CMPE 322/327 - Theory of Computation

Week 1: Central Concepts of Automata Theory & Mathematical Preliminaries

Burak Ekici

February 21-25, 2022

Outline

Definition (Sets)

- A set is a collection of objects

$$
\begin{array}{lll}
A & = & \{1, 2, 3\} \\
B & = & \{\text{bicycle, bus, train, airplane}\} \\
1 & \in & A \\
\text{ship} & \notin & B
\end{array}
$$

1 is an element of the set A

ship is not an element of the set B

Example (Representation of Sets)

$$
\begin{array}{lll}
C & = & \{a, b, c, d, e, f, g, h, i, j, k\} \\
C & = & \{a, b, \cdots, k\} \\
S & = & \{2, 4, 6, \cdots\} \\
S & := & \{j \in \mathbb{Z} \mid j > 0 \text{ and } j = 2k \text{ for some } k > 0\} \\
S & := & \{j \mid j \text{ is a positive and even integer}\}
\end{array}
$$

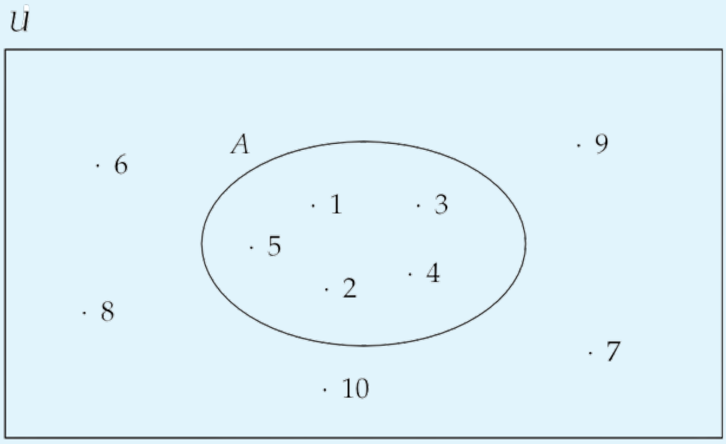
C is a finite set

S is an infinite set

Definition (Diagrammatic Representation of Sets (Venn Diagrams))

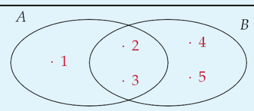
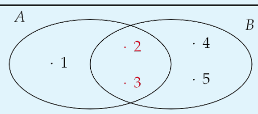
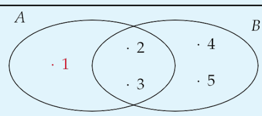
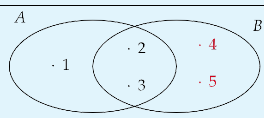
$A = \{1, 2, 3, 4, 5\}$

$U = \{1, 2, \cdots, 10\}$ U is a universal set (set of all elements under consideration)



Definition (Basic Set Operations)

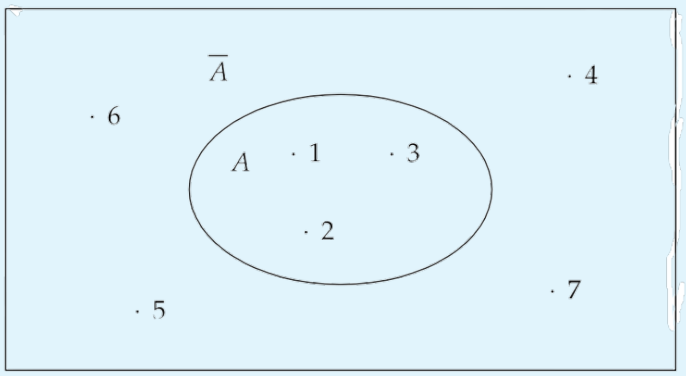
$A = \{1, 2, 3\}$ $B = \{2, 3, 4, 5\}$

Operation	Notation	Venn Diagram
Union	$A \cup B := \{x \mid x \in A \vee x \in B\} = \{1, 2, 3, 4, 5\}$	
Intersection	$A \cap B := \{x \mid x \in A \wedge x \in B\} = \{2, 3\}$	
Difference	$A - B := \{x \mid x \in A \wedge x \notin B\} = \{1\}$	
	$B - A := \{x \mid x \in B \wedge x \notin A\} = \{4, 5\}$	

Definition (Basic Set Operations (cont'd))

$$\begin{array}{lll} U & = & \{1, 2, \cdots, 7\} \\ A & = & \{1, 2, 3\} \\ \overline{A} & := & \{x \mid x \notin A \wedge x \in U\} = \{4, 5, 6, 7\} \end{array}$$

