Lecture Notes on Discrete Structures

Eleftherios Gkioulekas

These notes are constantly updated by the author. If you have not obtained this file from the author's website, it may be out of date. This notice includes the date of latest update to this file. If you are using these notes for a course, I would be very pleased to hear from you, in order to document for my University the impact of this work.
The main online lecture notes website is: https://faculty.utrgv.edu/eleftherios.gkioulekas/
You may contact the author at: drlf@hushmail.com
Last updated: April 24, 2021

Contents

DST1: Logic and sets

SETS AND LOGIC

The basic concepts that we work with are

a) Propositions $\longleftrightarrow$ Boolean Algebra
b) Sets $\longleftrightarrow$ Set Algebra
c) Predicates and quantifiers $\longleftrightarrow$ 1st-order logic

Propositions

• A proposition (or statement) p is an expression which is either TRUE or FALSE.

EXAMPLES

a) $3+5=8$ is a proposition with truth value T.
b) $1+1=3$ is a proposition with truth value F.
c) $2+(10-3)^2$ is an expression but is not a proposition.

• Given the statements p, q we define compound statements as follows

p	q	$p \vee q$	$p \wedge q$	$p \underline{\vee} q$	$\bar{p}$	$p \Rightarrow q$	$p \Leftrightarrow q$
T	T	T	T	F	F	T	T
T	F	T	F	T	F	F	F
F	T	T	F	T	T	T	F
F	F	F	F	F	T	T	T

- Interpretations

$p \vee q$	Disjunction	p is true or q is true (or both) at least one of p or q is true
$p \wedge q$	Conjunction	p is true and q is true
$p \veebar q$	Exclusive Disjunction	either p or q is true (but not both)
$\bar{p}$	Negation	p is false
$p \Rightarrow q$	Implication	if p is true then q is true p implies q p is true only if q is true
$p \Leftrightarrow q$	Equivalence	p is true if and only if q is true p is equivalent to q p, q have the same truth value

Note that if p is <u>false</u> we presume that the compound statement $p \Rightarrow q$ is TRUE regardless of the truth value of q. This is necessary to ensure that $p \Leftrightarrow q$ and $(p \Rightarrow q) \wedge (q \Rightarrow p)$ have the same truth table, as shown below:

p	q	$p \Leftrightarrow q$	$p \Rightarrow q$	$q \Rightarrow p$	$(p \Rightarrow q) \wedge (q \Rightarrow p)$
T	T	T	T	T	T
T	F	F	F	T	F
F	T	F	T	F	F
F	F	T	T	T	T

For example, statements of the form

$1+1=3 \Rightarrow 2=2$

$2+3=8 \Rightarrow 3=2$

are TRUE even though the corresponding hypotheses are false.

Boolean algebra

- A boolean expression is an abstract expression that involves:

a) propositions, represented by lower-case letters (e.g. p, q, r, etc)

b) Boolean operations: $\wedge$ (conjunction), $\vee$ (disjunction), $\veebar$ (exclusive disjunction), $\bar{\ }$ (negation), $\Rightarrow$ (implication), $\Leftrightarrow$ (equivalence)

c) T: a propostion with truth value fixed at TRUE.

d) F: a proposition with truth value fixed at FALSE

e) Parenthesis, to prioritize the order of boolean operations.

- Given two boolean expressions P, Q:

$P \equiv Q$: P and Q have the same truth table

P tautology $\Leftrightarrow P \equiv T$

P contradiction $\Leftrightarrow P \equiv F$

The above are an example of "metalogic", i.e. logic about logic!

- With the above terminology we can use truth tables to establish the following properties of Boolean Algebra:

- Commutative

$p \wedge q \equiv q \wedge p$

$p \vee q \equiv q \vee p$

- Associative

$p \wedge (q \wedge r) \equiv (p \wedge q) \wedge r$

$p \vee (q \vee r) \equiv (p \vee q) \vee r$

- Distributive

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

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

- Reductions $\longrightarrow$ These properties allow us to rewrite all boolean expressions in terms of conjunction, disjunction, and negation.

$p \underline{\vee} q \equiv (p \wedge \bar{q}) \vee (\bar{p} \wedge q)$

$p \Rightarrow q \equiv \bar{p} \vee q$

$p \Leftrightarrow q \equiv (p \Rightarrow q) \wedge (q \Rightarrow p)$

- Negations:

$\overline{p \wedge q} \equiv \bar{p} \vee \bar{q}$

$\overline{p \vee q} \equiv \bar{p} \wedge \bar{q}$

$\Big\}$ De Morgan's laws

and it follows that

$\overline{p \Rightarrow q} \equiv \overline{\bar{p} \vee q} \equiv p \wedge \bar{q}$

and

$\overline{p \Leftrightarrow q} \equiv \overline{(p \Rightarrow q) \wedge (q \Rightarrow p)} \equiv \overline{(p \Rightarrow q)} \vee \overline{(q \Rightarrow p)}$

$\equiv (p \wedge \bar{q}) \vee (\bar{p} \wedge q)$

- Relationship between equivalence and exclusive disjunction:

$\overline{p \Leftrightarrow q} \equiv p \underline{\vee} q$

$\overline{p \underline{\vee} q} \equiv p \Leftrightarrow q$

$\hookrightarrow$ The above properties are established via truth tables, as in the following example.

EXAMPLE

Use truth tables to show that $\overline{p \wedge q} \equiv \bar{p} \vee \bar{q}$.

Solution

We note that

p	q	$p \wedge q$	$\overline{p \wedge q}$
T	T	T	F
T	F	F	T
F	T	F	T
F	F	F	T

and

p	q	$\bar{p}$	$\bar{q}$	$\bar{p} \vee \bar{q}$
T	T	F	F	F
T	F	F	T	T
F	T	T	F	T
F	F	T	T	T

It follows that $\overline{p \wedge q} \equiv \bar{p} \vee \bar{q}$ $\square$

<u>Methodology</u> : To show that a boolean expression is a tautology via boolean algebra

- 1 Use the reduction formulas to rewrite the boolean expression in terms of $\wedge$ (conjunction), $\vee$ (disjunction), $^-$ (negation).
- 2 Use the De Morgan laws to reduce all negations down to individual statements
- 3 Simplify using the associative, distributive properties in addition to the following self-evident statements:

$p \vee F \equiv p$	$p \wedge T \equiv p$	$p \vee \bar{p} \equiv T$
$p \wedge F \equiv F$	$p \vee T \equiv T$	$p \wedge \bar{p} \equiv F$

EXAMPLE

Show that $[p \wedge (p \Rightarrow q)] \Rightarrow q$ is a tautology.

<u>Solution</u>

$$S \equiv [p \wedge (p \Rightarrow q)] \Rightarrow q \equiv \overline{[p \wedge (p \Rightarrow q)]} \vee q \equiv$$
$$\equiv [\bar{p} \vee \overline{(p \Rightarrow q)}] \vee q \equiv [\bar{p} \vee (p \wedge \bar{q})] \vee q \equiv$$
$$\equiv [(\bar{p} \vee p) \wedge (\bar{p} \vee \bar{q})] \vee q \equiv [T \wedge (\bar{p} \vee \bar{q})] \vee q \equiv$$
$$\equiv (\bar{p} \vee \bar{q}) \vee q \equiv \bar{p} \vee (\bar{q} \vee q) \equiv \bar{p} \vee T \equiv T$$

and therefore $[p \wedge (p \Rightarrow q)] \Rightarrow q$ is a tautology.

EXERCISES

(1) Evaluate the truth value of the following statements

a) $3+7=10 \vee 1+3=4$

b) $2+1=4 \vee 1+3=5$

c) $3\neq 4 \wedge 1+1=2$

d) $2+5=8 \wedge 3+3=6$

e) $1+4=5 \Rightarrow 3=2$

f) $3+2=0 \Rightarrow 5=6$

g) $1=2 \Rightarrow 3=3$

h) $2+3=5 \Leftrightarrow 1+1=2$

i) $3+1=2+2 \Leftrightarrow 1=0$

(2) In the following compound statements replace with letters (e.g. p, q, r, ...) the simple constituent statements and write the structure of the compound statements in terms of the letters you introdued

a) 30 is a multiple of 6 and divisible by 5

b) 5 is either an even or an odd number

c) If $ab=0$, then $a=0$ or $b=0$.

d) 8 is not a prime number

e) The triangles $\overset{\triangle}{ABC}$ and $\overset{\triangle}{DEF}$ are similar if and only if $\hat{A}=\hat{D}$ and $\hat{B}=\hat{E}$ and $\hat{C}=\hat{F}$.

(3) Show that the following expressions are tautologies using truth tables

a) $[\bar{p} \wedge (p \vee q)] \Rightarrow q$

b) $\overline{(p \Rightarrow q)} \Leftrightarrow (p \wedge \bar{q})$

c) $\overline{(p \Leftrightarrow q)} \Leftrightarrow (\bar{p} \Leftrightarrow q)$

d $\overline{(p \Leftrightarrow q)} \Leftrightarrow (p \Leftrightarrow \bar{q})$

(4) Show that the following expressions are tautologies using boolean algebra.

a) $(p \wedge q) \Rightarrow q$

b) $p \Rightarrow (p \vee q)$

c) $[\bar{q} \wedge (p \Rightarrow q)] \Rightarrow \bar{p}$

d) $(p \underline{\vee} q) \Rightarrow (p \vee q)$

e) $(\bar{p} \wedge (\bar{q} \Rightarrow p)) \Rightarrow q$

(5) Write the expressions of the previous exercise in English

▶ Methodology: Application to inequalities.

We note that:

$\overline{x<a} \Leftrightarrow x \geq a$	$\overline{x>a} \Leftrightarrow x \leq a$
$\overline{x \leq a} \Leftrightarrow x > a$	$\overline{x \geq a} \Leftrightarrow x < a$

- Weak inequalities are defined via disjunction from strong inequalities:

  $a \leq b \Leftrightarrow (a<b \vee a=b)$

  $a \geq b \Leftrightarrow (a>b \vee a=b)$
- Composite inequalities are equivalent to conjunction of elementary inequalities. For example:

  $$a<b<c \Leftrightarrow a<b \wedge b<c \Leftrightarrow \begin{cases} a<b \\ b<c \end{cases}$$

  The braces notation is used to represent conjunction.
- We can use the above, in conjunction with boolean algebra to negate expressions involving inequalities

EXAMPLE

Negate the statement

$p: \; 0<|x-x_0|<\delta \Rightarrow 0<|y-y_0|<\varepsilon$

Solution

$$\overline{p} \equiv \overline{0<|x-x_0|<\delta \Rightarrow 0<|y-y_0|<\varepsilon}$$
$$\equiv 0<|x-x_0|<\delta \wedge \overline{0<|y-y_0|<\varepsilon}$$
$$\equiv 0<|x-x_0|<\delta \wedge (\overline{0<|y-y_0| \wedge |y-y_0|<\varepsilon})$$
$$\equiv 0<|x-x_0|<\delta \wedge (\overline{0<|y-y_0|} \vee \overline{|y-y_0|<\varepsilon})$$
$$\equiv 0<|x-x_0|<\delta \wedge (0\geq|y-y_0| \vee |y-y_0|\geq\varepsilon)$$
$$\equiv 0<|x-x_0|<\delta \wedge (y=y_0 \vee |y-y_0|\geq\varepsilon)$$

EXERCISES

(6) Write and simplify the negation to the following statements.

a) $3x<x^2+1<5$

b) $\begin{cases} 2x+y>3 \\ x-y<1 \end{cases}$

c) $2x<1 \Leftrightarrow y>2$

d) $a<b<c \Leftrightarrow b+c+d>2$

e) $x+1<y \veebar x^2<2y<3x+5$

f) $a<b \Rightarrow (c<d \veebar c>e)$

g) $\begin{cases} x<1 \\ y\leq 2 \end{cases} \vee \begin{cases} x\geq 3 \\ y\geq 1 \end{cases}$

h) $\begin{cases} x>2 \veebar y<3 \\ z\leq 1 \end{cases}$

i) $ab\geq c \Rightarrow \begin{cases} b\geq d \\ a\leq d \end{cases}$

j) $\begin{cases} x\geq 1 \vee y<3 \\ z>y\geq x \end{cases}$

Sets – Definitions

- A set is an unordered collection of an arbitrary number of elements. A set can be an element of another set.

notation: $x \in A$: the element x belongs to A

$x \notin A$: the element x does NOT belong to A.

We also introduce the following abbreviations:

$x, y \in A \Leftrightarrow (x \in A \wedge y \in A)$

$x, y, z \in A \Leftrightarrow (x \in A \wedge y \in A \wedge z \in A)$

and so on.

Definition of sets

- Sets can be defined by providing a belonging condition i.e. a boolean expression $P(x)$ involving a variable x such that

$$x \in A \Leftrightarrow P(x)$$

is a tautology.

e.g. The set with elements $1, 2, 3$ can be defined by the belonging condition

$$x \in A \Leftrightarrow (x = 1 \vee x = 2 \vee x = 3)$$

Equivalently we write $A = \{1, 2, 3\}$.

- The empty set $\emptyset$ is a set that contains no elements. A formal definition is:

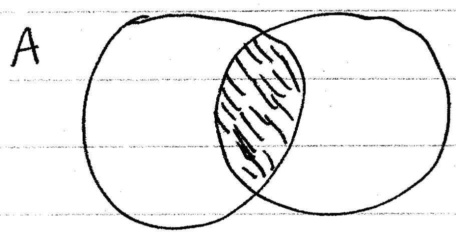
$$x \in \emptyset \Leftrightarrow F$$

▶ Operations with sets

Let A, B be two sets. We use belonging conditions to define:

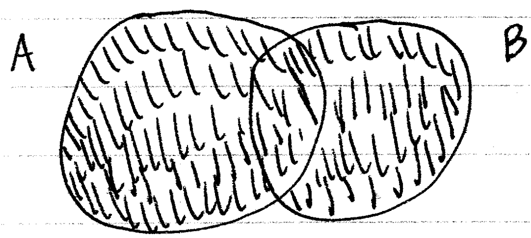
1) **Intersection** $A \cap B$

$$x \in A \cap B \Leftrightarrow x \in A \wedge x \in B$$



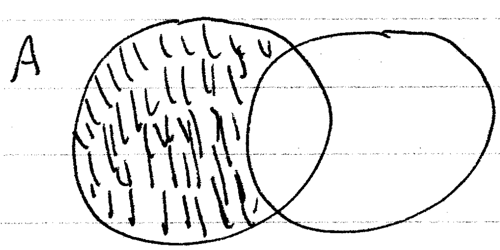
2) **Union** $A \cup B$

$$x \in A \cup B \Leftrightarrow x \in A \vee x \in B$$



3) **Difference** $A - B$

$$x \in A - B \Leftrightarrow x \in A \wedge x \notin B$$



▸ Relations between sets

a) Set equality : $A = B$ (i.e. "A is equal to B") means that the sets A, B have the same elements. A formal definition requires using metalogic:

$$\boxed{\begin{array}{l} A = B \Leftrightarrow [(x \in A \Leftrightarrow x \in B) \equiv T] \\ A \neq B \Leftrightarrow \overline{A = B} \end{array}}$$

↳ For any arbitrary boolean expression $P(x)$ we use the notation
$\forall x : P(x)$
as equivalent to $P(x) \equiv T$. In English; this statement reads: "For all x, $P(x)$ is true".
We may therefore rewrite the above definition as

$$\boxed{A = B \Leftrightarrow \forall x : (x \in A \Leftrightarrow x \in B)}$$

This is an example of the fundamental universal quantified statement. Later we will use set equality to define the 3 types of quantified statements that are regularly used in practice. The quantifier $\forall x$ runs over the class V of all elements that can ever be defined within a rigorous set theoretic axiomatic framework (e.g. ZFC).

b) <u>Subset</u> : $A \subseteq B$ means that all elements of A also belong to B (i.e. A is a subset of B).

The formal definition is:

$$\boxed{\begin{aligned} A \subseteq B &\Leftrightarrow [(x \in A \Rightarrow x \in B) \equiv T] \\ &\Leftrightarrow \forall x : (x \in A \Rightarrow x \in B) \\ A \not\subseteq B &\Leftrightarrow \overline{A \subseteq B} \end{aligned}}$$

Note that $x \in A \Rightarrow x \in A$ and $F \Rightarrow x \in A$ are obvious tautologies and therefore $A \subseteq A$ and $\emptyset \subseteq A$ are always true.

c) <u>Strict subset</u> : $A \subset B$ ("A is a strict subset of B") is defined as:

$$\boxed{\begin{aligned} A \subset B &\Leftrightarrow (A \subseteq B \wedge A \neq B) \\ A \not\subset B &\Leftrightarrow \overline{A \subset B} \end{aligned}}$$

▶ <u>Power set</u>

Given a set A, the power set $\mathcal{P}(A)$ is the set of all subsets of A. We define $\mathcal{P}(A)$ via the following belonging conditions:

$$\boxed{X \in \mathcal{P}(A) \Leftrightarrow X \subseteq A}$$

Note that for all sets A: $\emptyset \in \mathcal{P}(A) \wedge A \in \mathcal{P}(A)$.

EXAMPLES

$A = \{a, b\} \Rightarrow \mathcal{P}(A) = \{\emptyset, \{a\}, \{b\}, \{a, b\}\}$
$A = \{a, b, c\} \Rightarrow \mathcal{P}(A) = \{\emptyset, \{a\}, \{b\}, \{c\}, \{a, b\}, \{b, c\},$
$\{c, a\}, \{a, b, c\}\}$
$\mathcal{P}(\emptyset) = \{\emptyset\}$
$\mathcal{P}(\{a\}) = \{\emptyset, \{a\}\}$
$\hookrightarrow$ Note that $\emptyset$ and A always belong to $\mathcal{P}(A)$.

▸ Number sets

We define the following number sets.

a) Natural numbers
$\mathbb{N} = \{0, 1, 2, 3, \dots\}$
$\mathbb{N}^* = \{1, 2, 3, \dots\}$ $\quad [n] = \{1, 2, 3, \dots, n\}$

b) Integers (from Zahl in German)
$\mathbb{Z} = \{0, 1, -1, 2, -2, 3, -3, \dots\}$
$\mathbb{Z}^* = \{1, -1, 2, -2, 3, -3, \dots\}$

c) Rational numbers
$\mathbb{Q}$ contains all rational numbers
$\mathbb{Q}^* = \mathbb{Q} - \{0\}$

d) Real numbers
$\mathbb{R}$ contains all real numbers; $\mathbb{R}^* = \mathbb{R} - \{0\}$.

Remarks

a) Cantor proposed that starting from the empty set, with set operations, we can represent natural numbers as sets. Then, all other number sets can be constructed from $\mathbb{N}$. Cantor's construction was to define

$$0 = \emptyset$$
$$1 = \{0\} = \{\emptyset\}$$
$$2 = \{0,1\} = \{\emptyset, \{\emptyset\}\}$$
$$3 = \{0,1,2\} = \{\emptyset, \{\emptyset\}, \{\emptyset, \{\emptyset\}\}\}$$

etc.

Equivalently, Cantor's construction can be represented recursively as:

$$\begin{cases} 0 = \emptyset \\ (n+1) = n \cup \{n\} \end{cases}$$

Then, a "transfinite induction" step is used to round up all natural numbers to build $\mathbb{N}$.

b) The set $\mathbb{Q}$ of the rational numbers can be defined from $\mathbb{N}$ and $\mathbb{Z}$ using definition by mapping, to be explained later.

c) Constructing $\mathbb{R}$ from $\mathbb{Q}$ is a non-trivial problem, and many approaches exist.

EXAMPLES

a) Given $A=([6]-[3])\cap[5]$ and $B=([7]-[4])\cup[2]$
list the elements of $C=A-B$

Solutions

Since

$$A=([6]-[3])\cap[5]=$$
$$=(\{1,2,3,4,5,6\}-\{1,2,3\})\cap\{1,2,3,4,5\}=$$
$$=\{4,5,6\}\cap\{1,2,3,4,5\}=\{4,5\}$$

and

$$B=([7]-[4])\cup[2]=$$
$$=(\{1,2,3,4,5,6,7\}-\{1,2,3,4\})\cup\{1,2\}$$
$$=\{5,6,7\}\cup\{1,2\}=\{1,2,5,6,7\}$$

it follows that

$$A-B=\{4,5\}-\{1,2,5,6,7\}=\{4\}$$

b) List the elements of $A=\mathcal{P}([6]-([2]\cup[4]))$.

Solution

$$A=\mathcal{P}([6]-([2]\cup[4]))=$$
$$=\mathcal{P}(\{1,2,3,4,5,6\}-(\{1,2\}\cup\{1,2,3,4\}))$$
$$=\mathcal{P}(\{1,2,3,4,5,6\}-\{1,2,3,4\})$$
$$=\mathcal{P}(\{5,6\})=\{\emptyset,\{5\},\{6\},\{5,6\}\}$$

c) List the elements of $A = \mathcal{P}(\mathcal{P}(\{1\}))$

<u>Solution</u>

$$\begin{aligned} A &= \mathcal{P}(\mathcal{P}(\{1\})) \\ &= \mathcal{P}(\{\emptyset, \{1\}\}) \\ &= \{\emptyset, \{\emptyset\}, \{\{1\}\}, \{\emptyset, \{1\}\}\} \end{aligned}$$

EXERCISES

(7) List the elements of $A \cap B$, $A \cup B$, $A - B$, $B - A$ for the following choices of A and B:

a) $A = [6] - [3]$ and $B = [8] - [5]$

b) $A = [3] \cup [5]$ and $B = [4] \cap [2]$

c) $A = [3] \cap [2]$ and $B = [2] - [6]$

(8) List the elements of the following sets

a) $\mathcal{P}([2])$

b) $\mathcal{P}([5] - [4])$

c) $\mathcal{P}([3] - [6])$

d) $\mathcal{P}(([5] - [2]) \cap [4])$

e) $\mathcal{P}(([6] \cap [4]) - [2])$

f) $\mathcal{P}(\mathcal{P}(\emptyset))$

g) $\mathcal{P}(\{1\})$

(9) Which of the following statements is TRUE?

a) $\mathbb{N} \subseteq \mathbb{N}$

b) $\mathbb{N} \subset \mathbb{Z}$

c) $\mathbb{Z} \subseteq \mathbb{N}$

d) $\mathbb{N} \cap \mathbb{Z} = \mathbb{N}$

e) $\mathbb{N} \cap \mathbb{Z} = \mathbb{Z}$

f) $\mathbb{N} \cup \mathbb{Z} = \mathbb{N}$

g) $\mathbb{N} \cup \mathbb{Z} = \mathbb{Z}$

h) $[3] \cap [5] \subseteq [4]$

i) $[4] - [2] \subset [3]$

j) $[2] \cup [6] \subset [6]$

k) $[3] \cap [5] \subseteq [3]$

l) $1 \in \emptyset$

m) $\emptyset \in \mathcal{P}(\emptyset)$

n) $\emptyset \notin \mathcal{P}(\mathcal{P}(\emptyset))$

Proving set properties

Set properties can be proved via logic as follows:

a) Set operations can be reduced using the following tautologies:

$$x \in A \cap B \Leftrightarrow x \in A \wedge x \in B$$
$$x \in A \cup B \Leftrightarrow x \in A \vee x \in B$$
$$x \in A - B \Leftrightarrow x \in A \wedge x \notin B$$

b) To show that $A = B$ it is sufficient to show that
$x \in A \Leftrightarrow x \in B$.

This can be done with

1) Direct proof:

$$x \in A \Leftrightarrow p_1(x) \Leftrightarrow p_2(x) \Leftrightarrow \cdots \Leftrightarrow p_n(x) \Leftrightarrow x \in B$$

2) Separate forward / converse proof

($\Rightarrow$): Assume that $x \in A$. Then:

$$x \in A \Rightarrow p_1(x) \Rightarrow p_2(x) \Rightarrow \cdots \Rightarrow p_n(x) \Rightarrow x \in B$$

($\Leftarrow$): Assume that $x \in B$. Then

$$x \in B \Rightarrow q_1(x) \Rightarrow q_2(x) \Rightarrow \cdots \Rightarrow q_n(x) \Rightarrow x \in A$$

From the above: $\begin{cases} A \subseteq B \\ B \subseteq A \end{cases} \Rightarrow A = B.$

c) To show $A \subseteq B$ it is sufficient to show that
$x \in A \Rightarrow x \in B$

This requires only the forward argument.

d) To show $A = \emptyset$, it is sufficient to show that

$x \in A \Rightarrow F$

where F is a contradiction (i.e. a universally false statement). The converse statement $F \Rightarrow x \in A$ is also needed, but it is a tautology so it does not require a proof.

↳ For unidirectional arguments (i.e. using "$\Rightarrow$" steps instead of "$\Leftrightarrow$") we are allowed the following additional manipulations:

$p \Rightarrow p \vee q$ (where q is an arbitrary statement)

$p \wedge q \Rightarrow p$

i.e.: we can always ADD an arbitrary statement q using logical OR (disjunction), and from a statement $p \wedge q$ involving the logical AND (conjunction) of multiple statements we can remove any statement we want. However these manipulations are <u>not</u> reversible. More generally:

$p \Rightarrow p \vee q_1 \vee q_2 \vee \cdots \vee q_n$

$p \wedge q_1 \wedge q_2 \wedge \cdots \wedge q_n \Rightarrow p$

EXAMPLES

a) Show that: $G-(A\cap B)=(G-A)\cup(G-B)$.

Solution

Since,

$$
\begin{aligned}
x\in G-(A\cap B) &\Leftrightarrow x\in G \wedge \underline{x\notin A\cap B} \Leftrightarrow \\
&\Leftrightarrow x\in G\wedge \overline{(x\in A\wedge x\in B)} \Leftrightarrow \\
&\Leftrightarrow x\in G\wedge (x\notin A\vee x\notin B)\Leftrightarrow \\
&\Leftrightarrow (x\in G\wedge x\notin A)\vee(x\in G\wedge x\notin B)\Leftrightarrow \\
&\Leftrightarrow x\in G-A\vee x\in G-B \\
&\Leftrightarrow x\in (G-A)\cup(G-B)
\end{aligned}
$$

it follows that $G-(A\cap B)=(G-A)\cup(G-B)$ $\square$

b) Show that: $A\cap B\subseteq A\cup B$.

Solution

Since,

$$
\begin{aligned}
x\in A\cap B &\Rightarrow x\in A\wedge x\in B \\
&\Rightarrow x\in A \qquad \text{(remark: converse not true)} \\
&\Rightarrow x\in A\vee x\in B \qquad \text{(remark: converse not true)} \\
&\Rightarrow x\in A\cup B
\end{aligned}
$$

it follows that $A\cap B\subseteq A\cup B$ $\square$

↳ The 2nd and 3rd steps cannot be reversed because they are based on the tautologies $p\wedge q\Rightarrow p$ and $p\Rightarrow p\vee q$. The other steps can be reversed, but the proof does not require us to exercise that possibility

c) Show that: $(A-B)\cap B=\emptyset$

<u>Solution</u>

Since,

$$\begin{aligned} x\in(A-B)\cap B &\Rightarrow x\in A-B \wedge x\in B \\ &\Rightarrow (x\in A \wedge x\notin B)\wedge x\in B \\ &\Rightarrow x\in A \wedge (x\notin B \wedge x\in B) \\ &\Rightarrow x\in A \wedge F \\ &\Rightarrow F \end{aligned}$$

and therefore $(A-B)\cap B=\emptyset$.

EXERCISES

(10) Show the following set identities, given sets A, B, C, D.

a) $C-(C-A) = A \cap C$

b) $(A-B) \cup A = A$

c) $A \cap (B-C) = (A \cap B)-(A \cap C)$

d) $(A-B) \cap (B-A) = \emptyset$

e) $(A-C) \cap (B-C) = (A \cap B) - C$

f) $(B-A) \cap (A \cap B) = \emptyset$

g) $(A \cup B) - B = A - (A \cap B) = A - B$

h) $A-(B-C) = (A-B) \cup (A \cap C)$

i) $(A-B)-C = A-(B \cup C)$

j) $(A-B) \cap (C-D) = (A \cap C) - (B \cup D)$.

▼ Predicates and quantified statements

- A predicate $p(x)$ is a statement about x which is TRUE or FALSE depending on the value of x.
- Assume that $x \in U$ where U is some universal set. Then the truth set of $p(x)$ is the set of all $x \in U$ for which $p(x)$ is true, and is denoted as:

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

  The belonging condition for the truth set A is given by

  $$x \in A \Leftrightarrow x \in U \wedge p(x)$$

▶ Remark: In algebra, equations, inequalities, systems of equations, systems of inequalities are examples of predicates. For example, consider the predicate consisting of a quadratic equation:

$p(x)$: $x^2+3x+2=0$

Solving an equation is equivalent to finding the corresponding truth set:

$$x^2+3x+2=0 \Leftrightarrow (x+1)(x+2)=0 \Leftrightarrow x+1=0 \vee x+2=0 \Leftrightarrow$$
$$\Leftrightarrow x=-1 \vee x=-2 \Leftrightarrow x \in \{-1,-2\}.$$

It follows that

$$S = \{x \in \mathbb{R} \mid x^2+3x+2=0\} = \{-1,-2\}$$

For systems of equations and systems of inequalities we use braces as an abbreviation for conjunction. For example,

$\begin{cases} x+y=3 \\ x-y=2 \end{cases}$ is equivalent to $x+y=3 \wedge x-y=2$.

▶ Quantified statements

Let A be a set and $p(x)$ a predicate. Then, we define:

1) The universal quantifier $\forall$
$$\boxed{(\forall x \in A : p(x)) \Longleftrightarrow \{x \in A \mid p(x)\} = A}$$

interpretation: "For all $x \in A$, the statement $p(x)$ is true."

2) The existential quantifier $\exists$
$$\boxed{(\exists x \in A : p(x)) \Longleftrightarrow \{x \in A \mid p(x)\} \neq \emptyset}$$

interpretation: There exists some $x \in A$ such that $p(x)$ is true
There is at least one $x \in A$ such that $p(x)$ is true

3) The unique-existential quantifier $\exists!$
$$\boxed{(\exists!\, x \in A : p(x)) \Longleftrightarrow \exists y \in A : \{x \in A \mid p(x)\} = \{y\}}$$

interpretation: There is a unique $x \in A$ such that $p(x)$ is true.
There is one and only one $x \in A$ such that $p(x)$ is true.

↳ An equivalent definition of the unique-existential quantifier $\exists!$ reads:

$$\boxed{(\exists!\, x \in A : p(x)) \Longleftrightarrow \begin{cases} \forall x_1, x_2 \in A : ((p(x_1) \land p(x_2)) \Longrightarrow x_1 = x_2) \\ \exists x \in A : p(x) \end{cases}}$$

Remarks

a) If A is a finite set, then there is a direct correspondance between quantifiers and boolean operations:

$\forall \longleftrightarrow$ generalizes conjunction (i.e. $p \wedge q$)

$\exists \longleftrightarrow$ generalizes disjunction (i.e $p \vee q$)

$\exists! \longleftrightarrow$ generalizes exclusive disjunction (i.e. $p \veebar q$)

For example, for $A = \{a, b, c\}$

$(\forall x \in A : p(x)) \Leftrightarrow p(a) \wedge p(b) \wedge p(c)$

$(\exists x \in A : p(x)) \Leftrightarrow p(a) \vee p(b) \vee p(c)$

Thus, quantifiers function like "summation operators" for conjunction, disjunction, and exclusive disjunction.

b) In a statement of the form $\forall x \in A : p(x)$, the variable x is local, i.e. it exists only inside the quantifier to formulate the statement $p(x)$. However, x does not exist outside the overall statement. Likewise, for the other two quantifiers.

c) Quantifiers can be nested

$\forall x \in A : \exists y \in B : \forall z \in C : p(x, y, z)$

(i.e. for all $x \in A$, there is some $y \in B$ such that for all $z \in C$ we have $p(x, y, z)$)

We also use the following abbreviations:

$\forall x, y \in A : p(x, y) \Leftrightarrow \forall x \in A : \forall y \in A : p(x, y)$

$\exists x, y \in A : p(x, y) \Leftrightarrow \exists x \in A : \exists y \in A : p(x, y)$

and likewise for multiple variables.

▶ Negation of quantified statements

The universal and existential quantified statements can be negated by the following generalization of De Morgan's law:

$$\boxed{\begin{aligned} &\overline{\forall x \in A : p(x)} \iff \exists x \in A : \overline{p(x)} \\ &\overline{\exists x \in A : p(x)} \iff \forall x \in A : \overline{p(x)} \end{aligned}}$$

▶ Quantified statements and limits in Analysis

Historically, quantified statements were introduced to state precisely and concisely the definition of limits in analysis, as well as many other definitions and theorems.

For example, the standard definition of a limit can be written as

$$\lim_{x \to x_0} f(x) = \ell \iff \forall \varepsilon \in (0, +\infty) : \exists \delta \in (0, +\infty) : \forall x \in A :$$
$$: (0 < |x - x_0| < \delta \implies |f(x) - \ell| < \varepsilon)$$

It is standard convention in analysis to replace $\varepsilon \in (0, +\infty)$ with $\varepsilon > 0$ and $\delta \in (0, +\infty)$ with $\delta > 0$ and rewrite the above definition as:

$$\lim_{x \to x_0} f(x) = \ell \iff$$
$$\iff \forall \varepsilon > 0 : \exists \delta > 0 : \forall x \in A : (0 < |x - x_0| < \delta \implies |f(x) - \ell| < \varepsilon)$$

Translated in English: " $\lim_{x \to x_0} f(x) = \ell$ if and only if for all $\varepsilon > 0$, there is some $\delta > 0$ such that for all $x \in A$, if $0 < |x - x_0| < \delta$ then $|f(x) - \ell| < \varepsilon$ ".

Using the negation property we can rewrite the definition for $\lim\limits_{x\to x_0} f(x) \neq \ell$ as follows:

$$\lim_{x\to x_0} f(x) \neq \ell \Leftrightarrow$$

$$\Leftrightarrow \overline{\forall \varepsilon>0 : \exists \delta>0 : \forall x\in A : (0<|x-x_0|<\delta \Rightarrow |f(x)-\ell|<\varepsilon)}$$

$$\Leftrightarrow \exists \varepsilon>0 : \overline{\exists \delta>0 : \forall x\in A : (0<|x-x_0|<\delta \Rightarrow |f(x)-\ell|<\varepsilon)}$$

$$\Leftrightarrow \exists \varepsilon>0 : \forall \delta>0 : \overline{\forall x\in A : (0<|x-x_0|<\delta \Rightarrow |f(x)-\ell|<\varepsilon)}$$

$$\Leftrightarrow \exists \varepsilon>0 : \forall \delta>0 : \exists x\in A : \overline{(0<|x-x_0|<\delta \Rightarrow |f(x)-\ell|<\varepsilon)}$$

$$\Leftrightarrow \exists \varepsilon>0 : \forall \delta>0 : \exists x\in A : (0<|x-x_0|<\delta \wedge |f(x)-\ell|\geq\varepsilon).$$

Translated in English: "$\lim\limits_{x\to x_0} f(x) \neq \ell$ if and only if there is some $\varepsilon>0$ such that for all $\delta>0$, there is some $x\in A$ such that $0<|x-x_0|<\delta$ and $|f(x)-\ell|\geq\varepsilon$".

