R is symmetric provided that if $(x, y) \in R$, then $(y, x) \in R$.
  - R is transitive provided that if $(x, y) \in R$ and $(y, z) \in R$, then $(x, z) \in R$.

Combining Relations

- The relations R_1 and R_2 from X to Y are subsets of $X \times Y$. We can use operations on subsets to create new relations:
  - $R_1 \cup R_2$, $R_1 \cap R_2$, $R_1 - R_2$.
- We can compose relations R from X to Y and S from Y to Z to get a new relation $S \circ R$ from X to Z by defining:
  - $S \circ R = \{(a, c) | \exists b \in Y : aRb \wedge bSc\} \subseteq X \times Z$.
- Example: Let X be the set of all people (dead or alive) and the relation R be defined by xRy if x is a parent of y.
  - $R^2 = R \circ R$. x is the grandparent of y.
  - $R^3 = R^2 \circ R$. x is the great-grandparent of y.

Equivalence Relations and Classes

- If the relation R on the non-empty set X is reflexive, symmetric and transitive, then R is an equivalence relation on X.
- If R is an equivalence relation on X and $x \in X$, then the set: $[x] = \{y \in X | (x, y) \in R\}$ is the equivalence class of x.

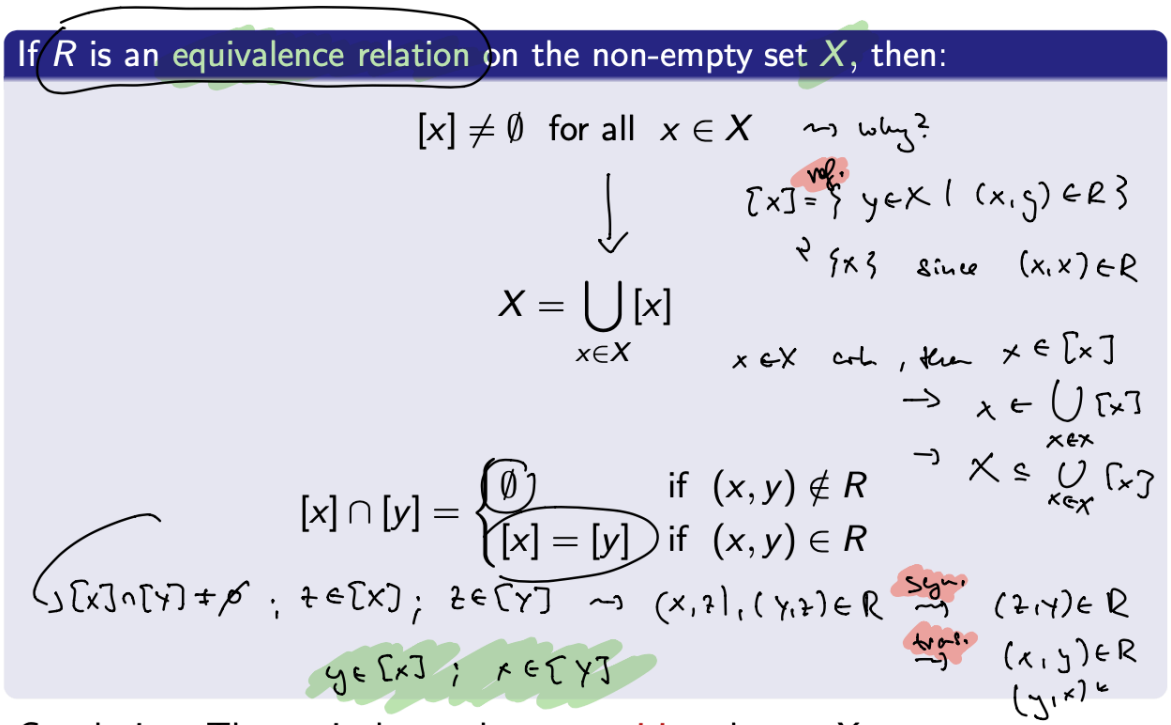
$$(m, n) \in R \text{ if and only if } 3 \mid (m - n)$$

There are three equivalence classes:

$\{\ldots, -6, -3, 0, 3, 6, \ldots\} = [0]$
$\{\ldots, -5, -2, 1, 4, 7, \ldots\} = [1]$
$\{\ldots, -4, -1, 2, 5, 8, \ldots\} = [2]$

$[0] = \{k \in \mathbb{Z} \mid 3 \mid (k-0)\}$
$= \{k \in \mathbb{Z} \mid 3 \mid k\}$
$= \{\ldots, -6, -3, 0, 3, 6, \ldots\}$

$[1] = \{k \in \mathbb{Z} \mid 3 \mid (k-1)\}$
$= \{\ldots, -2, 1, 4, 7, \ldots\}$



Conclusion: The equivalence classes partition the set X.