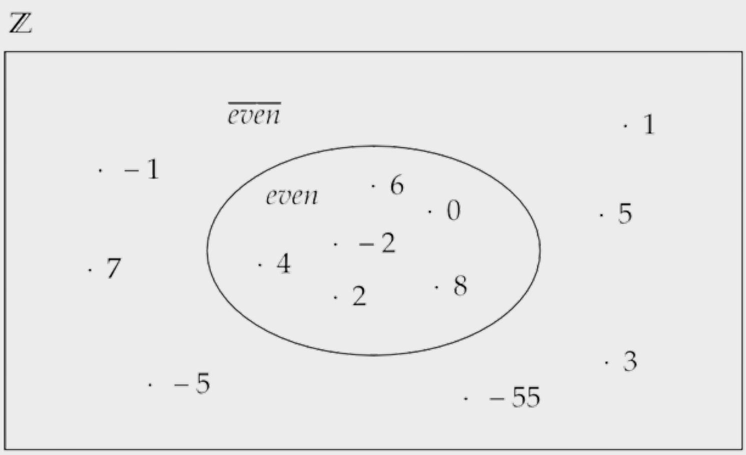
$\overline{A}$ is the complement of A with respect to U



$$\overline{(\overline{A})} = A$$

Example (Complement)

- The complement set of even integers $\overline{\{\text{even integers}\}}$:



Theorem

$\overline{(\overline{A})} = A$

Proof.

$$\begin{aligned} \overline{(\overline{A})} &:= \{x \mid x \notin \overline{A} \text{ and } x \in U\} && \text{by definition of complement} \\ &= \{x \mid x \in A \text{ and } x \in U\} \\ &= A \end{aligned}$$

Theorem (De Morgan Laws)

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

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

Theorem

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

Proof.

$$
\begin{array}{lll}
\overline{A \cup B} & := & \{x \mid x \notin (A \cup B)\} \quad \text{by definition of complement} \\
& = & \{x \mid x \notin A \text{ and } x \notin B\} \\
& = & \{x \mid x \in \overline{A} \text{ and } x \in \overline{B}\} \\
& = & \overline{A} \cap \overline{B} \quad \text{by definition of intersection}
\end{array}
$$

Sets Relations Functions Graphs Trees Proof Techniques Alphabets & Strings Languages

Theorem

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

Proof.

$$
\begin{array}{lll}
\overline{A \cap B} & := & \{x \mid x \notin (A \cap B)\} \quad \text{by definition of complement} \\
& = & \{x \mid x \notin A \text{ or } x \notin B\} \\
& = & \{x \mid x \in \overline{A} \text{ or } x \in \overline{B}\} \\
& = & \overline{A} \cup \overline{B} \quad \text{by definition of union}
\end{array}
$$

Theorem

$\overline{A} - \overline{B} = B - A$

Proof.

$$
\begin{array}{lcll}
\overline{A} - \overline{B} & := & \{x \mid x \in \overline{A} \text{ and } x \notin \overline{B}\} & \text{by definition of complement} \\
& = & \{x \mid x \notin A \text{ and } x \in B\} & \\
& = & \{x \mid x \in B \text{ and } x \notin A\} & \\
& = & B - A & \text{by definition of difference}
\end{array}
$$

Theorem

$\overline{B} - \overline{A} = A - B$

Proof.

$$
\begin{array}{lcll}
\overline{B} - \overline{A} & := & \{x \mid x \in \overline{B} \text{ and } x \notin \overline{A}\} & \text{by definition of complement} \\
& = & \{x \mid x \notin B \text{ and } x \in A\} & \\
& = & \{x \mid x \in A \text{ and } x \notin B\} & \\
& = & A - B & \text{by definition of difference}
\end{array}
$$

Definitions (Empty (Null) Set)

- The empty set, denoted $\varnothing$ (or $\{\}$), is the unique set having no elements
- It satisfies following properties:

$$\begin{aligned} S \cup \varnothing &= S \\ S \cap \varnothing &= \varnothing \\ S - \varnothing &= S \\ \varnothing - S &= \varnothing \\ \overline{\varnothing} &= U \end{aligned}$$

Definitions (Subsets)

- A set A is a subset of a set B if all elements of A are also elements of B; B is then called a superset of A

$$A = \{1, 2, 3, 4, 5\} \qquad B = \{1, 2, 3, 4, 5\} \qquad A \subseteq B$$

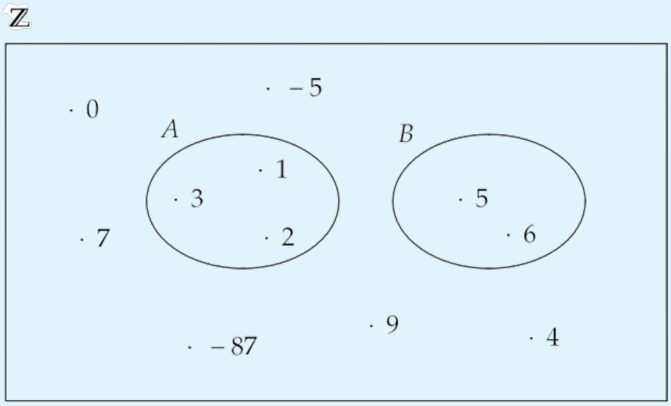
- A subset A of some set B is called a proper subset if A is not the same as B (i.e. there exists at least one element in B that does not appear in A)

$$A = \{1, 2, 3\} \qquad B = \{1, 2, 3, 4, 5\} \qquad A \subset B$$

Definition (Disjoint Sets)

