

# CS341 Automata Theory

## Mathematical Preliminaries: A Review<sup>1</sup>

### 1 Sets

A **set** is a collection of objects. For example, the collection of all students in the CS341 course is a set. The collection of the four letters  $a$ ,  $b$ ,  $c$  and  $d$  is a set, which we may call  $S$ ; we write  $S = \{a, b, c, d\}$ . There are various ways of describing sets:

- “*Explicit*” enumeration of all the objects in the set.  
Examples:  $S = \{a, b, c\}$ ,  $T = \{1, 2, 3, 4, 5, 6, 7, 8, 9, 10\}$ .
- “*Implicit*” enumeration of the objects in the set.  
Examples:  $A = \{0, 1, 2, 3, \dots\}$ , which represents the set of all non-negative integers.  
 $B = \{\dots, -2, -1, 0, 1, 2, \dots\}$ , which represents the set of all integers.
- *Formal description*.  
Examples:  $P = \{n : n \text{ is even}\}$ .  
 $Q = \{4n + 5 : n = 0, 1, 2, \dots\}$ .  
 $S = \{n : n \geq 3 \text{ and } \forall 1 < m < n, n \not\equiv 0 \pmod{m}\}$ .  
 $T = \{(p, q, r, s) : p \rightarrow q \text{ or } r \rightarrow s \text{ or } p \leftrightarrow s\}$ .
- “*Precise*” verbal description.  
Examples:  $A$  is the set of all students in CS341.  
 $B$  is the set of all leap years in the second century.

Some special symbols denote special sets.

- $\mathbb{N}$  denotes the set of all the Naturals  $(1, 2, 3, \dots)$ .
- $\mathbb{Z}$  denotes the set of all the Integers  $(\dots, -2, -1, 0, 1, 2, \dots)$ .
- $\mathbb{R}$  denotes the set of all the Reals.
- $\mathbb{Q}$  denotes the set of all the Rationals.

The objects comprising a set are called its **elements** or **members**. For example, 7 is a member of  $\mathbb{R}$ ; in symbols,  $7 \in \mathbb{R}$ . Sometimes we simply say  $a$  is **in** the set  $S$ , or that  $S$  **contains**  $a$ . On the other hand, if  $b$  is not an element of  $S$ , we write  $b \notin S$ .

In a set we do not distinguish repetitions of the elements. Moreover, the elements of a set are unordered. Thus,  $\{a, a, a\}$  is the same set as  $\{a\}$ , and  $\{a, b, c\}$  is the same set as  $\{b, a, c\}$ . *Two sets  $A$  and  $B$  are equal if and only if they have the same elements*; we write  $A = B$ .

---

<sup>1</sup>This material was prepared by Luay Nakhleh.

The elements of a set need not be related in any way, and in particular, need not be of the same type. For examples, the set  $\{10, \text{automata}, \{1, 29\}, (x, y, z)\}$  contains four elements: one is an integer, one is a string, one is a set and one is a list. A set which contains only one element is called a **singleton**. The set that contains no elements is called the **empty set** and is denoted by  $\emptyset$ . Any set other than the empty set is said to be **nonempty**.

A set  $S$  is a **subset** of a set  $T$ , denoted by  $S \subseteq T$ , if every element  $a$  of  $S$  is also an element of  $T$ , i.e.,

$$S \subseteq T: \forall a \in S, a \in T.$$

For example,  $\mathbb{N} \subseteq \mathbb{Z}$ . However,  $\mathbb{R} \not\subseteq \mathbb{N}$ . If  $S \subseteq T$  and  $S \neq T$ , we say that  $S$  is a **proper subset** of  $T$ , and write  $S \subset T$ .

One way of proving  $S = T$  is by showing that  $S \subseteq T$  and  $T \subseteq S$ . The **cardinality** of a set  $S$ , denoted by  $|S|$ , is the number of elements in  $S$ . For example,  $|\{2, 3, 9\}| = 3$ . If the set has an infinite number of elements, then its cardinality is  $\infty$ . Otherwise, the set  $S$  is finite.  $|\mathbb{N}| = \infty$ , for example.