Partitions

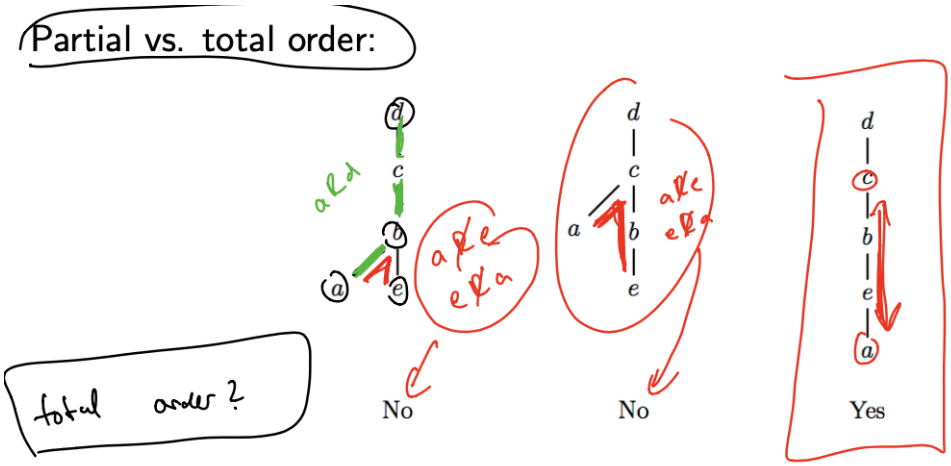
- A set $\{S_1, S_2, \dots\}$ is a partition of S if:
  - $S_i \neq \emptyset$ for all i.
  - $S = S_1 \cup S_2 \cup \dots$
  - $S_i \cap S_j = \emptyset$ whenever $i \neq j$.
- For example: Let E be the set of all even integers, and O be the set of all odd integers. Then $\{E, O\}$ is a partition of $\mathbb{Z}$.
- We saw that: an equivalence relation on X gives a partition of X. Conversely: a partition of X gives an equivalence relation on X.
  - Suppose $\{X_1, X_2, \dots\}$ is a partition of X. Then the relation R defined by $x \,\mathrm{R}\, y \iff \exists i$ such that $x \in X_i$ and $y \in X_i$ is an equivalence relation.
  - In other words: x is related to y if and only if x and y lie in the same set X_i of the partition.

Anti-Symmetric

- Let R be a relation on the set X.
  - The relation R is symmetric if and only if: For all $a, b \in X$, $(a, b) \in R$ implies $(b, a) \in R$.
  - The relation R is anti-symmetric if and only if: For all $a, b \in X$, $(a, b) \in R$ and $(b, a) \in R$ implies $a = b$.
- Anti-symmetric does not equal non symmetric.

Partial and Total Orders

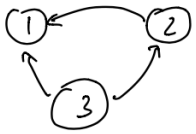
- A relation R on a set X which is reflexive, transitive and anti-symmetric is called a partial order on X.
- If in addition, for all $a, b \in X$, aRb or bRa, then R is called a total order on X.



Representing Relations

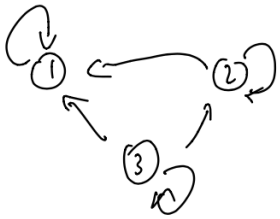
- Most relations (especially for infinite sets) are defined implicitly. For small sets, we can describe them using digraphs (directed graphs).

Example: $X = \{1, 2, 3\}$, and xRy if and only if $x > y$.



$R = \{(2,1), (3,1), (3,2)\}$

Example: $X = \{1, 2, 3\}$, and xRy if and only if $x \geq y$.



$R = \{(1,1), (2,2), (3,3), (2,1), (3,1), (3,2)\}$

Making a Relation Reflexive

- Recall that R is reflexive if $\forall x \in X, (x, x) \in R$. Suppose that $S \subseteq X \times X$ in some relation. The smallest reflexive relation containing S is $S \cup \{(x, x) | x \in X\}$.
- This is the reflexive closure $ref(S)$ of S. $\Delta = \{(x, x) | x \in X\}$ is the diagonal relation.
- For a digraph representation: add a loop at each vertex.