EXERCISES

(10) Write the following statements symbolically using quantifiers

a) Every real number is equal to itself.

b) There is a real number x such that $2x = 3(1-x)$.

c) The equation $x^2+4x+4=0$ has a unique solution on $\mathbb{R}$.

d) For every real number x, there is a natural number n such that $n > x$.

e) For every real number x, there is a complex number z such that $x - z^2 = 0$.

f) For every real number x, there is a unique real number y such that $x + y = 0$.

g) For all $\varepsilon > 0$, there is a $\delta > 0$ such that for all real numbers x, if $x_0 - \delta < x < x_0 + \delta$ then $f(x) > 1/\varepsilon$.

h) There is a real number b such that for all natural numbers n we have $a_n < b$.

i) For all $\varepsilon > 0$, there is a natural number n_0 such that for any two natural numbers n_1 and n_2, if $n_1 > n_0$ and $n_2 > n_0$ then we have $|a_{n_1} - a_{n_2}| < \varepsilon$.

j) For any $M > 0$, there is a natural number n_0 such that for any other natural number n, if $n > n_0$ then $a_n > M$.

(11) Write the negations of the statements of the previous exercise, first using quantifier notation, and then in English.

▶ Quantified statements and Euclidean geometry

Quantified statements can be used to encode Hilbert's axioms of Euclidean geometry. Let $\mathbb{P}$ be the set of all <u>points</u> on a <u>plane</u>. Let $\mathbb{L} \subseteq \mathcal{P}(\mathbb{P})$ be the set of all <u>lines</u> of the plane $\mathbb{P}$. Then we can restate some of Hilbert's axioms as follows:

1) For every two points A, B there is a unique line (ℓ) passing through them

$\forall A \in \mathbb{P} : \forall B \in \mathbb{P} - \{A\} : \exists ! (\ell) \in \mathbb{L} : A, B \in (\ell)$

2) There are at least two points on every line

$\forall (\ell) \in \mathbb{L} : \exists A, B \in \mathbb{P} : (A \neq B \wedge A, B \in (\ell))$

3) There exist at least three points that do not all lie on the same line

$\exists A, B, C \in \mathbb{P} : \forall (\ell) \in \mathbb{L} : \overline{(A, B, C \in (\ell))}$

↳ To eliminate the negation, we note that

$$\begin{aligned} \overline{A, B, C \in (\ell)} &\Leftrightarrow \overline{A \in (\ell) \wedge B \in (\ell) \wedge C \in (\ell)} \\ &\Leftrightarrow \overline{A \in (\ell)} \vee \overline{B \in (\ell)} \vee \overline{C \in (\ell)} \\ &\Leftrightarrow A \notin (\ell) \vee B \notin (\ell) \vee C \notin (\ell) \end{aligned}$$

and therefore the above statement can be rewritten as:

$\exists A, B, C \in \mathbb{P} : \forall (\ell) \in \mathbb{L} : (A \notin (\ell) \vee B \notin (\ell) \vee C \notin (\ell))$.

EXERCISES

(12) In Hilbert's axiomatic formulation of Euclidean Geometry he introduced the statement $A*B*C$ to represent "B is between A and C". This allows defining the line segment AC as

$$AC = \{B \in \mathbb{P} \mid A*B*C\} \cup \{A, C\}$$

Write the following Hilbert axioms using quantified statements.

a) If B is between A and C, then the points A, B, C lie on the same line and B is between C and A.

b) For any points B, D, there are points A, C, E such that B is between A and D, C is between B and D, and D is between B and E.

c) For any three points A, B, C on a line, there exists no more than one point that lies between the other two points.

d) For any line (ℓ) and any point A not on (ℓ), there is exactly one line (ℓ_0) passing through A that is parallel to (ℓ).

(13) Let $A, B \in \mathbb{P}$ be two points and $(\ell) \in \mathbb{L}$ be a line. Write the following statements using quantifiers and set notation.

a) For any points A, B and any line (ℓ), A, B are on the same side of line (ℓ) (notation $A*B*(\ell)$) if and only if AB does not intersect with the line (ℓ).

6) For any 3 points A, B, C and any line (ℓ), if A, B are on the same side of the line (ℓ) and B, C are on the same side of (ℓ), then A, C are on the same side of (ℓ).

▼ Indexed set collections

- Let I be a set. An indexed collection of sets $\{A_\alpha\}_{\alpha\in I}$ represents a collection of sets such that for every $\alpha \in I$, there is a corresponding set A_α. In this context, we say that I is the index set of the collection.
- Let $\{A_\alpha\}_{\alpha\in I}$ be an indexed collection of sets. We define:

$$\boxed{\begin{aligned} x \in \bigcup_{\alpha\in I} A_\alpha &\iff \exists \alpha \in I : x \in A_\alpha \\ x \in \bigcap_{\alpha\in I} A_\alpha &\iff \forall \alpha \in I : x \in A_\alpha \end{aligned}}$$

- The corresponding negation of this definition reads:

$$\boxed{\begin{aligned} x \notin \bigcup_{\alpha\in I} A_\alpha &\iff \forall \alpha \in I : x \notin A_\alpha \\ x \notin \bigcap_{\alpha\in I} A_\alpha &\iff \exists \alpha \in I : x \notin A_\alpha \end{aligned}}$$

- For proofs requiring us to "juggle" with quantified statements, the following factorization rules are helpful.

- **Associative property**

$$\boxed{\begin{aligned} p \wedge (\forall x \in A : q(x)) &\Leftrightarrow \forall x \in A : (p \wedge q(x)) \\ p \vee (\exists x \in A : q(x)) &\Leftrightarrow \exists x \in A : (p \vee q(x)) \end{aligned}}$$

- **Distributive property**

$$\boxed{\begin{aligned} p \vee (\forall x \in A : q(x)) &\Leftrightarrow \forall x \in A : (p \vee q(x)) \\ p \wedge (\exists x \in A : q(x)) &\Leftrightarrow \exists x \in A : (p \wedge q(x)) \end{aligned}}$$

↳ Recall that
a) $\forall$ represents an infinite string of $\wedge$
b) $\exists$ represents an infinite string of $\vee$
and note that p is not dependent on the quantifier variable x, although it could be dependent on other variables (not shown).

- **Exchange property**

$$\boxed{\begin{aligned} \forall x \in A : \forall y \in B : p(x,y) &\Leftrightarrow \forall y \in B : \forall x \in A : p(x,y) \\ \exists x \in A : \exists y \in B : p(x,y) &\Leftrightarrow \exists y \in B : \exists x \in A : p(x,y) \end{aligned}}$$

↳ We can exchange similar quantifiers but not opposite quantifiers.

▶ Diagonalization

$$\forall x \in A : (p(x) \wedge q(x)) \iff \begin{cases} \forall x \in A : p(x) \\ \forall x \in A : q(x) \end{cases}$$

$$\exists x \in A : (p(x) \vee q(x)) \iff (\exists x \in A : p(x)) \vee (\exists x \in A : q(x))$$

▶ Rearrangement

$$\forall x \in A \cup B : p(x) \iff \begin{cases} \forall x \in A : p(x) \\ \forall x \in B : p(x) \end{cases}$$

$$\exists x \in A \cup B : p(x) \iff (\exists x \in A : p(x)) \vee (\exists x \in B : p(x))$$

▶ Extraction / Extension

$$\begin{cases} \exists x \in A : p(x) \\ A \subseteq B \end{cases} \Rightarrow \exists x \in B : p(x) \qquad \longleftarrow \text{Extension}$$

$$\begin{cases} \forall x \in B : p(x) \\ A \subseteq B \end{cases} \Rightarrow \forall x \in A : p(x) \qquad \longleftarrow \text{Extraction}$$

EXAMPLES

a) Show that: $\bigcup\limits_{a\in I} (B - A_a) = B - \left(\bigcap\limits_{a\in I} A_a\right)$

Proof

$$
\begin{aligned}
x \in \bigcup_{a\in I} (B - A_a) &\Leftrightarrow \exists a \in I : x \in B - A_a \Leftrightarrow \\
&\Leftrightarrow \exists a \in I : (x \in B \wedge x \notin A_a) \Leftrightarrow \\
&\Leftrightarrow x \in B \wedge (\exists a \in I : x \notin A_a) \Leftrightarrow \\
&\Leftrightarrow x \in B \wedge \overline{(\forall a \in I : x \in A_a)} \Leftrightarrow \quad (*) \\
&\Leftrightarrow x \in B \wedge x \notin \bigcap_{a\in I} A_a \Leftrightarrow \\
&\Leftrightarrow x \in B - \left(\bigcap_{a\in I} A_a\right)
\end{aligned}
$$

therefore: $\bigcup\limits_{a\in I} (B - A_a) = B - \left(\bigcap\limits_{a\in I} A_a\right)$. ◻

b) Show that: $\left(\bigcap\limits_{a\in I} A_a\right) - \left(\bigcup\limits_{a\in I} B_a\right) = \bigcap\limits_{a\in I} \bigcap\limits_{b\in I} (A_a - B_b)$

Proof

$$
\begin{aligned}
x \in \left(\bigcap_{a\in I} A_a\right) - \left(\bigcup_{a\in I} B_a\right) &\Leftrightarrow \\
&\Leftrightarrow x \in \bigcap_{a\in I} A_a \wedge x \notin \bigcup_{b\in I} B_b \Leftrightarrow
\end{aligned}
$$

$$\Leftrightarrow (\forall a \in I : x \in A_a) \wedge \overline{(\exists b \in I : x \in B_b)} \Leftrightarrow$$

$$\Leftrightarrow (\forall a \in I : x \in A_a) \wedge (\forall b \in I : x \notin B_b) \Leftrightarrow \quad (*)$$

$$\Leftrightarrow \forall a \in I : (x \in A_a \wedge (\forall b \in I : x \notin B_b)) \Leftrightarrow \quad (*)$$

$$\Leftrightarrow \forall a \in I : (\forall b \in I : (x \in A_a \wedge x \notin B_b)) \Leftrightarrow$$

$$\Leftrightarrow \forall a \in I : \forall b \in I : \ x \in A_a - B_b \Leftrightarrow$$

$$\Leftrightarrow \forall a \in I : \left(x \in \bigcap_{b \in I} (A_a - B_b) \right) \Leftrightarrow$$

$$\Leftrightarrow x \in \bigcap_{a \in I} \bigcap_{b \in I} (A_a - B_b).$$

therefore: $\left(\bigcap_{a \in I} A_a \right) - \left(\bigcup_{a \in I} B_a \right) = \bigcap_{a \in I} \bigcap_{b \in I} (A_a - B_b)$. □

↳ We label the use of the associative/distributive properties for quantifiers with (*).

EXERCISES

(14) Let I be an index set and let $\{A_a\}_{a\in I}$, $\{B_a\}_{a\in I}$ be two indexed collections of sets. Prove that:

a) $C - \bigcap_{a\in I} A_a = \bigcup_{a\in I} (C - A_a)$

b) $C - \bigcup_{a\in I} A_a = \bigcap_{a\in I} (C - A_a)$

c) $C \cap \bigcup_{a\in I} A_a = \bigcup_{a\in I} (C \cap A_a)$

d) $C \cup \bigcap_{a\in I} A_a = \bigcap_{a\in I} (C \cup A_a)$

e) $\left[\bigcap_{a\in I} A_a\right] \cup \left[\bigcap_{a\in I} B_a\right] = \bigcap_{a\in I} \bigcap_{b\in I} (A_a \cup B_b)$

f) $\left[\bigcup_{a\in I} A_a\right] \cap \left[\bigcup_{a\in I} B_a\right] = \bigcup_{a\in I} \bigcup_{b\in I} (A_a \cap B_b)$

g) $\left[\bigcap_{a\in I} A_a\right] - C = \bigcap_{a\in I} (A_a - C)$

h) $\left[\bigcup_{a\in I} A_a\right] - C = \bigcup_{a\in I} (A_a - C)$.

Defining sets by description

The fundamental method for defining a set A is by providing a belonging condition of the form

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

where $p(x)$ is a predicate about x. That said, there are 3 general methods for defining sets in practice, and we have already encountered the first two:

1) <u>By listing</u>: $A = \{a_1, a_2, a_3, \ldots, a_n\}$

The corresponding belonging condition is:

$$x \in A \Leftrightarrow x = a_1 \vee x = a_2 \vee x = a_3 \vee \cdots \vee x = a_n$$

Note that the order by which elements are listed makes no difference.

2) <u>By selection</u>: $A = \{x \in U \mid p(x)\}$

with U a universal set and $p(x)$ a predicate about x. A contains all elements of U that satisfy $p(x)$.

The corresponding belonging condition is:

$$x \in A \Leftrightarrow x \in U \wedge p(x).$$

This condition can be rewritten as a quantified statement as:

$$\forall x \in U : (x \in A \Leftrightarrow p(x)).$$

▸ <u>example</u>

Definition by selection is oftentimes used to define solution sets. For example, the solution set of the inequality $3x-1 < x^2$ can be written as:

$$S = \{x \in \mathbb{R} \mid 3x - 1 < x^2\}$$

▸ <u>example</u>

Definition by selection can be used to define intervals:

$[a,b] = \{x \in \mathbb{R} \mid a \leq x \leq b\}$

$(a,b) = \{x \in \mathbb{R} \mid a < x < b\}$

and so on.

3) <u>By mapping</u>: $A = \{\varphi(x) \mid x \in U \land p(x)\}$

where U is a universal set, $p(x)$ is a predicate, and $\varphi(x)$ an expression that generates some new element from x. The belonging condition of A is:

$$x \in A \iff \exists a \in U : (p(a) \land \varphi(a) = x).$$

- The elements of A are generated as follows: for each $a \in U$ we test if it satisfies $p(a)$. If it does, then we add the element $\varphi(a)$ to the set A.
- Similar definitions can be made over expressions that use multiple variables. For example:

$A = \{\varphi(a,b) \mid a \in U_1 \land b \in U_2 \land p(a,b)\}$

has belonging condition

$x \in A \iff \exists a \in U_1 : \exists b \in U_2 : (p(a,b) \land \varphi(a,b) = x)$

and

$A = \{\varphi(a,b,c) \mid a \in U_1 \land b \in U_2 \land c \in U_3 \land p(a,b,c)\}$

has belonging condition

$x \in A \iff \exists a \in U_1 : \exists b \in U_2 : \exists c \in U_3 : (p(a,b,c) \land \varphi(a,b,c) = x)$

and so on.

- Another generalization is to include multiple expressions φ_1, φ_2, etc. For example:

$$A = \{\varphi_1(a), \varphi_2(a) \mid a \in U \wedge p(a)\}$$

has belonging condition

$$x \in A \Leftrightarrow \exists a \in U : (p(a) \wedge (\varphi_1(a) = x \vee \varphi_2(a) = x))$$

- We can also have a definition using both multiple variables and multiple expressions. For example

$$A = \{\varphi_1(a,b), \varphi_2(a,b) \mid a \in U_1 \wedge b \in U_2 \wedge p(a,b)\}$$

has belonging condition

$$x \in A \Leftrightarrow \exists a \in U_1 : \exists b \in U_2 : (p(a,b) \wedge (\varphi_1(a,b) = x \vee \varphi_2(a,b) = x))$$

EXAMPLES

a) Set of odd/even numbers

Recall that we defined the set of natural numbers:

$\mathbb{N} = \{0, 1, 2, 3, \dots\}$

We can define:

$A = \{2x \mid x \in \mathbb{N}\} = \{0, 2, 4, 6, \dots\}$

$B = \{2x+1 \mid x \in \mathbb{N}\} = \{1, 3, 5, 7, \dots\}$

The corresponding belonging condition is:

$x \in A \Leftrightarrow \exists a \in \mathbb{N} : x = 2a$

$x \in B \Leftrightarrow \exists a \in \mathbb{N} : x = 2a + 1$

and since $A \subseteq \mathbb{N}$ and $B \subseteq \mathbb{N}$, the definition of A, B can be rewritten using "definition by selection" as:

$A = \{x \in \mathbb{N} \mid \exists a \in \mathbb{N} : x = 2a\}$

$B = \{x \in \mathbb{N} \mid \exists b \in \mathbb{N} : x = 2a + 1\}$

b) The sets $\mathbb{Z}, \mathbb{Q}$

The set of integers $\mathbb{Z}$ and the set of rational numbers $\mathbb{Q}$ can be defined descriptively as:

$\mathbb{Z} = \mathbb{N} \cup \{-x \mid x \in \mathbb{N}\}$

$\mathbb{Q} = \{a/b \mid a, b \in \mathbb{Z} \wedge b \neq 0\}$

The corresponding belonging condition is:

$x \in \mathbb{Z} \Leftrightarrow x \in \mathbb{N} \vee (\exists a \in \mathbb{N} : x = -a)$

$x \in \mathbb{Q} \Leftrightarrow \exists a, b \in \mathbb{Z} : (b \neq 0 \wedge x = a/b)$

c) The sets $\mathbb{C}$ and $\mathbb{I}$

The set of complex numbers $\mathbb{C}$ and the set of imaginary numbers $\mathbb{I}$ can be defined descriptively from the set of real numbers $\mathbb{R}$ as:

$$\mathbb{C} = \{a+bi \mid a, b \in \mathbb{R}\}$$
$$\mathbb{I} = \{bi \mid b \in \mathbb{R}\}$$

The corresponding belonging conditions are:

$$z \in \mathbb{C} \Leftrightarrow \exists a, b \in \mathbb{R} : z = a+bi$$
$$z \in \mathbb{I} \Leftrightarrow \exists b \in \mathbb{R} : z = bi$$

d) Write the belonging condition and it's negation for the set

$$A = \{a^2+b^2 \mid a \in \mathbb{R} \wedge b \in \mathbb{Q} \wedge a+b<10\}$$

Solution

The belonging condition for A is:

$$x \in A \Leftrightarrow \exists a \in \mathbb{R} : \exists b \in \mathbb{Q} : (a+b<10 \wedge x = a^2+b^2)$$

The corresponding negation is:

$$\begin{aligned}
x \notin A &\Leftrightarrow \overline{\exists a \in \mathbb{R} : \exists b \in \mathbb{Q} : (a+b<10 \wedge x = a^2+b^2)} \\
&\Leftrightarrow \forall a \in \mathbb{R} : \overline{\exists b \in \mathbb{Q} : (a+b<10 \wedge x = a^2+b^2)} \\
&\Leftrightarrow \forall a \in \mathbb{R} : \forall b \in \mathbb{Q} : \overline{(a+b<10 \wedge x = a^2+b^2)} \\
&\Leftrightarrow \forall a \in \mathbb{R} : \forall b \in \mathbb{Q} : (\overline{a+b<10} \vee \overline{x = a^2+b^2}) \\
&\Leftrightarrow \forall a \in \mathbb{R} : \forall b \in \mathbb{Q} : (a+b \geq 10 \vee x \neq a^2+b^2)
\end{aligned}$$

$\hookrightarrow$ Recall the following negation rules.

$$\overline{p \wedge q} \equiv \bar{p} \vee \bar{q} \qquad \overline{p \Leftrightarrow q} \equiv p \veebar q$$
$$\overline{p \vee q} \equiv \bar{p} \wedge \bar{q} \qquad \overline{p \veebar q} \equiv p \Leftrightarrow q$$
$$\overline{p \Rightarrow q} \equiv p \wedge \bar{q}$$

→ Be careful not to confuse set definitions <u>by mapping</u> with set definitions <u>by description.</u> Here's an example of set definition by description.

e) Write the belonging condition and its negation for
$A = \{x \in \mathbb{R} \mid \exists y \in \mathbb{R} : 2y^2 + y = x + 1\}$

<u>Solution</u>

The belonging condition of A is:

$\forall x \in \mathbb{R} : (x \in A \Leftrightarrow \exists y \in \mathbb{R} : 2y^2 + y = x + 1)$

The negation, in <u>detail</u> is <u>derived</u> as follows:

$$\begin{aligned}
\forall x \in \mathbb{R} : (x \notin A &\Leftrightarrow \overline{\exists y \in \mathbb{R} : 2y^2 + y = x + 1}\ ) \\
&\Leftrightarrow \forall y \in \mathbb{R} : \overline{2y^2 + y = x + 1}\ ) \\
&\Leftrightarrow \forall y \in \mathbb{R} : 2y^2 + y \neq x + 1).
\end{aligned}$$

and therefore:

$\forall x \in \mathbb{R} : (x \notin A \Leftrightarrow \forall y \in \mathbb{R} : 2y^2 + y \neq x + 1).$

EXERCISES

(15) Write the belonging condition and its negation for the following sets, using quantifiers

a) $A = \{x^2+1 \mid x \in \mathbb{Q} \wedge 2x < 1\}$

b) $A = \{3x+1 \mid x \in \mathbb{Z} \wedge x \text{ prime number}\}$

c) $A = \{x \in \mathbb{R} \mid x^2+3x > 0\}$

d) $A = \{a^3+b^3+c^3 \mid a,b \in \mathbb{R} \wedge c \in \mathbb{Q} \wedge a+b+c=0\}$

e) $A = \{x \in \mathbb{R} \mid x^2+2x < 0 \vee 3x+1 > -4\}$

f) $A = \{a^2-b^2 \mid a \in \mathbb{N} \wedge b \in \mathbb{R} \wedge a+b > 5\}$

g) $A = \{x \in \mathbb{Z} \mid \exists a \in \mathbb{Z} : x = 3a\}$

h) $A = \{ab \mid a,b \in \mathbb{R} \wedge (a+b > 2 \vee a-b < -3)\}$

i) $A = \{x \in \mathbb{R} \mid \exists y \in \mathbb{R} : y^2+y = x\}$

j) $A = \{x \in \mathbb{R} \mid \forall y \in \mathbb{R} : x < y^2+1\}$

k) $A = \{a+b \mid a,b \in \mathbb{R} \wedge (ab > 1 \Rightarrow a^2+b^2 > 2)\}$

l) $A = \{abc \mid a,b,c \in \mathbb{R} \wedge (a+b > 2 \veebar a-c < 3)\}$

m) $A = \{2a+3b \mid a,b \in \mathbb{R} \wedge ab > 1 \wedge a-b < 0\}$

n) $A = \{a^2b,\ a+b \mid a \in \mathbb{Z} \wedge b \in \mathbb{Q} \wedge a-b = 3\}$

o) $A = \{3k,\ 3k+1 \mid k \in \mathbb{Z} \wedge k^2-10 > 0\}$

p) $A = \{ab, bc, ca \mid a,b,c \in \mathbb{N} \wedge a^2+b^2+c^2 < 100\}$

q) $A = \{a+b,\ a+3b \mid a,b \in \mathbb{Z} \wedge (a-b > 0 \Rightarrow a-3b > 0)\}$

Proof methodology with sets

We now consider proofs with sets that involve statements that are more complex that basic set identities.

▶ Methodology: Dealing with sets

- For proofs involving sets, we use:

  $x \in A \cap B \Leftrightarrow x \in A \land x \in B$

  $x \in A \cup B \Leftrightarrow x \in A \lor x \in B$

  $x \in A - B \Leftrightarrow x \in A \land x \notin B$

  $A \subseteq B \Leftrightarrow \forall x \in A : x \in B$

  $A = B \Leftrightarrow A \subseteq B \land B \subseteq A$

  $z \in \{x \in A \mid p(x)\} \Leftrightarrow z \in A \land p(z)$

  $z \in \{\varphi(x) \mid x \in A \land p(x)\} \Leftrightarrow \exists x \in A : (p(x) \land \varphi(x) = z)$

- If $A = B$ is given as an assumption (or previously proved) we can deduce:

  $x \in A \Leftrightarrow x \in B$

  $x \in A \Rightarrow x \in B$

  $x \in B \Rightarrow x \in A$

  or, in general, replace $x \in A$ with $x \in B$ and vice versa in any Boolean expression.

- If $A \subseteq B$ is given as an assumption (or previously proved) we can deduce

  $x \in A \Rightarrow x \in B$

  or, in general, replace $x \in A$ with $x \in B$ in any boolean expression.

▸ <u>Methodology</u>: Extension / Extraction

In a deductive argument we can ADD arbitrary statements with logical OR (disjunction) or remove statements connected with logical AND (conjunction):

$$p \Rightarrow p \vee q_1 \vee q_2 \vee \cdots \vee q_n \qquad \text{(extension)}$$

$$p \wedge q_1 \wedge q_2 \wedge \cdots \wedge q_n \Rightarrow p \qquad \text{(extraction)}$$

The corresponding generalization to quantified statements reads:

$$\boxed{\begin{array}{l} \begin{cases} \forall x \in A : p(x) \\ x_0 \in A \end{cases} \Rightarrow p(x_0) \\ \hline \begin{cases} x_0 \in A \\ p(x_0) \end{cases} \Rightarrow \exists x \in A : p(x) \end{array}} \qquad \begin{array}{l} \text{(extraction)} \\ \\ \text{(extension)} \end{array}$$

▸ <u>Methodology</u>: General proofwriting

① → To prove $\boxed{\forall x \in A : p(x)}$

Let $x \in A$ be given.
[Prove $p(x)$]
It follows that $\forall x \in A : p(x)$.

② → To prove $\boxed{\exists x \in A : p(x)}$

▸ 1st method
[Define some $x_0 \in A$]
[Prove $p(x_0)$]
It follows that $\exists x \in A : p(x)$

↳ Note that x_0 can be indirectly defined by deducing a statement of the form $\exists x \in B : q(x)$ via a theorem or by constructing it from other variables that have been indirectly defined via existential statements.

▶ <u>2nd method</u>

[Prove $p(x) \Leftrightarrow \cdots \Leftrightarrow \cdots \Leftrightarrow x \in S$]

[Choose a specific $x_0 \in S$]

[Prove $x_0 \in A$]

[Prove $p(x_0)$]

It follows that $\exists x \in A : p(x)$.

(3) → To prove $\boxed{p \Rightarrow q}$

▶ <u>Direct method</u>

Assume p is true

[Prove q]

▶ <u>Contrapositive method</u>

We will show that $\bar{q} \Rightarrow \bar{p}$

Assume $\bar{q}$ is true

[Prove $\bar{p}$]

From the above, it follows that $p \Rightarrow q$.

▶ <u>Contradiction method</u>

Assume p is true

To show q, we assume $\bar{q}$, and will derive a contradictio

[Prove r, using $p \wedge \bar{q}$]
[Prove $\bar{r}$] $\longleftarrow$ Contradiction
It follows that q is true

(4) $\rightarrow$ To prove $\boxed{p \Leftrightarrow q}$

($\Rightarrow$): [Prove $p \Rightarrow q$]
($\Leftarrow$): [Prove $q \Rightarrow p$]

▶ 2nd method: Occasionally, it is possible to use a direct argument of the form
$p \Leftrightarrow r_1 \Leftrightarrow r_2 \Leftrightarrow \cdots \Leftrightarrow r_n \Leftrightarrow q$
as long as every step can be justified in both directions.

(5) $\rightarrow$ To prove $\boxed{p \vee q \Rightarrow r}$

▶ Proof by cases
Assume that $p \vee q$. We distinguish between the following cases.
Case 1: Assume that p is true.
[Prove r]
Case 2: Assume that q is true
[Prove r]
From the above it follows that r is true.

▶ Contrapositive
We will show that $\bar{r} \Rightarrow \bar{p} \wedge \bar{q}$. Assume that $\bar{r}$ true

[Prove $\bar{p}$]
[Prove $\bar{q}$]
From the above, it follows that $p \vee q \Rightarrow r$

↳ Proof by cases is used when the hypothesis takes the form $p \vee q$ (or more generally $p_1 \vee p_2 \vee p_3 \vee \cdots \vee p_n$) and we do not really know which of the statements in the disjunction is true. However, for the individual cases we can use any of the proof techniques under 3.

↳ The skeletal structure of any proof combines the above elements as is appropriate.

EXAMPLES

a) Show that $B \subseteq A \Rightarrow A \cup B = A$

Proof

Assume that $B \subseteq A$.

($\Rightarrow$): Let $x \in A \cup B$ be given. Then:

$$\begin{aligned} x \in A \cup B &\Rightarrow x \in A \vee x \in B \\ &\Rightarrow x \in A \vee x \in A \quad [\text{via } B \subseteq A] \\ &\Rightarrow x \in A \end{aligned}$$

($\Leftarrow$): Let $x \in A$ be given. Then:

$$\begin{aligned} x \in A &\Rightarrow x \in A \vee x \in B \\ &\Rightarrow x \in A \cup B \end{aligned}$$

From the above, it follows that

$$\begin{cases} \forall x \in A \cup B : x \in A \\ \forall x \in A : x \in A \cup B \end{cases} \Rightarrow \begin{cases} A \cup B \subseteq A \\ A \subseteq A \cup B \end{cases} \Rightarrow A \cup B = A.$$

↳ Note the following:

a) We declare our assumptions.

b) The structure of the proof is to show

$$\begin{cases} \forall x \in A \cup B : x \in A \\ \forall x \in A : x \in A \cup B \end{cases}$$

from which we deduce the statement $A \cup B = A$. This is the general structure of a proof intended to show that two sets are equal.

b) Show that $A \cup B = A \Rightarrow B \subseteq A$.

Solution

Assume that $A \cup B = A$. Let $x \in B$ be given. Then:

$$\begin{aligned} x \in B &\Rightarrow x \in A \lor x \in B \\ &\Rightarrow x \in A \cup B \\ &\Rightarrow x \in A \qquad [\text{via } A \cup B = A] \end{aligned}$$

It follows that

$$(\forall x \in B : x \in A) \Rightarrow B \subseteq A.$$

In the context of proving set properties, contradiction proofs often arise when working with statements involving the empty set.

c) Show that $(A-B)-C = \emptyset \Rightarrow A \subseteq B \cup C$.

Solution

Assume that $(A-B)-C = \emptyset$. To show $A \subseteq B \cup C$, we assume that $A \not\subseteq B \cup C$ and will derive a contradiction.

Since,

$$\begin{aligned} A \not\subseteq B \cup C &\Rightarrow \overline{\forall x \in A : x \in B \cup C} \\ &\Rightarrow \exists x \in A : x \notin B \cup C \end{aligned}$$

Choose an $x_0 \in A$ such that $\overline{x_0 \notin B \cup C}$. Then,

$$\begin{aligned} x_0 \in A \land x_0 \notin B \cup C &\Rightarrow x_0 \in A \land \overline{(x_0 \in B \lor x_0 \in C)} \\ &\Rightarrow x_0 \in A \land (x_0 \notin B \land x_0 \notin C) \\ &\Rightarrow (x_0 \in A \land x_0 \notin B) \land x_0 \notin C \\ &\Rightarrow x_0 \in A - B \land x_0 \notin C \\ &\Rightarrow x_0 \in (A-B) - C \\ &\Rightarrow x_0 \in \emptyset \end{aligned}$$

This is a contradiction, since $x_0 \notin \emptyset$. It follows that $A \subseteq B \cup C$.

d) Show that $\mathcal{P}(A) \cup \mathcal{P}(B) \subseteq \mathcal{P}(A \cup B)$.

Solution

Let $X \in \mathcal{P}(A) \cup \mathcal{P}(B)$ be given. It is sufficient to show that $\forall y \in X : y \in A \cup B$. We note that

$$X \in \mathcal{P}(A) \cup \mathcal{P}(B) \Rightarrow X \in \mathcal{P}(A) \vee X \in \mathcal{P}(B) \Rightarrow$$
$$\Rightarrow X \subseteq A \vee X \subseteq B.$$

We distinguish between the following cases.

Case 1: Assume that $X \subseteq A$. Let $y \in X$ be given. Then:

$$\begin{aligned} y \in X &\Rightarrow y \in A \quad [\text{via } X \subseteq A] \\ &\Rightarrow y \in A \vee y \in B \\ &\Rightarrow y \in A \cup B. \end{aligned}$$

Case 2: Assume that $X \subseteq B$. Let $y \in X$ be given. Then

$$\begin{aligned} y \in X &\Rightarrow y \in B \quad [\text{via } X \subseteq B] \\ &\Rightarrow y \in A \vee y \in B \\ &\Rightarrow y \in A \cup B \end{aligned}$$

In both cases we obtain:

$$\begin{aligned} (\forall y \in X : y \in A \cup B) &\Rightarrow X \subseteq A \cup B \\ &\Rightarrow X \in \mathcal{P}(A \cup B). \end{aligned}$$

From the above argument, we have shown that

$$(\forall X \in \mathcal{P}(A) \cup \mathcal{P}(B) : X \in \mathcal{P}(A \cup B)) \Rightarrow \mathcal{P}(A) \cup \mathcal{P}(B) \subseteq \mathcal{P}(A \cup B).$$

EXERCISES

(16) Prove that

a) $A \cup B = A \cap B \Rightarrow A = B$

b) $\begin{cases} A \cup B = A \cup C \\ A \cap B = A \cap C \end{cases} \Rightarrow B = C$

(Hint: Distinguish between the cases $x \in A$ and $x \notin A$)

c) $\begin{cases} A \cup B \subseteq C \\ B \cup C \subseteq A \\ C \cup A \subseteq B \end{cases} \Rightarrow A = B = C$

d) $A \cup B = \emptyset \Rightarrow A = \emptyset \wedge B = \emptyset$

e) $A - B = \emptyset \wedge B - A = \emptyset \Rightarrow A = B$

f) $A - (B - C) = \emptyset \Rightarrow A - B = \emptyset \wedge A \cap C = \emptyset$

g) $(A - C) \cap (B - C) = \emptyset \Rightarrow A \cap B \subseteq C$

h) $(A - B) \cap (C - D) = \emptyset \Rightarrow A \cap C \subseteq B \cup D$.

(17) Prove the following equivalences

a) $(B - A) \cup A = B \Leftrightarrow A \subseteq B$

b) $B - (B - A) = A \Leftrightarrow A \subseteq B$

c) $A \cup B = B \Leftrightarrow A \subseteq B$

d) $A \cap B = A \Leftrightarrow A \subseteq B$

e) $A - B = \emptyset \Leftrightarrow A \subseteq B$

(18) Prove that

a) $\mathcal{P}(A) \cap \mathcal{P}(B) = \mathcal{P}(A \cap B)$

b) $\mathcal{P}(A-B) \subseteq \mathcal{P}(A) - \mathcal{P}(B)$

c) $A \cap B = \emptyset \Rightarrow \mathcal{P}(A-B) = \mathcal{P}(A) - \mathcal{P}(B)$

d) $A \subseteq B \Rightarrow \mathcal{P}(A) \subseteq \mathcal{P}(B)$

(19) Prove that

a) $\bigcap_{a\in I} A_a = \bigcup_{a\in I} A_a \Rightarrow \forall a,b \in I: A_a = A_b.$

b) $\bigcup_{a\in I} A_a = \emptyset \Rightarrow \forall a \in I: A_a = \emptyset$

c) $I \subseteq K \Rightarrow \bigcap_{a\in K} A_a \subseteq \bigcap_{a\in I} A_a$

d) $I \subseteq K \Rightarrow \bigcup_{a\in I} A_a \subseteq \bigcup_{a\in K} A_a$

e) $\bigcap_{a\in I} \mathcal{P}(A_a) = \mathcal{P}\left(\bigcap_{a\in I} A_a\right)$

f) $\bigcup_{a\in I} \mathcal{P}(A_a) \subseteq \mathcal{P}\left(\bigcup_{a\in I} A_a\right)$

g) $(\forall a,b \in I: A_a \cap A_b = \emptyset) \Rightarrow \bigcup_{a\in I} \mathcal{P}(A_a) = \mathcal{P}\left(\bigcup_{a\in I} A_a\right)$

DST2: Basic number theory

BASIC NUMBER THEORY

Modulo arithmetic

We recall the following set definitions

a) The set of natural numbers

$\mathbb{N} = \{0, 1, 2, 3, \dots\}$

$\mathbb{N}^* = \mathbb{N} - \{0\} = \{1, 2, 3, \dots\}$

b) The set of integers

$\mathbb{Z} = \{0, 1, -1, 2, -2, 3, -3, \dots\}$

$\mathbb{Z}^* = \mathbb{Z} - \{0\} = \{1, -1, 2, -2, 3, -3, \dots\}$

We now use those to define divisibility and modulo equivalence.