- Two sets A and B are called disjoint if they have no common element

$$A = \{1,2,3\} \qquad B = \{5,6\} \qquad A \cap B = \emptyset$$



Definitions (Power Sets)

- A power set of some set S (denoted 2^S) is the set of all subsets of S

$$\begin{aligned} S &= \{a,b,c\} \\ 2^S &= \{\emptyset, \{a\}, \{b\}, \{c\}, \{a,b\}, \{a,c\}, \{b,c\}, \{a,b,c\}\} \end{aligned}$$

- Observe that the number of elements in 2^S amount to the 2 to the number of elements in S:

$$|2^S| = 2^{|S|}$$

Definition (Cartesian Product of Sets)

The Cartesian product of two sets A and B, denoted $A \times B$, is the set of all ordered pairs (a, b) such that $a \in A$ and $b \in B$. That formally is

$$A \times B := \{(a, b) \mid a \in A \text{ and } b \in B\}$$

Example

$$A = \{2, 4\} \qquad B = \{2, 3, 5\}$$
$$A \times B = \{(2, 2), (2, 3), (2, 5), (4, 2), (4, 3), (4, 5)\}$$

- Remark also that Cartesian products generalize (to more than two sets)

$$A_1 \times A_2 \times \cdots \times A_n.$$

Formalism (Frege's Theory)

Frege's Theory has two axioms:

- The Axiom of Unrestricted Comprehension: $\exists x \forall z [z \in x \leftrightarrow \phi(z)]$
- The Axiom of Extensionality: $\forall x \forall y [x = y \leftrightarrow (\forall z (z \in x \leftrightarrow z \in y))]$

Theorem

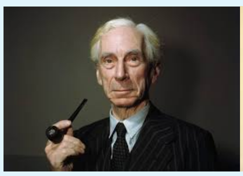
Frege's Theory is inconsistent

Proof.

1. $\phi(z) := z \notin z$
2. By Unrestricted Comprehension, we have:

$$\exists x \forall z [z \in x \leftrightarrow z \notin z]$$

3. $x \in x \leftrightarrow x \notin x$ – Russell's Paradox– This is not a pipe




The Axiom of Unrestricted Comprehension
$\exists x \forall z [z \in x \longleftrightarrow \phi(z)]$

The Axiom of Restricted Comprehension
$\forall y \exists x \forall z [z \in x \longleftrightarrow (z \in y \text{ and } \phi(z))]$

Remarks

1. Given some set y, the axiom of restricted comprehension only guarantees the existence of the subset x consisting of those elements of y that satisfy ϕ
2. Impossible to construct the set of all sets satisfying certain property
3. Axioms of Pairing, Extensionality and Foundation avoids having $\forall x,\ x \in x$
4. ZFC := Axioms of Restricted Comprehension, Pairing, Extensionality, Foundation + 6 other axioms
5. We silently consider sets in ZFC within the scope of this course (to avoid Russell-like paradoxes)

Outline

Definition (Binary Relations)

A binary relation R over sets A and B is a subset of the Cartesian product $A \times B$

$$R \subseteq A \times B$$

Example

$$\begin{aligned} M_5 &:= \{(m, n) \mid (m, n) \in \mathbb{N} \times \mathbb{N} \text{ and } m \equiv_5 n\} \\ M_5 &= \{(0, 0), (0, 5), (0, 10), \ldots, (5, 0), (5, 5), (5, 10), \ldots\} \end{aligned}$$

Definition (Equivalence Relations)

A binary relation R over some set A ($R \subseteq A \times A$) is said to be an equivalence relation if and only if it is *reflexive*, *symmetric* and *transitive* such that

$$\begin{array}{ll} \forall a \in A,\ (a, a) \in R & \text{reflexivity} \\ \forall a \in A,\ \forall b \in A,\ (a, b) \in R \implies (b, a) \in R & \text{symmetry} \\ \forall a \in A,\ \forall b \in A,\ \forall c \in A,\ \big((a, b) \in R\ \wedge\ (b, c) \in R\big) \implies (a, c) \in R & \text{transitivity} \end{array}$$

Theorem

M_5 is an equivalence relation.

Proof.