Making a Relation Symmetric

- Recall that R is symmetric if $(x, y) \in R$ implies that $(y, x) \in R$. Suppose that $S \subseteq X \times X$ is some relation. The smallest symmetric relation containing S is $S \cup \{(y, x) | (x, y) \in S\}$.
- This is the symmetric closure $sym(S)$ of S. $S^{-1} = \{(y, x) | (x, y) \in S\}$ is the inverse relation to S.

- For a digraph representation: add edges in the opposite direction.

Making a Relation Transitive

- For a digraph representation: If there is a path from x to y, add an edge from x to y.

Let R be a relation on a set X.

There is a path of length k from x to y in R if there are $x_0, \ldots, x_k$ such that $x_0 R x_1, \ldots, x_k R y$.

Claim: There is a path of length k from x to y if and only if $xR^k y$.

How to you prove this? Induction!

Base case: $k=1$; $xRy \hookrightarrow xR^1y$

Induction hypothesis: Suppose for $k \geq 1$: $\exists$ path of len. k from x to y $\hookrightarrow xR^k y$

Induction step: to prove: $xR^{k+1}y \Leftrightarrow \exists$ path of len. $k+1$ from x to y.

$\Rightarrow xR^{k+1}y \hookrightarrow \exists z \in X : xRz \wedge zR^k y$ → $\exists$ path of len. $k+1$ from z to y; $\exists$ path of len. 1 from x to z

$\hookrightarrow$ $\exists$ path of len. $k+1$ from x to y (Hence, the claim is true (!)) ← ANSWER SENTENCE

Define $R^\star = \bigcup_{k=1}^{\infty} R^k$. This is the connectivity closure.

Fact 1: The connectivity closure $R^\star$ is the transitive closure $\text{tra}(R)$.

Fact 2: If X has only n elements, then $R^\star = \bigcup_{k=1}^{n} R^k = \bigcup_{k=1}^{n+1} R^k$.

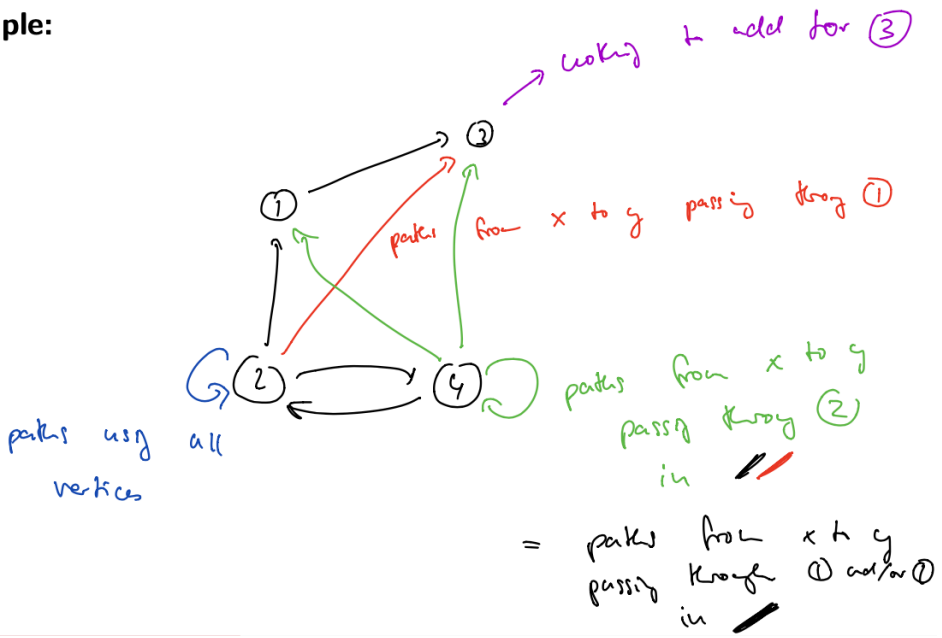
↳ for R relation on finite set, taking unions $R \cup R^2 \cup R^3 \cup \ldots$ will give us $\text{tra}(R)$ after finitely many steps.

