Theory of Languages and Automata

Chapter 0 - Introduction

Sharif University of Technology

References

- Main Reference
  - M. Sipser, ”Introduction to the Theory of Computation,” 3nd Ed., Cengage Learning , 2013.
- Additional References
  - P. Linz, “An Introduction to Formal Languages and Automata,” 3rd Ed., Jones and Barlett Publishers, Inc., 2001.
  - J.E. Hopcroft, R. Motwani and J.D. Ullman, “Introduction to Automata Theory, Languages, and Computation,” 2nd Ed., Addison-Wesley, 2001.
  - P.J. Denning, J.B. Dennnis, and J.E. Qualitz, “Machines, Languages, and Computation,” Prentice-Hall, Inc., 1978.
  - P.J. Cameron, “Sets, Logic and Categories,” Springer-Verlag, London limited, 1998.

Grading Policy

- Assignments 30-50%
- Exams 50-70%

Main Topics

- Computational Complexity Theory
- Computability Theory
- Automata Theory
- Mathematical Notions
- Alphabet
- Strings
- Languages

Computational Complexity Theory

- Computational Complexity theory focuses on classifying computational problems according to their resource usage and relating these classes to each other.

History of Computational Complexity Theory

- In 1965, Juris Hartmanis and Richard E. Stearns laid out the definitions of time complexity and space complexity and proved the hierarchy theorems.
- In 1971, Stephen Cook and Leonid Levin introduced the concept of NP-completeness and proved the existence of practically relevant problems that are NP-complete.
- In 1972, Richard Karp showed that 21 more important combinatorial and graph theoretical problems are also NP-complete.

Computability Theory

- Determining whether a problem is solvable by Computers.
- Classification of problems as solvable ones and unsolvable ones.

History of Computability Theory

- *On the Calculation with Hindu Numerals* written about 820, was principally responsible for spreading the Hindu-Arabic numeral system throughout the Middle East and Europe. It was translated into Latin as *Algoritmi de numero Indorum*. Al-Khwārizmī, rendered as (Latin) *Algoritmi*, led to the term "algorithm".

History of Computability Theory (cont.)

- In 1936, Alan Turing published a paper in which he proved that his "universal computing machine" would be capable of performing any conceivable mathematical computation if it were representable as an algorithm.
- According to the Church-Turing thesis , Turing machines and the lambda calculus are capable of computing anything that is computable.

Automata Theory

- **Automata theory** is the study of abstract machines and automata, as well as the computational problems that can be solved using them.
- Automata play a major role in theory of computation, compiler construction, artificial intelligence, parsing and formal verification.

History of Automata

- The ***Book of Ingenious Devices*** (Arabic: كتاب الحيل*Kitab al-Hiyal*, Persian: كتاب ترفندها*Ketab tarfandha*, literally: "The Book of Tricks") was a large illustrated work on mechanical devices, including automata, published in 850 by the three brothers of Persian descent, known as the Banu Musa (Ahmad, Muhammad and Hasan bin Musa ibn Shakir) working at the House of Wisdom (*Bayt al-Hikma*) in Baghdad, Iraq, under the Abbasid Caliphate.

Logic

- The science of the formal principles of reasoning

History of Logic

- Logic was known as 'dialectic' or 'analytic' in Ancient Greece. The word 'logic' (from the Greek *logos*, meaning discourse or sentence) does not appear in the modern sense until the commentaries of Alexander of Aphrodisias, writing in the third century A.D.
- While many cultures have employed intricate systems of reasoning, and logical methods are evident in all human thought, an explicit analysis of the principles of reasoning was developed only in three traditions: those of **China**, **India**, and **Greece**.

History of Logic (cont.)

- Although exact dates are uncertain, particularly in the case of **India**, it is possible that logic emerged in all three societies by the 4th century BC.
- The formally sophisticated treatment of modern logic descends from the Greek tradition, particularly Aristotelian logic, which was further developed by **Islamic Logicians** and then **medieval European logicians**.
- The work of **Frege** in the 19th century marked a radical departure from the Aristotlian leading to the rapid development of symbolic logic, later called mathematical logic.

Modern Logic

- **Descartes** proposed using algebra, especially techniques for solving for unknown quantities in equations, as a vehicle for scientific exploration.
- The idea of a calculus of reasoning was also developed by **Leibniz**. He was the first to formulate the notion of a broadly applicable system of mathematical logic.
- **Frege** in his 1879 work extended formal logic beyond propositional logic to include quantification to represent the "all", "some" propositions of Aristotelian logic.