We need to demonstrate that M_5 is reflexive, symmetric and transitive:

1. reflexivity: for every $m \in \mathbb{N}$, the remainder when divided by 5 is unique. Thus, $(m, m) \in M_5$ applies.
2. symmetry: If $(m, n) \in M_5$ then $m \equiv_5 n$, we consequently get $n \equiv_5 m$ and thus $(n, m) \in M_5$.
3. transitivity: from $(m, n) \in M_5$ and $(n, p) \in M_5$ we get $m \equiv_5 n$ and $n \equiv_5 p$, which is why $m \equiv_5 p$ and thus $(m, p) \in M_5$.

□

Outline

Definitions (Functions)

- A binary relation F over sets A and B is called a partial function if it is *right-unique* such that

$$\forall a \in A, \forall b_1 \in B, \forall b_2 \in B, \big((a, b_1) \in F \ \wedge \ (a, b_2) \in F\big) \implies b_1 = b_2 \qquad \text{right-unique}$$

- A partial function F over sets A and B is called a total function if it is *left-total* such that

$$\forall a \in A, \exists b \in B, (a, b) \in F \qquad \text{left-total}$$

Notation

- By convention, we write

$$\begin{array}{ll} F : A \twoheadrightarrow B & \text{if } F \subseteq A \times B \text{ is partial} \\ F : A \rightarrow B & \text{if } F \subseteq A \times B \text{ is total} \\ y = F(x) & \text{for } (x, y) \in F \end{array}$$

- In this lecture, the keyword “*function*” refers to “*total function*”.

Example (Functions)

A	$=$	$\{1, 2, 3\}$	B	$=$	$\{a, b, c, d\}$		
f	$:$	$A \twoheadrightarrow B$	f	$=$	$\{(2, d), (3, c)\}$	is f a function?	yes, f is a partial function
f	$\subseteq$	$A \times B$	f	$=$	$\{(2, d), (3, c), (2, a)\}$	is f a function?	no
f	$:$	$A \rightarrow B$	f	$=$	$\{(2, d), (3, c), (1, c)\}$	is f a function?	yes, f is a total function
f	$\subseteq$	$A \times B$	f	$=$	$\{(2, d), (3, c), (3, a)\}$	is f a function?	no
f	$:$	$A \twoheadrightarrow B$	f	$=$	$\{(1, a), (3, d)\}$	is f a function?	yes, f is a partial function

Lemma

The relation $f := \{(x, y) \mid (x, y) \in \mathbb{N} \times \mathbb{N} \text{ and } y = x + 1 \text{ for all } x \geqslant 10\}$ is a partial function.

Proof.

We are supposed to show that f is right-unique but not left-total:

1. right-unique: for all $a \geq 10$, from $(a, b_1) \in f$ and $(a, b_2) \in f$, we obtain $b_1 = a + 1$ and $b_2 = a + 1$. It is then obvious that $b_1 = b_2$. Therefore, f obeys right-uniqueness.
2. left-total: $\forall a \in \mathbb{N}$, $0 \leq a < 10$, $\nexists b \in \mathbb{N}$, $(a, b) \in f$. Thus, f does not satisfy left-totality.

$\square$

Lemma

The relation $f := \{(x, y) \mid (x, y) \in \mathbb{N} \times \mathbb{N} \text{ and } y = x + 1\}$ is a total function.

Proof.

We are supposed to show that f is right-unique and left-total:

1. right-unique: for all a, from $(a, b_1) \in f$ and $(a, b_2) \in f$, we obtain $b_1 = a + 1$ and $b_2 = a + 1$. It is then obvious that $b_1 = b_2$. Therefore, f obeys right-uniqueness.
2. left-total: $\forall a \in \mathbb{N}$, there exists $b = a + 1$ such that $(a, a + 1) \in f$. This gives $a + 1 = a + 1$ which definitely holds. Thus, f does satisfy left-totality.

$\square$

Definitions (Injection & Surjection)

- A function $f: A \to B$ is an injection (or one-to-one) if

$$\forall a_1 \in A, \forall a_2 \in A, f(a_1) = f(a_2) \implies a_1 = a_2 \quad \text{or}$$
$$\forall a_1 \in A, \forall a_2 \in A, a_1 \neq a_2 \implies f(a_1) \neq f(a_2) \quad \text{by logical contra-position}$$

- A function $f: A \to B$ is a surjection (or onto) if

$$\forall b \in B, \exists a \in A, b = f(a)$$

- A function $f: A \to B$ is a bijection (or both one-to-one and onto) if

$$\forall b \in B, \exists! a \in A, b = f(a)$$

Theorem

$\exists f: \mathbb{N} \to \mathbb{Z}$, f is a bijection.

Proof.