Warshall's Algorithm to Compute Transitive Closure

- Input: Relation R on set with n elements.
- Output: Transitive closure $\text{tra}(R)$.
- Main idea: A path exists between vertices x and y if and only if:
  - there is an edge from x to y; or
  - there is a path from x to y going through vertex 1; or

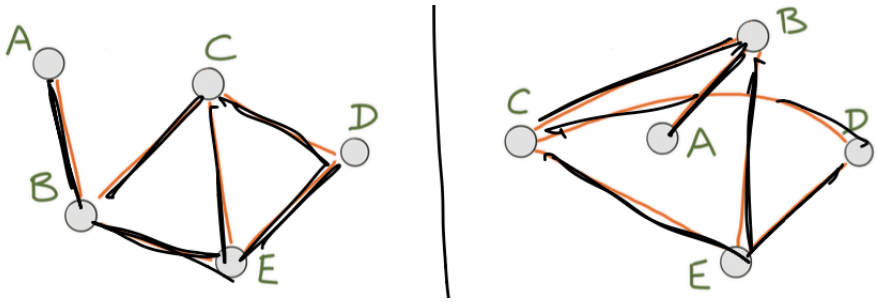
- there is a path from x to y going through vertex 1 and/or 2; or
- there is a path from x to y going through vertex 1 and/or 2 and/or 3; or ...
- there is a path from x to y going through any of the other $n-2$ vertices.

Example:



Graphs

- A graph G consists of two finite sets:
  - a non-empty set $V(G)$ of vertices;
  - a (possibly empty) set $E(G)$ of edges, where each edge is associated with a set $\{v, w\} \subseteq V(G)$. The vertices v and w are called the endpoints of the edge.
- The graph below has five vertices $V(G) = \{A, B, C, D, E\}$. There are six edges, and their endpoints are: $\{A, B\}, \{B, C\}, \{B, E\}, \{C, E\}, \{C, D\}, \{D, E\}$.



- A graph G is defined purely in terms of sets $V(G)$ and $E(G)$. The drawing is just a visual aid.
  - Finding a "good" drawing for a complex graph is difficult.
- An edge may have endpoints $\{v, v\} = \{v\}$. We call such an edge a loop.
- Two edges may have the same endpoints $\{v, w\}$. We call these parallel edges.
- A graph with no loops or parallel edges is called a simple graph.
- An edge e and a vertex v are incident if v is an endpoints of e.
- Vertices u, v are adjacent if there is an edge with endpoints $\{u, v\}$. A vertex u is adjacent to itself if there is a loop with endpoints $\{u\}$.
- The degree of a vertex v is the number of edges incident with v, where we count each loop twice. We write this as deg(v). Informally, deg(v) counts the ends of edges that meet v.

The Handshake Theorem

- Let G be a graph with n vertices $V(G) = \{v_1, \ldots, v_n\}$. Then:

$$\sum_{i=1}^{n} deg(v_i) = deg(v_1) + \ldots + deg(v_n) = 2 \cdot |E(G)|$$

  - In particular, in any graph, the sum of all vertex degrees must be even.
  - In any graph, the number of vertices of odd degree is even.

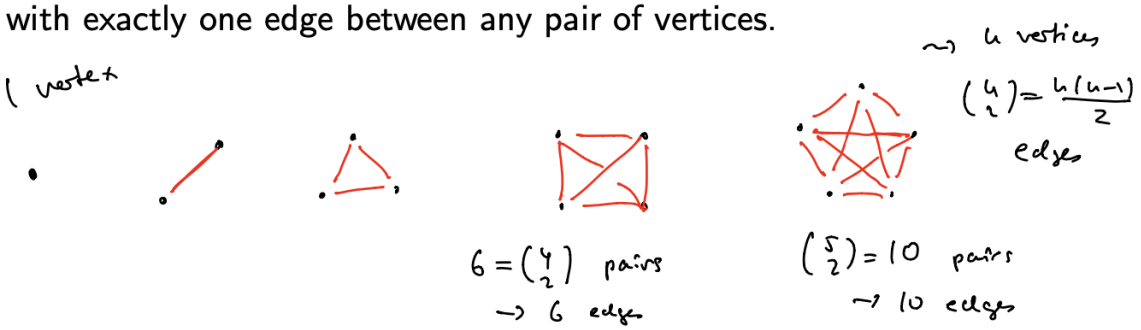
Directed Graphs

- Let G be a digraph (directed graph) and $v \in V(G)$. $(a, b) \in E(G)$ is an ordered pair, where $E(G)$ is a relation on $V(G)$ - $E(G) \subseteq V(G) \times V(G)$.
- The in-degree $\deg^-(v)$ is the number of edges terminating in v.
- The out-degree $deg^+(v)$ is the number of edges starting in v.
- The handshake theorem is as follows: Let G be a directed graph with n vertices $V(G) = \{v_1, \ldots, v_n\}$. Then:

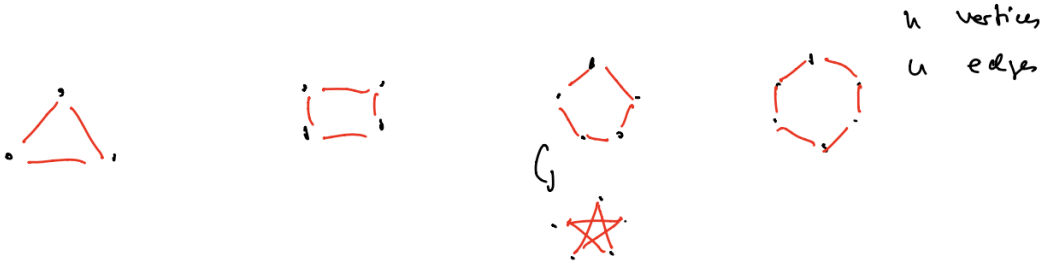
$$\sum_{i=1}^{n} deg^-(v_i) = \sum_{i=1}^{n} deg^+(v_i) = |E(G)|$$

Types of Graphs

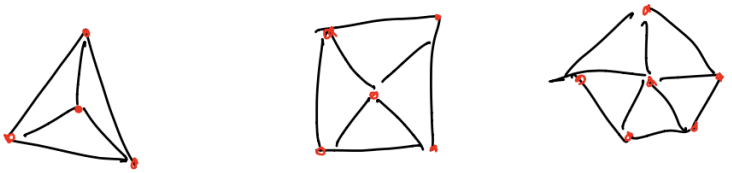
Complete graphs: The complete graph on n vertices is a simple graph with exactly one edge between any pair of vertices.



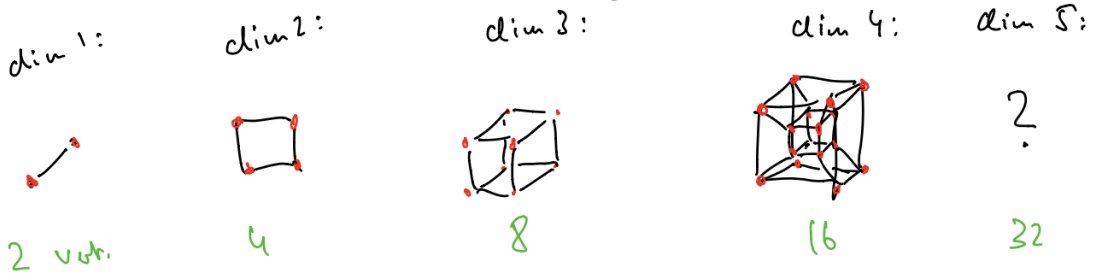
Cycles: A cycle C_n for $n \geq 3$ is a graph that looks like a loop:



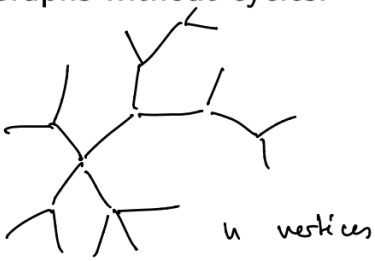
Wheels: You get a wheel W_n from the cycle C_n by adding a vertex that connects to each of the vertices:



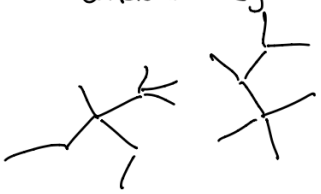
Cubes: An n–cube Q_n consists of the edges of an n–dimensional cube:



Trees: ~~Graphs~~ Connected Graphs without cycles.

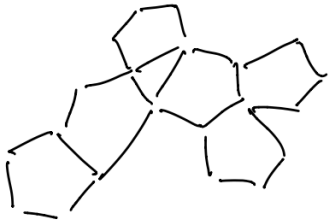


Possibly disc. graphs without cycles: Forests



trees = n − |E|

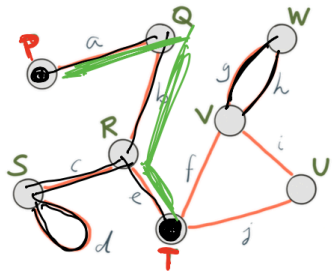
Cactus graphs:



Paths

- Let G be a graph, and let $x, y \in V(G)$. A path in G from x to y is an alternating sequence of vertices and edges $v_0, e_1, v_1, e_2, v_2, \ldots, v_{n-1}, e_n, v_n$ where $v_0 = x$, $v_n = y$ and each edge e_i has endpoints $\{v_{i-1}, v_i\}$.
- Informally, we "walk" through the graph from vertex x to vertex y, following edges $e_1, e_2, \ldots, e_n$ in turn.
- Remark: A path in a digraph has the extra requirement that e_i is directed from v_{i-1} to v_i.