## 1.1 Operations on sets

- UNION:  $A \cup B = \{a : a \in A \text{ or } a \in B\}$ .
- INTERSECTION:  $A \cap B = \{a : a \in A \text{ and } a \in B\}$ .
- DIFFERENCE:  $A - B = \{a : a \in A \text{ and } b \notin B\}$ .
- COMPLEMENT: (always taken with respect to a “universal” set,  $U$ )  $\bar{A} = U - A$ .

## 1.2 Properties of set operations

If  $A$ ,  $B$ , and  $C$  are sets, the following laws hold.

Idempotency	$A \cup A = A$ $A \cap A = A$
Commutativity	$A \cup B = B \cup A$ $A \cap B = B \cap A$
Associativity	$(A \cup B) \cup C = A \cup (B \cup C)$ $(A \cap B) \cap C = A \cap (B \cap C)$
Distributivity	$A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$ $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$
Absorption	$A \cap (A \cup B) = A$ $A \cup (A \cap B) = A$
DeMorgan's Laws	$A - (B \cup C) = (A - B) \cap (A - C)$ $A - (B \cap C) = (A - B) \cup (A - C)$

Two sets are said to be **disjoint** if they have no elements in common, i.e., if their intersection is  $\emptyset$ .

It is possible to form intersections and unions of more than two sets. If  $S$  is a collection of sets, we write  $\bigcup S$  for the set whose elements are the elements of the sets in  $S$ . For example, if  $S = \{\{n\} : n \in \mathbb{N}\}$ , then  $\bigcup S = \mathbb{N}$ . In general,

$$\bigcup S = \{x : x \in P \text{ for some set } P \in S\}.$$

Similarly,

$$\bigcap S = \{x : x \in P \text{ for each set } P \in S\}.$$

The collection of all subsets of a set  $S$  is itself a set, called the **power set** of  $S$  and denoted by  $2^S$ . For example,

$$2^{\{1,2\}} = \{\emptyset, \{1\}, \{2\}, \{1,2\}\}.$$

A **partition** of a nonempty set  $S$  is a subset  $\Pi$  of  $2^S$  such that  $\emptyset \notin \Pi$ , and such that each element of  $S$  is in one and only one set in  $\Pi$ . That is,  $\Pi$  is a partition of  $S$  if  $\Pi$  is a set of subsets of  $S$  such that

1. each element of  $\Pi$  is nonempty.
2. distinct members of  $\Pi$  are disjoint.
3.  $\bigcup \Pi = S$ .

For example,  $\{\{1,2\}, \{3,4\}\}$  is a partition of  $S = \{1,2,3,4\}$ , whereas  $\{\{1,2\}, \{1,3,4\}\}$  is not a partition of  $S$ .

## 2 Relations and Functions

As mentioned before, elements of a set are unordered. To distinguish between elements of a set, we have to “order” them. We begin by introducing **ordered pairs**. We write the ordered pair of two objects  $a$  and  $b$  as  $(a, b)$ . The ordered pair  $(a, b)$  is not the same as the set  $\{a, b\}$ . First, the order matters:  $(a, b)$  is different from  $(b, a)$ , whereas  $\{a, b\} = \{b, a\}$ . Second, the two components of an ordered pair need not be distinct;  $(7, 7)$  is a valid ordered pair. Note that the two ordered pairs  $(a, b)$  and  $(c, d)$  are equal only when  $a = c$  and  $b = d$ .

The **Cartesian product** of two sets  $A$  and  $B$ , denoted  $A \times B$ , is the set of all ordered pairs  $(a, b)$  with  $a \in A$  and  $b \in B$ , i.e.,  $A \times B = \{(a, b) : a \in A \text{ and } b \in B\}$ .

A **binary relation** on two sets  $A$  and  $B$  is a subset of  $A \times B$ . For example,  $\{(i, j) : i, j \in \mathbb{N} \text{ and } i < j\}$  is the *less-than* relation.

More generally, let  $n$  be any natural number. Then, if  $a_1, \dots, a_n$  are  $n$  objects, not necessarily distinct,  $(a_1, \dots, a_n)$  is an **ordered  $n$ -tuple**.