Modern Logic (cont.)

- A logic is a language of formulas.
- A formula is a finite sequence of symbols with a syntax and semantics.
- A logic can have a formal system.
- A formal system consists of a set of axioms and rules of inference.

Logic Types of Interest

- Propositional Logic
- Predicate Logic

Syntax of Propositional Logic

- Let {p0, p1, ..} be a countable set of *propositional variables,*
- $\{\wedge, \vee, \neg, \rightarrow, \leftrightarrow\}$ be a finite set of *connectives*,
- Also, there are left and right brackets,
- A propositional variable is a formula,
- If φ and ψ are formulas, then so are $(\neg\varphi)$, $(\varphi \wedge \psi)$, $(\varphi \vee \psi)$, $(\varphi \rightarrow \psi)$, and $(\varphi \leftrightarrow \psi)$.

Semantics of Propositional Logic

- Any formula, which involves the propositional variables $p_0,\ldots,p_n$, can be used to define a function of n variables, that is, a function from the set $\{T, F\}^n$ to $\{T, F\}$. This function is often represented as the truth table of the formula and is defined to be the **semantics** of that formula.

Examples

ψ	φ	(¬φ)	(φ ∧ ψ)	(φ ∨ ψ)	(φ → ψ)	(φ ↔ ψ)
T	T	F	T	T	T	T
F	T	F	F	T	F	F
T	F	T	F	T	T	F
F	F	T	F	F	T	T

Other Definitions

- A formulae is a *tautology* if it is always true.
  - $(P \vee (\neg P))$ is a *tautology*.
- A formulae is a *contradiction* if it is never true.
  - $(P \wedge (\neg P))$ is a *contradiction*.
- A formulae is a *contingency* if it is sometimes true.
  - $(P \rightarrow (\neg P))$ is a *contingency*.

Formal System

A *formal system* includes the following:

- An *alphabet* A, a set of symbols.
- A set of *formulae*, each of which is a string of symbols from A.
- A set of *axioms*, each axiom being a formula.
- A set of *rules of inference*, each of which takes as 'input' a finite sequence of formulae and produces as output a formula.

Proof

- A *proof* in a formal system is just a finite sequence of formulae such that each formula in the sequence either is an axiom or is obtained from earlier formulae by applying a rule of inference.

Theorem

- A *theorem* of the formal system is just the last formula in a proof.
- Example:

  For any formula φ, the formula (φ→φ) is a theorem of the propositional logic.

A Formal System for Propositional Logic

- There are three 'schemes' of axioms, namely:
  - (A1) $(\varphi \to (\psi \to \varphi))$
  - (A2) $(\varphi \to (\psi \to \theta)) \to ((\varphi \to \psi) \to (\varphi \to \theta))$
  - (A3) $(((\neg\varphi) \to (\neg\psi)) \to (\psi \to \varphi))$
- Each of these formulas is an axiom, for all choices of formulae φ, ψ, θ.
- There is only one rule of inference, namely Modus Ponens: From φ and $(\varphi \to \psi)$, infer ψ.

Example of a Proof

- Using (A2), taking φ, ψ, θ to be φ, $(\varphi\to\varphi)$ and φ respectively
  $((\varphi\to((\varphi\to\varphi)\to\varphi))\to((\varphi\to(\varphi\to\varphi))\to(\varphi \to\varphi)))$
- Using (A1), taking φ, ψ to be φ and $(\varphi\to\varphi)$ respectively
  $(\varphi\to((\varphi\to\varphi)\to\varphi))$
- Using Modus Ponens
  $((\varphi\to(\varphi\to\varphi))\to(\varphi \to\varphi))$
- Using (A1), taking φ, ψ to be φ and φ respectively
  $(\varphi\to(\varphi\to\varphi))$
- Using Modus Ponens
  $(\varphi\to\varphi)$

Soundness & Completeness

- A formal system is said to be sound if all theorems in that system are tautology.
- A formal system is said to be complete if all tautologies in that system are theorems.