We pick f to be

$$f(a) := \begin{cases} \dfrac{a}{2} & \text{if } a \text{ in even} \\ \dfrac{-(a+1)}{2} & \text{if } a \text{ is odd} \end{cases}$$

1. f is an inversion:
$\forall a_1, a_2 \in \mathbb{N}, f(a_1) = f(a_2) \implies a_1 = a_2$.
Given $f(a_1) = f(a_2)$
   - case 1: $f(a_1) = f(a_2) \geqslant 0$
   a_1 and a_2 are even.
   $f(a_1) = \dfrac{a_1}{2} = \dfrac{a_2}{2} = f(a_2) \implies a_1 = a_2$
   - case 2: $f(a_1) = f(a_2) < 0$
   a_1 and a_2 are odd.
   $f(a_1) = \dfrac{-(a_1+1)}{2} = \dfrac{-(a_2+1)}{2} =$
   $(a_1+1) = (a_2+1) = f(a_2) \implies a_1 = a_2$

2. f is a surjection: $\forall b \in \mathbb{Z}, \exists a \in \mathbb{N}, f(a) = b$
   - case 1: $f(a) \geqslant 0$
   a is even.
   pick $a := 2b$, $f(a) = f(2b) = \dfrac{2b}{2} = b$
   - case 2: $f(a) < 0$
   a is odd.
   pick $a := -2b - 1$,
   $f(a) = f(-2b-1) = \dfrac{-(-2b-1+1)}{2} = b$

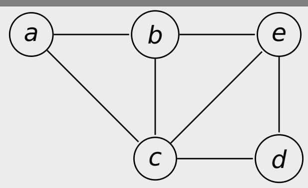
Outline

Definitions (Graphs)

- An undirected graph G is a pair of sets (V, E) such that
  - V is a non-empty (but finite) set of vertices
  - E is an unordered set of vertex pairs, namely $E \subseteq V \times V$

Example

$$
\begin{aligned}
G &= (V, E) \\
V &= \{a, b, c, d, e\} \\
E &= \{(a, b), (a, c), (b, c), (b, e), (c, d), (c, e), (e, d)\}
\end{aligned}
$$



Definitions (Graphs (cont'd))

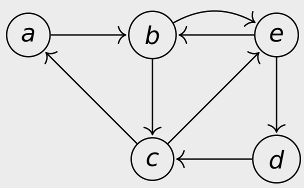
This page has a lot of overlays. Please refer to the original slides (w1.pdf) to monitor the whole content.

Definitions (Graphs (cont'd))

- An directed graph G is a pair of sets (V, E) such that
  - V is a non-empty (but finite) set of vertices
  - E is an ordered set of vertex pairs, namely $E \subseteq V \times V$

Example

$$\begin{aligned} G &= (V, E) \\ V &= \{a, b, c, d, e\} \\ E &= \{(a, b), (b, c), (b, e), (c, a), (c, e), (d, c), (e, b), (e, d)\} \end{aligned}$$



Outline

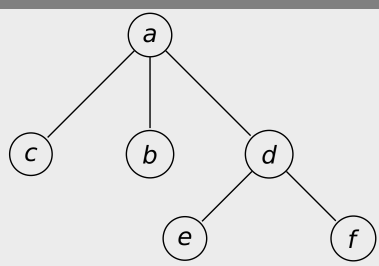
Sets Relations Functions Graphs Trees Proof Techniques Alphabets & Strings Languages

Definitions (Trees)

A tree is an undirected, acyclic, connected graph.

Example

$$
\begin{aligned}
T &= (V, E) \\
V &= \{a, b, c, d, e, f\} \\
E &= \{(a, b), (a, c), (a, d), (d, e), (d, f)\}
\end{aligned}
$$



Definitions (Trees (cont'd))

This page has a lot of overlays. Please refer to the original slides (w1.pdf) to monitor the whole content.

Definitions (Binary Trees)

A binary tree is a tree structure in which each node has at most two children.

Example

This page has a lot of overlays. Please refer to the original slides (w1.pdf) to monitor the whole content.

Outline

Definitions (Proof by Contradiction)

Suppose we want to prove that some property P holds:

1. we assume that P is false
2. then we arrive at an obviously false consequence
3. therefore, statement P must be true

Theorem

$\sqrt{2}$ is irrational.

Proof.

1. Assume that $\sqrt{2}$ is a rational number.
2. Therefore, there must exists some integers m and n with no common factors such that $\sqrt{2} = \frac{m}{n}$.
3. $2 = \frac{m^2}{n^2}$ gives $m^2 = 2n^2$. This yields that m^2 is even thus m is even.
4. Take $m = 2k$ for some integer k.
5. The equality in item 3 implies $4k^2 = 2n^2$ thus $2k^2 = n^2$. Obviously n^2 and so n are both even.
6. Take $n = 2l$ for some integer l.
7. Infer from items 4 and 6 that m and n has 2 as a common factor which contradicts with the fact in item 2.
8. $\sqrt{2}$ cannot be rational. □

Definitions (Proof by Mathematical Induction)

Suppose we want to prove that some property $P(n)$ holds for every single natural number n:

1. base case: prove that the statement $P(n)$ is true for $n = 0$, namely $P(0)$ holds.
2. step case: given that the statement $P(n)$ is true for some natural number $n = k$, prove that it also holds for its successor, $n = k + 1$. This amounts in second order logic to:

$$\forall P : \mathbb{N} \to \mathbb{B},\ (\underbrace{P(0)}_{\text{base case}} \wedge \underbrace{(\forall k \in \mathbb{N},\ \overbrace{P(k)}^{\text{IH}} \implies P(k+1))}_{\text{step case}}) \implies (\forall n \in \mathbb{N}, P(n))$$

prove $P(0)$ and the step case	plug $P(0)$ into the step case, and get $P(1)$
have $P(1)$	plug $P(1)$ into the step case, and get $P(2)$
have $P(2)$	plug $P(2)$ into the step case, and get $P(3)$
$\vdots$	$\vdots$

Theorem

Given a set A with k members. The power-set $P(A)$ has 2^k members. Namely, $|P(A)| = 2^k$.

Proof.

We argue by mathematical induction over the cardinality k of A.

1. Base case: $k = 0 \iff A = \varnothing \iff |P(\varnothing)| = 1 = 2^0$
2. Step case: Given : $|A| = k$ such that $k \geqslant 0$ $A = \{1, 2, 3, \cdots, k\}$ IH: $|P(A)| = 2^k$
   Show : $|P(A \cup \{p\})| = 2^{k+1}$

   By injecting p in A, we newly introduce

$\binom{k}{0}$	# of 1-element subset	$\{p\}$
$\binom{k}{1}$	# of 2-element subsets	$\{1,p\}, \{2,p\}, \cdots, \{k,p\}$
$\binom{k}{2}$	# of 3-element subsets	$\{1,2,p\}, \{1,3,p\}, \cdots, \{1,k,p\}, \cdots, \{k-1,k,p\}$
$\vdots$	$\vdots$	$\vdots$
$\binom{k}{k}$	# of (k+1)-element subset	$\{1,2,3,\cdots,k,p\}$

   It is provable (again by mathematical induction) that $\binom{k}{0} + \binom{k}{1} + \binom{k}{2} + \cdots + \binom{k}{k} = 2^k$.

   Therefore, $|P(A \cup \{p\})| = |P(A)| + \#$ of new subsets $= 2^k + 2^k = 2 \cdot 2^k = 2^{k+1}$

□

Theorem

A binary tree of height n has less than 2^{n+1} leaves.

Proof.

Let $L(i)$ be the maximum number of leaves of any subtree at height i. We argue by mathematical induction on the height n:

1. base case $n = 0$: $L(0) < 2^{0+1}$. Due to the fact that $L(0) = 1$, we get $1 < 2$ which trivially holds.
2. step case $n = k$: given the induction hypothesis (IH) $L(k) < 2^{k+1}$, we need to show that $L(k+1) < 2^{k+2}$. Observe that either of $L(k+1) = 2L(k)$ and $L(k+1) < 2L(k)$ holds (this needs to be explicitly proven but we skip the proof here).

$L(k+1) = 2L(k)$

$L(k)$	$<$	2^{k+1}	by IH
$2L(k)$	$<$	2^{k+2}	by arithmetic
$L(k+1)$	$<$	2^{k+2}	by observation

$L(k+1) < 2L(k)$

$L(k)$	$<$	2^{k+1}	by IH
$2L(k)$	$<$	2^{k+2}	by arithmetic
$L(k+1)$	$<$	$2L(k)$	by observation
$L(k+1)$	$<$	2^{k+2}	by transitivity of $<$

□

Outline

Definitions (Alphabets & Strings)

- An alphabet is a finite, nonempty set of symbols

Σ_T	$=$	$\{a, b\}$	A two set
Σ_L	$=$	$\{a, b, \ldots, z\}$	A set of all lowercase letters

- A string is a finite *sequence* of symbols (characters or letters) over some arbitrary alphabet Σ
  - "*abbbbbba*" is a string over the alphabet Σ_T
  - "*cat*", "*dog*", etc. are strings over the alphabet Σ_L

Example (Alphabets & Strings)

- $\Sigma_1 = \{0, 1\}$ – the alphabet of Binary numbers
  - $0, 1, 01, 11, 0110, 1010, 11100010101110$ are a few strings over Σ_1
- $\Sigma_2 = \{0, 1, 2, \ldots, 9\}$ – the alphabet of decimal numbers
  - $102345, 567463386, 109576, 3$ are strings over Σ_2
- $\Sigma_3 = \{1\}$ – the alphabet of unary numbers
  - $1, 11, 111, 11111$ are strings over Σ_3

Definitions (Length of a String)

- The length of a string w (denoted $|w|$) is the number of letters appearing in the corresponding sequence

$$\begin{array}{ll} w = a_1a_2a_3\cdots a_n & |w| = n \\ u = abba & |u| = 4 \\ v = aa & |v| = 2 \\ z = a & |z| = 1 \end{array}$$