<u>Def</u>: Let $a, b \in \mathbb{Z}$ be given. We say that <u>a divides b</u> (i.e. $a|b$) if and only if there is some integer k such that $b = ak$:

$$\forall a, b \in \mathbb{Z} : (a|b \Leftrightarrow \exists k \in \mathbb{Z} : b = ak)$$

<u>Def</u>: (modulo equivalence).

$$\forall a, b, m \in \mathbb{Z} : (a \equiv b \pmod{m} \Leftrightarrow m|(a-b))$$

<u>Def</u>: Let $a \in \mathbb{Z}^*$. We define the set Δ_a of all divisors of a as:

$$\Delta_a = \{b \in \mathbb{Z} \mid (b|a)\}$$

EXAMPLES

a) Show that $17 \equiv 3 \mod 7$

Solution

$$17-3 = 14 = 7\cdot 2 \Rightarrow \exists k\in\mathbb{Z} : 17-3 = 7k \quad (\text{for } k=2)$$
$$\Rightarrow 7 \mid (17-3)$$
$$\Rightarrow 17 \equiv 3 \mod 7$$

b) Evaluate $\Delta_2, \Delta_4, \Delta_6$

Solution

$$\Delta_2 = \{ b\in\mathbb{Z} \mid (b \mid 2)\} = \{1,-1,2,-2\}$$
$$\Delta_4 = \{ b\in\mathbb{Z} \mid (b \mid 4)\} = \{1,-1,2,-2,4,-4\}$$
$$\Delta_6 = \{ b\in\mathbb{Z} \mid (b \mid 6)\} = \{1,-1,2,-2,3,-3,6,-6\}$$

→ Modulo arithmetic satisfies the reflexive, symmetric, and transitive properties.

c) Show that $\forall a,m\in\mathbb{Z} : a\equiv a \mod m$

Solution

Let $a,m\in\mathbb{Z}$ be given. Then:

$$a-a = 0 = 0m \Rightarrow \exists k\in\mathbb{Z} : a-a = km$$
$$\Rightarrow m \mid (a-a)$$
$$\Rightarrow a\equiv a \mod m$$

It follows that $\forall a,m\in\mathbb{Z} : a\equiv a \mod m$

d) Show that

$$\forall a,b,m \in \mathbb{Z} : (a \equiv b \bmod m \Rightarrow b \equiv a \bmod m)$$

<u>Solution</u>

Let $\underline{a,b,m \in \mathbb{Z}}$ be given. Assume that $\underline{a \equiv b \bmod m}$.

Then,

$$a \equiv b \bmod m \Rightarrow m \mid (a-b)$$
$$\Rightarrow \exists k \in \mathbb{Z} : a-b = mk$$

Choose a $k_0 \in \mathbb{Z}$ such that $a-b = mk_0$. Then:

$$b-a = -(a-b) = -mk_0 = m(-k_0) \Rightarrow$$
$$\Rightarrow \exists k \in \mathbb{Z} : b-a = mk \quad (\text{for } k = -k_0)$$
$$\Rightarrow m \mid (b-a)$$
$$\Rightarrow \underline{b \equiv a \bmod m}$$

We have thus shown that

$$\forall a,b,m \in \mathbb{Z} : (a \equiv b \bmod m \Rightarrow b \equiv a \bmod m).$$

e) Show that

$$\forall a,b,c,m \in \mathbb{Z} : \left(\begin{cases} a \equiv b \bmod m \\ b \equiv c \bmod m \end{cases} \Rightarrow a \equiv c \bmod m \right)$$

<u>Solution</u>

Let $\underline{a,b,c,m \in \mathbb{Z}}$ be given. Assume that $\underline{a \equiv b \bmod m}$ and $\underline{b \equiv c \bmod m}$. Then,

$$\begin{cases} a \equiv b \bmod m \\ b \equiv c \bmod m \end{cases} \Rightarrow \begin{cases} m \mid (a-b) \\ m \mid (b-c) \end{cases} \Rightarrow$$
$$\Rightarrow \begin{cases} \exists k \in \mathbb{Z} : a-b = mk \\ \exists \lambda \in \mathbb{Z} : b-c = m\lambda \end{cases}$$

Choose $k_0, \lambda_0 \in \mathbb{Z}$ such that $a-b = mk_0$ and $b-c = m\lambda_0$.
It follows that

$$a-c = (a-b)+(b-c) = mk_0 + m\lambda_0 = m(k_0+\lambda_0) \Rightarrow$$
$$\Rightarrow \exists \mu \in \mathbb{Z} : a-c = m\mu \quad (\text{for } \mu = k_0+\lambda_0)$$
$$\Rightarrow m \mid (a-c)$$
$$\Rightarrow \underline{a \equiv c \bmod m}$$

We have thus shown that

$$\forall a,b,c,m \in \mathbb{Z} : \left(\begin{cases} a \equiv b \bmod m \\ b \equiv c \bmod m \end{cases} \Rightarrow a \equiv c \bmod m \right).$$

f) Show that $\forall a,b \in \mathbb{Z} : \Delta_a \cap \Delta_b \subseteq \Delta_{ab}$.

<u>Solution</u>

Let $a,b \in \mathbb{Z}$ be given. Let $x \in \Delta_a \cap \Delta_b$ be given. Then:

$$x \in \Delta_a \cap \Delta_b \Rightarrow x \in \Delta_a \wedge x \in \Delta_b \Rightarrow$$
$$\Rightarrow x \mid a \wedge x \mid b \Rightarrow$$
$$\Rightarrow \begin{cases} \exists k \in \mathbb{Z} : a = kx \\ \exists \lambda \in \mathbb{Z} : b = \lambda x \end{cases}$$

Choose $k_0, \lambda_0 \in \mathbb{Z}$ such that $a = k_0 x$ and $b = \lambda_0 x$.
It follows that:

$$ab = (k_0 x)(\lambda_0 x) = x(k_0 \lambda_0 x) \Rightarrow$$
$$\Rightarrow \exists \mu \in \mathbb{Z} : ab = \mu x \quad (\text{for } \mu = k_0 \lambda_0 x)$$
$$\Rightarrow x \mid ab$$
$$\Rightarrow x \in \Delta_{ab}.$$

From the above argument:

$$\forall a,b \in \mathbb{Z} : \forall x \in \Delta_a \cap \Delta_b : x \in \Delta_{ab} \Rightarrow$$
$$\Rightarrow \forall a,b \in \mathbb{Z} : \Delta_a \cap \Delta_b \subseteq \Delta_{ab}.$$

Division theorem

The division theorem is useful in divisibility proofs, and we state it without proof:

$$\boxed{\forall a \in \mathbb{Z}^*: \forall b \in \mathbb{Z}: \exists! q, r \in \mathbb{Z}: \begin{cases} b = aq + r \\ 0 \leq r < |a| \end{cases}}$$

▶ <u>interpretation</u>: The division theorem establishes that when we divide two integers b with a we obtain a unique quotient q and remainder r with $0 \leq r < |a|$. such that the division identity $b = aq + r$ is satisfied.

▶ <u>notation</u>: The unique quotient q and remainder r are denoted as: $q = b \div a$ and $r = b \bmod a$.

↳ A convincing explanation of this result can be made in terms of the well-known long division algorithm from high school, which will always produce a unique quotient and remainder. A rigorous proof uses the well-ordering principle, which in axiomatic set theory requires the axiom of choice.

- Choosing the value of a yields the following useful corollaries:

For $a = 2$: $\forall b \in \mathbb{Z}: \exists! q \in \mathbb{Z}: (b = 2q \vee b = 2q+1)$

For $a = 3$: $\forall b \in \mathbb{Z}: \exists! q \in \mathbb{Z}: (b = 3q \vee b = 3q+1 \vee b = 3q+2)$

For $a = 4$: $\forall b \in \mathbb{Z}: \exists! q \in \mathbb{Z}: (b = 4q \vee b = 4q+1 \vee b = 4q+2 \vee b = 4q+3)$.

EXAMPLES

a) Show that $\forall x \in \mathbb{Z}: (x^2 \equiv 1 \pmod 2 \Rightarrow x^2 \equiv 1 \pmod 4)$

Solution

Let $x \in \mathbb{Z}$ be given and assume that $x^2 \equiv 1 \pmod 2$. Then,

$$x^2 \equiv 1 \pmod 2 \Rightarrow 2 \mid (x^2-1) \Rightarrow$$
$$\Rightarrow \exists k \in \mathbb{Z}: x^2 - 1 = 2k$$
$$\Rightarrow \exists k \in \mathbb{Z}: x^2 = 2k+1$$
$$\Rightarrow x^2 \bmod 2 = 1.$$

From the division theorem:

$\exists! k \in \mathbb{Z}: (x = 2k \vee x = 2k+1)$.

We distinguish between the following cases.

Case 1: Assume that $x = 2k$ for some $k \in \mathbb{Z}$.

Then $x^2 = (2k)^2 = 4k^2 = 2(2k^2) \Rightarrow$
$$\Rightarrow \exists \lambda \in \mathbb{Z}: x^2 = 2\lambda$$
$$\Rightarrow x^2 \bmod 2 = 0 \longleftarrow \text{Contradiction.}$$

therefore, this case does not materialize.

Case 2: Assume that $x = 2k+1$ for some $k \in \mathbb{Z}$.

Then $x^2 - 1 = (2k+1)^2 - 1 = (4k^2 + 4k + 1) - 1 = 4k^2 + 4k$
$$= 4(k^2+k) \Rightarrow$$
$$\Rightarrow \exists \lambda \in \mathbb{Z}: x^2 - 1 = 4\lambda \quad (\text{for } \lambda = k^2 + k)$$
$$\Rightarrow 4 \mid (x^2 - 1)$$
$$\Rightarrow x^2 \equiv 1 \pmod 4.$$

From the above argument, we find:

$\forall x \in \mathbb{Z}: (x^2 \equiv 1 \pmod 2 \Rightarrow x^2 \equiv 1 \pmod 4)$.

8) Show that $\forall x \in \mathbb{Z} : (x \not\equiv 0 \pmod 3 \Rightarrow x^2 \equiv 1 \pmod 3)$

<u>Solution</u>

Let $x \in \mathbb{Z}$ be given. <u>Assume</u> $x \not\equiv 0 \pmod 3$. Then:

$$x \not\equiv 0 \pmod 3 \Rightarrow \overline{3 \mid (x-0)} \Rightarrow \overline{3 \mid x} \Rightarrow$$
$$\Rightarrow \overline{\exists k \in \mathbb{Z} : x = 3k}$$
$$\Rightarrow \exists k \in \mathbb{Z} : (x = 3k+1 \lor x = 3k+2)$$

via the division theorem. We distinguish between the following cases

<u>Case 1</u>: Assume that $x = 3k+1$ for some $k \in \mathbb{Z}$. Then,

$$x^2 - 1 = (3k+1)^2 - 1 = (9k^2 + 6k + 1) - 1 = 9k^2 + 6k$$
$$= 3(3k^2 + 2k) \Rightarrow$$
$$\Rightarrow \exists \lambda \in \mathbb{Z} : x^2 - 1 = 3\lambda \quad (\text{for } \lambda = 3k^2 + 2k).$$

<u>Case 2</u>: Assume that $x = 3k+2$ for some $k \in \mathbb{Z}$. Then,

$$x^2 - 1 = (3k+2)^2 - 1 = (9k^2 + 12k + 4) - 1 =$$
$$= 9k^2 + 12k + 3 = 3(3k^2 + 4k + 1) \Rightarrow$$
$$\Rightarrow \exists \lambda \in \mathbb{Z} : x^2 - 1 = 3\lambda \quad (\text{for } \lambda = 3k^2 + 4k + 1)$$

In both cases, we have shown:

$$(\exists \lambda \in \mathbb{Z} : x^2 - 1 = 3\lambda) \Rightarrow 3 \mid (x^2 - 1)$$
$$\Rightarrow x^2 \equiv 1 \pmod 3$$

From the above argument:

$$\forall x \in \mathbb{Z} : (x \not\equiv 0 \pmod 3 \Rightarrow x^2 \equiv 1 \pmod 3)$$

EXERCISES

(1) Let $a, b \in \mathbb{Z}$ be given. Show that

a) $a \mid b \Rightarrow a^2 \mid b^2$

b) $a \mid b \wedge b \mid a \Rightarrow a = b \vee a = -b$

c) $a \not\equiv 0 \pmod 3 \wedge b \not\equiv 0 \pmod 3 \Rightarrow a^2 \equiv b^2 \pmod 3$

d) $2a^2 + 1 \equiv 0 \pmod 3 \Rightarrow a \not\equiv 0 \pmod 3$

e) $a^3 \equiv a \pmod 3$

f) $a^5 \equiv 5a^3 - 4a \pmod 5$

(2) Let $a, b, c \in \mathbb{Z}$ such that

$$c \equiv 0 \pmod 3 \wedge a + b + c \equiv 0 \pmod 3 \wedge 3a + b \equiv 0 \pmod 3$$

Show that $\forall x \in \mathbb{Z} : ax^2 + bx + c \equiv 0 \pmod 3$

(3) Let $a, b, x \in \mathbb{Z}$ be given. Show that

a) $\begin{cases} 2a + b \equiv 0 \pmod 4 \\ x \equiv 2 \pmod 4 \end{cases} \Rightarrow ax + b \equiv 0 \pmod 4$

b) $\begin{cases} 2a \equiv b \pmod 5 \\ x \equiv 3 \pmod 5 \end{cases} \Rightarrow ax^3 \equiv b \pmod 5$

(4) Let $a, b, c \in \mathbb{Z}$ be given such that

$$4 \mid c \wedge 4 \mid (a + b + c) \wedge 4 \mid (3a + b) \wedge 4 \mid (5a + b)$$

Show that $\forall x \in \mathbb{Z} : ax^2 + bx + c \equiv 0 \pmod 4$.

Method of induction

Let $a \in \mathbb{Z}$ and define $\mathbb{Z}_a = \{x \in \mathbb{Z} \mid x \geq a\}$. The method of induction can be used to prove statements of the form: $\forall x \in \mathbb{Z}_a : p(x)$.
It is based on Peano's theorem:

<u>Thm</u> : Let $a \in \mathbb{Z}$. Then:

$$\boxed{\left.\begin{array}{l} p(a) \text{ true} \\ \forall x \in \mathbb{Z}_a : (p(x) \Rightarrow p(x+1)) \end{array}\right\} \Rightarrow \forall x \in \mathbb{Z}_a : p(x)}$$

This theorem can be shown via the well-ordering principle.

▶ <u>Method</u> : To show $\forall x \in \mathbb{Z}_a : p(x)$ true

- 1 For $x = a$, show that $p(x)$ is true
- 2 Assume that for $x = k > a$, $p(k)$ is true
- 3 Show that $p(k+1)$ true
- 4 It follows that $\forall x \in \mathbb{Z}_a : p(x)$ true.

EXAMPLES

a) Show that $1+2+3+\cdots+n = \dfrac{n(n+1)}{2}$, $\forall n \in \mathbb{N}-\{0\}$

<u>Proof</u>

For $n=1$: LHS $=1$

$$\text{RHS} = \frac{n(n+1)}{2} = \frac{1\cdot 2}{2} = 1$$

thus the statement is true.

For $n=k$, assume that

$$1+2+3+\dots+k = \frac{k(k+1)}{2}$$

For $n=k+1$, we will show that

$$1+2+3+\dots+(k+1) = \frac{(k+1)(k+2)}{2}$$

Since:

$$1+2+3+\dots+(k+1) = [1+2+3+\dots+k] + (k+1) =$$

$$= \frac{k(k+1)}{2} + (k+1) = (k+1)(k/2+1)$$

$$= (k+1)\frac{k+2}{2} = \frac{(k+1)(k+2)}{2}$$

It follows that $\forall n \in \mathbb{N}-\{0\}: 1+2+3+\dots+n = \frac{n(n+1)}{2}$ $\square$

b) Show that $\forall n \in \mathbb{N}: 3 \mid (2^{2n}-1)$.

<u>Proof</u>

For $n=0$: $2^{2n}-1 = 2^0-1 = 1-1 = 0 = 3\cdot 0 \Rightarrow 3 \mid 2^{2n}-1$.

For $n=k$: assume that $3 \mid (2^{2k}-1)$.

For $n=k+1$: we will show that $3 \mid (2^{2(k+1)}-1)$.

Since $3 \mid (2^{2k}-1) \Rightarrow \exists a \in \mathbb{Z}: 2^{2k}-1 = 3a$

$\Rightarrow \exists a \in \mathbb{Z}: 2^{2k} = 3a+1$

Choose $a \in \mathbb{Z}$ such that $2^{2k} = 3a+1$. Then:

$2^{2(k+1)} - 1 = 2^{2k+2} - 1 = 2^{2k} \cdot 4 - 1 = 4(3a+1) - 1 =$

$= 12a + 4 - 1 = 12a + 3 = 3(4a+1) \Rightarrow$

$\Rightarrow \exists \mu \in \mathbb{Z}: 2^{2(k+1)} - 1 = 3\mu \quad (\text{for } \mu = 4a+1)$

$\Rightarrow 3 \mid 2^{2(k+1)} - 1$

It follows by induction that

$\forall n \in \mathbb{N}: 3 \mid (2^{2n}-1)$.

EXERCISES

⑤ Prove the following identities by induction

a) $\forall n \in \mathbb{N}^*: 1\cdot 2 + 2\cdot 3 + 3\cdot 4 + \dots + n(n+1) = (1/3)n(n+1)(n+2)$

b) $\forall n \in \mathbb{N}^*: 1+3+5+\dots+(2n+1) = (n+1)^2$

c) $\forall n \in \mathbb{N}^*: 2+4+6+\dots+2n = n(n+1)$

d) $\forall n \in \mathbb{N}^*: 1\cdot 2^2 + 2\cdot 3^2 + \dots + n(n+1)^2 = (1/12)n(n+1)(n+2)(3n+5)$

e) $\forall n \in \mathbb{N}^* - \{1\}: 1^3+3^3+5^3+\dots+(2n-1)^3 = n^2(2n^2-1)$

f) $\forall n \in \mathbb{N}^* - \{1\}: 2^3+4^3+6^3+\dots+(2n)^3 = 2n^2(n+1)^2$

g) $\forall n \in \mathbb{N}^* - \{1,2\}: 2+2^2+\dots+2^n = 2(2^n-1)$

h) $\forall n \in \mathbb{N}^* - \{1\}: \dfrac{1}{2} + \dfrac{1}{2^2} + \dfrac{1}{2^3} + \dots + \dfrac{1}{2^n} = 1 - \dfrac{1}{2^n}$

i) $\forall n \in \mathbb{N}^*: 1\cdot 5 + 2\cdot 5^2 + 3\cdot 5^3 + \dots + n5^n = \dfrac{5+(4n-1)5^{n+1}}{16}$

⑥ Show the following statements by induction

a) $\forall n \in \mathbb{N}^*: 4\cdot 8^n + 21n \equiv 4 \pmod{49}$

b) $\forall n \in \mathbb{N}^*: 2^{2n} + 15n \equiv 1 \pmod 9$

c) $\forall n \in \mathbb{N}^*: 7^{2n+1} \equiv 48n+7 \pmod{288}$

d) $\forall n \in \mathbb{N}^*: 5^n \equiv 1 \pmod 4$

e) $\forall n \in \mathbb{N}^*: 10^{n+1} \equiv 9n+10 \pmod{81}$

f) $\forall n \in \mathbb{N}^*: 7^{2n} \equiv 1-16n \pmod{64}$

g) $\forall n \in \mathbb{N}^*: 3^{2n} \equiv 2^n \pmod 7$

⑦ Show that

$\forall n \in \mathbb{N}^*: (1+\sqrt{2})^{2n} + (1-\sqrt{2})^{2n} \equiv 0 \pmod 2.$

DST3: Relations

RELATIONS AND FUNCTIONS

▼ Cartesian product

- An ordered pair (a,b) is defined as an ordered collection of two elements a and b such that it satisfies the axiom:
  $$(a_1,b_1)=(a_2,b_2) \Leftrightarrow a_1=a_2 \wedge b_1=b_2.$$
- Ordered pairs can be represented as sets:
  $$(a,b)=\{a,\{a,b\}\}$$
  Then ordered pair equality corresponds to set equality.
- Let A,B be two sets. We define the cartesian product $A\times B$ as:
  $$A\times B=\{(a,b)\mid a\in A \wedge b\in B\}$$
  The corresponding belonging condition is:
  $$x\in A\times B \Leftrightarrow \exists a\in A : \exists b\in B : x=(a,b).$$
  however, in practice we find it more useful to use the following statement
  $$(a,b)\in A\times B \Leftrightarrow a\in A \wedge b\in B.$$
- We also define $A^2=A\times A$.
- It is easy to see that
  $$\emptyset\times A=\emptyset$$
  $$A\times\emptyset=\emptyset.$$

EXAMPLES

a) For $A = \{1,2\}$ and $B = \{2,3\}$, evaluate $A \times B$, $B \times A$ and A^2.

Solution

$A \times B = \{1,2\} \times \{2,3\} =$
$= \{(1,2), (1,3), (2,2), (2,3)\}$

$B \times A = \{2,3\} \times \{1,2\} =$
$= \{(2,1), (2,2), (3,1), (3,2)\}$

$A^2 = A \times A = \{1,2\} \times \{1,2\} =$
$= \{(1,1), (1,2), (2,1), (2,2)\}$

b) Let A, B, C be sets. Show that
$A \times (B \cup C) = (A \times B) \cup (A \times C)$

Solution

Since,

$(x,y) \in A \times (B \cup C) \Leftrightarrow x \in A \wedge y \in B \cup C$
$\Leftrightarrow x \in A \wedge (y \in B \vee y \in C)$
$\Leftrightarrow (x \in A \wedge y \in B) \vee (x \in A \wedge y \in C)$
$\Leftrightarrow (x,y) \in A \times B \vee (x,y) \in A \times C$
$\Leftrightarrow (x,y) \in (A \times B) \cup (A \times C)$,

it follows that

$A \times (B \cup C) = (A \times B) \cup (A \times C)$.

c) Show that; for sets A, B, C:
$(C \neq \emptyset \wedge A \times C = B \times C) \Rightarrow A = B.$

Solution

Assume that $C \neq \emptyset$ and $A \times C = B \times C$.
Since $C \neq \emptyset$, choose a $y \in C$.
Let $x \in A$ be given. Then:

$$\begin{aligned} x \in A \wedge y \in C &\Rightarrow (x,y) \in A \times C && [\text{definition}] \\ &\Rightarrow (x,y) \in B \times C && [A \times C \subseteq B \times C] \\ &\Rightarrow x \in B \wedge y \in C && [\text{definition}] \\ &\Rightarrow x \in B \end{aligned}$$

and therefore:

$$(\forall x \in A : x \in B) \Rightarrow A \subseteq B. \qquad (1)$$

Let $x \in B$ be given. Then

$$\begin{aligned} x \in B \wedge y \in C &\Rightarrow (x,y) \in B \times C \\ &\Rightarrow (x,y) \in A \times C \\ &\Rightarrow x \in A \wedge y \in C \\ &\Rightarrow x \in A \end{aligned}$$

and therefore

$$(\forall x \in B : x \in A) \Rightarrow B \subseteq A. \qquad (2)$$

From (1) and (2): $A = B$.

d) Let A, B be sets with $A \neq \emptyset$ and $B \neq \emptyset$. Show that

$A \times B = B \times A \Rightarrow A = B$.

<u>Solution</u>

Assume that $A \neq \emptyset$ and $B \neq \emptyset$ and $A \times B = B \times A$.

Let $x \in A$ be given.

Since $B \neq \emptyset$, choose a $y \in B$. Then

$$
\begin{aligned}
x \in A \wedge y \in B &\Rightarrow (x,y) \in A \times B \\
&\Rightarrow (x,y) \in B \times A \quad [\text{via } A \times B \subseteq B \times A] \\
&\Rightarrow x \in B \wedge y \in A \\
&\Rightarrow x \in B.
\end{aligned}
$$

and therefore:

$(\forall x \in A : x \in B) \Rightarrow A \subseteq B. \qquad (1)$

Let $x \in B$ be given.

Since $A \neq \emptyset$, choose a $y \in A$. Then

$$
\begin{aligned}
x \in B \wedge y \in A &\Rightarrow (x,y) \in B \times A \\
&\Rightarrow (x,y) \in A \times B \quad [\text{via } B \times A \subseteq A \times B] \\
&\Rightarrow x \in A \wedge y \in B \\
&\Rightarrow x \in A.
\end{aligned}
$$

and therefore

$(\forall x \in B : x \in A) \Rightarrow B \subseteq A. \qquad (2)$

From (1) and (2): $A = B$.

e) Let $\{A_a\}_{a\in I}$, $\{B_a\}_{a\in I}$ be indexed set collections and let G be a set. Show that

$$C\times\left[\bigcup_{a\in I}(A_a - B_a)\right] \subseteq \bigcup_{a\in I}\left[(C\times A_a)-(C\times B_a)\right]$$

Solution

Since

$$(x,y)\in G\times\left[\bigcup_{a\in I}(A_a - B_a)\right] \Rightarrow$$

$$\Rightarrow x\in G \wedge y\in\bigcup_{a\in I}(A_a - B_a) \Rightarrow$$

$$\Rightarrow x\in G \wedge \exists a\in I: y\in A_a - B_a$$
$$\Rightarrow x\in G \wedge \exists a\in I: (y\in A_a \wedge y\notin B_a)$$
$$\Rightarrow \exists a\in I: (x\in G\wedge y\in A_a \wedge y\notin B_a)$$
$$\Rightarrow \exists a\in I: [(x\in G\wedge y\in A_a)\wedge \underline{\underline{(x\notin G \vee y\notin B_a)}}] \qquad (!!!)$$
$$\Rightarrow \exists a\in I: ((x,y)\in G\times A_a \wedge \overline{(x\in C\wedge y\in B_a)})$$
$$\Rightarrow \exists a\in I: ((x,y)\in G\times A_a \wedge (x,y)\notin G\times B_a)$$
$$\Rightarrow \exists a\in I: (x,y)\in (G\times A_a)-(G\times B_a)$$
$$\Rightarrow (x,y)\in\bigcup_{a\in I}[(G\times A_a)-(G\times B_a)]$$

it follows that:

$$C\times\left[\bigcup_{a\in I}(A_a - B_a)\right] \subseteq \bigcup_{a\in I}\left[(G\times A_a)-(G\times B_a)\right]$$

↳ Note that the (!!) step is valid but cannot be reversed.

EXERCISES

① Let $A = \{x \in \mathbb{Z} \mid 1 \leq x \leq 3\}$
$B = \{3x - 1 \mid x \in \mathbb{Z} \wedge 0 < x < 4\}$
List the elements of $A \times B$.

② Prove that for A, B, C sets
$$A \times (B \cap C) = (A \times B) \cap (A \times C)$$

③ Prove the following
a) $A \times B = \emptyset \Leftrightarrow A = \emptyset \vee B = \emptyset$
b) $(A \times B) \cap (C \times D) = (A \cap C) \times (B \cap D)$
c) $(A \times B) \cap (C \times D) = \emptyset \Leftrightarrow A \cap C = \emptyset \vee B \cap D = \emptyset$.

④ Prove the following.
a) $(A \times B) \cup (C \times D) \subseteq (A \cup C) \times (B \cup D)$
b) $\{p, q\} \subseteq A \Rightarrow (A \times \{p\}) \cup (\{q\} \times A) \subseteq A \times A$

⑤ Prove the following:
a) $A \times B = B \times A \Leftrightarrow A = \emptyset \vee B = \emptyset \vee A = B$
b) $A \neq \emptyset \neq B \wedge (A \times B) \cup (B \times A) = C \times C \Rightarrow A = B = C$.

⑥ Let $\{A_a\}_{a\in I}$ and $\{B_a\}_{a\in I}$ be indexed set collections and let C be a set. Prove the following:

a) $\left(\bigcup_{a\in I} A_a\right) \times C = \bigcup_{a\in I} (A_a \times C)$

b) $\left(\bigcap_{a\in I} A_a\right) \times C = \bigcap_{a\in I} (A_a \times C)$

c) $\bigcap_{a\in I} (A_a \times B_a) = \left(\bigcap_{a\in I} A_a\right) \times \left(\bigcap_{a\in I} B_a\right)$

⑦ Show that for A, B sets

$$\bigcup_{S\in\mathcal{P}(A)} \left[\bigcup_{T\in\mathcal{P}(B)} \{S \times T\}\right] \subseteq \mathcal{P}(A\times B)$$

▼ Relations

- Let A, B be two sets with $A \neq \emptyset$ and $B \neq \emptyset$. We define the set of all relations from A to B via the following belonging condition:

$$R \in \text{Rel}(A,B) \Leftrightarrow R \subseteq A \times B$$

- If $R \in \text{Rel}(A,B)$, we say that R is a <u>relation</u> from A to B.
- Let $R \in \text{Rel}(A,B)$ be a relation and let $x \in A$ and $y \in B$. Then we define the statements xRy and $x \not R y$ as follows:

$$\forall x \in A : \forall y \in B : (xRy \Leftrightarrow (x,y) \in R)$$
$$\forall x \in A : \forall y \in B : (x \not R y \Leftrightarrow (x,y) \notin R)$$

We say that:

xRy : x is related with y via relation R.

$x \not R y$: x is NOT related with y via relation R.

EXAMPLE

Let $A = \{a,b,c\}$ and $B = \{d,e,f,g,h\}$. Then

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

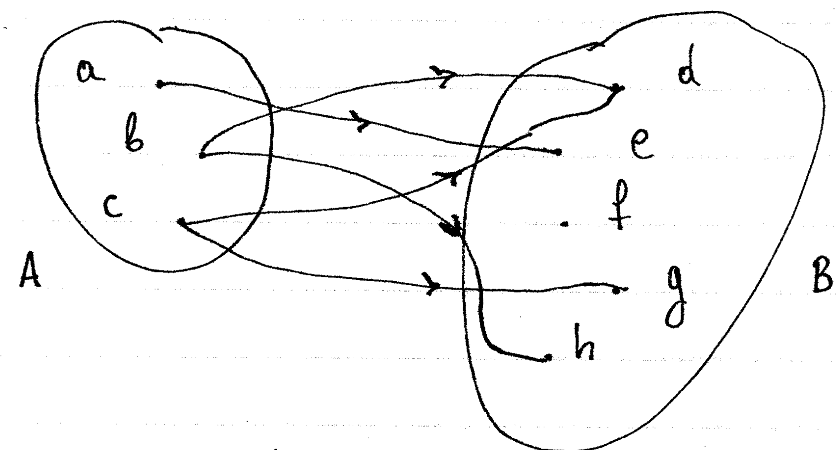
is a relation from A to B (i.e. $R \in \text{Rel}(A,B)$). Then

$(a,e) \in R \Rightarrow aRe \qquad (b,h) \in R \Rightarrow bRh$

$(b,d) \in R \Rightarrow bRd \qquad (c,d) \in R \Rightarrow cRd$

$(c,g) \in R \Rightarrow cRg$

↳ The relation R can be represented geometrically using a Venn diagram, as follows:



Each ordered pair (x,y) is represented by an arrow from x to y.

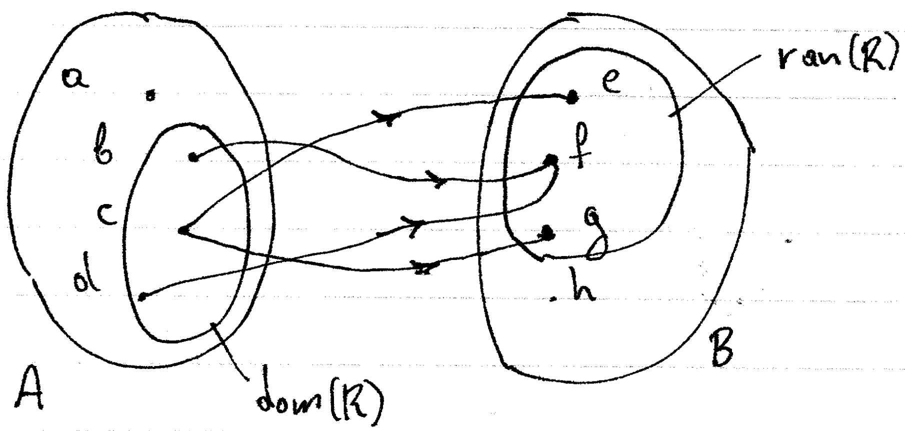
↳ Domain and range of a relation

- Let $R \in Rel(A,B)$ be a relation from A to B. We define the domain $dom(R)$ and range $ran(R)$ of R as:
$$dom(R) = \{x \in A \mid \exists y \in B : xRy\} \subseteq A$$
$$ran(R) = \{y \in B \mid \exists x \in A : xRy\} \subseteq B$$
- $dom(R)$ contains all the elements of A that are related with some element of B. In terms of Venn diagrams, $dom(R)$ has all the elements of A that have an outgoing arrow.
- $ran(R)$ contains all the elements of B that are related with some element of A. In terms of Venn diagrams,

ran(R) has all the elements of B that have an incoming arrow.

EXAMPLE

For $A=\{a,b,c,d\}$ and $B=\{e,f,g,h\}$, let $R \in Rel(A,B)$ be a relation from A to B with
$R=\{(b,f),(c,e),(d,f),(c,g)\}$.
Then: $dom(R)=\{b,c,d\}$ and
$ran(R)=\{e,f,g\}$



↦ Relations on A

We define $Rel(A)=Rel(A,A)$. Then:
$$R \in Rel(A) \Leftrightarrow R \subseteq A \times A$$
and we say that R is <u>a relation on A</u>.

▼ Equivalence relations

- Let $R \in \mathrm{Rel}(A)$ be a relation on A with $A \neq \emptyset$. We say that

  R reflexive $\iff \forall x \in A : xRx$

  R symmetric $\iff \forall x, y \in A : (xRy \Rightarrow yRx)$

  R transitive $\iff \forall x, y, z \in A : ((xRy \wedge yRz) \Rightarrow xRz)$

  and

$$R \text{ equivalence} \iff \begin{cases} R \text{ reflexive} \\ R \text{ symmetric} \\ R \text{ transitive} \end{cases}$$

EXAMPLES

a) Let $R \in Rel(A)$ be a relation on A. Show that
R reflexive $\Rightarrow$ $dom(R) = A$.