Predicate

- A proposition involving some variables, functions and relations
- Example

$$P(x) = \text{“}x > 3\text{”}$$

$$Q(x,y,z) = \text{“}x^2 + y^2 = z^2\text{”}$$

Quantifier

- Universal: "<u>for all</u>" $\forall$

  $\forall x\ P(x) \Leftrightarrow P(x_1) \wedge P(x_2) \wedge P(x_3) \wedge \ldots$

- Existential: "<u>there exists</u>" $\exists$

  $\exists x\ P(x) \Leftrightarrow P(x_1) \vee P(x_2) \vee P(x_3) \vee \ldots$

- Combinations:

  $\forall x\ \exists y\ \ y > x$

Quantifiers: Negation

- $\neg (\forall x\ P(x)) \Leftrightarrow \exists x\ \neg P(x)$
- $\neg (\exists x\ P(x)) \Leftrightarrow \forall x\ \neg P(x)$
- $\neg \exists x\ \forall y\ P(x,y) \Leftrightarrow \forall x\ \exists y\ \neg P(x,y)$
- $\neg \forall x\ \exists y\ P(x,y) \Leftrightarrow \exists x\ \forall y\ \neg P(x,y)$

Rules of Inference

1. Direct Proof
2. Indirect Proof
3. Proof by Contradiction
4. Proof by Cases
5. Induction

Direct Proof

- If the two propositions (premises) p and $p \rightarrow q$ are theorems, we may deduce that the proposition q is also a theorem.
- This fundamental rule of inference is called ***modus ponens*** by logicians.

$$\begin{array}{l} p \rightarrow q \\ \underline{p \qquad\quad} \\ \therefore q \end{array}$$

Indirect Proof

- Proves $p \to q$ by instead proving the contra-positive, $\sim q \to \sim p$

$$
\begin{array}{l}
p \to q \\
\underline{\sim q \qquad} \\
\therefore \sim p
\end{array}
$$

Proof by Contradiction

- The rule of inference used is that from theorems p and $\neg q \rightarrow \neg p$, we may deduce theorem q.

$$\begin{array}{l} p \\ \underline{\sim q \rightarrow \sim p} \\ \therefore q \end{array}$$

Proof by Cases

- You want to prove $p \rightarrow q$
- P can be decomposed to some cases:

$$p \leftrightarrow p_1 \vee p_2 \vee \ldots \vee p_n$$

- Independently prove the n implications given by

$$p_i \rightarrow q \text{ for } 1 \leq i \leq n.$$

Proof by Induction

- To prove

$$\forall x P(x)$$

1. Proof *P(0)*,
2. Proof $\forall x [P(x) \rightarrow P(x+1)]$

Constructive Existence Proof

- Want to prove that

$$\exists x\, P(x)$$

- Find an a and then prove that $P(a)$ is true

Non-Constructive Existence Proof

- Want to prove that

$$\exists x\, P(x)$$

- You cannot find an a that $P(a)$ is true
- Then, you can use proof by contradiction:

$$\sim \exists x\, P(x) \leftrightarrow \forall x \sim P(x) \rightarrow F$$

Intuitive (Naïve) Set Theory

There are three basic concepts in set theory:

- Membership
- Extension
- Abstraction

Membership

- *Membership* is a relation that holds between a set and an object
- $x \in A$ to mean "the object x is a member of the set A", or "x belongs to A". The negation of this assertion is written $x \notin A$ as an abbreviation for the proposition $\neg(x \in A)$
- One way to specify a set is to list its elements. For example, the set $A = \{a, b, c\}$ consists of three elements. For this set A, it is true that $a \in A$ but $d \notin A$

Extension

- The *concept* of extension is that two sets are identical if and if only if they contain the same elements. Thus we write $A=B$ to mean

$$\forall x\ [x \in A \Leftrightarrow x \in B]$$

Abstraction

Each property defines a set, and each set defines a property

- If $p(x)$ is a property then we can define a set A

$$A = \{x \mid p(x)\}$$

- If A is a set then we can define a predicate $p(x)$

$$p(x) = x \in A$$

Intuitive versus **axiomatic** set theory

- The theory of set built on the intuitive concept of membership, extension, and abstraction is known as *intuitive (naïve) set theory*.
- As an *axiomatic theory* of sets, it is not entirely satisfactory, because the principle of abstraction leads to contradictions when applied to certain simple predicates.

Russell’s Paradox

Let *p(X)* be a predicate defined as

$$P(X) = (X \notin X)$$

Define set *R* as

$$R=\{X|P(X)\}$$

- Is *P(R)* true?
- Is *P(R)* false?

Gottlob Frege's Comments

- The logician Gottlob Frege was the first to develop mathematics on the foundation of set theory. He learned of Russell Paradox while his work was in press, and wrote, "A scientist can hardly meet with anything more undesirable than to have the foundation give way just as the work is finished. In this position I was put by a letter from Mr. Bertrand Russell as the work was nearly through the press."