Example: Consider the following graph G.



Vertices: $V(G) = \{P, Q, R, S, T, U, V, W\}$
Edges: $E(G) = \{a, b, c, d, e, f, g, h, i, j\}$
An example of a path from T to P is:

T e R c S d S c R b Q a P

T e R b Q a P

shortest path

- A graph G is called connected if, for all vertices $x, y \in V(G)$, there is a path from x to y. Otherwise G is called disconnected.
- A path is called a circuit if it doesn't repeat an edge, and it starts and ends at the same vertices.

Eulerian Circuit

- A Eulerian circuit is a path that starts and ends at the same vertex and uses every edge exactly once.
  - If we ignore any isolated vertices (vertices with degree 0), then the remaining graph must be connected. Every edge and vertex is part of the Eulerian circuit.
  - The degree of every vertex must be even.
- Let G be a connected graph. Then G has a Eulerian circuit if and only if every vertex of G has an even degree.
- If two circuits pass through the same vertex v, we can splice them together at v. Splicing circuits eventually forms a single large Eulerian circuit.
- The bridges of Konigsberg do not have Eulerian trail since four vertices have uneven degrees.

Hamiltonian Circuits

- Eulerian circuit is a circuit using every edge exactly once.
- A Hamiltonian circuit or Hamiltonian cycle is a circuit using every vertex exactly once (except for the start and end vertex which appears twice).
  - A Hamiltonian circuit can be verified in $O(n)$, however, it is a NP-complete problem to solve.

Graph Isomorphism

- Two graphs $G_1 = (V_1, E_1)$ and $G_2 = (V_2, E_2)$ are said to be isomorphic (or equal), written $G_1 \cong G_2$, if there exists a bijective function $\phi : V_1 \to V_2$ such that:

$$\phi(E_1) = \{\{\phi(v_1), \phi(v_2)\} | \{v_1, v_2\} \in E_1\} = E_2$$

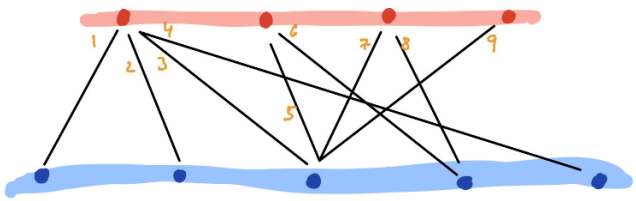
- Given two graphs G_1, G_2, are they isomorphic (does $G_1 \cong G_2$ hold)? The question is known as the graph isomorphism problem. It is famous for being "probably not difficult" to answer in general.

Representing Graphs with Matrices

- Let G be a graph with n vertices, and suppose we label these vertices $V(G) = \{1, 2, \ldots, n\}$.
- The adjacency matrix of G is the $n \times n$ matrix $A = (a_{i,j})$, where each entry $a_{i,j}$ is the number of edges with endpoints $\{i, j\}$ counted with multiplicity.
- Let G be a graph with vertices $V(G) = \{1, 2, \ldots, n\}$ and adjacency matrix A. Then the number of paths of length k from vertex i to vertex j is the entry in row i, column j of the kth power $A^k = A \cdot A \cdot \ldots \cdot A$.
    - Note that every loop edge contributes 2 to the respective diagonal entry of the adjacency matrix.
    - In the above theorem, this means we count paths with loop edges twice, once per orientation of the loop edge $\to$ a single path with k loop edges is counted 2^k times.
    - However: in our definition of a path, these paths are all considered to be the same.
    - We can remove this ambiguity by only adding 1 per loop edge on the diagonal entry.

Bipartite Graphs

- The simple graph G is bipartite if it has at least two vertices and satisfies one (and hence all) of the following equivalent conditions:
    1. The set of vertices $V(G)$ has a partition $\{V_1, V_2\}$ such that every edge is of the form $\{v_1, v_2\}$ where $v_k \in V_k$.
    2. The vertices can be coloured with two colours such that no two adjacent vertices have the same colour.
    3. Every circuit in G has even length.



- An example application of bipartite graphs is timetabelling, where V_1 = students and V_2 = units of study.