Solution

Assume that R is reflexive. Since
$dom(R) = \{x \in A \mid \exists y \in A : xRy\} \subseteq A \Rightarrow \underline{dom(R) \subseteq A}$ (1)
it is sufficient to show that $\forall x \in A : x \in dom(R)$.
Let $x \in A$ be given. Then:

$$\begin{aligned} R \text{ reflexive} &\Rightarrow xRx \\ &\Rightarrow \exists y \in A : xRy \\ &\Rightarrow x \in dom(R) \qquad [\text{via } x \in A] \end{aligned}$$

It follows that

$$\forall x \in A : x \in dom(R) \Rightarrow A \subseteq dom(R) \qquad (2)$$

From Eq.(1) and Eq.(2):

$$\left\{ \begin{array}{l} dom(R) \subseteq A \\ A \subseteq dom(R) \end{array} \right. \Rightarrow dom(R) = A.$$

b) Let $R \in Rel(A)$ be a relation on A. We define
R circular $\Leftrightarrow \forall x, y, z \in A : ((xRy \wedge yRz) \Rightarrow zRx)$

Show that:

$$\left\{ \begin{array}{l} R \text{ transitive} \\ R \text{ symmetric} \end{array} \right. \Rightarrow R \text{ circular}$$

Solution

Assume that R is transitive and symmetric.
Let $\underline{x, y, z \in A}$ be given and assume that $\underline{xRy \land yRz}$.
Then,

$$\left.\begin{matrix} xRy \\ yRz \end{matrix}\right\} \Rightarrow xRz \qquad [R \text{ is transitive}]$$

$$\Rightarrow \underline{zRx} \qquad [R \text{ is symmetric}]$$

From the above argument, it follows that
$\forall x, y, z \in A : ((xRy \land yRz) \Rightarrow zRx)$
$\Rightarrow R$ circular.

EXERCISES

(8) Show that the following relations are equivalences

a) $R \in Rel(\mathbb{Z})$ with $aRb \Leftrightarrow a+2b \equiv 0 \pmod{3}$

b) $R \in Rel(\mathbb{Z})$ with $aRb \Leftrightarrow a^3 \equiv b^3 \pmod{4}$

c) $R \in Rel(\mathbb{Z})$ with $aRb \Leftrightarrow 2a+3b \equiv 0 \pmod{5}$

(9) Show that the following relations on $\mathbb{R}^* \times \mathbb{R}^*$ are equivalences

a) $(x_1,y_1)R(x_2,y_2) \Leftrightarrow x_1y_2 - x_2y_1 = 0$

b) $(x_1,y_1)R(x_2,y_2) \Leftrightarrow \exists \lambda \in \mathbb{R}^* : (x_1 = \lambda x_2 \land y_1 = \lambda y_2)$

(Recall that $\mathbb{R}^* = \mathbb{R} - \{0\}$).

(10) Let $R \in Rel(A)$ be a relation on A. Show that

a) R reflexive $\Rightarrow$ $ran(R) = A$

b) R symmetric $\Rightarrow$ $dom(R) = ran(R)$

c) $(R$ circular $\land R$ symmetric$) \Rightarrow R$ transitive

d) R equivalence $\Leftrightarrow$ $(R$ reflexive $\land R$ circular$)$

$\hookrightarrow$ We use the definition

R circular $\Leftrightarrow \forall x,y,z \in A : ((xRy \land yRz) \Rightarrow zRx)$

(11) Let $R \in Rel(A)$. Write the definition, using quantifiers, for the following statements:

a) R is not reflexive

b) R is not symmetric

c) R is not transitive

d) R is not circular.

DST4: Mappings and Cardinality

MAPPINGS AND FUNCTIONS

▼ Basic Definitions

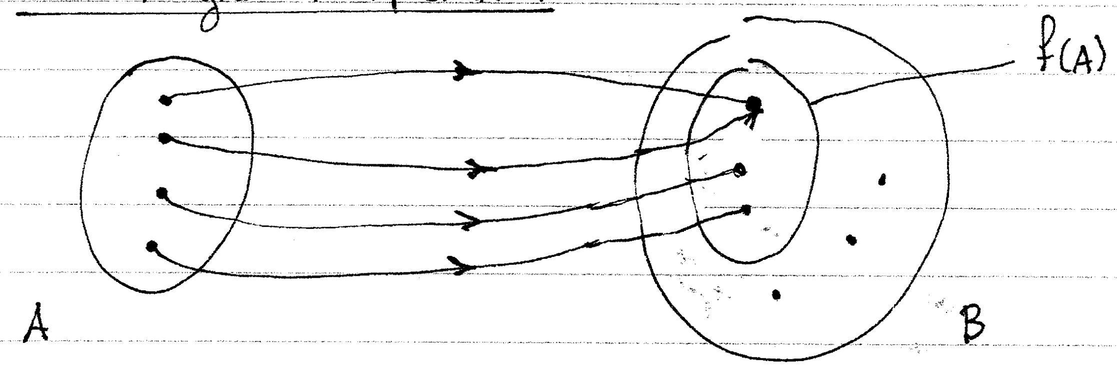
• Let A, B be two arbitrary sets. We say that f is a mapping that maps A to B (notation: $f: A \to B$) if and only if the following conditions are satisfied:

a) f is a relation $f \in Rel(A, B)$

b) $\forall x \in A : \exists y \in B : (x, y) \in f$

c) $\forall (x_1, y_1), (x_2, y_2) \in f : (x_1 = x_2 \Rightarrow y_1 = y_2)$

▶ Venn Diagram interpretation



Conditions (b) and (c) above have the following interpretations:

b) All elements of A have an outgoing arrow to some element of B

c) No element of A can have more than one outgoing arrow

Note that there are no restrictions on where the arrows go to as long as they go to some element of B.

▶ Special cases

- We denote the set of all mappings $f: A \to B$ as
$\mathrm{Map}(A,B) = \{f \in \mathrm{Rel}(A,B) \mid f: A \to B\}$
- For $A \subseteq \mathbb{R}$ we define the set of all <u>functions</u> with domain A:
  $F(A) = \mathrm{Map}(A, \mathbb{R})$.
- Also relevant are the following definitions
  $F(\mathbb{N})$ = the set of all real-valued <u>sequences</u>
  $\mathrm{Map}(\mathbb{R}^n, \mathbb{R})$ = the set of all <u>scalar fields</u>
  $\mathrm{Map}(\mathbb{R}^n, \mathbb{R}^n)$ = the set of all <u>vector fields</u>

▶ $f(x)$ notation

For every element $x \in A$, there is a unique $y \in B$ such that $(x,y) \in f$. We denote this unique y as <u>$y = f(x)$.</u>

EXAMPLE

For $f = \{(1,7), (2,5), (3,7)\}$, it follows that
$f(1) = 7$
$f(2) = 5$
$f(3) = 7$.

▶ $f(S)$ notation

Let $f: A \to B$ and let $S \subseteq A$. We define the <u>image $f(S)$</u> of S as follows:

$$f(\mathcal{S}) = \{f(x) \mid x \in \mathcal{S}\}$$

The belonging condition corresponding to $f(\mathcal{S})$ is given by

$$y \in f(\mathcal{S}) \iff \exists x \in \mathcal{S} : \ y = f(x)$$

EXAMPLE

For $f = \{(1,7),(2,5),(3,7)\}$, it follows that

$$f(\{1,2\}) = \{5,7\}$$
$$f(\{1,3\}) = \{7\}$$
$$f(\{1,2,3\}) = \{5,7\}$$
$$f(\varnothing) = \varnothing.$$

EXAMPLES

a) Let $f: A \to B$ be given and let $S \subseteq A$ and $T \subseteq A$.
Show that $f(S \cup T) = f(S) \cup f(T)$.

Solution

($\Rightarrow$): Let $y \in f(S \cup T)$ be given. Then

$y \in f(S \cup T) \Rightarrow \exists x \in S \cup T : f(x) = y$.

Choose $x_0 \in S \cup T$ such that $f(x_0) = y$.

Since $x_0 \in S \cup T \Rightarrow x_0 \in S \vee x_0 \in T$, we distinguish between the following cases:

Case 1: Assume that $x_0 \in S$. Then

$$\begin{cases} x_0 \in S \\ f(x_0) = y \end{cases} \Rightarrow \exists x \in S : y = f(x) \Rightarrow y \in f(S)$$

$$\Rightarrow y \in f(S) \vee y \in f(T) \Rightarrow y \in f(S) \cup f(T).$$

Case 2: Assume that $x_0 \in T$. Then

$$\begin{cases} x_0 \in T \\ f(x_0) = y \end{cases} \Rightarrow \exists x \in T : y = f(x) \Rightarrow y \in f(T)$$

$$\Rightarrow y \in f(S) \vee y \in f(T) \Rightarrow y \in f(S) \cup f(T).$$

In both cases we find $y \in f(S) \cup f(T)$ and therefore

$\forall y \in f(S \cup T) : y \in f(S) \cup f(T)$. (1)

($\Leftarrow$): Let $y \in f(S) \cup f(T)$ be given. Then:

$y \in f(S) \cup f(T) \Rightarrow y \in f(S) \vee y \in f(T) \Rightarrow$

$\Rightarrow (\exists x \in S : y = f(x)) \vee (\exists x \in T : y = f(x))$

We distinguish between the following two cases:

<u>Case 1</u> : Assume that $\exists x \in S : y = f(x)$.

Choose $x_0 \in S$ such that $y = f(x_0)$. Then:

$$\begin{cases} x_0 \in S \\ y = f(x_0) \end{cases} \Rightarrow \begin{cases} x_0 \in S \lor x_0 \in T \\ y = f(x_0) \end{cases} \Rightarrow \begin{cases} x_0 \in S \cup T \\ y = f(x_0) \end{cases} \Rightarrow$$

$$\Rightarrow \exists x \in S \cup T : y = f(x)$$

$$\Rightarrow y \in f(S \cup T).$$

<u>Case 2</u> : Assume that $\exists x \in T : y = f(x)$.

Choose $x_0 \in T$ such that $y = f(x_0)$. Then:

$$\begin{cases} x_0 \in T \\ y = f(x_0) \end{cases} \Rightarrow \begin{cases} x_0 \in S \lor x_0 \in T \\ y = f(x_0) \end{cases} \Rightarrow \begin{cases} x_0 \in S \cup T \\ y = f(x_0) \end{cases} \Rightarrow$$

$$\Rightarrow \exists x \in S \cup T : y = f(x)$$

$$\Rightarrow y \in f(S \cup T).$$

In both cases we find $y \in f(S \cup T)$ and therefore

$$\forall y \in f(S) \cup f(T) : y \in f(S \cup T). \qquad (2)$$

From Eq.(1) and Eq.(2):

$$\begin{cases} \forall y \in f(S \cup T) : y \in f(S) \cup f(T) \\ \forall y \in f(S) \cup f(T) : y \in f(S \cup T) \end{cases} \Rightarrow \begin{cases} f(S \cup T) \subseteq f(S) \cup f(T) \\ f(S) \cup f(T) \subseteq f(S \cup T) \end{cases}$$

$$\Rightarrow f(S \cup T) = f(S) \cup f(T).$$

b) Let $f: A \to B$ be given. Use a counterexample to explain why we cannot prove that for $S \subseteq A$ and $T \subseteq A$ we have $f(S \cap T) = f(S) \cap f(T)$.

Solution

Consider the mapping

$f = \{(a,x),(b,x),(c,y),(d,y)\}$

and define $S = \{b,c\}$ and $T = \{a,d\}$.

Then:

$$f(S \cap T) = f(\{b,c\} \cap \{a,d\}) = f(\emptyset) = \emptyset \qquad (1)$$

but

$f(b) = x \wedge f(c) = y \Rightarrow f(S) = f(\{b,c\}) = \{x,y\}$

$f(a) = x \wedge f(d) = y \Rightarrow f(T) = f(\{a,d\}) = \{x,y\}$

and therefore

$$f(S) \cap f(T) = \{x,y\} \cap \{x,y\} = \{x,y\} \qquad (2)$$

From Eq. (1) and Eq. (2):

$f(S \cap T) \neq f(S) \cap f(T)$

↳ Proof by counterexample can be very challenging. The statement $f(S \cap T) = f(S) \cap f(T)$ can be true for some choices of S, T and false for other choices of S, T. Can you find alternate choices for S, T for which the statement is true?

EXERCISES

① Let $f: A \to B$ be given, and let $S \subseteq A$ and $T \subseteq A$. Show that

a) $f(S \cap T) \subseteq f(S) \cap f(T)$

b) $f(S) - f(T) \subseteq f(S - T)$

② Find a counterexample of an $f: A \to B$ and $S \subseteq A$ and $T \subseteq A$ such that the following statements are false:

a) $f(S \cap T) = f(S) \cap f(T)$

b) $f(S) - f(T) = f(S - T)$

→ We will later show that these statements can be proved if additional assumptions about f are introduced.

③ Let $f: A \to B$ be given and let S_a such that $\forall a \in I : S_a \subseteq A$ with I an index set. Show that

a) $f\left(\bigcup_{a \in I} S_a\right) = \bigcup_{a \in I} f(S_a)$

b) $f\left(\bigcap_{a \in I} S_a\right) \subseteq \bigcap_{a \in I} f(S_a)$

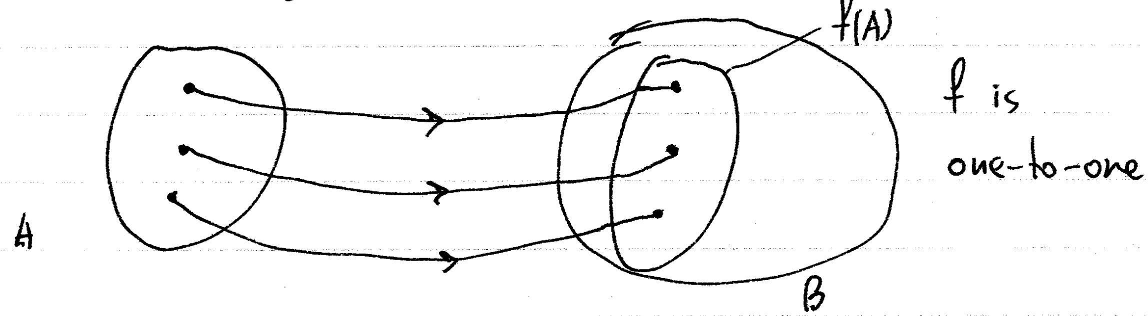
One-to-one and onto mappings

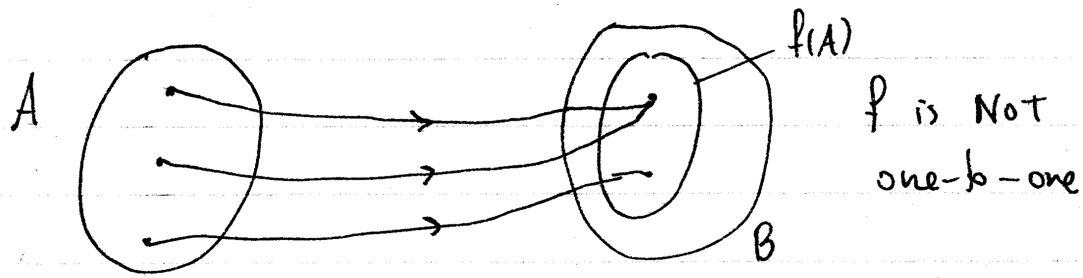
- Let $f: A \to B$ be given. We say that

$$
\boxed{
\begin{array}{l}
f \text{ one-to-one} \Leftrightarrow \forall x_1, x_2 \in A : (f(x_1) = f(x_2) \Rightarrow x_1 = x_2) \\
f \text{ onto} \Leftrightarrow f(A) = B \\
f \text{ Bijection} \Leftrightarrow f \text{ one-to-one} \land f \text{ onto}
\end{array}
}
$$

▸ Remarks

a) In a one-to-one mapping, every point in the range $f(A)$ receives only one incoming arrow.





This interpetation becomes clear in terms of the negation of the one-to-one definition. Since $\overline{p \Rightarrow q} \equiv p \land \overline{q}$:

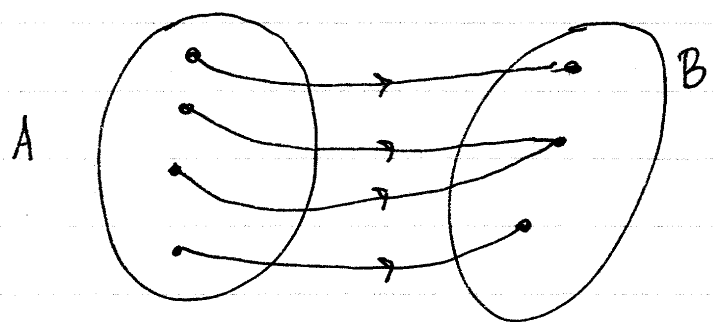
$$
\boxed{f \text{ NOT one-to-one} \Leftrightarrow \exists x_1, x_2 \in A : (f(x_1) = f(x_2) \land x_1 \neq x_2)}
$$

b) From the definition of $f(A)$, we always have $f(A) \subseteq B$.
It follows that the "onto" definition can be rewritten as:

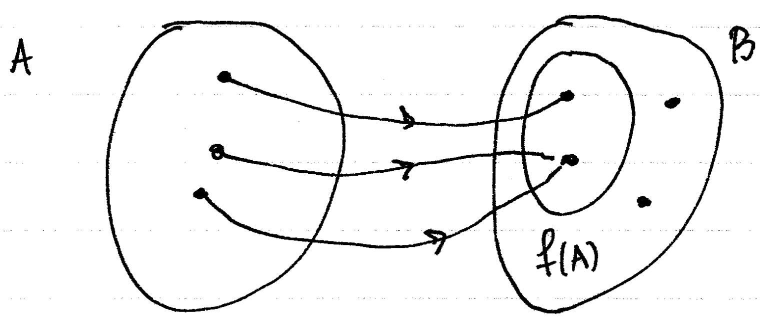
$$f \text{ onto} \Leftrightarrow f(A) = B \Leftrightarrow f(A) \subseteq B \wedge B \subseteq f(A)$$
$$\Leftrightarrow B \subseteq f(A) \Leftrightarrow \forall y \in B: y \in f(A) \Leftrightarrow$$
$$\Leftrightarrow \forall y \in B: \exists x \in A: f(x) = y$$

This gives the following interpretation:
"f is onto if and only if for every element y of B, there is an element $x \in A$ such that $f(x) = y$"
or equivalently
"f is onto if and only if every element in B has at least one incoming arrow".
In summary:

$$\boxed{\begin{array}{l} f \text{ onto} \Leftrightarrow \forall y \in B: \exists x \in A: f(x) = y \\ f \text{ not onto} \Leftrightarrow \exists y \in B: \forall x \in A: f(x) \neq y \end{array}}$$



f onto



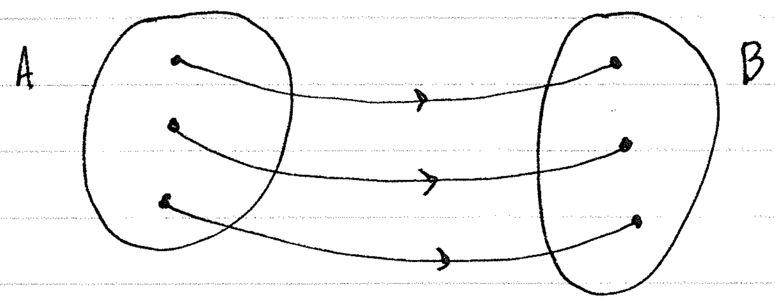
f not onto

c) If f is a bijection then both conditions are satisfied:

(1) one-to-one: No element of B has more than 1 incoming arrow

(2) onto: Every element of B has at least one incoming arrow.

Combining the two conditions together we see that f is a bijection if and only if every element of B has exactly 1 incoming arrow:



f bijection

▶ Inverse mapping

If $f: A \to B$ is a bijection, then reversing all arrows gives the bijection $f^{-1}: B \to A$ such that for all $x \in A$ and for all $y \in B$,

$$(y = f^{-1}(x) \Leftrightarrow f(y) = x)$$

We say that f^{-1} is the inverse mapping of f. Note that

$\forall x \in A: f^{-1}(f(x)) = x$

$\forall x \in B: f(f^{-1}(x)) = x$

▶ Methodology

To derive statements of the form $A=B \Rightarrow C=D$ we use the following properties of real numbers

1) We can add/cancel any number to both sides of an equation:
$$\forall a,x,y \in \mathbb{R}: (x=y \Leftrightarrow a+x=a+y)$$
2) We can always add or multiply two equations
$$\forall a,b,x,y \in \mathbb{R}: (a=b \land x=y \Rightarrow a+x=b+y)$$
$$\forall a,b,x,y \in \mathbb{R}: (a=b \land x=y \Rightarrow ax=by)$$
3) We can multiply any number to both sides of an equation:
$$\forall a,x,y \in \mathbb{R}: (x=y \Rightarrow ax=ay)$$
However the converse does not work for $a=0$. With the restriction $a\neq 0$ we have:
$$\forall x,y \in \mathbb{R}: \forall a \in \mathbb{R}-\{0\}: (x=y \Leftrightarrow ax=ay)$$
4) We can raise both sides of an equation to any integer power:
$$\forall x,y \in \mathbb{R}: \forall n \in \mathbb{N}: (x=y \Rightarrow x^n=y^n)$$
In general, the converse does not work. However, if we require $n\neq 0$ and distinguish between odd and even powers, we have:
$$\forall x,y \in \mathbb{R}: \forall n \in \mathbb{Z}: (x^{2n+1}=y^{2n+1} \Leftrightarrow x=y)$$
$$\forall x,y \in \mathbb{R}: \forall n \in \mathbb{Z}-\{0\}: (x^{2n}=y^{2n} \Leftrightarrow x=y \lor x=-y)$$
5) Factored equation:
$$\forall a,b \in \mathbb{R}: (ab=0 \Leftrightarrow a=0 \lor b=0)$$

EXAMPLES

a) Consider the function

$$\forall x \in \mathbb{R} - \{a\} : f(x) = \frac{x}{x-a}$$

Show that $a \neq 0 \Rightarrow f$ one-to-one.

Solution

Assume that $a \neq 0$. Let $x_1, x_2 \in \mathbb{R} - \{a\}$ be given such that $f(x_1) = f(x_2)$. Then

$$\underline{f(x_1) = f(x_2)} \Rightarrow \frac{x_1}{x_1 - a} = \frac{x_2}{x_2 - a} \Rightarrow$$

$$\Rightarrow (x_1 - a)(x_2 - a)\frac{x_1}{x_1 - a} = (x_1 - a)(x_2 - a)\frac{x_2}{x_2 - a} \Rightarrow$$

$$\Rightarrow x_1(x_2 - a) = x_2(x_1 - a) \Rightarrow x_1 x_2 - a x_1 = x_1 x_2 - a x_2$$

$$\Rightarrow \left.\begin{array}{c} -a x_1 = -a x_2 \\ a \neq 0 \end{array}\right\} \Rightarrow \underline{x_1 = x_2}$$

It follows that

$\forall x_1, x_2 \in \mathbb{R} - \{a\} : (f(x_1) = f(x_2) \Rightarrow x_1 = x_2)$

$\Rightarrow f$ one-to-one.

$\hookrightarrow$ Note that to cancel $-a$ in $-a x_1 = -a x_2$ we need the assumption $a \neq 0$, otherwise the cancellation cannot be justified.

b) Consider the function $f(x) = 2x^2 + 6x - 7,\ \forall x \in \mathbb{R}$

Show that f is not one-to-one.

Solution

Solve $f(x) = -7 \Leftrightarrow 2x^2 + 6x - 7 = -7 \Leftrightarrow 2x^2 + 6x = 0 \Leftrightarrow$

$\Leftrightarrow 2x(x+3) = 0 \Leftrightarrow 2x = 0 \vee x + 3 = 0$

$\Leftrightarrow x = 0 \vee x = -3$

It follows that

$f(0) = f(-3) = -7 \wedge 0 \neq -3 \Rightarrow$

$\Rightarrow \exists x_1, x_2 \in \mathbb{R} : f(x_1) = f(x_2) \wedge x_1 \neq x_2$

$\Rightarrow f$ not one-to-one.

c) Let $f: A \to B$ be given and let $S \subseteq A$ and $T \subseteq A$.

Show that

f one-to-one $\Rightarrow f(S \cap T) = f(S) \cap f(T)$.

Solution

Assume that f is one-to-one.

($\Rightarrow$): Let $y \in f(S \cap T)$ be given. Then,

$y \in f(S \cap T) \Rightarrow \exists x \in S \cap T : f(x) = y$

Choose $x_0 \in S \cap T$ such that $f(x_0) = y$. It follows that

$$\begin{cases} x_0 \in S \cap T \\ f(x_0) = y \end{cases} \Rightarrow \begin{cases} x_0 \in S \land x_0 \in T \\ f(x_0) = y \end{cases} \Rightarrow$$

$$\Rightarrow \begin{cases} x_0 \in S \\ f(x_0) = y \end{cases} \land \begin{cases} x_0 \in T \\ f(x_0) = y \end{cases} \Rightarrow$$

$$\Rightarrow (\exists x \in S : f(x) = y) \land (\exists x \in T : f(x) = y)$$

$$\Rightarrow y \in f(S) \land y \in f(T) \Rightarrow$$

$$\Rightarrow y \in f(S) \cap f(T).$$

($\Leftarrow$): Let $y \in f(S) \cap f(T)$ be given. Then:

$y \in f(S) \cap f(T) \Rightarrow y \in f(S) \land y \in f(T) \Rightarrow$

$$\Rightarrow \begin{cases} \exists x \in S : f(x) = y \\ \exists x \in T : f(x) = y. \end{cases}$$

Choose $x_1 \in S$ and $x_2 \in T$ such that $f(x_1) = y$ and $f(x_2) = y$.

Then:

$$\begin{cases} f(x_1) = y = f(x_2) \\ f \text{ one-to-one} \end{cases} \Rightarrow x_1 = x_2 \in T \Rightarrow x_1 \in T.$$

and therefore:

$$\begin{cases} x_1 \in S \wedge x_1 \in T \\ f(x_1) = y \end{cases} \Rightarrow \begin{cases} x_1 \in S \cap T \\ f(x_1) = y \end{cases} \Rightarrow$$

$$\Rightarrow \exists x \in S \cap T : f(x) = y$$

$$\Rightarrow \underline{y \in f(S \cap T)}$$

From the above argument we have:

$$\begin{cases} \forall : y \in f(S \cap T) : y \in f(S) \cap f(T) \\ \forall y \in f(S) \cap f(T) : y \in f(S \cap T) \end{cases} \Rightarrow$$

$$\Rightarrow \begin{cases} f(S \cap T) \subseteq f(S) \cap f(T) \\ f(S) \cap f(T) \subseteq f(S \cap T) \end{cases} \Rightarrow$$

$$\Rightarrow f(S \cap T) = f(S) \cap f(T).$$

EXERCISES

(4) Show that the following functions are one-to-one

a) $\forall x \in \mathbb{R}: f(x) = 3x^5 + 2$

b) $\forall x \in (0, +\infty): f(x) = 2x^2 + 5$

c) $\forall x \in \mathbb{R}: f(x) = ax + b$ with $a, b \in \mathbb{R} \wedge a \neq 0$

d) $\forall x \in \mathbb{R}: f(x) = (2x^3 + 1)^5$

e) $\forall x \in \mathbb{R} - \{0\}: f(x) = a/x$ with $a \in \mathbb{R} \wedge a \neq 0$

f) $\forall x \in \mathbb{R} - \{-d/c\}: f(x) = \dfrac{ax+b}{cx+d}$ with $a, b, c, d \in \mathbb{R} \wedge ad - bc \neq 0$

(5) Show that for $\forall x \in \mathbb{R}: f(x) = ax^2 + bx + c$ with $a, b, c \in \mathbb{R}$ and $a \neq 0$ is <u>not</u> one-to-one.

(6) Let $f: A \to B$ be given and let $S \subseteq A$ and $T \subseteq A$. Show that

$$f \text{ one-to-one} \Rightarrow f(S - T) = f(S) - f(T).$$

(7) Let $f: A \to B$ be given and let S_a be a set collection such that $\forall a \in I: S_a \subseteq A$, with I an index set. Show that

$$f \text{ one-to-one} \Rightarrow f\left(\bigcap_{a \in I} S_a\right) = \bigcap_{a \in I} f(S_a)$$

Cardinality

- Given two finite sets A, B, if there is a bijection $f: A \to B$ then A and B have to have the same number of elements. Cantor proposed extending his observation to infinite sets according to the following definitions:

Def: Let A, B be two sets. We say that $A \sim B \Leftrightarrow \exists f \in \text{Map}(A,B) : f$ bijection

- The statement $A \sim B$ reads "A, B are equipotent", or "A and B have the same cardinality".
- Recall the definition
  $[n] = \{x \in \mathbb{N}^* \mid x \leq n\} = \{1, 2, 3, \dots, n\}$
  Based on that, we introduce the following characterizations:

$$\boxed{\begin{aligned}
&A \text{ finite set} \Leftrightarrow A = \varnothing \vee (\exists n \in \mathbb{N}^* : A \sim [n]) \\
&A \text{ infinite set} \Leftrightarrow A \text{ not finite set} \\
&\qquad\qquad\qquad \Leftrightarrow A \neq \varnothing \wedge (\forall n \in \mathbb{N}^* : \overline{A \sim [n]}) \\
&A \text{ countable set} \Leftrightarrow \exists B \in \mathcal{P}(\mathbb{N}) : A \sim B \\
&A \text{ countably infinite} \Leftrightarrow A \sim \mathbb{N} \\
&A \text{ uncountable} \Leftrightarrow A \text{ not countable}
\end{aligned}}$$

- A relative comparison of sets in terms of cardinality is:
  $$\text{finite} \leq \text{count}\underbrace{\text{able} \leq \text{countably infinite} < \text{uncountable}}_{\text{infinite}}$$

It should be stressed that since $\emptyset, \mathbb{N} \in \mathcal{P}(\mathbb{N})$ and
$\forall n \in \mathbb{N}^* : [n] \in \mathcal{P}(\mathbb{N})$
it follows that

A finite $\Rightarrow$ A countable

A countably infinite $\Rightarrow$ A countable

However, the converse statements do not hold.

▸ <u>Interpretation</u>: A countably infinite set contains an infinite number of elements, however the existence of some bijection $f: A \to \mathbb{N}$ allows us to enumerate each element of A by assigning it to a unique natural number from $\mathbb{N}$.

▸ <u>$\mathbb{Z}$ and $\mathbb{Q}$ are countable</u>

Recall that

$$\mathbb{Z} = \mathbb{N} \cup \{-x \mid x \in \mathbb{N}^*\} = \{0, 1, -1, 2, -2, 3, -3, \dots\}$$
$$\mathbb{Q} = \{(a/b) \mid a, b \in \mathbb{Z} \wedge b \neq 0\}$$

with $\mathbb{Z}$ the set of integers and $\mathbb{Q}$ the set of rational numbers. The remarkable insight of Cantor is that even though $\mathbb{Z}$ and $\mathbb{Q}$ contain "more numbers" than $\mathbb{N}$, in the sense that $\mathbb{N} \subset \mathbb{Z} \subset \mathbb{Q}$, from the standpoint of cardinality, we can show that $\mathbb{Z} \sim \mathbb{N}$ and $\mathbb{Q} \sim \mathbb{N}$. Equivalently, we can show that

$$\begin{cases} \mathbb{Z} \text{ countably infinite} \\ \mathbb{Q} \text{ countably infinite} \end{cases}$$

▸ <u>$\mathbb{R}$ is uncountable</u>

With some additional theory we can show that the set $\mathbb{R}$ of all real numbers satisfies the following statements:

$$\begin{cases} \mathbb{R} \text{ is uncountable} \\ \mathbb{R} \sim \mathcal{P}(\mathbb{N}) \end{cases}$$

Proof of $\mathbb{Z} \sim \mathbb{N}$ ($\mathbb{Z}$ is countably infinite)

We define the mapping $f : \mathbb{Z} \to \mathbb{N}$ such that

$$\forall x \in \mathbb{Z} : f(x) = \begin{cases} 2x-1 & , \text{if } x > 0 \\ -2x & , \text{if } x \leq 0 \end{cases}$$

and note that

$$f = \{(0,0), (1,1), (-1,2), (2,3), (-2,4), (3,5), (-3,6), \ldots\}$$

which indicates that f is a bijection. To prove that, we show that f is one-to-one and that f is onto.

- one-to-one : Sufficient to show that
  $\forall x_1, x_2 \in \mathbb{Z} : (f(x_1) = f(x_2) \Rightarrow x_1 = x_2)$

Let $x_1, x_2 \in \mathbb{Z}$ be given and assume that $f(x_1) = f(x_2)$.

We distinguish between the following cases.

Case 1 : Assume that $f(x_1) = -2x_1$ and $f(x_2) = -2x_2$. Then,

$f(x_1) = f(x_2) \Rightarrow -2x_1 = -2x_2 \Rightarrow x_1 = x_2$.

Case 2: Assume that $f(x_1) = 2x_1 - 1$ and $f(x_2) = 2x_2 - 1$. Then

$f(x_1) = f(x_2) \Rightarrow 2x_1 - 1 = 2x_2 - 1 \Rightarrow 2x_1 = 2x_2 \Rightarrow x_1 = x_2$

Case 3 : Assume that $f(x_1) = 2x_1 - 1$ and $f(x_2) = -2x_2$. Then

$$f(x_1) = f(x_2) \Rightarrow 2x_1 - 1 = -2x_2 \Rightarrow 2x_1 + 2x_2 = 1 \Rightarrow$$
$$\Rightarrow 2(x_1 + x_2) = 1 \Rightarrow x_1 + x_2 = 1/2$$

This is a contradiction, because

$$x_1, x_2 \in \mathbb{Z} \Rightarrow x_1 + x_2 \in \mathbb{Z} \Rightarrow x_1 + x_2 \neq 1/2$$

therefore case 3 does not materialize.

From the above cases we conclude that $x_1 = x_2$ and therefore:

$$\forall x_1, x_2 \in \mathbb{Z} : (f(x_1) = f(x_2) \Rightarrow x_1 = x_2) \qquad (1)$$

• <u>Onto:</u> Sufficient to show that $\forall y \in \mathbb{N} : \exists x \in \mathbb{Z} : f(x) = y$.
Let $\underline{y \in \mathbb{N}}$ be given. From the division theorem we have:
$\exists k \in \mathbb{N} : (y = 2k \lor y = 2k+1)$
Choose a $k \in \mathbb{N}$ such that $y = 2k \lor y = 2k+1$ and distinguish between the following cases.