Power Set

- The set of all subsets of a given set A is known as the *power set* of A, and is denoted by $P(A)$:

$$P(A) = \{B \mid B \subseteq A\}$$

Set Operation-Union

- The *union* of two sets *A* and *B* is

$$A \cup B = \{x \mid (x \in A) \vee (x \in B)\}$$

and consists of those elements in at least one of *A* and *B*.

- If $A_1, \ldots, A_n$ constitute a family of sets, their *union* is

$$\bigcup_{i=1}^{n} A_i = (A_1 \cup \ldots \cup A_n)$$
$$= \{x \mid x \in A_i \text{ for some i, } 1 \leq i \leq n\}$$

Set Operation-Intersection

- The *intersection* of two sets *A* and *B* is

$$A \cap B = \{x \mid (x \in A) \wedge (x \in B)\}$$

and consists of those elements in at least one of *A* and *B*.

- If $A_1, \ldots, A_n$ constitute a family of sets, their *intersection* is

$$\bigcap_{i=1}^{n} A_i = (A_1 \cap \ldots \cap A_n)$$
$$= \{x \mid x \in A_i \text{ for all } i,\ 1 \leq i \leq n\}$$

Set Operation-Complement

- The *complement* of a set A is a set A^c defined as:

$$A^c = \{x \mid x \notin A\}$$

- The *complement of a set B with respect to A*, also denoted as A-B, is defined as:

$$A\text{-}B = \{x \in A \mid x \notin B\}$$

Ordered Pairs and n-tuples

- An *ordered pair* of elements is written

$$(x,y)$$

where x is known as the *first element*, and y is known as the *second element*.

- An *n-tuple* is an ordered sequence of elements

$$(x_1, x_2, \ldots, x_n)$$

And is a generalization of an ordered pair.

Ordered Sets and Set Products

- By Cartesian product of two sets *A* and *B*, we mean the set

$$A \times B = \{(x,y) \mid x \in A \text{ , } y \in B\}$$

- Similarly,

$$A_1 \times A_2 \times \ldots A_n = \{x_1 \in A_1,\ x_2 \in A_2,\ \ldots,\ x_n \in A_n\}$$

Relations

A relation ρ between sets A and B is a subset of $A \times B$:

$$\rho \subseteq A \times B$$

- The *domain* of ρ is defined as

$$D_\rho = \{x \in A \mid \text{for some } y \in B, (x,y) \in \rho\}$$

- The *range* of ρ is defined as

$$R_\rho = \{y \in B \mid \text{for some } x \in A, (x,y) \in \rho\}$$

- If $\rho \subseteq A \times A$, then ρ is called a "relation on A".

Types of Relations on Sets

- A relation ρ is *reflexive* if

  $$(x,x) \in \rho, \text{ for each } x \in A$$

- A relation ρ is *symmetric* if, for all $x,y \in A$

  $$(x,y) \in \rho \text{ implies } (y,x) \in \rho$$

- A relation ρ is *antisymmetric* if, for all $x,y \in A$

  $$(x,y) \in \rho \text{ and } (y,x) \in \rho \text{ implies } x=y$$

- A relation ρ is *transitive* if, for all $x,y,z \in A$

  $$(x,y) \in \rho \text{ and } (y,z) \in \rho \text{ implies } (x,z) \in \rho$$

Partial Order and Equivalence Relations

- A relation ρ on a set A is called a *partial ordering* of A if ρ is reflexive, anti-symmetric, and transitive.
- A relation ρ on a set A is called an *equivalence* relation if ρ is reflexive, symmetric, and transitive.

Total Ordering

- A relation ρ on a set A is a *total ordering* if ρ is a partial ordering and, for each pair of elements (x,y) in $A \times A$ at least one of $(x,y) \in \rho$ or $(y,x) \in \rho$ is true.

Inverse Relation

- For any relation $\rho \subseteq A \times B$, the *inverse* of ρ is defined by

$$\rho^{-1} = \{(y,x) \mid (x,y) \in \rho\}$$

- *If D_ρ and R_ρ* are the domain and range of ρ, then

$$D_{\rho^{-1}} = R_\rho \text{ and } R_{\rho^{-1}} = D_\rho$$

Equivalence

- Let $\rho \subseteq A \times A$ be an equivalence relation on A. The *equivalence class* of an element x is defined as

$$[x] = \{y \in A \mid (x,y) \in \rho\}$$

- An equivalence relation on a set partitions the set.