- The string with length zero is called the empty string, and denoted ε

$$|\varepsilon| = 0$$

Definitions (String Operations)

- String concatenation is the binary operation of joining strings end-to-end

$$w = a_1 a_2 \cdots a_n \qquad v = b_1 b_2 \cdots b_m \qquad \begin{array}{l} wv = a_1 a_2 \cdots a_n b_1 b_2 \cdots b_m \\ |wv| = |w| + |v| = n + m \end{array}$$

- String reversal

$$w = a_1 a_2 \cdots a_n \qquad \begin{array}{l} w^R = a_n \cdots a_2 a_1 \\ |w^R| = |w| = n \end{array}$$

Sets Relations Functions Graphs Trees Proof Techniques Alphabets & Strings Languages

Definition (Substring)

A substring of some arbitrary string is indeed a *consecutive subsequence of letters* in the corresponding sequence

String	Substring
$\underline{abb}ab$	abb
$\underline{abba}b$	$abba$
$ab\underline{b}ab$	b
$a\underline{bbab}$	$bbab$
$\vdots$	$\vdots$

Definition (Powers of an Alphabet)

Σ^i is the set of all strings over Σ with the length i. That formally is

$$\Sigma^{i+1} := \{vw \mid w \in \Sigma^i \text{ and } v \in \Sigma\} \text{ for each } i > 0.$$

Example

$$\begin{array}{lcl}
\Sigma & = & \{0,1\} \\
\Sigma^0 & = & \{\varepsilon\} \\
\Sigma^1 & = & \{0,1\} \\
\Sigma^2 & = & \{00,01,10,11\} \\
\Sigma^3 & = & \{000,001,010,011,100,101,110,111\} \\
\vdots & & \qquad\qquad\vdots
\end{array}$$

Sets Relations Functions Graphs Trees Proof Techniques Alphabets & Strings Languages

Definition (The Kleene Star *)

The Kleene star Σ^* is the set of all strings over the alphabet Σ. That formally is

$$\Sigma^* := \bigcup_{i \geq 0}^{\infty} \Sigma^i = \Sigma^0 \cup \Sigma^1 \cup \Sigma^2 \cup \Sigma^3 \cdots$$

Example

$$\begin{array}{lcl}
\Sigma & = & \{0,1\} \\
\Sigma^* & = & \{\varepsilon,0,1,00,01,10,11,000,001,010,011,100,101,110,111\ldots\}
\end{array}$$

Definition (The Kleene Plus $^+$)

The Kleene plus Σ^+ omits the Σ^0 term in the definition of the Kleene star. That formally is

$$\Sigma^+ := \Sigma^* \setminus \Sigma^0 = \bigcup_{i \geq 1}^{\infty} \Sigma^i = \Sigma^1 \cup \Sigma^2 \cup \Sigma^3 \cdots$$

Example

$$\begin{aligned} \Sigma &= \{0, 1\} \\ \Sigma^+ &= \{0, 1, 00, 01, 10, 11, 000, 001, 010, 011, 100, 101, 110, 111 \ldots\} \end{aligned}$$

Outline

Definition (Language)

- Any subset of the set Σ^* for some alphabet Σ is called a language

$$\begin{aligned}
\Sigma &= \{0, 1\} \\
\Sigma^* &= \{\varepsilon, 0, 1, 00, 01, 10, 11, 000, 001, 010, 011, 100, 101, 110, 111 \ldots\}
\end{aligned}$$

$$\begin{aligned}
\mathcal{L}_1 &= \{\} \\
\mathcal{L}_2 &= \{\varepsilon\} \\
\mathcal{L}_3 &= \{0, 00, 001\} \\
\mathcal{L}_4 &= \{\varepsilon, 0110, 1010, 00, 01, 000000\} \\
&\vdots \qquad \vdots
\end{aligned}$$

Example (Language)

- Let $\mathcal{L}$ be the language of all strings w over the alphabet $\Sigma = \{a, b\}$ such that $w = a^n b^n$ for some $n \geq 0$. That, in set comprehension notation, is $\mathcal{L} := \{w | w \in \Sigma^* \text{ and } w = a^n b^n \text{ for some } n \geq 0\}$.

$$\begin{aligned}
\varepsilon &\in \mathcal{L} \\
ab &\in \mathcal{L} \\
aabb &\in \mathcal{L} \\
aaaaabbbbb &\in \mathcal{L} \\
bbabb &\notin \mathcal{L} \\
abb &\notin \mathcal{L} \\
\vdots &\quad \vdots
\end{aligned}$$

Example (Language)

- A *prime number* is a number $x \geq 1$ that is divided (with reminder 0) only by 1 and itself. Let $\mathcal{L}$ be the set of prime numbers defined over the alphabet $\Sigma = \{0, 1, 2, \ldots, 9\}$. Namely, $\mathcal{L} := \{x \mid x \in \Sigma^+ \text{ and } x \text{ is prime}\}$.