<u>Case 1</u>: Assume that $y = 2k$. Then:

$$k \in \mathbb{N} \Rightarrow k \geq 0 \Rightarrow -k \leq 0 \Rightarrow f(-k) = -2(-k) = 2k = y \Rightarrow$$
$$\Rightarrow \underline{\exists x \in \mathbb{Z} : f(x) = y} \quad (\text{for } x = -k)$$

<u>Case 2</u>: Assume that $y = 2k+1$. Then:

$$k \in \mathbb{N} \Rightarrow k \geq 0 \Rightarrow k+1 > 0 \Rightarrow$$
$$\Rightarrow f(k+1) = 2(k+1) - 1 = 2k+2-1 = 2k+1 = y \Rightarrow$$
$$\Rightarrow \underline{\exists x \in \mathbb{Z} : f(x) = y.} \quad (\text{for } x = k+1)$$

From the above argument, in all cases, we find that

$$(\forall y \in \mathbb{N} : \exists x \in \mathbb{Z} : f(x) = y) \Rightarrow \forall y \in \mathbb{N} : y \in f(\mathbb{Z}) \Rightarrow$$
$$\Rightarrow \mathbb{N} \subseteq f(\mathbb{Z}) \Rightarrow$$
$$\Rightarrow f(\mathbb{Z}) = \mathbb{N} \quad (2)$$

From Eq. (1) and Eq. (2).

$$\begin{cases} \forall x_1, x_2 \in \mathbb{Z} : (f(x_1) = f(x_2) \Rightarrow x_1 = x_2) \\ f(\mathbb{Z}) = \mathbb{N} \end{cases} \Rightarrow$$
$$\Rightarrow \begin{cases} f \text{ one-to-one} \\ f \text{ onto} \end{cases} \Rightarrow f : \mathbb{Z} \to \mathbb{N} \text{ bijection}$$
$$\Rightarrow \mathbb{Z} \sim \mathbb{N} \Rightarrow \mathbb{Z} \text{ countably infinite.}$$

Sketch of proof that $\mathbb{Q} \sim \mathbb{N}$

A bijection $f: \mathbb{Q} \to \mathbb{N}$ can be constructed via the process of diagonalization, originally proposed by Cantor. We will explain this process and the overall argument informally, for the sake of clarity. We sequence the rational numbers using the diagonalizing pattern shown in the table below, making sure to ship any numbers previously encountered in an equivalent fractional representation:

	0	1	2	3	4	...
1	0/1	1/1	2/1	3/1	4/1	...
2	0/2	1/2	2/2	3/2	4/2	...
3	0/3	1/3	2/3	3/3	4/3	...
4	0/4	1/4	2/4	3/4	4/4	...
5	0/5	...	...	...	...	
⋮	⋮					

Consequently, we sequence the rational numbers of $\mathbb{Q}$ as follows:

$\underline{0/1}$, $\underline{1/1}$, $0/2$, $\underline{2/1}$, $\underline{1/2}$, $0/3$, $\underline{3/1}$, $2/2$, $\underline{1/3}$, $0/4$, $\underline{4/1}$, $\underline{3/2}$, $\underline{2/3}$, $\underline{1/4}$, $0/5$, etc.

where we have underlined all rational numbers that appear for the first time and thus are not being skipped. We can thus define a bijection $f: \mathbb{Z} \to \mathbb{Q}$

with the initial assignments:

$f(0) = 0/1 = 0$ $\quad f(4) = 3/1$ $\quad f(8) = 2/3$

$f(1) = 1/1 = 1$ $\quad f(5) = 1/3$ $\quad f(9) = 1/4$

$f(2) = 2/1 = 2$ $\quad f(6) = 4/1$

$f(3) = 1/2$ $\quad f(7) = 3/2$ $\quad$ etc.

The algorithm for generating this bijection is as follows:

for $a = 0, 1, 2, 3, 4, \ldots$
  for $b = 0, 1, 2, \ldots, a$
    if it has not occured previously then add
    the number $(a-b)/(b+1)$ to the sequence.
  end for
end for.

To account for negative rational numbers, we extend the definition by the algorithm above as follows:

$\forall x \in \mathbb{N}^* : f(-x) = -f(x)$

and that completes the bijection $f: \mathbb{Z} \to \mathbb{Q}$. Skipping numbers that occured previously ensures that f is one-to-one. It is also clear that any rational number will be reached by this algorithm with a finite numbers of steps, which ensures that f is onto. Thus, it follows that

$$
\begin{aligned}
f: \mathbb{Z} \to \mathbb{Q} \text{ bijection} &\Rightarrow \mathbb{Q} \sim \mathbb{Z} && [\text{definition}] \\
&\Rightarrow \mathbb{Q} \sim \mathbb{N} && [\text{via } \mathbb{Z} \sim \mathbb{N}] \\
&\Rightarrow \mathbb{Q} \text{ countable} && \square
\end{aligned}
$$

EXAMPLE – APPLICATION

↱ The following problem is also a necessary first step towards proving that $\mathbb{R}$ is uncountable.

Show that $\boxed{\mathbb{R} \sim (0,1)}$

Solution

Define $\forall x \in \mathbb{R} : f(x) = (1/2) + (1/\pi) \operatorname{Arctan}(x)$.

We will show that $f : \mathbb{R} \to (0,1)$ is a bijection.

• Onto : Sufficient to show $\begin{cases} \forall y \in f(\mathbb{R}) : y \in (0,1) \\ \forall y \in (0,1) : y \in f(\mathbb{R}) \end{cases}$

($\Rightarrow$): Let $y \in f(\mathbb{R})$ be given. Then

$y \in f(\mathbb{R}) \Rightarrow \exists x \in \mathbb{R} : f(x) = y$.

Choose $x_0 \in \mathbb{R}$ such that $f(x_0) = y$. Then,

$$-\pi/2 < \operatorname{Arctan}(x_0) < \pi/2 \Rightarrow$$
$$\Rightarrow -1/2 < (1/\pi)\operatorname{Arctan}(x_0) < 1/2 \Rightarrow$$
$$\Rightarrow 0 < (1/2) + (1/\pi)\operatorname{Arctan}(x_0) < 1 \Rightarrow$$
$$\Rightarrow 0 < f(x_0) < 1 \Rightarrow 0 < y < 1 \Rightarrow \underline{y \in (0,1)}$$

It follows that $\forall y \in f(\mathbb{R}) : y \in (0,1)$. (1)

($\Leftarrow$): Let $y \in (0,1)$ be given. Then, we note that

$$f(x) = y \Leftrightarrow (1/2) + (1/\pi)\operatorname{Arctan}(x) = y \Leftrightarrow$$
$$\Leftrightarrow (1/\pi)\operatorname{Arctan}(x) = y - 1/2$$
$$\Leftrightarrow \operatorname{Arctan}(x) = \pi(y - 1/2) \qquad (2)$$

and also that

$$y \in (0,1) \Rightarrow 0 < y < 1 \Rightarrow -1/2 < y - 1/2 < 1/2 \Rightarrow$$
$$\Rightarrow -\pi/2 < \pi(y - 1/2) < \pi/2 \Rightarrow \tan \text{ is defined at } \pi(y - 1/2).$$

Now we can define $x_0 = \tan(\pi(y-1/2))$ and conclude that

$$\operatorname{Arctan}(x_0) = \operatorname{Arctan}(\tan(\pi(y-1/2))) = \pi(y-1/2) \overset{(2)}{\Rightarrow}$$
$$\Rightarrow f(x_0) = y \Rightarrow \exists x \in \mathbb{R} : f(x) = y \Rightarrow$$
$$\Rightarrow \underline{y \in f(\mathbb{R})}$$

and therefore,

$$\forall y \in (0,1) : y \in f(\mathbb{R}) \qquad (3)$$

From Eq.(2) and Eq.(3):

$$\begin{cases} \forall y \in f(\mathbb{R}) : y \in (0,1) \\ \forall y \in (0,1) : y \in f(\mathbb{R}) \end{cases} \Rightarrow \begin{cases} f(\mathbb{R}) \subseteq (0,1) \\ (0,1) \subseteq f(\mathbb{R}) \end{cases} \Rightarrow f(\mathbb{R}) = (0,1)$$
$$\Rightarrow f \text{ onto.} \qquad (4)$$

• <u>One-to-one</u>

Let $\underline{x_1, x_2 \in \mathbb{R}}$ be given and assume that $\underline{f(x_1) = f(x_2)}$. Then,

$$f(x_1) = f(x_2) \Rightarrow (1/2) + (1/\pi)\operatorname{Arctan}(x_1) = (1/2) + (1/\pi)\operatorname{Arctan}(x_2) \Rightarrow$$
$$\Rightarrow (1/\pi)\operatorname{Arctan}(x_1) = (1/\pi)\operatorname{Arctan}(x_2) \Rightarrow$$
$$\Rightarrow \operatorname{Arctan}(x_1) = \operatorname{Arctan}(x_2) \Rightarrow$$
$$\Rightarrow \tan(\operatorname{Arctan}(x_1)) = \tan(\operatorname{Arctan}(x_2))$$
$$\Rightarrow \underline{x_1 = x_2}$$

and therefore, we have

$$\forall x_1, x_2 \in \mathbb{R} : (f(x_1) = f(x_2) \Rightarrow x_1 = x_2)$$
$$\Rightarrow f \text{ one-to-one} \qquad (5)$$

From Eq.(4) and Eq.(5):

$$\begin{cases} f \text{ onto} \\ f \text{ one-to-one} \end{cases} \Rightarrow f : \mathbb{R} \to (0,1) \text{ bijection} \Rightarrow \mathbb{R} \sim (0,1).$$

EXERCISES

(8) Learn the proofs for the following statements

a) $\mathbb{Z}$ is countable

b) $\mathbb{Q}$ is countable

c) $\mathbb{R} \sim (0,1)$

(9) Let A, B be two sets. Show that

A countable $\wedge$ B countable $\Rightarrow A \cup B$ countable.

(10) Let A_a with $a \in \mathbb{N}$ be a set collection. Show that:

a) $(\forall a \in \mathbb{N} : A_a \text{ finite}) \Rightarrow \bigcup_{a \in \mathbb{N}} A_a$ countable

b) Use part (a) to show that

$$(\forall a \in \mathbb{N} : A_a \sim \mathbb{N}) \Rightarrow \bigcup_{a \in \mathbb{N}} A_a \sim \mathbb{N}$$

(11) Given 3 sets A, B, C show that the set equivalence satisfies the reflexive, symmetric, and transitive properties.

a) $A \sim A$

b) $A \sim B \Rightarrow B \sim A$

c) $A \sim B \wedge B \sim C \Rightarrow A \sim C$

(12) Let $a, b, c, d \in \mathbb{R}$ with $a < b$ and $c < d$ and consider the intervals

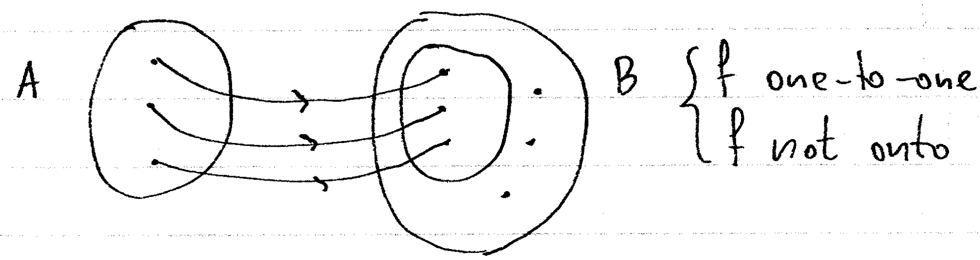
$$[a,b] = \{x \in \mathbb{R} \mid a \leq x \leq b\}$$

$$[c,d] = \{x \in \mathbb{R} \mid c \leq x \leq d\}$$

Construct a bijection to show that $[a,b] \sim [c,d]$.

▼ Cardinality inequalities

If we can define a mapping $f: A \to B$ which is one-to-one but not necessarily onto, then from an intuitive standpoint the only conclusion that can be drawn is that either A, B are of "equal cardinality" or "B has greater cardinality than A", as illustrated by the following figure:



Consequently, we propose the following definitions:

$$\boxed{\begin{array}{l} A \leqslant B \Leftrightarrow \exists f \in \mathrm{Map}(A, B) : f \text{ one-to-one} \\ A < B \Leftrightarrow A \leqslant B \wedge A \not\sim B \end{array}}$$

Note that it is easy to show that:

$$A \sim B \wedge B \sim C \Rightarrow A \sim C$$
$$A \leqslant B \wedge B \leqslant C \Rightarrow A \leqslant C$$
$$A \subseteq B \Rightarrow A \leqslant B$$

which are left as homework problems. Starting from Cantor, the following two major theorems will be used to show that $\mathbb{R} \sim \mathcal{P}(\mathbb{N})$ and $\mathbb{R}$ uncountable.

① → Cantor's theorem

$$\boxed{\underline{\text{Thm}} : \text{For any set } A, \; A < \mathcal{P}(A)}$$

Proof

Define $f_0 : A \to \mathcal{P}(A)$ such that $\forall x \in A : f_0(x) = \{x\}$. Then:

$\forall x_1, x_2 \in A : (\{x_1\} = \{x_2\} \Rightarrow x_1 = x_2) \Rightarrow$

$\Rightarrow \forall x_1, x_2 \in A : (f_0(x_1) = f_0(x_2) \Rightarrow x_1 = x_2)$

$\Rightarrow f_0$ one-to-one $\Rightarrow$

$\Rightarrow \exists f \in \text{Map}(A, \mathcal{P}(A)) : f$ one-to-one (for $f = f_0$)

$\Rightarrow A \leqslant \mathcal{P}(A)$. (1)

To show that $A \nsim \mathcal{P}(A)$, assume that $A \sim \mathcal{P}(A)$. Then

$A \sim \mathcal{P}(A) \Rightarrow \exists f \in \text{Map}(A, \mathcal{P}(A)) : f$ bijection

Choose an $f \in \text{Map}(A, \mathcal{P}(A))$ such that ~~for~~ $f : A \to \mathcal{P}(A)$ is a bijection. We define a set of "bad elements"

$B = \{x \in A \mid x \notin f(x)\} \subseteq A \Rightarrow B \in \mathcal{P}(A)$.

and note that

f bijection $\Rightarrow f$ onto $\Rightarrow f(A) = \mathcal{P}(A) \Rightarrow \mathcal{P}(A) \subseteq f(A)$

$\Rightarrow \forall y \in \mathcal{P}(A) : y \in f(A) \Rightarrow$

$\Rightarrow \forall y \in \mathcal{P}(A) : \exists x \in A : f(x) = y$

Let $y = B$ and choose a $b \in A$ such that $f(b) = B$.

We distinguish between the following cases.

Case 1 : Assume that $b \in B$. Then

$b \in B \Rightarrow b \in \{x \in A \mid x \notin f(x)\} \Rightarrow$

$\Rightarrow b \in A \land b \notin f(b) \Rightarrow b \notin f(b) \Rightarrow b \notin B$

which is a contradiction, therefore case 1 does not materialize.

Case 2 : Assume that $b \notin B$. We also now, by definition, that $b \in A$, and therefore:

$$\begin{cases} b \in A \\ b \notin B \end{cases} \Rightarrow \begin{cases} b \in A \\ b \notin f(b) \end{cases} \Rightarrow b \in \{x \in A \mid x \notin f(x)\} \Rightarrow$$
$$\Rightarrow b \in B$$

which is also a contradiction.

Since none of the possible cases materialize, it follows that

$A \not\sim \mathcal{P}(A)$. (2)

From Eq.(1) and Eq.(2):

$$\begin{cases} A \not\sim \mathcal{P}(A) \\ A \leqslant \mathcal{P}(A) \end{cases} \Rightarrow A < \mathcal{P}(A).$$

② → Schroeder – Bernstein theorem

$$\boxed{\begin{array}{l}\underline{\text{Thm}}: \text{ Let } A, B \text{ be two sets. Then} \\ \qquad A \leqslant B \wedge B \leqslant A \Rightarrow A \sim B\end{array}}$$

Proof

Assume that $A \leqslant B$ and $B \leqslant A$. Then

$$\begin{cases} A \leqslant B \\ B \leqslant A \end{cases} \Rightarrow \begin{cases} \exists f \in \mathrm{Map}(A,B): f \text{ one-to-one} \\ \exists g \in \mathrm{Map}(B,A): g \text{ one-to-one} \end{cases} \quad (1)$$

Choose $f \in \mathrm{Map}(A,B)$ and $g \in \mathrm{Map}(B,A)$ such that f, g are one-to-one.

Define $C_0 = A - g(B)$ and distinguish between the following two cases.

Case 1: Assume that $C_0 = \emptyset$. By construction, we have

$g \in \mathrm{Map}(B,A) \Rightarrow g(B) \subseteq A$.

We will show that $A \subseteq g(B)$. (2)

Let $x \in A$ be given. To show that $x \in g(B)$, assume that $x \notin g(B)$ in order to derive a contradiction. It follows that

$$\begin{cases} x \in A \\ x \notin g(B) \end{cases} \Rightarrow x \in A - g(B) \Rightarrow x \in C_0 \Rightarrow x \in \emptyset$$

which is a contradiction. We conclude that $\underline{x \in g(B)}$

We have thus shown that

$\forall x \in A: x \in g(B) \Rightarrow A \subseteq g(B)$ (3)

From Eq. (1), Eq. (2), Eq. (3) we conclude that:

$$\left\{\begin{aligned} &A \subseteq g(B) \wedge g(B) \subseteq A \\ &g \text{ one-to-one} \end{aligned}\right. \Rightarrow \left\{\begin{aligned} &g(B) = A \\ &g \text{ one-to-one} \end{aligned}\right. \Rightarrow$$

$$\Rightarrow \left\{\begin{aligned} &g \text{ onto} \\ &g \text{ one-to-one} \end{aligned}\right. \Rightarrow g : B \to A \text{ bijection}$$

$$\Rightarrow B \sim A \Rightarrow \underline{A \sim B}.$$

<u>Case 2</u> : Assume that $C_0 \neq \emptyset$. Then we define by recursion

$$\forall n \in \mathbb{N} : C_{n+1} = g(f(C_n)) = g(\{f(x) \mid x \in C_n\}) =$$
$$= \{g(f(x)) \mid x \in C_n\}$$

We construct the needed bijection $h : A \to B$ by the following definition:

$$\forall x \in A : h(x) = \begin{cases} f(x) & , \text{ if } \exists n \in \mathbb{N} : x \in C_n \\ g^{-1}(x) & , \text{ if } \forall n \in \mathbb{N} : x \notin C_n \end{cases}$$

Since we do not know if g is a bijection, we need to prove that $A - \bigcup_{n \in \mathbb{N}} C_n \subseteq g(B)$ to ensure that $g^{-1}(x)$ has a unique evaluation.

To show the claim, let $x \in A - \bigcup_{n \in \mathbb{N}} C_n$ be given. Then:

$$x \in A - \bigcup_{n \in \mathbb{N}} C_n \Rightarrow x \in A \wedge x \notin \bigcup_{n \in \mathbb{N}} C_n \Rightarrow x \notin \bigcup_{n \in \mathbb{N}} C_n \Rightarrow$$

$$\Rightarrow \overline{\exists n \in \mathbb{N} : x \in C_n} \Rightarrow$$

$$\Rightarrow \forall n \in \mathbb{N} : x \notin C_n \Rightarrow x \notin C_0.$$

To show that $x \in g(B)$, assume that $x \notin g(B)$. Then

$$\left\{\begin{aligned} &x \in A \\ &x \notin g(B) \end{aligned}\right. \Rightarrow x \in A - g(B) \Rightarrow x \in C_0$$

which is a contradiction, since we previously showed that $x \notin C_0$.

We conclude that

$$\forall x \in A - \bigcup_{n\in\mathbb{N}} C_n : x \in g(B) \Rightarrow A - \bigcup_{n\in\mathbb{N}} C_n \subseteq g(B)$$

which proves the claim.

• We will show that <u>h is one-to-one.</u>

Let $\underline{x_1, x_2 \in A}$ be given and assume that $\underline{h(x_1) = h(x_2)}$.

We distinguish between the following subcases.

<u>Case A</u>: Assume that $\begin{cases} \exists n \in \mathbb{N} : x_1 \in C_n \\ \exists n \in \mathbb{N} : x_2 \in C_n \end{cases}$

Then $h(x_1) = h(x_2) \Rightarrow f(x_1) = f(x_2)$ [definition of h]

$\Rightarrow \underline{x_1 = x_2}$ [f one-to-one]

<u>Case B</u>: Assume that $\begin{cases} \forall n \in \mathbb{N} : x_1 \notin C_n \\ \forall n \in \mathbb{N} : x_2 \notin C_n \end{cases}$. Then,

$$h(x_1) = h(x_2) \Rightarrow g^{-1}(x_1) = g^{-2}(x_2) \Rightarrow \text{[definition of h]}$$

$$\Rightarrow g(g^{-1}(x_1)) = g(g^{-1}(x_2)) \Rightarrow$$

$$\Rightarrow \underline{x_1 = x_2}.$$

<u>Case C</u>: Assume that $\begin{cases} \exists n \in \mathbb{N} : x_1 \in C_n \\ \forall n \in \mathbb{N} : x_2 \notin C_n \end{cases}$

Choose $n_0 \in \mathbb{N}$ such that $x_1 \in C_{n_0}$. We note that

$$\begin{cases} x_2 \in A \\ \forall n \in \mathbb{N} : x_2 \notin C_n \end{cases} \Rightarrow x_2 \in A - \bigcup_{n\in\mathbb{N}} C_n \Rightarrow g^{-1}(x_2) \text{ is defined}$$

and therefore:

$$\begin{aligned} x_2 &= g(g^{-1}(x_2)) \\ &= g(h(x_2)) && \text{[Definition of } h(x) \text{ - 2nd case]} \\ &= g(h(x_1)) && \text{[Hypothesis } h(x_1) = h(x_2)\text{]} \\ &= g(f(x_1)) && \text{[Definition of } h(x) \text{ - 1st case]} \end{aligned}$$

$\Rightarrow \exists x \in C_{n_0} : g(f(x)) = x_2 \Rightarrow$

$\Rightarrow x_2 \in \{g(f(x)) \mid x \in C_{n_0}\}$

$\Rightarrow x_2 \in g(f(C_{n_0}))$

$\Rightarrow x_2 \in C_{n_0+1}$

This is a contradiction because

$(\forall n \in \mathbb{N} : x_2 \notin C_n) \Rightarrow x_2 \notin C_{n_0+1}$

therefore Case C does not materialize. In all of the above cases we conclude that $x_1 = x_2$ and therefore:

$\forall x_1, x_2 \in A : (h(x_1) = h(x_2) \Rightarrow x_1 = x_2)$

$\Rightarrow h$ one-to-one. (4)

•2 We will show that $h(A) = B$.

By definition, we have $h(A) \subseteq B$, so it is sufficient to show that $\forall y \in B : y \in h(A)$. Let $\underline{y \in B}$ be given. We distinguish between the following cases.

<u>Case 1</u> : Assume that $\exists n \in \mathbb{N} : y \in f(C_n)$.

Choose $n_0 \in \mathbb{N}$ such that $y \in f(C_{n_0})$. Since

$h(C_{n_0}) = \{h(x) \mid x \in C_{n_0}\}$

$= \{f(x) \mid x \in C_{n_0}\}$ [Definition of $h(x)$ – 1st case]

$= f(C_{n_0})$

it follows that

$y \in f(C_{n_0}) \Rightarrow y \in h(C_{n_0})$ [Because $h(C_{n_0}) = f(C_{n_0})$]

$\Rightarrow \underline{y \in h(A)}$ [because $C_{n_0} \subseteq A$]

<u>Case 2</u> : Assume that $\forall n \in \mathbb{N} : y \notin f(C_n)$.

We claim that $\forall n \in \mathbb{N} : g(y) \notin C_n$.

To show the claim, we note that:

$$\forall n \in \mathbb{N} : y \notin f(C_n) \Rightarrow \forall n \in \mathbb{N} : g(y) \notin g(f(C_n))$$
$$\Rightarrow \forall n \in \mathbb{N} : g(y) \notin C_{n+1}$$
$$\Rightarrow \forall n \in \mathbb{N}^* : g(y) \notin C_n \quad (5)$$

For $n=0$, to show that $g(y) \notin C_0$, we will assume that $g(y) \in C_0$ and derive a contradiction. Then:

$$g(y) \in C_0 \Rightarrow g(y) \in A - g(B)$$
$$\Rightarrow g(y) \in A \wedge g(y) \notin g(B)$$
$$\Rightarrow g(y) \notin g(B)$$

which is a contradiction because

$$y \in B \Rightarrow g(y) \in g(B)$$

It follows that $g(y) \notin C_0 \quad (6)$

From Eq.(5) and Eq.(6) we prove the claim. It follows that

$$h(g(y)) = g^{-1}(g(y)) \quad [\text{because } \forall n \in \mathbb{N} : g(y) \notin C_n]$$
$$= y \Rightarrow$$
$$\Rightarrow \exists x \in A : y = h(x) \quad (\text{for } x = g(y))$$
$$\Rightarrow \underline{y \in h(A)}$$

From the above argument we have:

$$\begin{cases} h(A) \subseteq B \\ \forall y \in B : y \in h(A) \end{cases} \Rightarrow \begin{cases} h(A) \subseteq B \\ B \subseteq h(A) \end{cases} \Rightarrow h(A) = B \Rightarrow \underline{h \text{ onto}} \quad (7)$$

From Eq.(4) and Eq.(7):

$$\begin{cases} h \text{ one-to-one} \\ h \text{ onto} \end{cases} \Rightarrow h : A \to B \text{ bijection}$$

$$\Rightarrow A \sim B \qquad \square$$

③ → Uncountability of $\mathbb{R}$

The Schroeder–Bernstein theorem can be used to derive the following characterization for the cardinality of $\mathbb{R}$:

$$\boxed{\mathbb{R} \sim \mathcal{P}(\mathbb{N})}$$

Once this result is established, we can use Cantor's theorem to argue that:

$$\begin{cases} \mathbb{R} \sim \mathcal{P}(\mathbb{N}) \\ \mathcal{P}(\mathbb{N}) > \mathbb{N} \end{cases} \Rightarrow \mathbb{R} > \mathbb{N} \Rightarrow \mathbb{R} \text{ uncountable}$$

The argument below uses the previous result that $\mathbb{R} \sim (0,1)$.

► Proof of $\mathbb{R} \sim \mathcal{P}(\mathbb{N})$

It is sufficient to show that $\mathcal{P}(\mathbb{N}) \leqslant \mathbb{R} \wedge \mathbb{R} \leqslant \mathcal{P}(\mathbb{N})$.

•$_1$ Proof of $\mathcal{P}(\mathbb{N}) \leqslant \mathbb{R}$.

We define a mapping $f : \mathcal{P}(\mathbb{N}) \to [0,1]$ as follows. Given $X \in \mathcal{P}(\mathbb{N})$ we define $f(X)$ via the expansion

$$f(X) = (0.a_0 a_1 a_2 \cdots)_{10} =$$

$$= \sum_{n=0}^{+\infty} a_n 10^{-n-1}$$

with

$$\forall n \in \mathbb{N} : \ a_n = \begin{cases} 1 & , \text{if } n \in X \\ 0 & , \text{if } n \notin X \end{cases}$$

To show that f is one-to-one, it is necessary to define it using a base representation that is greater than binary (i.e. base 2) while restricting the digits used to 0 and 1. This way, a number that terminates with an infinite sequence of 1s (e.g. 0.101111...) will not have an second alternate representation, as it would have in the binary system. We may therefore now argue as follows: Let $X_1, X_2 \in \mathcal{P}(\mathbb{N})$ be given and assume that $f(X_1) = f(X_2)$. Define the sequences (a_n) and (b_n) via the decimal representations:

$$f(X_1) = 0.a_0a_1a_2\dots = \sum_{n=0}^{+\infty} a_n \cdot 10^{-n-1}$$

$$f(X_2) = 0.b_0b_1b_2\dots = \sum_{n=0}^{+\infty} b_n \cdot 10^{-n-1}$$

We note that

$$f(X_1) = f(X_2) \Rightarrow 0.a_0a_1a_2\dots = 0.b_0b_1b_2\dots \Rightarrow$$
$$\Rightarrow \forall n \in \mathbb{N} : a_n = b_n.$$

We use this result to show that

$$\begin{aligned} n \in X_1 &\Leftrightarrow a_n = 1 && [\text{definition of } a_n] \\ &\Leftrightarrow b_n = 1 && [\text{via } a_n = b_n] \\ &\Leftrightarrow n \in X_2 && [\text{definition of } b_n] \end{aligned}$$

It follows that $\underline{X_1 = X_2}$. We have thus shown that

$$\forall X_1, X_2 \in \mathcal{P}(\mathbb{N}) : (f(X_1) = f(X_2) \Rightarrow X_1 = X_2)$$
$$\Rightarrow f \text{ one-to-one} \Rightarrow \mathcal{P}(\mathbb{N}) \leq [0,1]$$

We also have: $[0,1] \subseteq \mathbb{R} \Rightarrow [0,1] \leq \mathbb{R}$

and therefore

$$\begin{cases} \mathcal{P}(\mathbb{N}) \leqslant [0,1] \\ [0,1] \leqslant \mathbb{R} \end{cases} \Rightarrow \underline{\mathcal{P}(\mathbb{N}) \leqslant \mathbb{R}.} \qquad (1)$$

•2 Proof of $\mathbb{R} \leqslant \mathcal{P}(\mathbb{N})$.

We define a mapping $g : [0,1] \to \mathcal{P}(\mathbb{N})$ as follows.

Let $x \in [0,1]$ be given with binary representation

$$x = (0.a_0 a_1 a_2 \cdots)_2 = \sum_{n=0}^{+\infty} a_n 2^{-n-1}$$

To ensure uniqueness, we do not allow terminating the binary representation of x with an infinite sequence of 1s except for $x = 1$ (represented as $x = (0.1111\cdots)_2$)

Define $g(x) = \{n \in \mathbb{N} \mid a_n = 1\}$

Let $\underline{x_1, x_2 \in [0,1]}$ be given and assume that $\underline{g(x_1) = g(x_2)}$.

Define the sequences (a_n) and (b_n) via the unique binary representations (as explained above)

$$x_1 = (0.a_0 a_1 a_2 \cdots)_2$$
$$x_2 = (0.b_0 b_1 b_2 \cdots)_2$$

To show that $x_1 = x_2$, we assume that $x_1 \neq x_2$ and derive a contradiction. Then, we have

$$x_1 \neq x_2 \Rightarrow \underline{(0.a_0 a_1 a_2 \cdots)_2} \neq (0.b_0 b_1 b_2 \cdots)_2$$
$$\Rightarrow \overline{\forall n \in \mathbb{N} : a_n = b_n}$$
$$\Rightarrow \exists n \in \mathbb{N} : a_n \neq b_n$$

Choose $n_0 \in \mathbb{N}$ such that $a_{n_0} \neq b_{n_0}$. It follows that

$$a_{n_0} \neq b_{n_0} \Rightarrow \begin{cases} a_{n_0} = 1 \\ b_{n_0} = 0 \end{cases} \veebar \begin{cases} a_{n_0} = 0 \\ b_{n_0} = 1 \end{cases} \Rightarrow$$

$$\Rightarrow \begin{cases} n_0 \in g(x_1) \\ n_0 \notin g(x_2) \end{cases} \veebar \begin{cases} n_0 \notin g(x_1) \\ n_0 \in g(x_2) \end{cases} \Rightarrow$$

$$\Rightarrow \underline{(\exists n \in g(x_1) : n \notin g(x_2))} \vee \underline{(\exists n \in g(x_2) : n \notin g(x_1))}$$

$$\Rightarrow \overline{(\forall n \in g(x_1) : n \in g(x_2))} \veebar \overline{(\forall n \in g(x_2) : n \in g(x_1))}$$

$$\Rightarrow g(x_1) \not\subseteq g(x_2) \veebar g(x_2) \not\subseteq g(x_1)$$

which is a contradiction because

$$g(x_1) = g(x_2) \Rightarrow \begin{cases} g(x_1) \subseteq g(x_2) \\ g(x_2) \subseteq g(x_1) \end{cases}$$

We have thus shown that $\underline{x_1 = x_2}$

From the above argument we have show that

$$\forall x_1, x_2 \in [0,1] : (g(x_1) = g(x_2) \Rightarrow x_1 = x_2)$$

$$\Rightarrow g \text{ one-to-one} \Rightarrow [0,1] \leqslant \mathcal{P}(\mathbb{N}).$$

and therefore:

$$\begin{aligned} \mathbb{R} &\sim (0,1) && [\text{previous result}] \\ &\leqslant [0,1] && [\text{via } (0,1) \subseteq [0,1]] \\ &\leqslant \mathcal{P}(\mathbb{N}) && [\text{above proof}] \end{aligned}$$

$$\Rightarrow \underline{\mathbb{R} \leqslant \mathcal{P}(\mathbb{N})} \qquad (2)$$

From Eq. (1) and Eq. (2) via the Schroeder-Bernstein theorem, it follows that

$$\begin{cases} \mathcal{P}(\mathbb{N}) \leqslant \mathbb{R} \\ \mathbb{R} \leqslant \mathcal{P}(\mathbb{N}) \end{cases} \Rightarrow \mathbb{R} \sim \mathcal{P}(\mathbb{N}). \qquad \square$$

EXERCISES

(13) Study the proofs for

a) The Cantor theorem

b) The Schroder - Bernstein theorem

c) The statement $\mathbb{R} \sim \mathcal{P}(\mathbb{N})$.

(14) Use Exercise 9 and the previous results that $\mathbb{Q} \sim \mathbb{N}$ and $\mathbb{R} \sim \mathcal{P}(\mathbb{N})$ to show that $\mathbb{R} - \mathbb{Q}$ (the set of irrational numbers) is uncountable.
(Hint: Use proof by contradiction)

(15) Show that, given 3 sets A, B, C, we have:

a) $A \preccurlyeq B \wedge B \preccurlyeq C \Rightarrow A \preccurlyeq C$

b) $(A \preccurlyeq B \preccurlyeq C \wedge A \sim C) \Rightarrow (B \sim C \wedge A \sim B)$

c) $A \sim B \wedge B \subseteq C \Rightarrow A \preccurlyeq C$.

(16) Consider the sets

$$\mathbb{R}_+^* = \{x \in \mathbb{R} \mid x > 0\}$$
$$\mathbb{R}_-^* = \{x \in \mathbb{R} \mid x < 0\}$$

Use the Schroder - Bernstein theorem to show that $\mathbb{R} \sim \mathbb{R}_+^*$ and $\mathbb{R} \sim \mathbb{R}_-^*$

(Hint: The needed one-to-one mappings can be constructed using the exponential function)

(Another hint: It is sufficient to show $\mathbb{R}_+^* \succcurlyeq \mathbb{R}$ and $\mathbb{R}_-^* \succcurlyeq \mathbb{R}$).

(17) Use Exercise 16 to show that given two sets A, B we have:

$A \sim \mathbb{R} \wedge B \sim \mathbb{R} \Rightarrow A \cup B \sim \mathbb{R}$.

(Hint: Distinguish between the following cases. For case 1 assume that $A \cap B = \emptyset$. For Case 2 assume that $A \cap B \neq \emptyset$. Define $B_1 = B - A$, show that $A \cup B = A \cup B_1$ and use Case 1 and the Schroeder–Bernstein theorem to show that $A \cup B_1 \sim \mathbb{R}$).

(18) Use the Schoeder–Bernstein theorem to show that $\mathbb{R} \times \mathbb{R} \sim \mathbb{R}$.