Functions

A relation $f \subseteq A \times B$ *is a* *function* if it has the property:
for all x,y,z, $(x,y) \in f$ and $(x,z) \in f$ implies $y = z$

- If $f \subseteq A \times B$ is a function, we write

$$f: A \rightarrow B$$

and say that f maps A into B. We use the common notation

$$Y = f(x)$$

to mean $(x,y) \in f$.

Functions (cont.)

- As before, the *domain* of f is the set

$$D_f = \{x \in A \mid \text{for some } y \in B, (x,y) \in f\}$$

and the *range* of f is the set

$$R_f = \{y \in B \mid \text{for some } x \in A, (x,y) \in f\}$$

- If $D_f \subseteq A$, we say the function is a *partial function*; *if* $D_f = A$, we say that f is a *total function*.

Functions (cont.)

- If $x \in D_f$, we say that f is *defined* at x; otherwise f is *undefined* at x.
- If $R_f = B$, we say that f maps D_f *onto* B.
- If a function f has the property

  for all x,y,z, $f(x) = z$ and $f(y) = z$ implies $x = y$

then f is a *one-to-one function*.

If $f: A \rightarrow B$ *is* a one-to-one function, f gives a *one-to one correspondence* between elements of its domain and range.

Functions (cont.)

- Let X be a set and suppose A $\subseteq X$. Define *function*

$$C_A: X \rightarrow \{0,1\}$$

such that $C_A(x) = 1$, if $x \in A$; $C_A(x) = 0$, otherwise. $C_A(x)$ is called the *characteristic function of set A with respect to set X.*

- *If $f \subseteq A \times B$* is a function, then the inverse of f is the set

$$f^{-1} = \{(y,x) \mid (x,y) \in f\}$$

f^{-1} is a function if and only if f is one-to-one.

Functions (cont.)

- Let $f\colon A \rightarrow B$ be a function, and suppose that $X \subseteq A$. Then the set

$$Y = f(X) = \{y \in B \mid y = f(x) \text{ for some } x \in X\}$$

is known as the *image* of X under f.

- Similarly, the *inverse image* of a set Y included in the range of f is

$$f^{-1}(Y) = \{x \in A \mid y = f(x) \text{ for some } y \in Y\}$$

Cardinality

- Two sets A and B are of *equal cardinality*, written as

$$|A| = |B|$$

  if and only if there is a one-to-one function $f: A \to B$ that maps A *onto* B.
- We write

$$|A| \leq |B|$$

  if B includes a subset C such that $|A| = |C|$.
- If $|A| \leq |B|$ and $|A| \neq |B|$, then A has cardinality *less than* that of B, and we write

$$|A| < |B|$$

Cardinality (cont.)

- Let $J = \{1, 2, \ldots\}$ and $J_n = \{1, 2, \ldots, n\}$.
- A *sequence* on a set X is a function $f: J \rightarrow X$. A sequence may be written as

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

  However, we often use the simpler notation

$$x_1, x_2, x_3, \ldots., \qquad x_i \in X$$

- A *finite sequence of length n* on X is a function $f: J_n \rightarrow X$, usually written as

$$x_1, x_2, x_3, \ldots., x_n, \qquad x_i \in X$$

- The *sequence of length zero* is the function $f: \emptyset \rightarrow X$.

Finite and Infinite Sets

- A set A is *finite* if $|A| = |J_n|$ for some integer $n \geq 0$, in which case we say that A has cardinality n.
- A set is *infinite* if it is not finite.
- A set X is *denumerable* if $|X| = |J|$.
- A set is *countable* if it is either finite or denumerable.
- A set is *uncountable* if it is not countable.

Some Properties

- **Proposition:** Every subset of J is countable. Consequently, each subset of any denumerable set is countable.
- **Proposition:** A function $f: J \to Y$ has a countable range. Hence any function on a countable domain has a countable range.
- **Proposition:** The set $J \times J$ is denumerable. Therefore, $A \times B$ is countable for arbitrary countable sets A and B.

Some Properties (cont.)

- **Proposition:** The set $A \cup B$ is countable whenever A and B are countable sets.
- **Proposition:** Every infinite set X is at least denumerable ; that is $|X| \geq |J|$.
- **Proposition:** The set of all infinite sequence on $\{0, 1\}$ is uncountable.

Some Properties (cont.)

- **Proposition: (Schröder-Bernestein Theorem)**

  For any set A and B, if $|A| \geq |B|$ and $|B| \geq |A|$, then $|A| = |B|$.