$$\begin{array}{lcl} 2 & \in & \mathcal{L} \\ 13 & \in & \mathcal{L} \\ 17 & \in & \mathcal{L} \\ 23 & \in & \mathcal{L} \\ 4 & \notin & \mathcal{L} \\ 12 & \notin & \mathcal{L} \\ \vdots & & \vdots \end{array}$$

Example (Language)

Alphabet	Language
$\Sigma = \{0, 1, 2, \ldots 9\}$	$\mathcal{L}_E := \{x \mid x \in \Sigma^+ \text{ and } x \text{ is even}\}$ $\mathcal{L}_E = \{0, 2, 4, 6, 8, 10, \ldots\}$
$\Sigma = \{0, 1, 2, \ldots 9\}$	$\mathcal{L}_O := \{x \mid x \in \Sigma^+ \text{ and } x \text{ is odd}\}$ $\mathcal{L}_O = \{1, 3, 5, 7, 9, 11, \ldots\}$
$\Sigma = \{1, +, =\}$	$\mathcal{L}_A := \{x + y = z \in \Sigma^+ \mid x = 1^n, y = 1^m, z = 1^k$ $n + m = k, n \geq 1, \text{ and } m \geq 1\}$ $\mathcal{L}_A = \{1 + 11 = 111, 11 + 111 = 11111, \ldots\}$
$\Sigma = \{1, \#\}$	$\mathcal{L}_S := \{x\#y \in \Sigma^+ \mid x = 1^n, y = 1^m, m = n^2 \text{ and } n \geq 1\}$ $\mathcal{L}_S = \{1\#1, 11\#1111, 111\#111111111, \ldots\}$
$\vdots$	$\vdots$

Remarks (Languages)

- The empty language $\varnothing$ (or $\{\}$) and the language $\{\varepsilon\}$ are distinct, namely $\varnothing \neq \{\varepsilon\}$
- Languages do have sizes – number of elements –

$$\begin{aligned} |\varnothing| &= 0 \\ |\{\varepsilon\}| &= 1 \\ |\{a, aa, aab\}| &= 3 \\ |\{\varepsilon, aa, bb, abba, baba\}| &= 5 \end{aligned}$$

- Recall that $|\varepsilon| = 0$ which should not be confused with $|\{\varepsilon\}| = 1$

Definitions (Operations on Languages)

Let Σ be an alphabet and let $\mathcal{L}, \mathcal{L}_1, \mathcal{L}_2$ be languages over Σ.

- Concatenation $\mathcal{L}_1\mathcal{L}_2$ is defined as

$$\mathcal{L}_1\mathcal{L}_2 := \{xy \mid x \in \mathcal{L}_1 \ \wedge \ y \in \mathcal{L}_2\}$$

- Union is defined as

$$\mathcal{L}_1 \cup \mathcal{L}_2 := \{x \mid x \in \mathcal{L}_1 \ \vee \ x \in \mathcal{L}_2\}$$

- Intersection is defined as

$$\mathcal{L}_1 \cap \mathcal{L}_2 := \{x \mid x \in \mathcal{L}_1 \ \wedge \ x \in \mathcal{L}_2\}$$

- Kleene star (similarly Kleene plus) can be viewed as an operation defined as

$$\Sigma^* = \mathcal{L} := \{x \mid x = \varepsilon \ \vee \ x \in \mathcal{L} \ \vee \ x \in \mathcal{L}\mathcal{L} \ \vee \ x \in \mathcal{L}\mathcal{L}\mathcal{L} \ \vee \ \ldots\}$$

Example (Operations on Languages)

$$\begin{array}{lcl} \Sigma & = & \{a, b, c, d\} \\ \mathcal{L}_1 & = & \{a, ab, c, d, \varepsilon\} \\ \mathcal{L}_2 & = & \{d\} \\ \mathcal{L}_3 & := & \mathcal{L}_1\mathcal{L}_2 \end{array}$$

- Which of the following strings are not in $\mathcal{L}_3$? a, abd, cd, d?

$$\begin{array}{lcl} \Sigma & = & \{a, b, c, d\} \\ \mathcal{L}_1 & = & \{a, ab, c, d, \varepsilon\} \\ \mathcal{L}_2 & = & \{d\} \\ \mathcal{L}_3 & := & \mathcal{L}_1 \cup \mathcal{L}_2 \end{array}$$

- Which of the following strings are not in $\mathcal{L}_3$? a, abd, cd, d?

Remarks (Automata Theoretic Problems)

- A problem in automata theory is always in the form of the question

  whether a given string is a member of some particular language $\mathcal{L}$:

  given a string $w \in \Sigma^*$, the problem is to decide whether or not $w \in \mathcal{L}$
- The idea is to build automatons which help in solving such decision problems out

Thanks! & Questions?