(Hint: Use binary or decimal representations to show that $[0,1] \times [0,1] \sim [0,1]$ by defining one-to-one mappings $f: [0,1] \times [0,1] \to [0,1]$ and $g: [0,1] \to [0,1] \times [0,1]$. Then uplift this result to the statement $\mathbb{R} \times \mathbb{R} \sim \mathbb{R}$).

Cardinal numbers

- To introduce the concept of cardinality and cardinal numbers, we note first that

$$\forall n, m \in \mathbb{N}^* : \left(\begin{cases} A \sim [n] \\ A \sim [m] \end{cases} \Rightarrow n = m \right)$$

Thus, for finite sets A, we can define a unique integer $|A|$ such that $A \sim [|A|]$.

- $|A|$ is the number of elements in A and we call it the <u>cardinality</u> of A.
- Cantor proposed introducing "transfinite cardinal numbers" to denote the cardinality $|A|$ of infinite sets. A key requirement of this cardinal number arithmetic is that it should satisfy:

$$A \sim B \Leftrightarrow |A| = |B|$$
$$A \prec B \Leftrightarrow |A| \leq |B|$$
$$A \precneqq B \Leftrightarrow |A| < |B|$$

The Schroeder-Bernstein theorem ensures self-consistent behaviour of inequalities in cardinal arithmetic.

- Since $\mathbb{N} \sim \mathbb{Z} \sim \mathbb{Q}$, Cantor introduced the cardinal number $\aleph_0$ to represent the cardinality of countably infinite sets. Consequently, we may write

$$|\mathbb{N}| = |\mathbb{Z}| = |\mathbb{Q}| = \aleph_0$$

- <u>Aleph sequence</u>: Cantor proposed defining a sequence of cardinalities $\aleph_1, \aleph_2, \aleph_3, \ldots$ as follows.

Let $\mathbb{V}$ be the set of <u>all sets</u> that exist. We define:

$|A| = \aleph_1 \Leftrightarrow \forall B_1 \in \mathbb{V}: \overline{\mathbb{N} < B_1 < A}$

$|A| = \aleph_2 \Leftrightarrow \forall B_1, B_2 \in \mathbb{V}: \overline{\mathbb{N} < B_1 < B_2 < A}$

$|A| = \aleph_3 \Leftrightarrow \forall B_1, B_2, B_3 \in \mathbb{V}: \overline{\mathbb{N} < B_1 < B_2 < B_3 < A}$

etc.

- <u>Beth sequence</u>: Another sequence of cardinal numbers is the beth sequence. It is based on the Cantor theorem that tells us that $A < \mathcal{P}(A)$. The beth sequence is defined as follows:

  $\beth_0 = \aleph_0 = |\mathbb{N}| = |\mathbb{Z}| = |\mathbb{Q}|$

  $\beth_1 = |\mathcal{P}(\mathbb{N})| = |\mathbb{R}|$

  $\beth_2 = |\mathcal{P}(\mathcal{P}(\mathbb{N}))|$

  $\beth_3 = |\mathcal{P}(\mathcal{P}(\mathcal{P}(\mathbb{N})))|$

  etc.

- <u>Continuum hypothesis</u>: With the above definitions, Cantor posed the question of whether the aleph and beth sequences coincide. This leads to two questions:

a) Continuum Hypothesis: The claim that $\beth_1 = \aleph_1$.

b) General Continuum Hypothesis: The claim that $\beth_a = \aleph_a$ for all a.

It was later found that these hypotheses are undecidable, i.e. it can neither be proved true or false. The underlying problem is that for the case of infinite sets, the mechanism for generating the powerset $\mathcal{P}(A)$ of an infinite set A is not precisely given. As a result, we have no way of deducing the correct "size" of $\mathcal{P}(\mathbb{N})$, $\mathcal{P}(\mathcal{P}(\mathbb{N}))$, etc.

DST5: Basic graphs

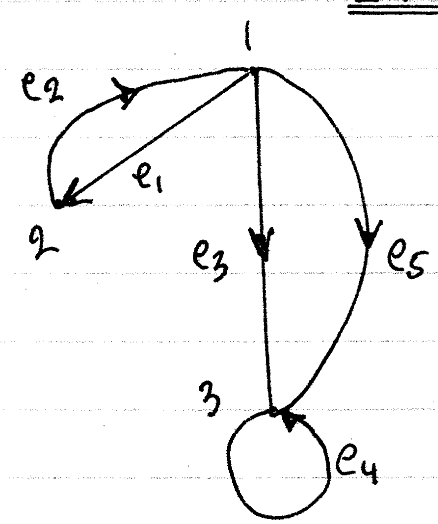
BASIC GRAPH THEORY

Directed Graphs

Def: A directed graph G is an object that consists of
a) A set of vertices $V(G)$
b) A set of edges $E(G)$
c) An incidence mapping $\psi_G : E(G) \to V(G) \times V(G)$ that maps every edge $e \in E(G)$ to a unique pair of vertices $(u_1, u_2) \in V(G) \times V(G)$.

- Graphical representation: Each vertex $u \in V(G)$ is represented as a point on a plane. Each edge $e \in E(G)$ with $\psi_G(e) = (u_1, u_2)$ is represented as an arrow from u_1 to u_2. If $\psi_G(e) = (u,u)$ then the edge is a loop and is represented by an arrow that begins at u and loops back to terminate at u.

EXAMPLE



$V(G) = \{1,2,3\}$
$E(G) = \{e_1, e_2, e_3, e_4, e_5\}$
$\psi_G(e_1) = (1,2)$ $\psi_G(e_4) = (3,3)$
$\psi_G(e_2) = (2,1)$ $\psi_G(e_5) = (1,3)$
$\psi_G(e_3) = (1,3)$

↳ In this example note that $e_3 \neq e_5$ but nonetheless $\psi_G(e_3) = \psi_G(e_5)$

↳ ψ_G can be also represented as a set $\psi_G \subseteq E(G) \times (V(G) \times V(G))$ with

$$\psi_G = \{(e_1,(1,2)),(e_2,(2,1)),(e_3,(1,3)),(e_4,(3,3)),(e_5,(1,3))\}.$$

▸ Successor vertices

Def: Let G be a graph and let $u_1, u_2 \in V(G)$ be two vertices. We say that

u_2 is successor of $u_1 \Leftrightarrow \exists e \in E(G) : \psi_G(e) = (u_1, u_2)$

• notation: The set of all successors of a vertex $u \in V(G)$ is denoted as:

$$succ(u) = \{w \in V(G) \mid \exists e \in E(G) : \psi_G(e) = (u,w)\}$$

EXAMPLE

In the previous example:

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

$succ(2) = \{1\}$

$succ(3) = \{3\}$

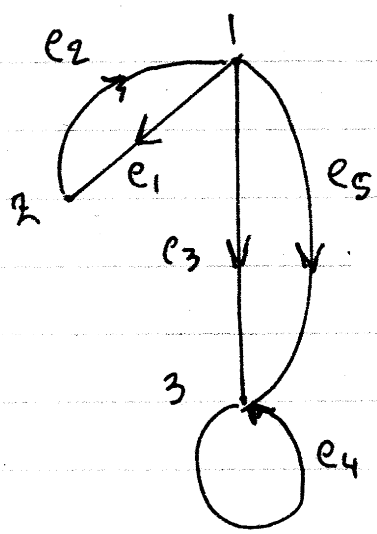
▸ Adjacency matrix

Let G be a graph with $|V(G)| = n$ (i.e. with n vertices labeled as $V(G) = \{u_1, u_2, u_3, \dots, u_n\}$). The adjacency matrix $A(G) \in M_n(\mathbb{R})$ is an $n \times n$ square matrix such that

$$\forall a, b \in [n] : [A(G)]_{ab} = |\{e \in E(G) \mid \psi_G(e) = (u_a, u_b)\}|$$

EXAMPLE

For the graph in the previous example:



$$A(G) = \begin{bmatrix} A_{11} & A_{12} & A_{13} \\ A_{21} & A_{22} & A_{23} \\ A_{31} & A_{32} & A_{33} \end{bmatrix}$$

$$= \begin{bmatrix} 0 & 1 & 2 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

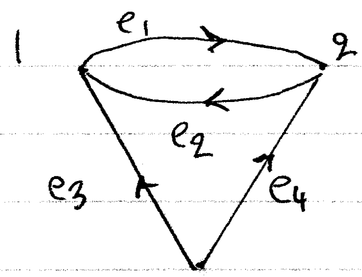
↳ Adjacency matrices make it easy to define the concept of a simple graph. We say that a graph G is <u>simple</u> if and only if it contains no loops and no double or multiple edges. A rigorous definition is:

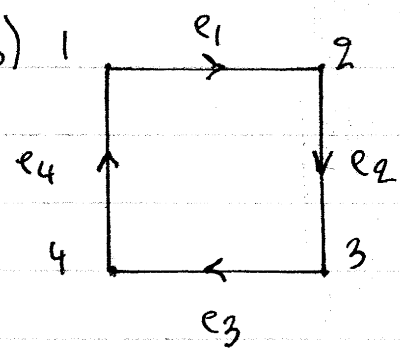
$$\boxed{G \text{ simple} \Leftrightarrow \begin{cases} \forall a,b \in [|V(G)|] : A_{ab}(G) \in \{0,1\} \\ \forall a \in [|V(G)|] : A_{aa}(G) = 0 \end{cases}}$$

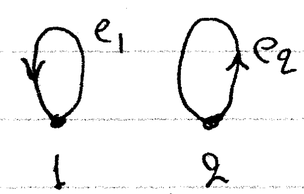
The first condition rules out multiple edges and the second condition rules out loops.

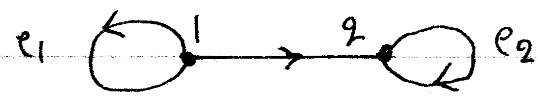
EXERCISES

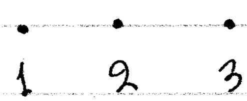
① Define the sets $V(G)$, $E(G)$, the mapping ψ_G, and the adjacency matrix $A(G)$ for the directed graphs shown below:

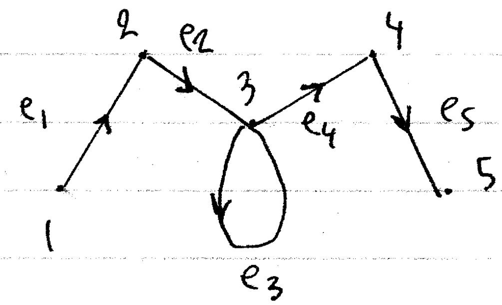
a) 

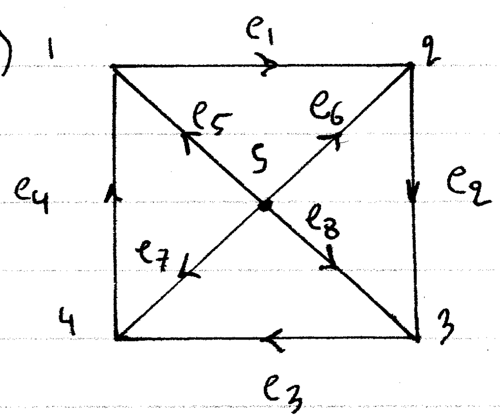
b) 

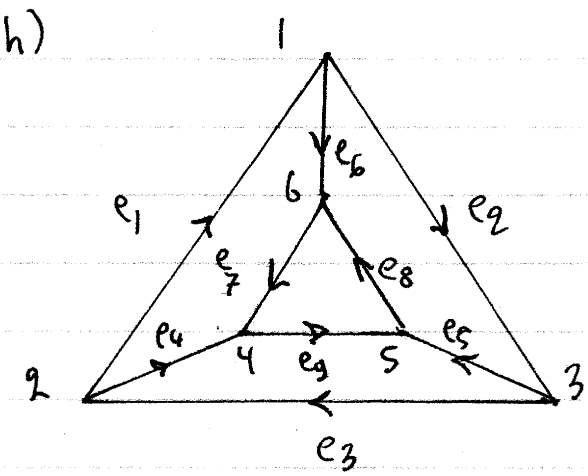
c) 

d) 

e) 

f) 

g) 

h) 

(2) Identify which of the above directed graphs ~~is~~ are simple.

(3) Draw the directed graphs G given by the following set theory definitions: and define the corresponding $A(G)$

a) $V(G) = \{1,2,3,4\}$

$E(G) = \{e_1, e_2, e_3, e_4, e_5, e_6\}$

$\psi_G(e_1) = (1,3)$ $\quad$ $\psi_G(e_4) = (2,4)$

$\psi_G(e_2) = (2,2)$ $\quad$ $\psi_G(e_5) = (4,3)$

$\psi_G(e_3) = (3,1)$ $\quad$ $\psi_G(e_6) = (1,1)$

b) $V(G) = \{2,3\}$, $E(G) = \emptyset$, $\psi_G = \emptyset$

c) $V(G) = \{1\}$, $E(G) = \{e_1\}$, $\psi_G(e_1) = (1,1)$

d) $V(G) = \{1,2,3\}$

$E(G) = \{e_1, e_2, e_3, e_4, e_5\}$

$\psi_G = \{(e_1, (1,1)), (e_2, (1,3)), (e_3, (2,3)), (e_4, (2,3)), (e_5, (3,3))\}$

e) $V(G) = \{1,2,3,4\}$

$E(G) = \{a, b, c, d, e, f, g, h\}$

$\psi_G(a) = (1,1)$ $\quad$ $\psi_G(e) = (3,3)$

$\psi_G(b) = (1,2)$ $\quad$ $\psi_G(f) = (3,4)$

$\psi_G(c) = (2,2)$ $\quad$ $\psi_G(g) = (4,4)$

$\psi_G(d) = (2,3)$ $\quad$ $\psi_G(h) = (4,1)$

f) $V(G) = \{1, 2\}$

$E(G) = \{e_1, e_2, e_3\}$

$\psi_G = \{(e_1, (1,1)), (e_2, (1,2)), (e_3, (1,1))\}$

▼ Walks

Def: Let G be a directed graph. A walk w is an n-tuple of the form

$$w = (u_0, e_1, u_1, e_2, u_2, \dots, e_n, u_n)$$

of alternating edges/vertices such that

$$\begin{cases} \forall a \in [n] : e_a \in E(G) \\ \forall a \in \{0\} \cup [n] : u_a \in V(G) \\ \forall a \in [n] : \psi_G(e_a) = (u_{a-1}, u_a) \end{cases}$$

▶ Terminology

$|w| = n \longleftarrow$ length of the walk

$\$(w) = u_0 \longleftarrow$ initial vertex

$t(w) = u_n \longleftarrow$ terminal vertex

$u_a(w) = u_a \longleftarrow$ the a^{th} vertex, counting from 0

$e_a(w) = e_a \longleftarrow$ the a^{th} edge, counting from 1

$W(G) \longleftarrow$ the set of all walks on G.

Def: Let G be a graph and choose two vertices $u, v \in V(G)$. We define

a) The set of all walks that begin with u and terminate at v:

$$W(G|u,v) = \{w \in W(G) \mid \$(w) = u \land t(w) = v\}$$

b) The set of all walks with length n that begin with u and terminate at v:

$$W_n(G|u,v) = \{w \in W(G) \mid \$(w) = u \land t(w) = v \land |w| = n\}$$

▸ Enumeration of walks

The set $W(G|u,v)$ has an infinite number of elements. However, the set $W_n(G|u,v)$ can be enumerated using the adjacency matrix according to the following statement.

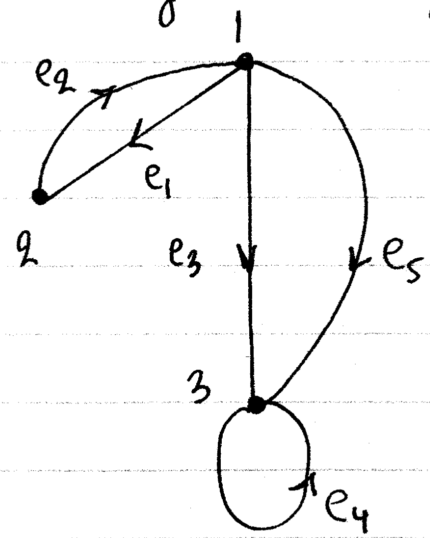
$$\boxed{\begin{array}{l}\underline{\text{Thm}}:\ \text{Let } G \text{ be a graph with } V(G) = \{u_1, u_2, \dots, u_m\} \\ \text{and corresponding adjacency matrix } A(G). \text{ Then} \\ \forall a,b \in [m] : |W_n(G|u_a, u_b)| = [A^n(G)]_{ab}\end{array}}$$

The n^{th} power $A^n(G)$ of the adjacency matrix is defined recursively as follows:

$$\boxed{\begin{array}{l}\forall a,b \in [m] : [A^1(G)]_{ab} = [A(G)]_{ab} \\ \forall a,b \in [m] : \forall k \in \mathbb{N}^* : [A^{k+1}(G)]_{ab} = \displaystyle\sum_{c=1}^{m} [A^k(G)]_{ac} [A(G)]_{cb}\end{array}}$$

EXAMPLE

Use the adjacency matrix to enumerate the walks with length 3 from vertex 1 to 3 for the following directed graph.



Solution

We have

$$A(G) = \begin{bmatrix} 0 & 1 & 2 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \Rightarrow$$

$$\Rightarrow A^2(G) = A(G)\,A(G) = \begin{bmatrix} 0 & 1 & 2 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 0 & 1 & 2 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} =$$

$$= \begin{bmatrix} 0\cdot 0 + 1\cdot 1 + 2\cdot 0 & 0\cdot 1 + 1\cdot 0 + 2\cdot 0 & 0\cdot 2 + 1\cdot 0 + 2\cdot 1 \\ 1\cdot 0 + 0\cdot 1 + 0\cdot 0 & 1\cdot 1 + 0\cdot 0 + 0\cdot 0 & 1\cdot 2 + 0\cdot 0 + 0\cdot 1 \\ 0\cdot 0 + 0\cdot 1 + 1\cdot 0 & 0\cdot 1 + 0\cdot 0 + 1\cdot 0 & 0\cdot 2 + 0\cdot 0 + 1\cdot 1 \end{bmatrix}$$

$$= \begin{bmatrix} 1 & 0 & 2 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{bmatrix} \Rightarrow$$

$$\Rightarrow A^3(G) = A^2(G)\,A(G) = \begin{bmatrix} 1 & 0 & 2 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 0 & 1 & 2 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} =$$

$$= \begin{bmatrix} 1\cdot0+0\cdot1+2\cdot0 & 1\cdot1+0\cdot0+2\cdot0 & 1\cdot2+0\cdot0+2\cdot1 \\ 0\cdot0+1\cdot1+2\cdot0 & 0\cdot1+1\cdot0+2\cdot0 & 0\cdot2+1\cdot0+2\cdot1 \\ 0\cdot0+0\cdot1+1\cdot0 & 0\cdot1+0\cdot0+1\cdot0 & 0\cdot2+0\cdot0+1\cdot1 \end{bmatrix}$$

$$= \begin{bmatrix} 0 & 1 & 4 \\ 1 & 0 & 2 \\ 0 & 0 & 1 \end{bmatrix} \Rightarrow$$

$$\Rightarrow |W_3(G|1,3)| = [A^3(G)]_{13} = 4$$

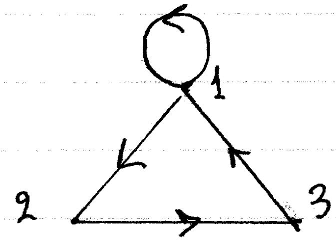
<u>Remark</u>: By inspection, the 4 walks from vertex 1 to vertex 3 with length 3 can be easily identified as follows:

$W_1 = (1, e_3, 3, e_4, 3, e_4, 3)$
$W_2 = (1, e_5, 3, e_4, 3, e_4, 3)$
$W_3 = (1, e_1, 2, e_2, 1, e_3, 3)$
$W_4 = (1, e_1, 2, e_2, 1, e_5, 3)$

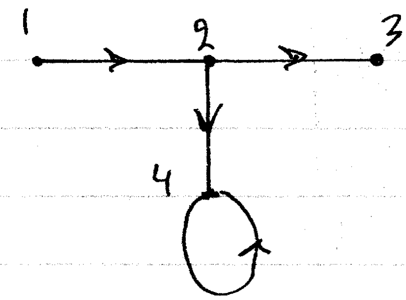
EXERCISES

(4) Enumerate the total number of open and closed walks of length 3 for the graphs shown below, using the adjacency matrix

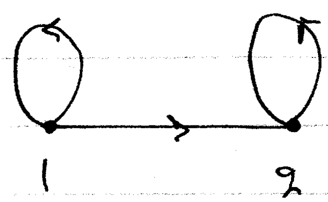
a)



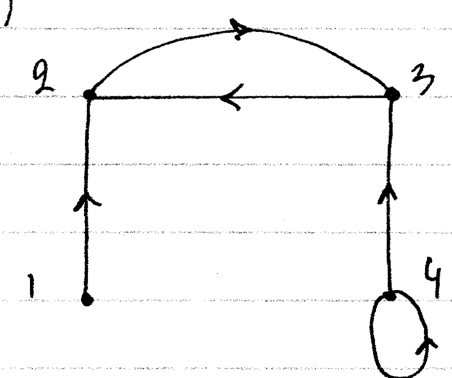
b)



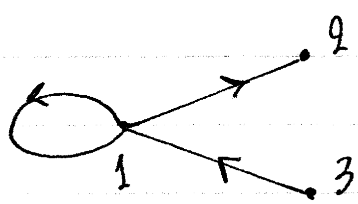
c)



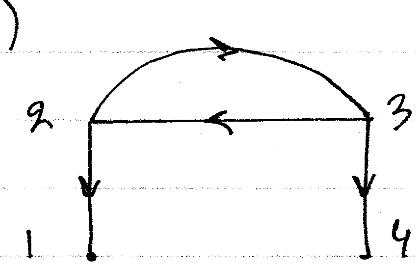
d)



e)



f)



DST6: Formal Languages and Automata

FORMAL LANGUAGES AND AUTOMATA

Languages

Intuitively, a language is defined as a set of strings. A string is defined as a finite ordered set of symbols that originate from a finite set Σ, which we call the alphabet. To provide a rigorous definition we recall that

$$\mathbb{N}^* = \{1, 2, 3, \dots\}$$

and that given a set A, via the Cartesian product, we define

$$A^2 = A \times A = \{(a, b) \mid a, b \in A\}$$
$$A^3 = A \times A \times A = \{(a, b, c) \mid a, b, c \in A\}$$
$$A^4 = A \times A \times A \times A = \{(a, b, c, d) \mid a, b, c, d \in A\}$$

etc.

The n^{th} case is defined as

$$A^n = \{(x_1, x_2, \dots, x_n) \mid \forall k \in [n] : x_k \in A\}$$

For purposes of the definitions below, we also define

$$A^0 = \{\emptyset\}$$

with $\emptyset$ the empty set.

Definition of strings and languages

From the above concepts, we define the notion of language rigorously as follows:

Def: Given a set Σ, we define

a) The star-closure Σ^* (also: Kleene closure) as

$$\Sigma^* = \bigcup_{n\in\mathbb{N}} \Sigma^n = \Sigma^0 \cup \Sigma^1 \cup \Sigma^2 \cup \Sigma^3 \cup \cdots$$

b) The positive closure Σ^+ as:

$$\Sigma^+ = \bigcup_{n\in\mathbb{N}-\{0\}} \Sigma^n = \Sigma^1 \cup \Sigma^2 \cup \Sigma^3 \cup \cdots$$

↳ The corresponding belonging conditions for Σ^* and Σ^+ are given by

$u \in \Sigma^* \Leftrightarrow \exists n \in \mathbb{N} : u \in \Sigma^n$

$u \in \Sigma^+ \Leftrightarrow \exists n \in \mathbb{N}-\{0\} : u \in \Sigma^n$

Def: Let L, Σ be two sets. We say that

L language with alphabet $\Sigma \Leftrightarrow L \subseteq \Sigma^*$
$\Leftrightarrow L \in \mathcal{P}(\Sigma^*)$

notation:

a) Strings are essentially n-tuples but we prefer to denote them as an aggregation of symbols

e.g. for the alphabet $\Sigma = \{a, b, c\}$ we denote

$u = abbc = (a, b, b, c)$

and note that since

$$abbc \in \Sigma^4 \Rightarrow \exists n \in \mathbb{N} : abbc \in \Sigma^n$$
$$\Rightarrow abbc \in \Sigma^*$$

b) We use power notation to represent repeating symbols.

e.g. $ab^2c^3b = abbcccb =$
$$= (a, b, b, c, c, c, b)$$

c) Given a string $u \in \Sigma^*$, the n^{th} symbol of u is represented as u_n.

e.g. For $u = ab^2c = abbc$, we have $u_2 = u_3 = b$ and $u_1 = a$ and $u_4 = c$

<u>Remark</u>: Note that $\Sigma^0 = \{\emptyset\}$, therefore for any alphabet Σ we have $\emptyset \in \Sigma^*$. In the context of formal languages, we define $\lambda = \emptyset$ with λ representing the <u>empty string</u> (or <u>null string</u>).

String properties and operations

> <u>Def</u>: Let Σ be an alphabet and let $u \in \Sigma^*$ be a string. We define:
>
> a) The length $|u|$ of u via:
> $|u| = m \Leftrightarrow u \in \Sigma^m$
>
> b) For any letter $a \in \Sigma$ of the alphabet, $n_a(u)$ is the number of occurences of the letter a in the string u, and we define it formally as:
> $n_a(u) = |\{k \in [|u|] \mid u_k = a\}|$

e.g. Consider $u = a^2 bad^2 b$. Then

$|u| = 2+1+1+2+1 = 7$ $\quad n_b(u) = 1+1 = 2$

$n_a(u) = 2+1 = 3$ $\quad n_d(u) = 2$

<u>Def</u>: (<u>string concatenation</u>)

a) Let Σ be an alphabet and let $u,v \in \Sigma^+$ be two non-null strings. We define the <u>concatenation</u> $uv \in \Sigma^+$ as follows:

$$\forall a \in [|u|+|v|] : (uv)_a = \begin{cases} u_a & , \text{if } 1 \leq a \leq |u| \\ v_{a-|u|} & , \text{if } |u| < a \leq |u|+|v| \end{cases}$$

b) To extend concatenation to strings in Σ^*, we also define:

$$\begin{cases} \forall u \in \Sigma^+ : \lambda u = u\lambda = u \\ \lambda\lambda = \lambda \end{cases}$$

<u>Remark</u>: An immediate consequence of this definition is that

$$\forall u,v \in \Sigma^* : |uv| = |u| + |v|$$

e.g. Consider $\Sigma = \{a,b,c,d\}$ and $u,v \in \Sigma^*$ with $u = ab^2 db$ and $v = bc^2 a$. Then:

$|u| = |ab^2 db| = 1+2+1+1 = 5$

$|v| = |bc^2 a| = 1+2+1 = 4$

$uv = (ab^2 db)(bc^2 a) = ab^2 db^2 c^2 a$

$|uv| = |u| + |v| = 5+4 = 9.$

Def: (String reversal)
Let Σ be an alphabet and let $u \in \Sigma^+$ be a non-null string.
a) We define the reverse string u^R via
$\forall a \in [|u|] : (u^R)_a = u_{|u|+1-a}$
b) We also define: $\lambda^R = \lambda$.

e.g. For $u = ab^2dac$, the reverse string is
$u^R = cadb^2a$

Language operations

Def: (Language concatenation)
Let Σ be an alphabet and consider two languages $L_1, L_2 \in \mathcal{P}(\Sigma^*)$. We define a new language $L_1L_2 \in \mathcal{P}(\Sigma^*)$ (the concatenation of L_1 and L_2) as:
$L_1L_2 = \{uv \mid u \in L_1 \land v \in L_2\}$

Def: (Language concatenation power)
Let Σ be an alphabet and let $L \in \mathcal{P}(\Sigma^*)$ be a language. We define L^n for all $n \in \mathbb{N}$ recursively as follows:

$$\begin{cases} L^0 = \{\lambda\} \\ L^1 = L \\ \forall n \in \mathbb{N}^* : L^{n+1} = LL^n \end{cases}$$

<u>Def:</u> (<u>Star-closure and positive closure of languages</u>)
Let Σ be an alphabet and let $L \in \mathcal{P}(\Sigma^*)$ be a language. We define the star-closure L^* and the positive closure L^+ of L as follows:

$$L^* = \bigcup_{n \in \mathbb{N}} L^n \quad \text{and} \quad L^+ = \bigcup_{n \in \mathbb{N} - \{0\}} L^n$$

EXAMPLES

a) Consider the languages
$L_1 = \{\lambda, ab, ac\}$ and $L_2 = \{b, da\}$
Evaluate L_1L_2, L_1^2, L_2^2.

Solution

$$L_1L_2 = \{\lambda, ab, ac\}\{b, da\} =$$
$$= \{\lambda b, \lambda da, abb, abda, acb, acda\} =$$
$$= \{b, da, ab^2, abda, acb, acda\}.$$
$$L_1^2 = L_1L_1 = \{\lambda, ab, ac\}\{\lambda, ab, ac\} =$$
$$= \{\lambda^2, \lambda ab, \lambda ac, ab\lambda, abab, abac, ac\lambda, acab, acac\}$$
$$= \{\lambda, ab, ac, abab, abac, acab, acac\}$$
$$L_2^2 = L_2L_2 = \{b, da\}\{b, da\} =$$
$$= \{bb, bda, dab, dada\} \quad \square$$

b) Let $L_1 = \{a^2b\}$ and $L_2 = \{ba\}$. Evaluate using set builder notation the languages $L_3 = L_1L_2^* \cup L_1^*L_2$ and $L_1^*L_2^*$.

Solution

Since
$$L_1^* = \{a^2b\}^* = \bigcup_{n\in\mathbb{N}} \{a^2b\}^n = \bigcup_{n\in\mathbb{N}} \{(a^2b)^n\} =$$
$$= \{(a^2b)^n \mid n \in \mathbb{N}\}$$
and

$$L_2^* = \{ba\}^* = \bigcup_{n\in\mathbb{N}} \{ba\}^n = \bigcup_{n\in\mathbb{N}} \{(ba)^n\} =$$

$$= \{(ba)^n \mid n\in\mathbb{N}\}$$

it follows that

$$L_3 = L_1L_2^* \cup L_1^*L_2 = \{a^2b\}\{(ba)^n \mid n\in\mathbb{N}\} \cup \{(a^2b)^n \mid n\in\mathbb{N}\}\{ba\}$$

$$= \{a^2b(ba)^n \mid n\in\mathbb{N}\} \cup \{(a^2b)^n ba \mid n\in\mathbb{N}\}$$

$$= \{a^2b(ba)^n, (a^2b)^n ba \mid n\in\mathbb{N}\}.$$

and

$$L_1^*L_2^* = \{(a^2b)^n \mid n\in\mathbb{N}\}\{(ba)^n \mid n\in\mathbb{N}\} =$$

$$= \{(a^2b)^n (ba)^m \mid n\in\mathbb{N} \land m\in\mathbb{N}\} \quad \square$$

c) Let $L = \{u\in\Sigma^* \mid n_a(u) < n_b(u)\}$ be a language on the alphabet $\Sigma = \{a,b\}$. Show that $L^2 \subseteq L$.

<u>Solution</u>

We note that

$$L^2 = LL = \{u\in\Sigma^* \mid n_a(u) < n_b(u)\}\{v\in\Sigma^* \mid n_a(v) < n_b(v)\}$$

$$= \{uv \mid u\in\Sigma^* \land v\in\Sigma^* \land n_a(u) < n_b(u) \land n_a(v) < n_b(v)\}$$

$$= \{uv \mid u,v\in\Sigma^* \land n_a(u) < n_b(u) \land n_a(v) < n_b(v)\}.$$

Let $w\in L^2$ be given. Then:

$$w\in L^2 \Rightarrow \exists u,v\in\Sigma^* : (n_a(u) < n_b(u) \land n_a(v) < n_b(v))$$

Choose $u,v\in\Sigma^*$ such that $n_a(u) < n_b(u)$ and $n_a(v) < n_b(v)$.

We have:

$$n_a(w) = n_a(uv) = n_a(u) + n_a(v)$$

$$< n_b(u) + n_a(v) \qquad [\text{via } n_a(u) < n_b(u)]$$

$$< n_b(u) + n_b(v) \qquad [\text{via } n_a(v) < n_b(v)]$$

$$= n_b(uv) = n_b(w) \Rightarrow$$

$$\Rightarrow n_a(w) < n_b(w) \Rightarrow w \in L$$

From the above argument we have

$$(\forall w \in L^2 : w \in L) \Rightarrow L^2 \subseteq L.$$

EXERCISES

① Let $L_1 = \{a, b, b^2\}$ and $L_2 = \{ab, a^2\}$ be languages.
a) Evaluate L_1L_2 and L_2L_1
b) Evaluate $L_1^2, L_1^3, L_2^2, L_2^3$.

② Let $L = \{abc, b\}$ be a language. Evaluate L^2, L^3, L^4.

③ Let $L_1 = \{b^2\}$ and $L_2 = \{a^3\}$. Use definition of set by mapping notation to define the following languages and write the corresponding belonging condition
a) L_1^* b) L_2^* c) $L_1^*L_2^*$ d) $(L_1L_2)^*$

④ Let $L_1 = \{ab^2\}$ and $L_2 = \{b^4\}$ and $L_3 = \{a^2\}$. Use definition of set by mapping notation to define the following languages and write the corresponding belonging condition.
a) $L_1L_2^* \cup L_1^*L_2$
b) $L_1^*L_2^*$
c) $(L_2L_3)^*$
d) $L_1^*(L_2^* \cup L_3^*)$
e) $(L_1^* \cup L_2^*)L_3^*$
f) $L_1(L_2L_3)^* \cup (L_1L_2)^*L_3$

⑤ Let $\Sigma = \{a, b\}$ be an alphabet and define
$L = \{u \in \Sigma^* \mid n_a(u) = n_b(u)\}$
Show that $L^* = L$

(Hint: A preliminary step is to show by induction that $\forall n \in \mathbb{N} - \{0,1\}: L^n \subseteq L$.)

⑥ Let $\Sigma = \{a,b\}$ be an alphabet and define
$L_1 = \{u \in \Sigma^* \mid n_a(u) = n_b(u) + 1\}$
$L_2 = \{u \in \Sigma^* \mid n_a(u) = n_b(u) + 2\}$

a) Show that $(\{a\}L_1) \cup (L_1\{a\}) \subseteq L_2$

b) Find a counterexample that disproves the claim $(\{a\}L_1) \cup (L_1\{a\}) = L_2$.

c) Show that $L_1^* = L_1$.

⑦ Let $\Sigma = \{a,b\}$ be an alphabet.

a) Show that Σ^* is countably infinite.

b) Is the set of all languages using alphabet $\Sigma = \{0,1\}$ countable or uncountable?