- **Proposition: (Cantor's Theorem)**

  For any set X,

$$|X| < |\mathcal{P}(X)|.$$

1-1 correspondence Q↔N

- Proof (dove-tailing):

	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	
6	$\frac{1}{6}$	$\frac{2}{6}$	$\frac{3}{6}$	$\frac{4}{6}$	$\frac{5}{6}$	$\frac{6}{6}$	$\cdots$
5	$\frac{1}{5}$	$\frac{2}{5}$	$\frac{3}{5}$	$\frac{4}{5}$	$\frac{5}{5}$	$\frac{6}{5}$	$\cdots$
4	$\frac{1}{4}$	$\frac{2}{4}$	$\frac{3}{4}$	$\frac{4}{4}$	$\frac{5}{4}$	$\frac{6}{4}$	$\cdots$
3	$\frac{1}{3}$	$\frac{2}{3}$	$\frac{3}{3}$	$\frac{4}{3}$	$\frac{5}{3}$	$\frac{6}{3}$	$\cdots$
2	$\frac{1}{2}$	$\frac{2}{2}$	$\frac{3}{2}$	$\frac{4}{2}$	$\frac{5}{2}$	$\frac{6}{2}$	$\cdots$
1	$\frac{1}{1}$	$\frac{2}{1}$	$\frac{3}{1}$	$\frac{4}{1}$	$\frac{5}{1}$	$\frac{6}{1}$	$\cdots$
	1	2	3	4	5	6	

Countable Sets

- Any subset of a countable set
- The set of integers, algebraic/rational numbers
- The union of two/finite number of countable sets
- Cartesian product of a finite number of countable sets
- The set of all finite subsets of N
- Set of binary strings

Diagonal Argument

$$r_1 = 0.d_{1,1}d_{1,2}d_{1,3}d_{1,4}d_{1,5}d_{1,6}d_{1,7}d_{1,8}\cdots$$
$$r_2 = 0.d_{2,1}d_{2,2}d_{2,3}d_{2,4}d_{2,5}d_{2,6}d_{2,7}d_{2,8}\cdots$$
$$r_3 = 0.d_{3,1}d_{3,2}d_{3,3}d_{3,4}d_{3,5}d_{3,6}d_{3,7}d_{3,8}\cdots$$
$$r_4 = 0.d_{4,1}d_{4,2}d_{4,3}d_{4,4}d_{4,5}d_{4,6}d_{4,7}d_{4,8}\cdots$$

Uncountable Sets

- **R**, $\mathbf{R}^2$, P(**N**)
- The intervals [0,1), [0, 1], (0, 1)
- The set of all real numbers
- The set of all functions from **N** to {0, 1}
- The set of functions **N** $\rightarrow$ **N**
- Any set having an uncountable subset

Transfinite Cardinal Numbers

- Cardinality of a *finite* set is simply the number of elements in the set.
- Cardinalities of *infinite* sets are not natural numbers, but are special objects called *transfinite cardinal numbers*.
- $\aleph_0 :\equiv |\mathbf{N}|$, is the *first transfinite cardinal* number.
- *continuum hypothesis* claims that $|\mathbf{R}| = \aleph_1$, the *second transfinite cardinal*.

Ordinal Numbers

- An **ordinal number**, or an **ordinal**, is a generalization of the concept of a natural number that is used to describe a way to arrange a (possibly infinite) collection of objects in order, one after another.
- Ordinals were introduced by Georg Cantor in 1883 in order to accommodate infinite sequences.

Well-Ordered Sets

- A *well-order* on a set X is a total order < on X having the property that every non-empty subset of X has a least element.
- For example, (N, <) is the simplest infinite well-ordered set. Any finite totally ordered set is well-ordered.

Ordinals

- **Definition:** Given a totally ordered set $(X, <)$, and an element $a \in X$, we define the *section* X_a to consist of all elements of X which are less than a:

$$X_a = \{x \in X : x < a\}.$$

- **Definition:** An *ordinal* is a well-ordered set $(X, <)$ with the property that $X_a = a$ for all $a \in X$. In other words, each element of X is the set of all its predecessors.

Ordinals (con.)

- Obviously, the least ordinal is (vacuously) = Ø. We take:
- 0= Ø
- 1={Ø}={0}
- 2 = {Ø,{Ø}} = {0,1}
- 3 = {Ø, {Ø},{Ø,{Ø}}} = {0,1,2}
- 4 = ... = (0,1,2,3}

  and so on. In general, we have

$$n = \{0,1,2, ..., n-1\}$$

Ordinals (con.)