A **function  $f$  from a set  $A$  to a set  $B$** , denoted by  $f : A \rightarrow B$ , is a binary relation  $R$  on  $A$  and  $B$  with the following special property: for each element  $a \in A$ , there is exactly

one ordered pair in  $R$  with first component  $a$ . We call  $A$  the **domain** of  $f$ , and the set  $\{f(a) : a \in A\} \subseteq B$  the **range** of  $f$ .

A function  $f : A \rightarrow B$  is **one-to-one** if for any two distinct elements  $a, b \in A$ ,  $f(a) \neq f(b)$ .

A function  $f : A \rightarrow B$  is **onto**  $B$  if each element of  $B$  is the image under  $f$  of some element in  $A$ .

A function  $f : A \rightarrow B$  is a **bijection** between  $A$  and  $B$  if it is both one-to-one and onto.

If  $Q$  and  $R$  are binary relations, then their **composition**  $Q \circ R$ , or simply  $QR$ , is  $\{(q, r) : \text{for some } c, (q, c) \in Q \text{ and } (c, r) \in R\}$ .

## 2.1 Special types of binary relations

We can define a binary relation on a single set; for example,  $R \subseteq A \times A$ . Certain properties of such relations are:

- *Reflexivity*: A binary relation  $R \subseteq A \times A$  is **reflexive** if  $(a, a) \in R$  for each  $a \in A$ .
- *Symmetry*: A binary relation  $R \subseteq A \times A$  is **symmetric** if  $(b, a) \in R$  whenever  $(a, b) \in R$ .
- *Antisymmetry*: A binary relation  $R \subseteq A \times A$  is **antisymmetric** if whenever  $(a, b) \in R$  and  $a$  and  $b$  are distinct, then  $(b, a) \notin R$ .
- *Transitivity*: A binary relation  $R \subseteq A \times A$  is **transitive** if whenever  $(a, b) \in R$  and  $(b, c) \in R$ , then  $(a, c) \in R$ .

A binary relation  $R \subseteq A \times A$  that is reflexive, symmetric and transitive is called an **equivalence relation**.

Example: The binary relation  $R_5 = \{(a, b) : a, b \text{ are nonnegative integers and } a \equiv_5 b\}$  is an equivalence relation:

- $R_5$  is reflexive, since  $a \equiv_5 a$  for each  $a$ , which means that  $(a, a) \in R_5$  for each  $a$ .
- $R_5$  is symmetric, since  $a \equiv_5 b$  implies that  $b \equiv_5 a$  (due to commutativity of  $\equiv_5$ ), which means that if  $(a, b) \in R_5$  then  $(b, a) \in R_5$ .
- $R_5$  is transitive, since if  $a \equiv_5 b$  and  $b \equiv_5 c$  then  $a \equiv_5 c$ , which means that if  $(a, b) \in R_5$  and  $(b, c) \in R_5$ , then  $(a, c) \in R_5$ .

An equivalence relation  $R \subseteq A \times A$  induces a set of **equivalence classes**. Each equivalence class contains all the elements  $a \in A$  that are related to each other under  $R$ . From each equivalence class, we choose (randomly) an element  $a$  to be the “representative” of the class. Formally,

$$[a] = \{b \in A : (a, b) \in R\}$$

Example: The equivalence classes of the relation  $R_5$  described above are:

- $[0] = \{0, 5, 10, 15, \dots\} = \{5n : n = 0, 1, 2, \dots\}$
- $[1] = \{1, 6, 11, 16, \dots\} = \{5n + 1 : n = 0, 1, 2, \dots\}$

- $[2] = \{2, 7, 12, 17, \dots\} = \{5n + 2 : n = 0, 1, 2, \dots\}$
- $[3] = \{3, 8, 13, 18, \dots\} = \{5n + 3 : n = 0, 1, 2, \dots\}$
- $[4] = \{4, 9, 14, 19, \dots\} = \{5n + 4 : n = 0, 1, 2, \dots\}$

The cardinality of an equivalence relation  $R$ , denoted by  $|R|$ , is the number of equivalence classes of  $R$ . For example,  $|R_5| = 5$ .

**Theorem 1** *Let  $R$  be an equivalence relation on a set  $A$ . Then the equivalence classes of  $R$  constitute a partition of  $A$ .*

A relation  $R \subseteq A \times A$  that is reflexive, antisymmetric, and transitive is called a **partial order**. For examples, the relation

$$R_{\leq} = \{(a, b) : a, b \in \mathbb{Z} \text{ and } a \leq b\}$$

is a partial order. Prove!