Grammars

Grammars provide a method for defining languages that can be more powerful than set builder notation. We begin with an alphabet set Σ. In grammar terminology Σ is called a set of terminal symbols. Then, we make the following definitions:

Def: A grammar G is defined as a 4-tuple $G = (V, \Sigma, S, P)$ where

a) V is a set of variables

b) Σ is the alphabet

c) $S \in V$ is the start variable

d) $P \subseteq (V \cup \Sigma)^+ \times (V \cup \Sigma)^*$ is a set of productions

such that

$$\begin{cases} V, \Sigma, P \text{ are finite sets} \\ V \cap \Sigma = \emptyset \\ V \neq \emptyset \wedge \Sigma \neq \emptyset \end{cases}$$

Remark: Production rules describe how the grammar is allowed to transform one string into another. Given the production rule $(x \to y) \in P$ with $x \in (V \cup \Sigma)^+$ and $y \in (V \cup \Sigma)^*$, the grammar is allowed to transform any string of the form $u = axb$ with $a, b \in (V \cup \Sigma)^*$ to $v = ayb$, in which case we write $u \underset{G}{\Rightarrow} v$.

We now give the formal definitions corresponding to the previous remark:

Def: Let $G=(V,\Sigma,S,P)$ be a grammar and let $u,v\in(V\cup\Sigma)^*$. We say that

a) $u \underset{G}{\Rightarrow} v \Leftrightarrow \exists a,b,y\in(V\cup\Sigma)^* : \exists x\in(V\cup\Sigma)^+ :$
$$: \begin{cases} u=axb \land v=ayb \\ (x\to y)\in P \end{cases}$$

b) $\begin{cases} u \overset{1}{\underset{G}{\Rightarrow}} v \Leftrightarrow u \underset{G}{\Rightarrow} v \\ \forall n\in\mathbb{N}^* : \left(u \overset{n+1}{\underset{G}{\Rightarrow}} v \Leftrightarrow \exists w\in(V\cup\Sigma)^* : (u \underset{G}{\Rightarrow} w \land w \overset{n}{\underset{G}{\Rightarrow}} v)\right) \end{cases}$

c) $u \overset{*}{\underset{G}{\Rightarrow}} v \Leftrightarrow \exists n\in\mathbb{N}^* : u \overset{n}{\underset{G}{\Rightarrow}} v$

Remark: $u \underset{G}{\Rightarrow} v$ means that u derives v with the application of exactly one production rule.
$u \overset{n}{\underset{G}{\Rightarrow}} v$ means that u derives v with the application of exactly n production rules. Finally, $u \overset{*}{\underset{G}{\Rightarrow}} v$ means that u derives v with the application of an arbitrary number of production rules.

Def: Let $G=(V,\Sigma,S,P)$ be a grammar. We define the language $\mathcal{L}(G)$ generated by the grammar G via
$$\mathcal{L}(G)=\{u\in\Sigma^* \mid S \overset{*}{\underset{G}{\Rightarrow}} u\}$$

The corresponding belonging condition is:

$$u \in \mathcal{L}(G) \Leftrightarrow u \in \Sigma^* \wedge S \xRightarrow[G]{*} u$$

<u>Remark</u>: It is easy to see that the relation "$\xRightarrow[G]{*}$" is transitive, in the sense that:

$$\boxed{\forall u, v, w \in (V \cup \Sigma)^* : ((u \xRightarrow[G]{*} v \wedge v \xRightarrow[G]{*} w) \Rightarrow (u \xRightarrow[G]{*} v))}$$

An immediate consequence is the following lemma:

$$\boxed{\forall u, v \in \Sigma^* : \left(\begin{cases} u \in \mathcal{L}(G) \\ u \xRightarrow[G]{*} v \end{cases} \Rightarrow v \in \mathcal{L}(G) \right)}$$

<u>notation</u>

If a grammar contains production rules of the form $x \to y_1$ and $x \to y_2$, we can rewrite them in condensed form as $x \to y_1 | y_2$. In general, given production rules $x \to y_1, x \to y_2, \ldots, x \to y_n$, we can rewrite them as $x \to y_1 | y_2 | \cdots | y_n$.

EXAMPLE

Consider the grammar $G = (V, \Sigma, S, P)$ with $V = \{S, A\}$ and $\Sigma = \{a, b\}$ and production rules

$$S \longrightarrow Ab$$
$$A \longrightarrow aAb$$
$$A \longrightarrow \lambda$$

Show that

a) $aabbb \in \mathcal{L}(G)$

b) $\mathcal{L}(G) = \{a^n b^{n+1} \mid n \in \mathbb{N}\}$.

<u>Solution</u>

a) Since:

$$\begin{aligned} S &\underset{G}{\Rightarrow} Ab && [\text{via } S \to Ab] \\ &\underset{G}{\Rightarrow} aAbb && [\text{via } A \to aAb] \\ &\underset{G}{\Rightarrow} aaAbbb && [\text{via } A \to aAb] \\ &\underset{G}{\Rightarrow} aa\lambda bbb && [\text{via } A \to \lambda] \\ &= aabbb \end{aligned}$$

it follows that $aabbb \in \mathcal{L}(G)$.

b) It is sufficient to show that

$$\begin{cases} \forall u \in \{a^n b^{n+1} \mid n \in \mathbb{N}\} : u \in \mathcal{L}(G) \\ \forall u \in \mathcal{L}(G) : u \in \{a^n b^{n+1} \mid n \in \mathbb{N}\} \end{cases}$$

<u>($\Rightarrow$)</u> : Let <u>$u \in \mathcal{L}(G)$</u> be given. It follows that $S \overset{*}{\underset{G}{\Rightarrow}} u$. Note that the first step has to be $S \longrightarrow Ab$. The next p steps have no choice but to be $A \to aAb$

Then the only way to terminate by eliminating A is to apply $A \to \lambda$. This results in the derivation

$$S \underset{G}{\Rightarrow} Ab \underset{G}{\overset{p}{\Rightarrow}} a^p A b^p b \underset{G}{\Rightarrow} a^p \lambda b^p b = a^p b^{p+1}.$$

It follows that if the derivation of u uses p steps $A \to aAb$ that

$$u = a^p b^{p+1} \Rightarrow \exists n \in \mathbb{N} : u = a^n b^{n+1}$$
$$\Rightarrow \underline{u \in \{a^n b^{n+1} \mid n \in \mathbb{N}\}}$$

$\underline{(\Leftarrow)}$: Let $u \in \{a^n b^{n+1} \mid n \in \mathbb{N}\}$ be given. Then

$u \in \{a^n b^{n+1} \mid n \in \mathbb{N}\} \Rightarrow \exists n \in \mathbb{N} : u = a^n b^{n+1}$

Choose some $n_0 \in \mathbb{N}$ such that $u = a^{n_0} b^{n_0+1}$.

We claim that $\forall n \in \mathbb{N} : S \underset{G}{\overset{*}{\Rightarrow}} a^n A b^{n+1}$. We prove the claim by induction:

For $n=0$: $u = a^0 b^{0+1} = \lambda b = b$ and since

$$\begin{aligned} S &\underset{G}{\Rightarrow} Ab && [\text{via } S \to Ab] \\ &\underset{G}{\Rightarrow} \lambda b && [\text{via } A \to \lambda] \\ &= b = a^0 b^1 \end{aligned}$$

we have: $S \underset{G}{\overset{*}{\Rightarrow}} a^0 b^1$.

For $n = k$, assume that $S \underset{G}{\overset{*}{\Rightarrow}} a^k A b^{k+1}$

For $n = k+1$, we will show that $S \underset{G}{\overset{*}{\Rightarrow}} a^{k+1} A b^{k+1}$.

Since

$$\begin{aligned} S &\underset{G}{\overset{*}{\Rightarrow}} a^k A b^{k+1} && [\text{induction hypothesis}] \\ &\underset{G}{\Rightarrow} a^k a A b b^{k+1} && [\text{via } A \to aAb] \\ &= a^{k+1} A b^{k+2} \end{aligned}$$

we have shown by induction that

$$\forall n \in \mathbb{N} : S \Rightarrow a^n A b^{n+1} \qquad (1)$$

It follows that

$$
\begin{aligned}
(S &\xRightarrow[G]{*} a^{n_0} A b^{n_0+1} && [\text{via Eq. (1)}] \\
&\xRightarrow[G]{*} a^{n_0} \lambda b^{n_0+1} && [\text{via } A \to \lambda] \\
&= a^{n_0} b^{n_0+1}) \Rightarrow (S \xRightarrow[G]{*} a^{n_0} b^{n_0+1}) \Rightarrow
\end{aligned}
$$

$$\Rightarrow u = a^{n_0} b^{n_0+1} \in \mathcal{L}(G).$$

From the above argument:

$$
\begin{cases}
\forall u \in \mathcal{L}(G) : u \in \{a^n b^{n+1} \mid n \in \mathbb{N}\} \\
\forall u \in \{a^n b^{n+1} \mid n \in \mathbb{N}\} : u \in \mathcal{L}(G)
\end{cases} \Rightarrow
$$

$$
\Rightarrow \begin{cases}
\mathcal{L}(G) \subseteq \{a^n b^{n+1} \mid n \in \mathbb{N}\} \\
\{a^n b^{n+1} \mid n \in \mathbb{N}\} \subseteq \mathcal{L}(G)
\end{cases}
$$

$$\Rightarrow \mathcal{L}(G) = \{a^n b^{n+1} \mid n \in \mathbb{N}\}$$

EXAMPLE

Consider the grammar $G=(V,\Sigma,S,P)$ with $V=\{S\}$ and $\Sigma=\{a,b\}$ and production rules

$$S \to SS \mid \lambda \mid aSb \mid bSa$$

Show that $\mathcal{L}(G)=\{u\in\Sigma^* \mid n_a(u)=n_b(u)\}$

Solution

($\Rightarrow$): Let $v\in\mathcal{L}(G)$ be given. We note that

a) the production rules $S\to aSb$ and $S\to bSa$ generate an equal number of "a" and "b"

b) the production rules $S\to SS$ and $S\to\lambda$ do not modify the number of "a" and "b".

It follows that

$n_a(v)=n_b(v) \Rightarrow v\in\{u\in\Sigma^* \mid n_a(u)=n_b(u)\}$.

($\Leftarrow$): It is sufficient to show that

$$\forall \mu\in\mathbb{N}: \forall v\in\Sigma^*: (n_a(v)=n_b(v)=\mu \Rightarrow v\in\mathcal{L}(G))$$

We use proof by induction of $\mu\in\mathbb{N}$.

For $\mu=0$: Let $v\in\Sigma^*$ be given such that $n_a(v)=n_b(v)=0$.

Then $|v|=n_a(v)+n_b(v)=0+0=0 \Rightarrow v=\lambda$

and therefore, via the production rule $S\to\lambda$:

$S \underset{G}{\Rightarrow} \lambda = v \Rightarrow v\in\mathcal{L}(G)$.

For $\mu=\mu_0>0$, assume that

$\forall \mu\in\mathbb{N}: \forall v\in\Sigma^*: (n_a(v)=n_b(v)\leq\mu_0 \Rightarrow v\in\mathcal{L}(G))$

For $\mu=\mu_0+1$, we will show that

$\forall v\in\Sigma^*: (n_a(v)=n_b(v)=\mu_0+1 \Rightarrow v\in\mathcal{L}(G))$

Let $v \in \Sigma^*$ be given and assume that $n_a(v) = n_b(v) = \mu_0 + 1$.
We distinguish between the following cases:

Case 1: Assume that $v = awb$ with $w \in \Sigma^*$. Then

$$\left.\begin{array}{l} n_a(w) = n_a(awb) - 1 = (\mu_0 + 1) - 1 = \mu_0 \\ n_b(w) = n_b(awb) - 1 = (\mu_0 + 1) - 1 = \mu_0 \end{array}\right\} \Rightarrow$$

$\Rightarrow n_a(w) = n_b(w) = \mu_0 \Rightarrow$

$\Rightarrow w \in \mathcal{L}(G)$ (via induction hypothesis).

It follows that

$$\begin{aligned} S &\underset{G}{\Longrightarrow} aSb && [\text{via } S \to aSb] \\ &\overset{*}{\underset{G}{\Longrightarrow}} awb && [\text{via } w \in \mathcal{L}(G)] \end{aligned}$$

and therefore $v = awb \in \mathcal{L}(G)$.

Case 2: Assume that $v = bwa$ with $w \in \Sigma^*$. Then

$$\left.\begin{array}{l} n_a(w) = n_a(bwa) - 1 = (\mu_0 + 1) - 1 = \mu_0 \\ n_b(w) = n_b(bwa) - 1 = (\mu_0 + 1) - 1 = \mu_0 \end{array}\right\} \Rightarrow$$

$\Rightarrow n_a(w) = n_b(w) = \mu_0 \Rightarrow$

$\rightarrow w \in \mathcal{L}(G)$ (via induction hypothesis)

It follows that

$$\begin{aligned} S &\underset{G}{\Longrightarrow} bSa && [\text{via } S \to bSa] \\ &\overset{*}{\underset{G}{\Longrightarrow}} bwa && [\text{via } w \in \mathcal{L}(G)] \end{aligned}$$

and therefore $v = bwa \in \mathcal{L}(G)$.

Case 3: Assume that with no loss of generality

$v = awa$ with $w \in \Sigma^*$

We claim that $\exists p, q \in \Sigma^* : (w = pq \land ap \in \mathcal{L}(G) \land qa \in \mathcal{L}(G))$

Define $\forall n \in [2\mu_0 + 2] : \Delta(n) = n_a(v_1 v_2 \cdots v_n) - n_b(v_1 v_2 \cdots v_n)$

We have:

$$\Delta(1) = n_a(v_1) - n_b(v_1) = n_a(a) - n_b(b) = 1 - 0 = 1 > 0 \qquad (1)$$

and

$$\Delta(2\mu_0+1) = n_a(v_1 v_2 \cdots v_{2\mu_0+1}) - n_b(v_1 v_2 \cdots v_{2\mu_0+1}) =$$
$$= [n_a(v) - n_a(a)] - [n_b(v) - n_b(a)] =$$
$$= [n_a(v) - n_b(v)] - [n_a(a) - n_b(a)] =$$
$$= 0 - [1 - 0] = -1 < 0 \qquad (2)$$

From Eq. (1) and Eq. (2):

$\exists m \in [2\mu_0+2] : \Delta(m) = 0$

Choose an $m \in [2\mu_0+2]$ such that $\Delta(m) = 0$ and define $p, q \in \Sigma^*$ such that

$ap = v_1 v_2 \cdots v_m$

$qa = v_{m+1} v_{m+2} \cdots v_{2\mu_0+2}$.

and note that $v = apqa$. Then:

$$n_a(ap) - n_b(ap) = n_a(v_1 v_2 \cdots v_m) - n_b(v_1 v_2 \cdots v_m)$$
$$= \Delta(m) = 0 \Rightarrow$$

$\Rightarrow n_a(ap) = n_b(ap) \Rightarrow ap \in \mathcal{L}(G)$. [via induction hypothesis]

and

$$n_a(qa) - n_b(qa) = [n_a(apqa) - n_a(ap)] - [n_b(apqa) - n_b(ap)]$$
$$= [n_a(v) - n_b(v)] - [n_a(ap) - n_b(ap)]$$
$$= 0 - 0 = 0 \Rightarrow$$

$\Rightarrow n_a(qa) = n_b(qa) \Rightarrow qa \in \mathcal{L}(G)$ [via induction hypothesis]

This argument proves the claim. It follows that

$$(S \underset{G}{\Rightarrow} SS \qquad [\text{via } S \to SS]$$
$$\overset{*}{\underset{G}{\Rightarrow}} apS \qquad [\text{via } ap \in \mathcal{L}(G)]$$
$$\overset{*}{\underset{G}{\Rightarrow}} apqa \qquad [\text{via } qa \in \mathcal{L}(G)]$$
$$= v) \Rightarrow (S \overset{*}{\underset{G}{\Rightarrow}} v) \Rightarrow v \in \mathcal{L}(G)$$

We have thus shown that

$\forall \mu \in \mathbb{N} : \forall v \in \Sigma^* : (n_a(v) = n_b(v) = \mu \Rightarrow v \in \mathcal{L}(G))$

Let $v \in \{u \in \Sigma^* \mid n_a(u) = n_b(u)\}$ be given. Then

$$v \in \{u \in \Sigma^* \mid n_a(u) = n_b(u)\} \Rightarrow \begin{cases} v \in \Sigma^* \\ n_a(v) = n_b(v) \in \mathbb{N} \end{cases} \Rightarrow$$

$$\Rightarrow v \in \mathcal{L}(G).$$

From the above argument we have:

$$\begin{cases} \forall v \in \mathcal{L}(G) : v \in \{u \in \Sigma^* \mid n_a(u) = n_b(u)\} \\ \forall v \in \{u \in \Sigma^* \mid n_a(u) = n_b(u)\} : v \in \mathcal{L}(G) \end{cases} \Rightarrow$$

$$\Rightarrow \begin{cases} \mathcal{L}(G) \subseteq \{u \in \Sigma^* \mid n_a(u) = n_b(u)\} \\ \{u \in \Sigma^* \mid n_a(u) = n_b(u)\} \subseteq \mathcal{L}(G) \end{cases} \Rightarrow$$

$$\Rightarrow \mathcal{L}(G) = \{u \in \Sigma^* \mid n_a(u) = n_b(u)\}$$

EXERCISES

(8) Consider a grammar $G = (V, \Sigma, S, P)$ with $V = \{S, A\}$ and $\Sigma = \{a, b\}$ and production rules

$$S \longrightarrow aA \mid \lambda$$
$$A \longrightarrow bS$$

Show that $\mathcal{L}(G) = \{(ab)^n \mid n \in \mathbb{N}\}$

(9) Consider a grammar $G = (V, \Sigma, S, P)$ with $V = \{S, A, B\}$ and $\Sigma = \{a\}$ and production rules

$$S \longrightarrow Aa$$
$$A \longrightarrow Ba$$
$$B \longrightarrow Aa \mid \lambda$$

Show that $\mathcal{L}(G) = \{a^n \mid n \in \mathbb{N} \wedge n \geq 2\}$

(10) Consider the grammar $G = (V, \Sigma, S, P)$ with $V = \{S\}$ and $\Sigma = \{a, b\}$ and production rules

$$S \longrightarrow aSa \mid bSb \mid \lambda$$

Show that $\mathcal{L}(G) = \{w^R w \mid w \in \Sigma^*\}$.

(11) Consider the grammar $G = (V, \Sigma, S, P)$ with $V = \{S, A\}$ and $\Sigma = \{a, b\}$ and production rules

$$S \longrightarrow aSb \mid A$$
$$A \longrightarrow Ab \mid \lambda$$

Show that $\mathcal{L}(G) = \{a^n b^m \mid n, m \in \mathbb{N} \wedge m \geq n \geq 1\}$.

(12) Consider the grammar $G=(V,\Sigma,S,P)$ with $V=\{S\}$ and $\Sigma=\{a,b\}$ and production rules

$$S \to aSbb \mid \lambda$$

Show that $\mathcal{L}(G)=\{a^n b^{2n} \mid n\in\mathbb{N}\}$.

(13) Consider the grammar $G=(V,\Sigma,S,P)$ with $V=\{S,A\}$ and $\Sigma=\{a,b\}$ and production rules

$$S \longrightarrow aaA$$
$$A \longrightarrow aBb$$
$$B \longrightarrow aBb \mid \lambda$$

Show that $\mathcal{L}(G)=\{a^{n+2} b^n \mid n\in\mathbb{N} \wedge n\geq 1\}$.

(14) Consider two grammars $G_1=(V_1,\Sigma,S_1,P_1)$ and $G_2=(V_2,\Sigma,S_2,P_2)$ that share the same alphabet and assume that $V_1\cap V_2=\emptyset$.

a) Let $G=(V,\Sigma,S,P)$ with $V=\{S\}\cup V_1\cup V_2$ and $P=\{S\to S_1S_2\}\cup P_1\cup P_2$. Show that $\mathcal{L}(G)=\mathcal{L}(G_1)\mathcal{L}(G_2)$.

b) Let $G=(V,\Sigma,S,P)$ with $V=\{S\}\cup V_1\cup V_2$ and $P=\{S\to S_1,\ S\to S_2\}\cup P_1\cup P_2$. Show that $\mathcal{L}(G)=\mathcal{L}(G_1)\cup\mathcal{L}(G_2)$.

(15) Consider the grammar $G=(V,\Sigma,S,P)$ with $V=\{S,A\}$ and $\Sigma=\{a,b\}$ and production rules

$$S \longrightarrow aSaSb \mid bSaSa \mid aSbSa \mid SS \mid \lambda$$

Show that $\mathcal{L}(G)=\{u\in\Sigma^* \mid n_a(u)=2n_b(u)\}$.

▼ Deterministic Finite Accepters

The deterministic finite accepter (dfa) is the mathematical model of a basic computational device that can accept or reject strings defined over an alphabet Σ

> **Def**: A deterministic finite accepter (dfa) M is defined as the 5-tuple $M = (Q, \Sigma, \delta, q_0, F)$ where
> a) Q is a finite set of internal states
> b) Σ is the input alphabet
> c) $\delta : Q \times \Sigma \longrightarrow Q$ is a transition function
> d) $q_0 \in Q$ is the initial state
> e) $F \subseteq Q$ is a set of final states.

notation: The set of all dfa machines that can be defined over a given alphabet Σ is denote as $dfa(\Sigma)$.

The computational action of a dfa M can be defined via the extended transition function as follows:

> **Def**: Let $M = (Q, \Sigma, \delta, q_0, F)$ be a dfa. The extended transition function $\delta^* : Q \times \Sigma^* \longrightarrow Q$ is defined recursively as follows:
> $\forall q \in Q : \delta^*(q, \lambda) = q$
> $\forall q \in Q : \forall u \in \Sigma^* : \forall a \in \Sigma : \delta^*(q, ua) = \delta(\delta^*(q, u), a)$

<u>interpretation</u>: To evaluate $\delta^*(q,u)$ with $q \in Q$ and $u \in \Sigma^*$, we begin with internal state q and process the string u character by character. For each processed character $a \in \Sigma$, the next state is given by $\delta(q,a)$. As we consume the string u, the dfa transitions from state to state. The resulting sequence of states is given by:

$$\begin{cases} q_1 = q \\ \forall n \in [|u|-1] : q_{n+1} = \delta(q_n, u_n) \end{cases}$$

The final state $q_{|u|}$ is returned by $\delta^*(q,u)$.

▸ <u>Language accepted by a dfa</u>

<u>Def</u>: Let $M = (Q, \Sigma, \delta, q_0, F)$ be a dfa. The language $\mathcal{L}(M)$ accepted by the dfa M is defined as

$$\mathcal{L}(M) = \{u \in \Sigma^* \mid \delta^*(q_0, u) \in F\}$$

<u>interpretation</u>: The belonging condition for $\mathcal{L}(M)$ is given by:

$$\forall u \in \Sigma^* : (u \in \mathcal{L}(M) \Leftrightarrow \delta^*(q_0, u) \in F)$$

This means that a string u is accepted by the dfa M if and only if, initializing M at the initial state q_0 and processing the string u results in a state $\delta^*(q_0,u)$ which is among the permitted final internal states in F. If the resulting state $\delta^*(q_0,u)$ is not among the states in F, then M rejects the string u.

Def: Let $L \in \mathcal{P}(\Sigma^*)$ be a language on Σ.
We say that:
L regular $\iff \exists M \in \text{dfa}(\Sigma): L = \mathcal{L}(M)$

▸ Graph representation of a dfa

Consider a dfa $M = (Q, \Sigma, \delta, q_0, F)$. We can represent M with a directed graph $G = \text{Graph}(M) = (V(G), E(G), \psi_G)$ such that:

a) The set of vertices is $V(G) = Q$. The final states in $F \subseteq Q$ are represented by specially labeled vertices, as shown in the example below:

$\text{(}q\text{)}$ for $q \notin F$

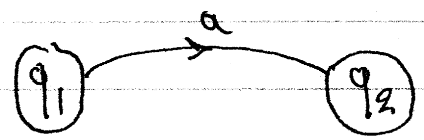
$\text{((}q\text{))}$ for $q \in F$

The initial vertex q_0 is indicated with an open-ended incoming arrow:

$\longrightarrow (q_0)$ for initial state q_0

b) The set of edges is
$E(G) = Q \times \Sigma = \{(q, a) \mid q \in Q \land a \in \Sigma\}$
and the corresponding incidence function $\psi_G : E(G) \to V(G) \times V(G)$ is given by
$\forall q \in Q : \forall a \in \Sigma : \psi_G((q, a)) = (q, \delta(q, a))$

In other words, every transition rule $\delta(q_1, a) = q_2$ defines an edge $e = (q_1, a)$ from vertex q_1 to q_2:



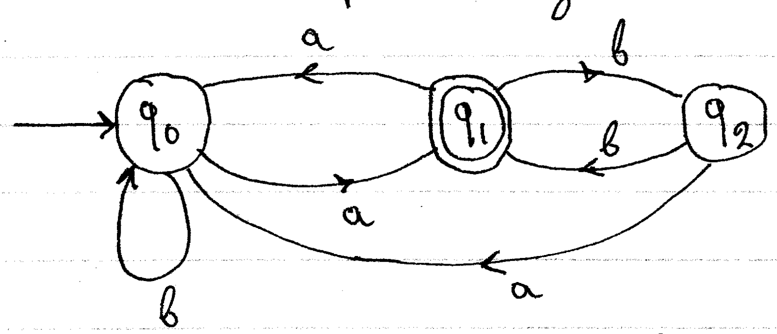
The outgoing edges from q_1 are distinguished by $a \in \Sigma$, so by convention each outgoing edge is distinguished by the character $a \in \Sigma$.

EXAMPLE - ILLUSTRATION

Consider the deterministic finite accepter $M = (Q, \Sigma, \delta, q_0, F)$ with $Q = \{q_0, q_1, q_2\}$ and $\Sigma = \{a, b\}$ and $F = \{q_1\}$ with transition function δ given by:

$$\delta(q_0, a) = q_1 \qquad \delta(q_1, b) = q_2$$
$$\delta(q_0, b) = q_0 \qquad \delta(q_2, a) = q_0$$
$$\delta(q_1, a) = q_0 \qquad \delta(q_2, b) = q_1$$

•1 This dfa is represented by the following graph:



•2 To show that $u = abbaa \in \mathcal{L}(M)$, we argue as follows:

$$\delta(q_0,a)=q_1 \Rightarrow \delta^*(q_0,ab)=\delta(q_1,b)=q_2 \Rightarrow$$
$$\Rightarrow \delta^*(q_0,abb)=\delta(q_2,b)=q_1 \Rightarrow$$
$$\Rightarrow \delta^*(q_0,abba)=\delta(q_1,a)=q_0 \Rightarrow$$
$$\Rightarrow \delta^*(q_0,abbaa)=\delta(q_0,a)=q_1 \in F$$
$$\Rightarrow \delta^*(q_0,abbaa)\in F \Rightarrow$$
$$\Rightarrow abbaa \in \mathcal{L}(M)$$

Remarks

a) Note that in order for a graph G, as shown in the above example, to represent a dfa, it is necessary for each vertex to have one outgoing edge for every element of Σ.

b) The string $abbaa$ defines a walk on the above graph given by the following alternating sequence of vertices/edges:

$$W=(q_0,(q_0,a),q_1,(q_1,b),q_2,(q_2,b),q_1,(q_1,a),$$
$$q_0,(q_0,a),q_1)$$

The walk traces out the string $\sigma(w)=abbaa$ and we can say that a string $u\in\Sigma^*$ is accepted by M if and only if there is a walk w from q_0 to an element of F such that $\sigma(w)=u$.

We now restate the above more formally as follows:

<u>Def</u>: Let $G=\mathrm{Graph}(M)$ be the directed graph representing the dfa $M\in \mathrm{dfa}(\Sigma)$. Let $w\in W(G)$ be a walk on G. The string $\sigma(w)\in\Sigma^*$ induced by the walk w is defined as:
$$\forall n\in[|w|]: [\sigma(w)]_n = a \Leftrightarrow \exists q\in Q: e_n(w)=(q,a)$$

Remark:

Recall the following notation:

$|w|$ = the length of the walk w (number of edges)

$e_n(w)$ = the n^{th} edge of walk w.

Also recall that for vertices $u, v \in V(G)$, we define

$W(G|u,v)$ = the set of all walks on G from u to v.

Thm: Let G be the directed graph $G = \mathrm{Graph}(M)$ representing the dfa $M = (Q, \Sigma, \delta, q_0, F)$. Then:
$\forall u \in \Sigma^* : (u \in \mathcal{L}(M) \Leftrightarrow \exists q \in F : \exists w \in W(G|q_0, q) : \sigma(w) = u$

Equivalently, we can state that

$$\mathcal{L}(M) = \{u \in \Sigma^* \mid \exists w \in \bigcup_{q \in F} W(G|q_0, q) : \sigma(w) = u\}$$

In words: A string $u \in \Sigma^*$ is accepted by the dfa M if and only if there is some walk w from the initial state q_0 to some final state $q \in F$ that induces the string u.

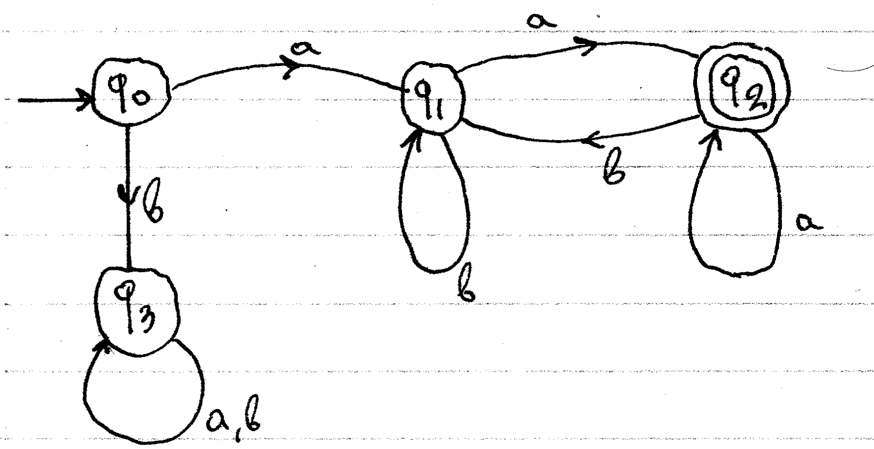
Methodology: The graph terminology is useful in constructing general arguments about the language $\mathcal{L}(M)$ accepted by some dfa $M \in \mathrm{dfa}(\Sigma)$, as shown in the following example.

EXAMPLE

Show that the language $L = \{awa \mid w \in \Sigma^*\}$ with $\Sigma = \{a, b\}$ is regular.

Solution

Consider the dfa given by the following graph M:



<u>$(\Rightarrow)$</u>: Let $u \in \mathcal{L}(M)$ be given. Then:

$$u \in \mathcal{L}(M) \Rightarrow \delta^*(q_0, u) \in F \Rightarrow \delta^*(q_0, u) \in \{q_2\} \Rightarrow$$
$$\Rightarrow \delta^*(q_0, u) = q_2$$

To show that $u \in L$, assume that $u \notin L$. Then:

$$u \notin L \Rightarrow u \notin \{awa \mid w \in \Sigma^*\} \Rightarrow$$
$$\Rightarrow \overline{\exists w \in \Sigma^* : u = awa} \Rightarrow$$
$$\Rightarrow \forall w \in \Sigma^* : u \neq awa$$

Given this restriction, we distinguish between the following cases.

<u>Case 1</u>: Assume that $u = bw$ with $w \in \Sigma^*$. Then:

$$\delta^*(q_0, u) = \delta^*(q_0, bw) = \delta^*(\delta^*(q_0, b), w) =$$
$$= \delta^*(q_3, w) = q_3 \neq q_2$$

which is a contradiction.

<u>Case 2</u>: Assume that $u = awb$ with $w \in \Sigma^*$. Then:

$$\delta^*(q_0, u) = \delta^*(q_0, awb) = \delta^*(q_1, wb) =$$
$$= \delta^*(\delta^*(q_1, w), b) \qquad (1)$$

We distinguish between the following subcases:

▶ <u>Case 2A</u>: Assume that $\delta^*(q_1, w) = q_1$. Then, from Eq. (1)

$$\delta^*(q_0, u) = \delta^*(\delta^*(q_1, w), b) = \delta^*(q_1, b) = q_1 \neq q_2$$

which is a contradiction.

▶ <u>Case 2B</u>: Assume that $\delta^*(q_1, w) = q_2$. Then, from Eq. (1)

$$\delta^*(q_0, u) = \delta^*(\delta^*(q_1, w), b) = \delta^*(q_2, b) = q_1 \neq q_2$$

which is a contradiction.

Since all possibilities lead to a contradiction, it follows that $u \in L$. We have thus shown that: $\forall u \in \mathcal{L}(M) : u \in L$.

<u>($\Leftarrow$)</u>: Let $u \in L$ be given. Then,

$$u \in L \Rightarrow u \in \{awa \mid w \in \Sigma^*\} \Rightarrow$$
$$\Rightarrow \exists w \in \Sigma^* : u = awa$$

Choose an $w \in \Sigma^*$ such that $u = awa$. Note that $\delta(q_0, a) = q_1$. From q_1, there are no walks from q_1 to q_3 and no walks from q_1 to q_0. It follows that

$$\delta^*(q_0, aw) \in \{q_1, q_2\}.$$

We distinguish between the following cases:

<u>Case 1</u>: Assume that $\delta^*(q_0, aw) = q_1$. Then:

$$\delta^*(q_0, awa) = \delta(\delta^*(q_0, aw), a) = \delta(q_1, a) = q_2 \in F \Rightarrow$$
$$\Rightarrow \delta^*(q_0, awa) \in F \Rightarrow u = awa \in \mathcal{L}(M).$$

Case 2: Assume that $\delta^*(q_0, aw) = q_2$. Then,

$\delta^*(q_0, awa) = \delta(\delta^*(q_0, aw), a) = \delta(q_2, a) = q_2 \in F \Rightarrow$

$\Rightarrow \delta^*(q_0, awa) \in F \Rightarrow u = awa \in \mathcal{L}(M)$.

We have thus shown that $\forall u \in L : u \in \mathcal{L}(M)$.

From the above argument:

$$\begin{cases} \forall u \in \mathcal{L}(M) : u \in L \\ \forall u \in L : u \in \mathcal{L}(M) \end{cases} \Rightarrow \begin{cases} \mathcal{L}(M) \subseteq L \\ L \subseteq \mathcal{L}(M) \end{cases} \Rightarrow \mathcal{L}(M) = L$$

$$\Rightarrow \exists M \in \mathrm{dfa}(\Sigma) : L = \mathcal{L}(M)$$

$$\Rightarrow L \text{ regular.}$$

▶ Recursion on extended transition function

Thm: Let $M=(Q,\Sigma,\delta,q_0,F)$ be a dfa with extended transition function $\delta^*: Q\times\Sigma^*\to Q$. Then
$\forall q\in Q: \forall u,v\in\Sigma^*: \delta^*(q,uv)=\delta^*(\delta^*(q,u),v)$

Proof

Let $q\in Q$ and let $u,v\in\Sigma^*$ be given. We use induction on the length of v.