- So, every natural number is an ordinal.
- But the ordinals continue after the natural numbers leave off. If ω denotes the smallest ordinal which is not a natural number, then ω is the set of natural numbers.
- Then the next ordinal after ω is $\omega \cup \{\omega\} = \omega + 1$, and so on .

Some Properties (con.)

Theorem:

- For ordinals x and y, the following are equivalent:

  (a) $x < y$;
  (b) $x \in y$;
  (c) $x \subset y$.

- O Moreover, exactly one of $x < y$, $x = y$, $y < x$ holds.

Limit Ordinals

- A non-zero ordinal λ is called a *limit ordinal* if it is the union of all its predecessors:

$$\lambda = \bigcup a_{\ a < \lambda}$$

- A successor ordinal is not a limit ordinal: for if $\lambda = a \cup \{a\}$ then the ordinals smaller than λ are all contained in a, and so is their union.

Limit Ordinals (con.)

Theorem:

- Any non-zero ordinal is either a successor ordinal or a limit ordinal.

Burali-Forti paradox

- The ordinals thus form a sequence of well-ordered sets, each contained in the next, which go on for ever. One variant of Russell's Paradox, known as the *Burali-Forti paradox*, is the following assertion:
- **Theorem** 2.5
  The ordinal numbers do not form a set.

Cardinal Numbers

Definition: A *cardinal* is an ordinal a with the property that there is no bijection between a and any section of a.

- Cardinal numbers measure the size of arbitrary sets.
- Note that, according to this definition, all finite ordinals (that is, all natural numbers) are cardinals.
- ω is a cardinal, since it is infinite but all its sections are finite. However, $\omega + 1$ is not a cardinal, since it is countable (i.e., has a bijection to its section ω).

Cardinality of a Set

- **Definition:** We denote the cardinal of the set X (the unique cardinal bijective with X) by |X|. Note that, if a is a cardinal, then $|a| = a$.

Alpha Notation for Cardinals

- Cantor introduced the *aleph notation* for infinite cardinals. (The letter ℵ, aleph, is the first letter of the Hebrew alphabet.) This is a function from ordinals to cardinals, defined by transfinite recursion as follows:
- $\aleph_0 = \omega$
- $\aleph_{a+1}$ is the smallest cardinal greater than $\aleph_a$
- if λ is a limit ordinal then

$$\aleph_\lambda = \bigcup_{\beta < \lambda} \aleph_\beta$$

Cardinal Arithmetics

For cardinals α and β, we define

- $\alpha + \beta = |(\alpha \times \{0\})| \cup |(\beta \times \{1\})|$
- $\alpha . \beta = | \alpha \times \beta|$
- $\alpha^{\beta} = | \alpha^{\beta} |$

where in the third (confusing) equation, on the right-hand side, A^B means the set of all functions from B to A.

Cardinal Arithmetics (con.)

- **Theorem:**

  (a) For any set X, $|\wp(X)| = 2^{|X|}$.
- (b) $|\mathbb{R}|=2^{\aleph_0}$.

Cantor’s Theorem

The *Cantor's Theorem* can be translated into the form

- **Theorem:** For any cardinal a, $2^a > a$.

The *Schroder-Bernstein Theorem* can be written in terms of cardinals as follows.

- **Theorem:**
  For any two sets X and Y, if $|X| \leq |Y|$ and $|Y| \leq |X|$ then $|X| = |Y|$.

Continuum Hypothesis

- By Cantor's Theorem, we have $2^{\aleph_a} \geq \aleph_{a+1}$ for any ordinal a.
- Do we have equality or not?
- The famous *Continuum Hypothesis* asserts that $2^{\aleph_0} = \aleph_1$.
- This was one of the problems posed in 1900 to the mathematical community by David Hilbert, to guide the development of mathematics in the twentieth century.

Continuum Hypothesis (con.)

- Godel proved $2^{\aleph_0} = \aleph_1$ cannot be disproved in ZFC.
- Thirty years later, however, Cohen, showed that $2^{\aleph_0} = \aleph_1$ cannot be proved in ZFC either. By a new technique known as *forcing*, he constructed a model of ZFC in which $2^{\aleph_0} = \aleph_2$.

Alphabet

- A finite and nonempty set of symbols (usually shown by $\sum$ or Γ).