A partial order  $R \subseteq A \times A$  is a **total order** if, for all  $a, b \in A$ , either  $(a, b) \in R$  or  $(b, a) \in R$ . For example, the relation  $R_{\leq}$  is a total order. Prove!

### 3 Closures

Let  $D$  be a set, let  $n \geq 0$ , and let  $R \subseteq D^{n+1}$  be an  $(n + 1)$ -ary relation on  $D$ . Then a subset  $B$  of  $D$  is said to be **closed under  $R$**  if  $b_{n+1} \in B$  whenever  $b_1, \dots, b_n \in B$  and  $(b_1, \dots, b_{n+1}) \in R$ . Any property of the form “the set  $B$  is closed under relations  $R_1, \dots, R_m$ ” is called a **closure property** of  $B$ .

**Theorem 2** *Let  $P$  be a closure property defined by relations  $R_1, \dots, R_m$  on a set  $D$  and let  $A \subseteq D$ . Then there is a unique minimal set  $B$  of which  $A$  is a subset and which has property  $P$ .*

We call  $B$  the **closure** of  $A$  under the relations  $R_1, \dots, R_m$ . A particular case of Theorem 2 that is of special importance is the formation of the **reflexive, transitive closure** of a binary relation  $R \subseteq A \times A$ . This relation, denoted by  $R^*$ , is the closure of  $R$  under the relations

$$Q = \{(a, b), (b, c), (a, c) : a, b, c \in A\}$$

$$S = \{(a, a) : a \in A\}$$

In other words,  $R^*$  is the minimal reflexive transitive relation of which  $R$  is a subset. Examples of closures:

- The set of all Integers is closed under addition and subtraction.
- The set of all Reals is closed under division by a nonzero number.
- The set of all Naturals is not closed under subtraction.

Example: The reflexive transitive closure  $R^*$  of the relation

$$R = \{(a, b), (a, c), (a, d), (d, c), (d, e)\}$$

is:  $R^* = R \cup \{(x, x) : x \in \{a, b, c, d, e\}\} \cup \{(a, e)\}$ .

## 4 Finite and infinite sets: Counting

Two sets  $A$  and  $B$  have the same cardinality (**equinumerous**) if there is a bijection  $f : A \rightarrow B$ . For example, the sets  $A = \{1, 2, 3, 4, 5\}$  and  $B = \{red, blue, green, yellow, grey\}$  have the same cardinality, since there is a bijection from  $A$  to  $B$  (find one.)

A set  $A$  is **finite** if there is a bijection from a set  $\{1, 2, \dots, n\}$ , for some  $n$ , to  $A$ . A set  $A$  is **infinite** if it is not finite. A set  $A$  is said to be **countably infinite** if there is a bijection from  $\mathbb{N}$  to  $A$ . A set  $A$  is said to be **countable** if it is finite or countably infinite. A set that is not countable is **uncountable**.

**The Pigeonhole Principle:** *If  $A$  and  $B$  are nonempty finite sets and  $|A| > |B|$ , then there is no one-to-one function from  $A$  to  $B$ .*

**The Diagonalization Principle:** *Let  $R$  be a binary relation on a set  $A$ , and let  $D$ , the diagonal set for  $R$ , be  $\{a : a \in A \text{ and } (a, a) \notin R\}$ . For each  $a \in A$ , let  $R_a = \{b : b \in A \text{ and } (a, b) \in R\}$ . Then  $D$  is distinct from each  $R_a$ .*

**Theorem 3** *The set  $2^{\mathbb{N}}$  is uncountable.*

*Proof.* Suppose that  $2^{\mathbb{N}}$  is countably infinite, that is, there is a bijection  $f : \mathbb{N} \rightarrow 2^{\mathbb{N}}$ . Then  $2^{\mathbb{N}}$  can be enumerated as

$$2^{\mathbb{N}} = \{S_0, S_1, S_2, \dots\},$$

where  $S_i = f(i)$  for each  $i \in \mathbb{N}$ . Now consider the set

$$D = \{n \in \mathbb{N} : n \notin S_n\}.$$

$D$  is a set of natural numbers, and therefore should be  $S_k$  for some natural number  $k$ . Now we ask if  $k \in S_k$ .