1 implies 2

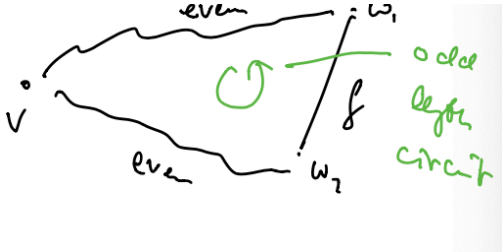
- Partition $V(G)$ into $\{V_1, V_2\}$ such that all edges are in $V_1 \times V_2$. Colour vertices in V_1 red and vertices in V_2 blue.
- 2-colouring of vertices such that no two adjacent vertices have the same colour.

2 implies 3

- 2-colouring of vertices such that no two adjacent vertices have the same colour. Along every circuit colours must alternate. All circuits have even length.

3 implies 1

- Every circuit in G has even length $\rightarrow V(G)$ has a partition $\{V_1, V_2\}$ such that every edge is of the form $\{v_1, v_2\}$ where $v_k \in V_k$.
  - Isolated vertices and multiple components is a trivial proof.
  - Hence, assume G is connected with at least one edge.
  - Fix $v \in V(G)$:
    - $V_1 = \{w \in V(G) | \exists \text{ path of odd length between v and w}\}$
    - $V_2 = \{w \in V(G) | \exists \text{ path of even length between v and w}\}$
  - $V_2 \neq \emptyset$ because $v \in V_2$.
  - $V_1 \neq \emptyset$ because G is connected and has at least two vertices.
  - $V_1 \cup V_2 = V(G)$ since G is connected.
  - $V_1 \cap V_2 = \emptyset$ because $w \in V_1 \cap V_2 \rightarrow G$ contains odd length circuit.
  - Moreover: edge with both endpoints in $V_1 (V_2)$ implies G contains an odd length circuit.



Hall's Marriage Theorem

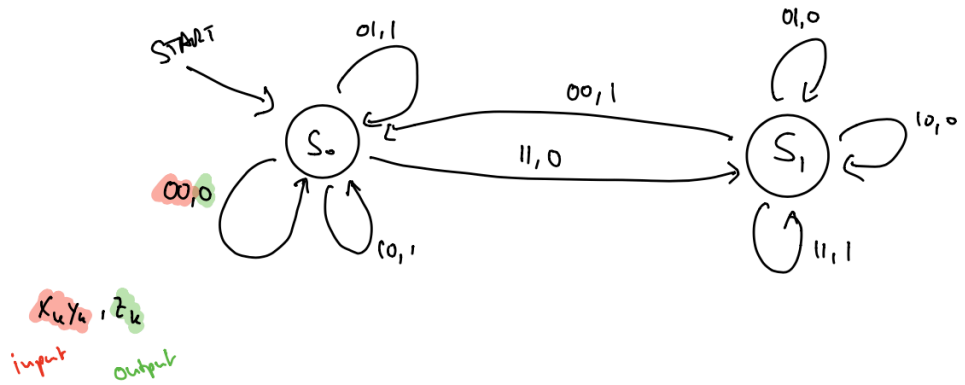
- Let G be a graph and $v \in V(G)$.
  - The neighbourhood $N(v)$ is the set of all vertices adjacent to v.
  - The neighbourhood of $A \subseteq V(G)$ is $N(A) = \bigcup_{v \in A} N(v)$.
- A matching in the bipartite graph G is a subset $M \subseteq E(G)$ with the property that no two edges in M share a vertex. A matching is a complete matching from V_1 to V_2 if every vertex in V_1 is incident with an edge in M. Equivalently, if $|M| = |V_1|$.
- **Hall's Marriage Theorem**: Let G be a bipartite graph with partition $\{V_1, V_2\}$ of the vertices. There is a complete matching from V_1 to V_2 if and only if $|A| \leq |N(A)|$ for all $A \subseteq V_1$.
  - $A \subset V_1$ such that $|N(A) < |A|$ is called a Hall violator. Prove that:

1. A complete matching M from V_1 to $V_2 \rightarrow$ no Hall violator.
2. There exists a Hall-violator or a complete matching (use maximal alternating paths).