- Example:

$$\sum = \{a, b, c, \dots, z\}$$

Strings

- A finite list of symbols from an alphabet
- Example: *house*
- If ω is a string over $\sum$, the *length* of ω, written $|\omega|$, is the number of symbols that it contains.
- The *empty string* (ε or λ) is the string of length zero.

$$\varepsilon\omega = \omega\varepsilon = \omega$$

- String z is a *substring* of ω if z appears consecutively within ω.

Operations on Strings

- Reverse of a string:

$$w = a_1a_2 \ldots a_n \qquad\qquad ababaaabbb$$

$$w^R = a_n \ldots a_2a_1 \qquad\qquad bbbaaababa$$

- Concatenation:

$$w = a_1a_2 \ldots a_n \qquad\qquad abba$$

$$v = b_1b_2 \ldots b_m \qquad\qquad bbbaaa$$

$$wv = a_1a_2 \ldots a_n\, b_1b_2 \ldots b_m \qquad\qquad abbabbbaaa$$

$$|wv| = |w| + |v|$$

$$w^n = \underbrace{ww \cdots w}_{n} \qquad\qquad w^0 = \lambda$$

$$(abba)^2 = abbaabba \qquad\qquad (abba)^0 = \lambda$$

Operations on Alphabet

- * :
  - The set of all strings that can be produced from ∑.

$$\Sigma = \{a, b\}$$

$$\Sigma^* = \{\lambda, a, b, aa, ab, ba, bb, aaa, \ldots\}$$

- + :
  - The set of all strings, excluding λ, that can be produced from ∑.
  - Suppose that **λ**={λ}. Then:

$$\Sigma^+ = \Sigma^* - \lambda$$

$$\Sigma^+ = \{a, b, aa, ab, ba, bb, aaa, aab, \ldots\}$$

Languages

- A set of strings
- Any language on the alphabet Σ is a subset of Σ^*.
- Examples:

$$\Sigma = \{a, b\}$$

$$L = \{a^n b^n : n \geq 0\}$$

$$\left.\begin{array}{l} \lambda \\ ab \\ aabb \\ aaaaabbbbb \end{array}\right\} \in L \qquad abb \notin L$$

$$\Sigma = \{0,1,2,\ldots,9\}$$

$$EVEN = \{0,2,4,6,\ldots\}$$

$$PRIMES = \{1,2,3,5,7,11,13,17,\ldots\}$$

Operations on Languages

Languages are a special kind of sets and operations on sets can be defined on them as well.

- Union
- Intersection
- Relative Complement
- Complement

λ, b, aa, ab, bb, aaa, …}

Operations on Languages (cont.)

- Reverse of a language

$$L^R = \{w^R : w \in L\}$$

- Example:

$$\{ab, aab, baba\}^R = \{ba, baa, abab\}$$

$$L = \{a^n b^n : n \geq 0\}$$

$$L^R = \{b^n a^n : n \geq 0\}$$

Operations on Languages (cont.)

- Concatenation

$$L_1L_2 = \{xy : x \in L_1, y \in L_2\}$$

$$L^n = \underbrace{LL\cdots L}_{n}$$

- Example:

$$\{a,ab,ba\}\{b,aa\} = \{ab,aaa,abb,abaa,bab,baaa\}$$

$$\{a,b\}^3 = \{a,b\}\{a,b\}\{a,b\} = \{aaa,aab,aba,abb,baa,bab,bba,bbb\}$$

Operations on Languages (cont.)

- *: Kleene *
  - The set of all strings that can be produced by concatenation of strings of a language.

$$L^* = L^0 \cup L^1 \cup L^2 \cdots$$

  - Example:

$$\{a,bb\}^* = \begin{Bmatrix} \lambda, \\ a, bb, \\ aa, abb, bba, bbbb, \\ aaa, aabb, abba, abbbb, \ldots \end{Bmatrix}$$

- +:
  - The set of all strings, excluding λ, that can be produced by concatenation of strings of a language.

Gödel Numbering

- Let $\sum$ be an alphabet containing n objects. Let $h: \sum \to J_n$ be an arbitrary one-to-one correspondence. Define function f as:

$$f: \sum^* \to N$$

such that $f(\varepsilon) = 0$; $f(w.v) = n\, f(w) + h(v)$, for $w \in \sum^*$ and $v \in \sum$.

f is called a *Gödel Numbering* of $\sum^*$.

- **Proposition:** $\sum^*$ is denumerable.
- **Proposition:** Any language on an alphabet is countable.