1. Suppose the answer is yes,  $k \in S_k$ . Since  $D = \{n : n \notin S_n\}$ , it follows that  $k \in D$ ; but  $D = S_k$ , so  $k \notin S_k$ , a contradiction.
2. Suppose the answer is no,  $k \notin S_k$ ; then  $k \in D$ . But  $D = S_k$ , so  $k \in S_k$ , another contradiction.

Since neither (1) nor (2) is possible, the assumption that  $D = S_k$  for some  $k$  must have been an error. Hence  $2^{\mathbb{N}}$  is uncountable.  $\square$ .

## 5 Strings and languages

An **alphabet** is a nonempty finite set of symbols, denoted by  $\Sigma$ . E.g.,  $\Sigma_1 = \{0, 1\}$ ,  $\Sigma_2 = \{a, b, c, \dots, x, y, z\}$ .

A **string**  $w$  over  $\Sigma$  is a finite sequence of symbols from  $\Sigma$ . The **empty string** is a string:

- that contains no symbols
- of length 0
- denoted by  $\varepsilon$
- defined over any alphabet.

The **length** of a string  $w$ , denoted by  $|w|$ , is the number of symbols in the string. E.g.,  $|abc| = 3$ ,  $|\varepsilon| = 0$ .

$\Sigma^*$  denotes the set of all strings over  $\Sigma$ . **String concatenation**

Given two strings  $w_1 = a_1a_2 \cdots a_m$  and  $w_2 = b_1b_2 \cdots b_n$ , the concatenation of  $w_1$  and  $w_2$ , denoted by  $w_1w_2$  is the string  $a_1a_2 \cdots a_mb_1 \cdots b_n$ .

- String concatenation is associative:  $(w_1w_2)w_3 = w_1(w_2w_3)$ .
- String concatenation is not commutative:  $ab \cdot aa \neq aa \cdot ab$ .
- $\forall w \in \Sigma^*, w\varepsilon = \varepsilon w = w$ .
- $\forall w_1, w_2, |w_1w_2| = |w_1| + |w_2|$ .

A string  $v$  is a **substring** of string  $w$  if and only if there are strings  $x$  and  $y$  such that  $w = xvy$ . Both  $x$  and  $y$  could be  $\varepsilon$ , so *every string is a substring of itself*; and taking  $x = w$  and  $v = y = \varepsilon$ , we see that  $\varepsilon$  is a substring of every string. If  $w = xv$  for some  $x$ , then  $v$  is a **suffix** of  $w$ ; if  $w = vy$  for some  $y$ , then  $v$  is a **prefix** of  $w$ .

For each string  $w$  and each natural number  $i$ , the string  $w^i$  is defined as

$$\begin{aligned}w^0 &= \varepsilon, \\w^{i+1} &= w^i \cdot w, \forall i \geq 0\end{aligned}$$

The **reversal** of a string  $w$ , denoted by  $w^R$ , is defined formally as follows:

1. If  $w$  is a string of length 0, then  $w^R = w = \varepsilon$ .
2. If  $w$  is a string of length  $n + 1 > 0$ , then  $w = ua$  for some  $a \in \Sigma$ , and  $w^R = au^R$ .

E.g.,  $(park)^R = krap$ ,  $(aba)^R = aba$ .

- $(wx)^R = x^Rw^R$ . (Prove it by induction on the length of  $x$ .)

A **language**  $L$  over  $\Sigma$  is a set of strings over  $\Sigma$ , i.e., any subset of  $\Sigma^*$ . Examples of languages over  $\Sigma = \{0, 1\}$ :

$$\underbrace{L_1 = \{0, 1\}, L_2 = \{\varepsilon, 00, 001, 10001\}}_{\text{finite languages}}$$

$$\underbrace{L_3 = \{w | \#_0 w = 2\}, L_4 = \{0^i 1^i | i \in \mathbb{N}\}}_{\text{Infinite languages}}$$

- An infinite language contains strings of any length. However, the length of each string is finite.
- The language  $L = \{\varepsilon\}$  is **not** the same as  $L = \emptyset$ .