Models of Computation

- A key property of a computer is that it is programmable.
- First programmable "machine": Analytical engine due to Charles Babbage (1837). First computer program: Computes Bernoulli numbers due to Ada Lovelace (1843).
- Modelling binary additon as a finite state machine:

$s_0 =$ "remember previous carry is 0"
$s_1 =$ "remember previous carry is 1"



Finite State Machines

- A finite state machine $M = (S, I, O, f, g, s_0)$ consists of:
  - a finite set S of states e.g. $\{S_0, S_1\}$
  - a finite input alphabet I e.g. $\{00, 01, 10, 11\}$
  - a finite output alphabet O e.g. $\{0, 1\}$
  - a transition function $f : S \times I \rightarrow S$ e.g. arrows of diagram above
  - an output function $g : S \times I \rightarrow O$ e.g. arrows of diagram, described differently
  - an initial state s_0 e.g. S_0

Formal Languages

- Communication with finite state machine through input/output.
- String of symbols (e.g. 0 and 1) with a certain structure.
- This is called a Formal language, which is essential for:
  - designing programming languages
  - building compilers
  - complexity theory (analysing difficulty of computational tasks)

- The way a speaker can generate new sentences can be described using either Functional syntax or behaviour theory (Piaget) or Generative syntax theory (Chomsky).
    - Chomsky's theory: Language is only partially learned.
        - There is a universal or generative grammar - principles valid for all languages that we are born with.
        - We can judge sentences correct because we possess an abstract system of unconscious knowledge about language.
- **Formal Languages**: Let A be a finite alphabet. A formal language L is a set of strings with symbols in A. The empty string is denoted λ.
    - E.g. $A = \{a, b, \ldots, z\}$, $L =$ all English words.

Grammars

- How to describe an infinite language with a finite amount of resources?
- Use grammars as formalised and categorised by Chomsky (1957).
- Idea: Grammar of language L will be dvice that generates all the grammatical sentences of L and none of the ungrammatical sentences.
- Aside: Chomsky aimed (and failed) to define natural languages. His work became an important and useful tool in theoretical computer science and mathematics.

Phase-Structure Grammar

- A phase-structuer grammar $G = (V, T, S, P)$ consists of:
    - a vocabulary (or alphabet) V,
    - a subset $T \subset V$ of terminal symbols,
    - a start symbol $S \in V$,
    - a finite set of productions P.
- The non-terminal symbols are $N = V - T$. Every production in P must have at least one non-terminal symbol on its left side.
- The language generated by G is the set $L(G)$ of all words in terminals that can be derived by S using the productions P.
- The set of all sentences (or words) over V is V^*. So $L(G) \subseteq T^* \subseteq V^*$.

$G = (V, T, S, P)$ where
$V = \{0, 1, 2, S, B, C, H\}$
$T = \{0, 1, 2\}$,

$$P = \{S \to 0BC,\ S \to 0SBC,$$

$$CB \to BC,$$

$$0B \to 01,\ 1B \to 11,\ 1C \to 12,\ 2C \to 22\}.$$

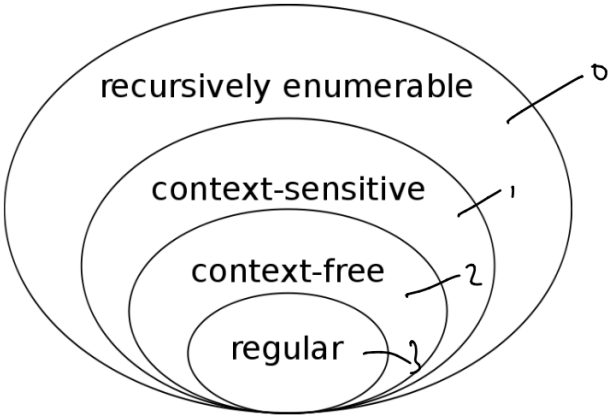
Let's derive a word in the language $L(G)$:

$S \to 0SBC \to 00SBCBC \to 000BCBCBC \to 000BBCCBC$

$\to 000BBBCCC \to 0001BBCCC \to 000111CCC \to 000111222$

$\to \ldots \to 000111222 \rightsquigarrow L(G) = \{0^n 1^n 2^n \mid n > 0\}$

Types of Grammars: The Chomsky Hierarchy



Type 0: no restrictions
Type 1: context-sensitive

$L = \{0^n 1^n 2^n \mid n > 0\}$ has a context-sensitive grammar

Type 2: context-free

$L = \{0^n 1^n \mid n > 0\}$ has a context-free grammar

Type 3: regular

$L = \{0^n 1^m \mid m, n > 0\}$ has a regular grammar

Type	Restrictions on productions
0	No restrictions
1	Either $lAr \to lwr$ where $A \in N$ non-terminal, and $l, r, w \in V^*$ arbitrary words over V, $w \neq \emptyset$; or $S \to \emptyset$ and S cannot be the right hand side of another production.
2	$A \to w$ where $A \in N$ non-terminal, and $w \in V^*$ arbitrary
3	$A \to aB$ or $A \to a$ where $A, B \in N$ non-terminal and $a \in T$ terminal; or $S \to \emptyset$.