Languages are sets. Therefore, we can define the following:

- UNION:  $L_1 \cup L_2 = \{w | w \in L_1 \text{ or } w \in L_2\}$
- INTERSECTION:  $L_1 \cap L_2 = \{w | w \in L_1 \text{ and } w \in L_2\}$
- DIFFERENCE:  $L_1 - L_2 = \{w | w \in L_1 \text{ and } w \notin L_2\}$
- COMPLEMENT:  $\bar{L} = \{w \in \Sigma^* | w \notin L\}$
- CARTESIAN PRODUCT:  $L_1 \times L_2 = \{(w_1, w_2) | w_1 \in L_1 \text{ and } w_2 \in L_2\}$
- CONCATENATION:  $L_1 \cdot L_2 = \{w_1 w_2 | w_1 \in L_1 \text{ and } w_2 \in L_2\}$
- KLEENE STAR:

$$L^0 = \{\varepsilon\}$$

$$L^i = \{w_1 \cdots w_i | w_t \in L, \forall 1 \leq t \leq i\}$$

$$L^* = \bigcup_{i=0}^{\infty} L^i$$

$$L^+ = \bigcup_{i=1}^{\infty} L^i$$

Notice that  $L^*$  and  $L^+$  are infinite unless  $L = \emptyset$  or  $L = \{\varepsilon\}$ .

**Theorem 4**  $\Sigma = \{a, b\}$ .  $\Sigma^*$  is countably infinite.

*Proof:* First, sort the strings over  $\Sigma$  in lexicographic order, i.e.,  $\varepsilon, a, b, aa, ab, ba, bb, aaa, \dots$ . Notice that this way we can enumerate all the strings over  $\Sigma$ .

Define:  $g(a) = 0$  and  $g(b) = 1$ . We now define the function  $f : \Sigma^* \rightarrow \mathbb{N}$  as follows:

$$f(a_1, a_2, \dots, a_n) = 2^n + g(a_1) \cdot 2^{n-1} + g(a_2) \cdot 2^{n-2} + \dots + g(a_n) \cdot 2^0$$

E.g.,  $f(\varepsilon) = 1$ ,  $f(a) = 2$ ,  $f(b) = 3$ , ...

$f$  is one-to-one and onto (*prove!*).

$\Rightarrow \Sigma^*$  is countably infinite.  $\square$

- Every infinite language  $L$  over an alphabet  $\Sigma$  is countable, since  $L \subseteq \Sigma^*$  and  $\Sigma^*$  is countable.
- Generalize the proof of Theorem 4 to alphabets of arbitrary finite sizes.

**Theorem 5** *The set of all the languages over  $\Sigma$  is uncountable.*

*Proof:* We proved that  $\Sigma^*$  is countable.

$\Rightarrow$  we can enumerate all the strings over  $\Sigma$  in some order:  $w_1, w_2, \dots$

Assume that the set of all the languages over  $\Sigma$ , denoted by  $2^{\Sigma^*}$ , is countable.

$\Rightarrow$  We can enumerate all the languages in  $2^{\Sigma^*}$  in some order:  $L_1, L_2, \dots$

We show that there exists a language  $L \subseteq 2^{\Sigma^*}$  that is not included in the mentioned enumeration.

Build the following matrix:

	L1	L2	L3	L4	•	•	•
w1	0	1	1	0			
w2	0	1	1	0			
w3	0	1	0	0	•	•	•
w4	0	0	0	0			
w5			•				
•			•				
•			•				

Figure 1: "Enumeration" of all languages and strings

$A$  is a binary matrix:

$$a_{i,j} = \begin{cases} 1, & \text{if } w_i \in L_j \\ 0, & \text{if } w_i \notin L_j \end{cases} \tag{1}$$

Define the language  $L = \{w_i | a_{ii} = 0\}$ .

It is easy to see that:

$\forall i: w_i \in L \Leftrightarrow w_i \notin L_i$  (since  $a_{ii} = 0$ ).  
 $\Rightarrow \forall i: L_i \neq L$ .  
 $\Rightarrow L$  is not any of the languages in the given enumeration.  
 $\Rightarrow 2^{\Sigma^*} \equiv$  the set of all the languages over  $\Sigma$ , is uncountable.  $\square$