For $|v|=0$, we have $v=\lambda$ and therefore

$$\delta^*(q,uv)=\delta^*(q,u\lambda)=\delta^*(q,u)$$
$$=\delta^*(\delta^*(q,u),\lambda)=\delta^*(\delta^*(q,u),v)$$

Assume that $\forall w\in\Sigma^*: (|w|\leq k \Rightarrow \delta^*(q,uw)=\delta^*(\delta^*(q,u),w))$

For $|v|=k+1$, we will show that

$$\delta^*(q,uv)=\delta^*(\delta^*(q,u),v)$$

Choose a $w\in\Sigma^*$ and $a\in\Sigma$ such that $v=wa$. Note that $|w|=|wa|-|a|=(k+1)-1=k$, so the induction hypothesis applies to w. Then

$$
\begin{aligned}
\delta^*(q,uv) &= \delta^*(q,uwa) && [\text{via } v=wa]\\
&= \delta(\delta^*(q,uw),a) && [\text{def of } \delta^*]\\
&= \delta(\delta^*(\delta^*(q,u),w),a) && [\text{induction hypothesis}]\\
&= \delta^*(\delta^*(q,u),wa) && [\text{def of } \delta^*]\\
&= \delta^*(\delta^*(q,u),v) && [\text{via } v=wa]
\end{aligned}
$$

From the above argument, we have shown that
$\forall q\in Q: \forall u,v\in\Sigma^*: \delta^*(q,uv)=\delta^*(\delta^*(q,u),v)$.

EXERCISES

(16) Let $\Sigma = \{a, b\}$. Construct a dfa M that accepts the following languages L and prove that $L = \mathcal{L}(M)$.

a) $L = \{u \in \Sigma^* \mid n_a(u) = 1\}$

b) $L = \{u \in \Sigma^* \mid n_a(u) \geqslant 1\}$

c) $L = \{u \in \Sigma^* \mid n_a(u) = 2\}$

d) $L = \{uvu \mid u, v \in \Sigma^* \wedge |u| = 2\}$

e) $L = \{a^n \mid n \in \mathbb{N} - \{3\}\}$

f) $L = \{abwb^2 \mid w \in \Sigma^*\}$

g) $L = \{a, aba, b\}$

h) $L = \{ab, abab\}$

(17) Use the identity $\delta^*(q, uv) = \delta^*(\delta^*(q, u), v)$ to show that

a) $\forall L \in \mathcal{P}(\Sigma^*) : (L \text{ regular} \Rightarrow L^2 \text{ regular})$

b) $\forall L_1, L_2 \in \mathcal{P}(\Sigma^*) : \left(\begin{cases} L_1 \text{ regular} \\ L_2 \text{ regular} \end{cases} \Rightarrow L_1 L_2 \text{ regular} \right)$

▼ Pigeonhole principle and non-regular languages

Intuitively, the pigeonhole principle states that if we put n objects (e.g. pigeons) in m boxes (e.g. pigeonholes) and if $n > m$, then at least one box has at least two objects in it. The principle still applies when m, n are infinite cardinalities and can be stated rigorously as follows:

> **Lemma (Pigeonhole principle)**
> Let A, B be two sets. Then
> $$\begin{cases} f \in \text{Map}(A,B) \\ A > B \end{cases} \Rightarrow \exists x, y \in A : (x \neq y \wedge f(x) = f(y))$$

Proof

Assume that $f \in \text{Map}(A,B)$ and $A > B$. To show a contradiction, assume that $\overline{\exists x,y \in A : (x \neq y \wedge f(x) = f(y))}$. Then,

$$\overline{\exists x,y \in A : (x \neq y \wedge f(x) = f(y))} \Rightarrow \forall x,y \in A : (f(x) \neq f(y) \vee x = y)$$
$$\Rightarrow \forall x,y \in A : (f(x) = f(y) \Rightarrow x = y) \quad [\text{via } p \Rightarrow q \equiv \bar{p} \vee q]$$
$$\Rightarrow f \text{ one-to-one} \Rightarrow A \leq B$$

and therefore:

$$A > B \wedge A \leq B \Rightarrow (A > B \vee A \sim B) \wedge A \leq B \quad [\text{Extension}]$$
$$\Rightarrow A \geq B \wedge A \leq B \quad [\text{Definition}]$$
$$\Rightarrow A \sim B \quad [\text{Schröder–Bernstein}]$$

On the other hand, by definition: $A > B \Rightarrow A \nsim B$, so we have a contradiction. It follows that

$$\exists x,y \in A : (x \neq y \wedge f(x) = f(y)).$$

EXAMPLE

Show that the language $L = \{a^n b^n \mid n \in \mathbb{N}\}$ is not regular.

Solution

To show that L is not regular, assume that L is regular in order to derive a contradiction. Choose $M \in \mathrm{dfa}(\{a, b\})$ such that $\mathcal{L}(M) = L$ with $M = (Q, \{a, b\}, \delta, q_0, F)$.

Define $f : \mathbb{N}^* \to Q$ given by:

$\forall n \in \mathbb{N}^* : f(n) = \delta^*(q_0, a^n)$

From the pigeonhole principle,

$$\left\{\begin{array}{l} \mathbb{N}^* \text{ countably infinite} \Rightarrow \mathbb{N}^* > Q \Rightarrow \\ Q \text{ finite} \end{array}\right.$$

$$\Rightarrow \exists n_1, n_2 \in \mathbb{N}^* : (n_1 \neq n_2 \wedge f(n_1) = f(n_2))$$

Choose $n_1 = k$ and $n_2 = \mu$ such that $k \neq \mu$ and $f(k) = f(\mu) = q \in Q$. We also note that

$a^k b^k \in L \Rightarrow \delta^*(q_0, a^k b^k) \in F$

Define $q_f = \delta^*(a^k b^k, q_0)$. It follows that

$$\begin{aligned} \delta^*(q_0, a^\mu b^k) &= \delta^*(\delta^*(q_0, a^\mu), b^k) = \\ &= \delta^*(f(\mu), b^k) = \\ &= \delta^*(f(k), b^k) = \\ &= \delta^*(\delta^*(q_0, a^k), b^k) \\ &= \delta^*(q_0, a^k b^k) \\ &= q_f \in F \Rightarrow \end{aligned}$$

$\Rightarrow \delta^*(q_0, a^\mu b^k) \in F \Rightarrow a^\mu b^k \in L \Rightarrow$

$\Rightarrow a^\mu b^k \in \{a^n b^n \mid n \in \mathbb{N}\} \Rightarrow \underline{\mu = k}.$ ← Contradiction

It follows that L is not regular.

Lemma: (Pumping Lemma)
Let Σ be an alphabet and let $L \in \mathcal{P}(\Sigma^*)$ be a language. Then:

$$L \text{ regular} \wedge L \text{ infinite} \Rightarrow$$

$$\Rightarrow \exists m \in \mathbb{N}^* : \forall w \in L : \left(|w| \geq m \Rightarrow \exists x,y,z \in \Sigma^* : \begin{cases} xyz = w \\ |xy| \leq m \\ |y| \geq 1 \\ \forall k \in \mathbb{N} : xy^k z \in L \end{cases} \right)$$

Proof

Assume that L regular $\wedge$ L infinite. Choose $M \in \text{dfa}(\Sigma)$ such that $L = \mathcal{L}(M)$. Let $n = |Q - \{q_0\}|$ be the number of non-initial internal states of M and write

$$Q = \{q_0, q_1, q_2, \ldots, q_n\}$$

Choose $\underline{m = n+1 \in \mathbb{N}^*}$. Let $w \in L$ be given and assume that $|w| \geq m$. Define $\ell = |w|$.

▶ Construction of x, y, z

Define $p_0, p_1, p_2, \ldots, p_\ell$ according to

$$\begin{cases} p_0 = q_0 \\ \forall k \in [\ell] : p_k = \delta(p_{k-1}, w_k) \end{cases}$$

and note that $p_0, p_1, \ldots, p_\ell$ are the internal states that the machine M goes through as it processes the string w.
Since $w \in L$, we have $p_\ell = \delta(q_0, w) = q_f \in F$.
From the pigeonhole principle it follows that

$$\exists a, b \in [n+1] : (a > b \wedge p_a = p_b)$$

Choose $a, b \in [n+1]$ such that $a > b$ and $p_a = p_b$.

Define $x, y, z \in \Sigma^*$ such that

$$\begin{cases} x = w_1 w_2 \cdots w_a \\ y = w_{a+1} w_{a+2} \cdots w_b \\ z = w_{b+1} w_{b+2} \cdots w_\ell \end{cases}$$

and note that

$$\begin{cases} \delta^*(q_0, x) = p_a \\ \delta^*(q_0, xy) = p_b = p_a \\ \delta^*(p_b, z) = q_\ell \in F \end{cases}$$

▶ Proof of claims

We have

$$|xy| = |x| + |y| = (a-1+1) + (b-(a+1)+1) =$$
$$= a + (b-a-1+1) = a + (b-a) = b \leq n+1 \quad (\text{via } b \in [n+1])$$

and

$$|y| = b - (a+1) + 1 = b-a-1+1 = b-a > 0 \quad (\text{via } b > 0)$$
$$\Rightarrow |y| > 0 \Rightarrow |y| \geq 1.$$

We claim that $\forall k \in \mathbb{N} : \delta^*(q_0, xy^k) = p_b$.

For $k=0$:

$$\delta^*(q_0, xy^k) = \delta^*(q_0, xy^0) = \delta^*(q_0, x) = p_a = p_b.$$

For $k = k_0$, assume that $\delta^*(q_0, xy^{k_0}) = p_b$.

For $k = k_0+1$, we have

$$\begin{aligned} \delta^*(q_0, xy^{k_0+1}) &= \delta^*(\delta^*(q_0, xy^{k_0}), y) = && [\text{Definition}] \\ &= \delta(p_b, y) = && [\text{induction hyp}] \\ &= \delta(p_a, y) = && [\text{via } p_a = p_b] \\ &= p_b \end{aligned}$$

We have thus shown the claim.

Let $\underline{k \in \mathbb{N}}$ be given. Then:

$$\begin{aligned} \delta^*(q_0, xy^k z) &= \delta^*(\delta^*(q_0, xy^k), z) \\ &= \delta^*(p_\ell, z) \\ &= q_f \in F \Rightarrow \end{aligned}$$

$\rightarrow \underline{xy^k z \in L}$

and therefore: $\forall k \in \mathbb{N} : xy^k z \in L$

From the above argument we obtain the lemma.

EXAMPLES

a) Let $\Sigma = \{a, b\}$. Show that
$L = \{ww^R \mid w \in \Sigma^*\}$
is not regular, via the pumping lemma

Solution

To show that L is not regular, assume that L is regular.
We also note that

$\forall n \in \mathbb{N}^* : a^{2n} = a^n a^n = a^n (a^n)^R \Rightarrow$
$\Rightarrow \forall n \in \mathbb{N}^* : a^{2n} \in L$
$\Rightarrow L$ infinite.

It follows that the pumping lemma applies. Choose $m \in \mathbb{N}^*$ such that

$$\forall w \in L : \left(|w| \geq m \Rightarrow \exists x, y, z \in L : \begin{cases} |xy| \leq m \wedge |y| \geq 1 \\ \forall k \in \mathbb{N} : xy^k z \in L \end{cases} \right)$$

Let $w = a^m b^{2m} a^m = a^m b^m b^m a^m = (a^m b^m)(a^m b^m)^R \in L$
and note that $|w| = m + 2m + m = 4m \geq m$. Choose $x, y, z \in L$ such that $|xy| \leq m$ and $|y| \geq 1$ and $\forall k \in \mathbb{N} : xy^k z \in L$.

Define $l_1 = |x|$ and $l_2 = |y|$. Since
$l_1 + l_2 = |x| + |y| = |xy| \leq m$
it follows that $x = a^{l_1}$ and $y = a^{l_2}$. Then:

$$xyz = w \Leftrightarrow a^{l_1} a^{l_2} z = a^m b^{2m} a^m \Leftrightarrow$$
$$\Leftrightarrow z = a^{m - l_1 - l_2} b^{2m} a^m$$

For $k = 2$: $xy^2 z \in L$. Note that:

$$xy^2z = a^{l_1}(a^{l_2})^2\left[a^{m-l_1-l_2}\,b^{2m}\,a^m\right] =$$
$$= a^{l_1+2l_2+m-l_1-l_2}\,b^{2m}\,a^m$$
$$= a^{m+l_2}\,b^m\,b^m\,a^m \in L \Rightarrow$$
$$\Rightarrow m+l_2 = m \Rightarrow l_2 = 0$$

which is a contradiction since $l_2 = |y| \geq 1$

It follows that L is not regular. $\square$

b) Let $\Sigma = \{a, b\}$ and $L = \{w \in \Sigma^* \mid n_a(w) < n_b(w)\}$.
Show that L is not regular.

Solution

To show that L is not regular, assume that L is regular. We also note that

$\forall n \in \mathbb{N}^* : (n_a(b^n) = 0 < n = n_b(b^n)) \Rightarrow$
$\Rightarrow \forall n \in \mathbb{N}^* : (b^n \in L) \Rightarrow L$ infinite.

It follows that the pumping lemma applies.
Choose $m \in \mathbb{N}^*$ such that

$$\forall w \in L : \left(|w| \geq m \Rightarrow \exists x, y, z \in \Sigma^* : \begin{cases} w = xyz \land |y| \geq 1 \land |xy| \leq m \\ \forall n \in \mathbb{N} : xy^n z \in L \end{cases} \right)$$

Choose $w = a^n b^{n+1} \in L$ with $n > m$, and note that
$|w| = |a^n b^{n+1}| = n + (n+1) = 2n+1 > 2n > n > m \Rightarrow |w| \geq m$.
Choose $x, y, z \in \Sigma^*$ such that $w = xyz$ and $|y| \geq 1$ and $|xy| \leq m$ and $\forall k \in \mathbb{N} : xy^k z \in L$.
Define $l_1 = |x|$ and $l_2 = |y|$. Since
$l_1 + l_2 = |x| + |y| = |xy| \leq m < n$, it follo
it follows that $x = a^{l_1}$ and $y = a^{l_2}$ and

$$xyz = a^n b^{n+1} \Leftrightarrow a^{l_1} a^{l_2} z = a^n b^{n+1} \Leftrightarrow$$
$$\Leftrightarrow z = a^{n-l_1-l_2} b^{n+1}$$

We then have:

$$xy^2 z = a^{l_1} (a^{l_2})^2 [a^{n-l_1-l_2} b^{n+1}] =$$
$$= a^{l_1 + 2l_2 + n - l_1 - l_2} b^{n+1} = a^{n+l_2} b^{n+1} \in L \Rightarrow$$

$$\Rightarrow n_a(a^{n+l_2} b^{n+1}) < n_b(a^{n+l_2} b^{n+1}) \Rightarrow$$
$$\Rightarrow n+l_2 < n+1 \Rightarrow l_2 < 1 \Rightarrow |y| < 1$$

which is a contradiction since $|y| \geq 1$.
We conclude that L is not regular.

EXERCISES

(18) Use the pigeonhole principle to show that the following languages are not regular.

a) $L = \{a^n b^{2n} \mid n \in \mathbb{N}\}$

b) $L = \{a^{2n} b^{3n+1} \mid n \in \mathbb{N}\}$

c) $L = \{(ab)^n \cdot a^k \mid n, k \in \mathbb{N} \wedge n > k\}$

(19) Use the pumping lemma to show that the following languages are not regular.

a) $L = \{a^n b^n \mid n \in \mathbb{N}\}$

b) $L = \{x^a y^b \mid a, b \in \mathbb{N} \wedge a < b\}$

c) $L = \{x^a y^b z^c \mid a, b, c \in \mathbb{N} \wedge a + b \leq c\}$

d) $L = \{u \in \{a, b\}^* \mid n_a(u) = n_b(u)\}$

e) $L = \{uu \mid u \in \{a, b\}^+\}$

f) $L = \{x^a y^b x^c \mid a, b, c \in \mathbb{N} \wedge a = c\}$

Non-deterministic finite accepter

The fundamental difference between a non-deterministic finite accepter (hereafter, nfa) and a dfa is that in an nfa, the transition function maps from the current state to a set of possible states, depending on the string input. A string is accepted if a walk from the initial state to some final state exists that traces out the string characters. By contrast, in a dfa, given the current state and string input, there is a unique next state, and a unique walk on the graph representation of the dfa as a function of the string input. In spite of the increased flexibility of the nfa, it is equally powerful to the dfa; a language is accepted by an nfa if and only if it is accepted by some dfa. The formal definitions for these concepts are as follows:

<u>Def</u>: A non-deterministic finite accepter (nfa) M is defined as the 5-tuple $M = (Q, \Sigma, \delta, q_0, F)$ where:

a) Q is a set of internal states

b) Σ is the alphabet set

c) $\delta : Q \times (\Sigma \cup \{\lambda\}) \to \mathcal{P}(Q)$ is a transition function

d) $q_0 \in Q$ is an initial state

e) $F \subseteq Q$ is a set of final states

<u>notation</u>: The set of all nfas that can be defined over a given alphabet is denoted as $nfa(\Sigma)$.

▶ **<u>Graph representation of nfa</u>**

A non-deterministic finite accepter can be represented by a directed graph according to the following definition.

> <u>Def</u>: Let $M = (Q, \Sigma, \delta, q_0, F)$ be an nfa. We define the directed graph $G = Graph(M)$ with
> a) Set of vertices $V(G) = Q$
> b) Set of edges $E(G) = \{(q_1, q_2, a) \in Q \times Q \times (\Sigma \cup \{\lambda\}) : q_2 \in \delta(q_1, a)\}$
> c) Incidence function $\psi_G : E(G) \to V(G) \times V(G)$ given by
> $\forall e = (q_1, q_2, a) \in E(G) : \psi_G(e) = (q_1, q_2)$

<u>notation</u>

a) The edge (q_1, q_2, a) represents a transition from q_1 to q_2 upon processing the character a. The character "a" labels the edge.

b) Otherwise, to draw G we use the same conventions used for directed graph representations of dfas.

<u>Remark</u>: Note that, unlike directed graphs representing dfas, directed graphs of nfas can have λ-edges.

<u>Def</u> : Let $G = Graph(M)$ be the directed graph representing the nfa $M = (Q, \Sigma, \delta, q_0, F)$. Let $w \in W(G)$ be a walk on G. The string $\sigma(w) \in (\Sigma \cup \{\lambda\})^*$ induced by the walk w is defined as

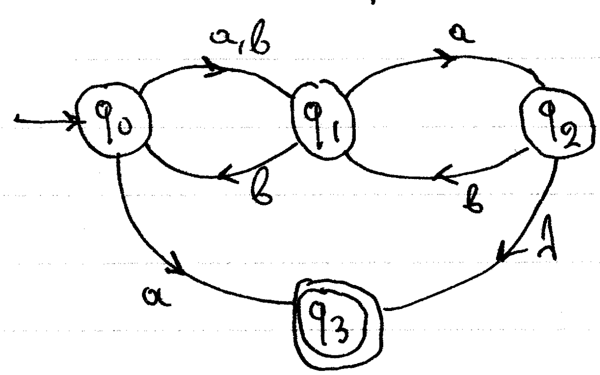
$$\forall n \in [|w|] : ([\sigma(w)]_n = a \Leftrightarrow \exists q_1, q_2 \in Q : e_n(w) = (q_1, q_2, a))$$

<u>Remark</u> : Because $\sigma(w) \in (\Sigma \cup \{\lambda\})^*$ may contain the null symbol λ, it can be simplified to a shorter string $\sigma^*(w) \in \Sigma^*$ by deleting all null symbols.

EXAMPLE

Consider the nfa represented by the following directed graph:

G:



Then $Q = \{q_0, q_1, q_2, q_3\}$ and $\Sigma = \{a, b\}$ and $F = \{q_3\}$ and δ is defined as

$\delta(q_0, \lambda) = \emptyset$ $\quad$ $\delta(q_1, \lambda) = \emptyset$ $\quad$ $\delta(q_2, \lambda) = \{q_3\}$

$\delta(q_0, a) = \{q_1, q_3\}$ $\quad$ $\delta(q_1, a) = \{q_2\}$ $\quad$ $\delta(q_2, a) = \emptyset$

$\delta(q_0, b) = \{q_1\}$ $\quad$ $\delta(q_1, b) = \{q_0\}$ $\quad$ $\delta(q_2, b) = \{q_1\}$

$\delta(q_3, \lambda) = \emptyset$ $\quad$ $\delta(q_3, a) = \emptyset$ $\quad$ $\delta(q_3, b) = \emptyset$

Consider the walk w given by
$w = (q_0, (q_0, q_1, b), q_1, (q_1, q_2, a), q_2, (q_2, q_3, \lambda), q_3)$
Then $\sigma(w) = ba\lambda$ and $\sigma^*(w) = ba$.

▸ The extended transition function

A unique feature of nfas is that the transition function δ allows transitions between internal states via the null string λ. This complicates the definition of the extended transition function δ^*. Given a string $u \in \Sigma^*$ and an initial internal state $q \in Q$, we wish to define $\delta^*(q, u)$ as the set of all states that can be reached from q following either null edges or edges that follow and consume the string u character by character. Because the nfa is non-deterministic, it is possible that from some state in Q, there are multiple possible next states corresponding to a given character in Σ. The first step is to define a mapping $\Lambda : \mathcal{P}(Q) \to \mathcal{P}(Q)$ that gives the set of all states that can be reached from some set of states $P \in \mathcal{P}(Q)$ following a finite sequence of null edges. Then the Λ function can be used to define the extended transition function δ^*. Both functions are defined recursively as follows:

<u>Def</u>: Let $M=(Q,\Sigma,\delta,q_0,F)$ be an nfa. We define:

a) The <u>λ-closure function</u> $\Lambda:\mathcal{P}(Q)\to\mathcal{P}(Q)$ is defined recursively as follows:

$$\begin{cases}\forall P\in\mathcal{P}(Q): \Lambda_0(P)=\{q\in Q\mid \exists q_1\in P: q\in\delta(q_1,\lambda)\}\\ \forall n\in\mathbb{N}^*:\forall P\in\mathcal{P}(Q): \Lambda_n(P)=\Lambda_0(\Lambda_{n-1}(P))\\ \forall P\in\mathcal{P}(Q): \Lambda(P)=\bigcup_{n\in\mathbb{N}}\Lambda_n(P)\end{cases}$$

b) The <u>extended transition function</u> $\delta^*:Q\times\Sigma^*\to\mathcal{P}(Q)$ is also defined recursively as follows:

$$\begin{cases}\forall q\in Q: \delta^*(q,\lambda)=\Lambda(\{q\})\\ \forall q\in Q:\forall u\in\Sigma^*:\forall a\in\Sigma: \delta^*(q,ua)=\Lambda\left(\bigcup_{p\in\delta^*(q,u)}\delta(p,a)\right)\end{cases}$$

▶ <u>Language accepted by an nfa</u>

<u>Def</u>: Let $M=(Q,\Sigma,\delta,q_0,F)$ be an nfa with extended transition function $\delta^*:Q\times\Sigma^*\to\mathcal{P}(Q)$. The language $\mathcal{L}(M)$ accepted by the nfa M is:

$$\mathcal{L}(M)=\{u\in\Sigma^*\mid \delta^*(q_0,u)\cap F\neq\varnothing\}$$

<u>Remark</u>: An equivalent way to define the language $\mathcal{L}(M)$ accepted by M is via the graph representation $G=\text{Graph}(M)$ of M. Intuitively, a string $u\in\Sigma^*$ is accepted by M, if and only if there is a walk from q_0 to some final state $q\in F$ such that $\sigma^*(w)=u$.

Using quantifiers, we write:

$$\boxed{u \in \mathcal{L}(M) \iff \exists q \in F : \exists w \in W(\mathrm{Graph}(M) \mid q_0, q) : \sigma^{+}(w) = u}$$

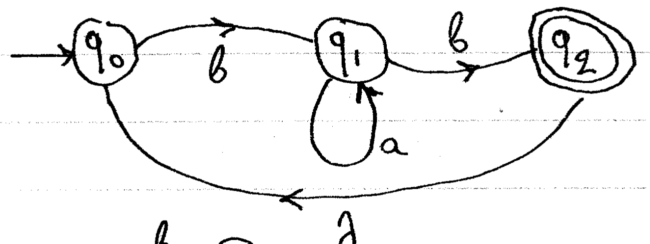
The belonging condition can be also rewritten using set builder notation as follows:

$$\begin{aligned} \mathcal{L}(M) &= \{ u \in \Sigma^* \mid \exists w \in \bigcup_{q \in F} W(\mathrm{Graph}(M) \mid q_0, q) : \sigma^*(w) = u \} \\ &= \{ \sigma^*(w) \mid w \in \bigcup_{q \in F} W(\mathrm{Graph}(M) \mid q_0, q) \} \end{aligned}$$

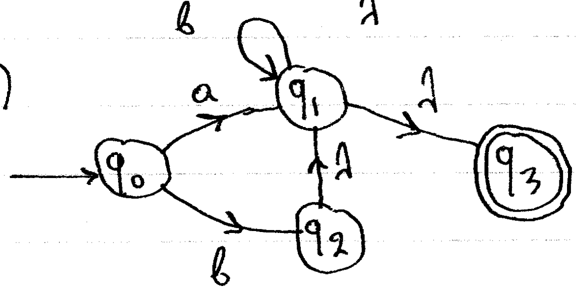
EXERCISES

(20) Write the formal definition for the following non-deterministic finite accepters:

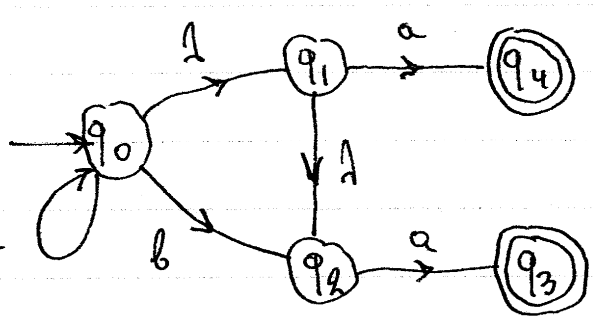
a)



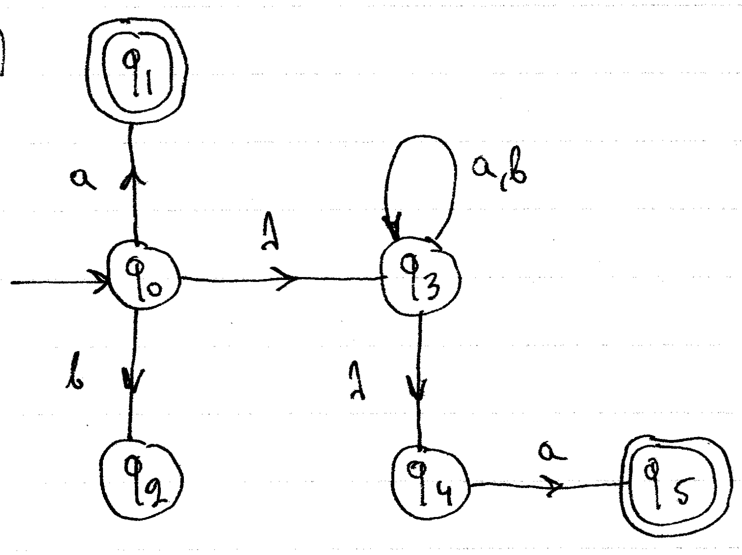
b)



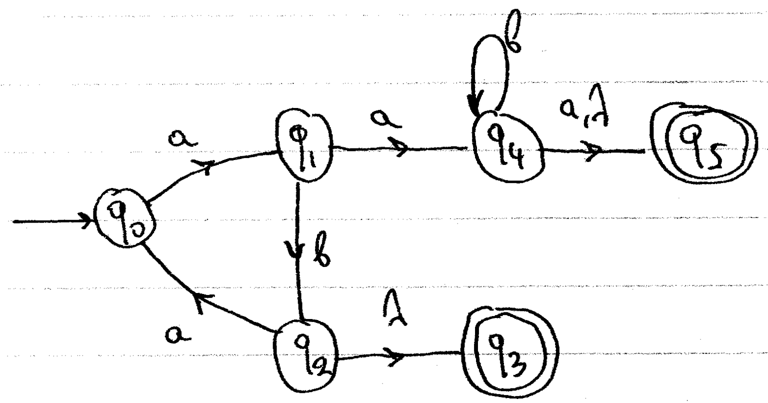
c)



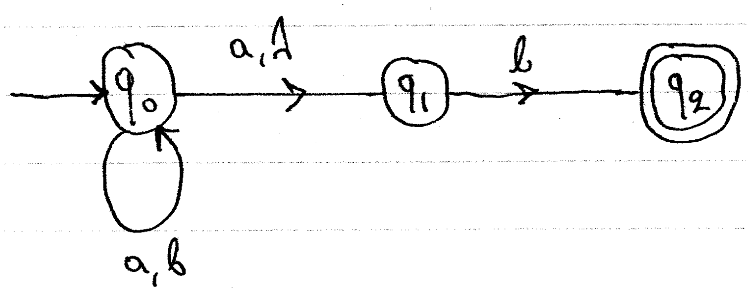
d)



e)



(21) Consider the non-deterministic finite accepter represented by



Evaluate

a) $\delta^*(q_0, aaa)$

b) $\delta^*(q_0, ba^2)$

c) $\delta^*(q_0, a^2bab)$

d) $\delta^*(q_0, a^n b^n)$ for $n \in \mathbb{N}$

e) $\delta^*(q_0, b^n a^n)$ for $n \in \mathbb{N}$

(Hint: For (d),(e) you will need to use method of induction as part of a wider argument)

(22) Given a alphabet Σ, show that for all $M \in nfa(\Sigma)$ there is at least one $M_0 \in nfa(\Sigma)$ that has exactly one final state such that $\mathcal{L}(M) = \mathcal{L}(M_0)$.

(23) Show that all finite languages are regular.
(Hint: Construct an appropriate nfa. Use induction on the cardinality of the language in question)

Equivalence of nfa and dfa

Let Σ be an alphabet. Given a deterministic finite accepter $M_2 \in dfa(\Sigma)$ with $M_2 = (Q_2, \Sigma, \delta_2, q_{02}, F_2)$ we can easily define an equivalent non-deterministic finite accepter $M_1 \in nfa(\Sigma)$ with $M_1 = (Q_1, \Sigma, \delta_1, q_{01}, F_1)$ such that $\mathcal{L}(M_1) = \mathcal{L}(M_2)$ by choosing:

$$\begin{cases} Q_1 = Q_2 \land q_{01} = q_{02} \land F_1 = F_2 \\ \forall q \in Q_1 : \forall a \in \Sigma : \delta_1(q,a) = \{\delta_2(q,a)\} \\ \forall q \in Q_1 : \delta_1(q,\lambda) = \{q\} \end{cases}$$

We can then claim that

$$\forall M_2 \in dfa(\Sigma) : \exists M_1 \in nfa(\Sigma) : \mathcal{L}(M_1) = \mathcal{L}(M_2)$$

We will now provide an algorithm for converting an nfa to a dfa, which in turn establishes the converse statement:

$$\forall M_1 \in nfa(\Sigma) : \exists M_2 \in dfa(\Sigma) : \mathcal{L}(M_1) = \mathcal{L}(M_2).$$

Algorithm: (nfa to dfa)

Let $M_1 \in nfa(\Sigma)$ be an nfa with $M_1 = (Q_1, \Sigma, \delta_1, q_{01}, F_1)$. We construct an equivalent dfa $M_2 \in dfa(\Sigma)$ with $M_2 = (Q_2, \Sigma, \delta_2, q_{02}, F_2)$ with $Q_2 \subseteq \mathcal{P}(Q_1)$ as follows:

a) Define $q_{02} = \{q_{01}\}$ (initial state) and define $Q_{2,0} = \{q_{02}\}$ as an initial approximation of Q_2.

b) Starting from $n=0$, assume we have worked our way to $Q_{2,n}$, and have defined δ partially. Find an element

$q \in Q_{2,n}$ and $a \in \Sigma$ for which $\delta_2(q,a)$ is undefined.

Define:

$$\begin{cases} \delta_2(q,a) = \bigcup_{s \in q} \delta_1^*(s,a) \\ Q_{2,n+1} = Q_{2,n} \cup \{\delta_2(q,a)\} \end{cases}$$

c) Repeat the previous step until

$\forall q \in Q_{2,n} : \forall a \in \Sigma : (\delta_2(q,a) \text{ has been defined})$

Then set $Q_2 = Q_{2,n}$ and note that δ_2 is completely defined as well.

d) Define the set of final states F_2 as follows:

$F_2 = \{q \in Q_2 \mid \exists s \in q : s \in F_1\}$

↳ While running the above algorithm, we note that it is possible to find $\delta_2(q,a) = \emptyset$ for some $q \in \mathcal{P}(Q_1)$ and $a \in \Sigma$. Then $\emptyset$ will be a state of the dfa and the above algorithm will then result in a definition of δ_2 such that

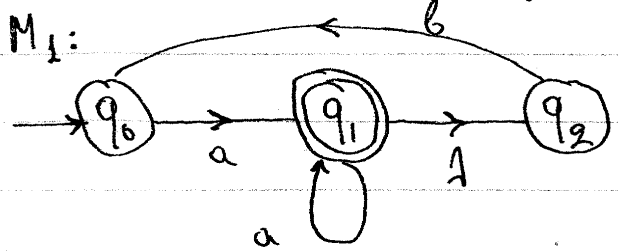
$\forall a \in \Sigma : \delta_2(\emptyset, a) = \emptyset$

It follows that if $\emptyset \in Q_2$, then the internal state $\emptyset$ will function as a "trap state" in that it is impossible to transition from $\emptyset$ to other states.

EXAMPLE

Define a dfa that is equivalent to the following nfa:

M_1:



Solution

We begin with $Q_{20} = \{\{q_0\}\}$. Then

$$\delta_2(\{q_0\}, a) = \bigcup_{\xi \in \{q_0\}} \delta_1^*(\xi, a) = \delta_1^*(\{q_0\}, a) = \{q_1, q_2\}$$

Let $Q_{21} = \{\{q_0\}, \{q_1, q_2\}\}$

$$\delta_2(\{q_0\}, b) = \bigcup_{\xi \in \{q_0\}} \delta_1^*(\xi, b) = \delta_1^*(q_0, b) = \emptyset$$

Let $Q_{22} = \{\{q_0\}, \{q_1, q_2\}, \emptyset\}$

$$\delta_2(\{q_1, q_2\}, a) = \bigcup_{\xi \in \{q_1, q_2\}} \delta_1^*(\xi, a) = \delta_1^*(q_1, a) \cup \delta_1^*(q_2, a)$$

$$= \{q_1, q_2\} \cup \emptyset = \{q_1, q_2\}.$$

$$\delta_2(\{q_1, q_2\}, b) = \bigcup_{\xi \in \{q_1, q_2\}} \delta^*(\xi, b) =$$

$$= \delta_1^*(q_1, b) \cup \delta_1^*(q_2, b) =$$

$$= \{q_0\} \cup \{q_0\} = \{q_0\}$$

$\delta_2(\emptyset, a) = \emptyset$

$\delta_2(\emptyset, b) = \emptyset$

At this point the algorithm terminates with

$Q_2 = Q_{22} = \{\{q_0\}, \{q_1, q_2\}, \emptyset\}$

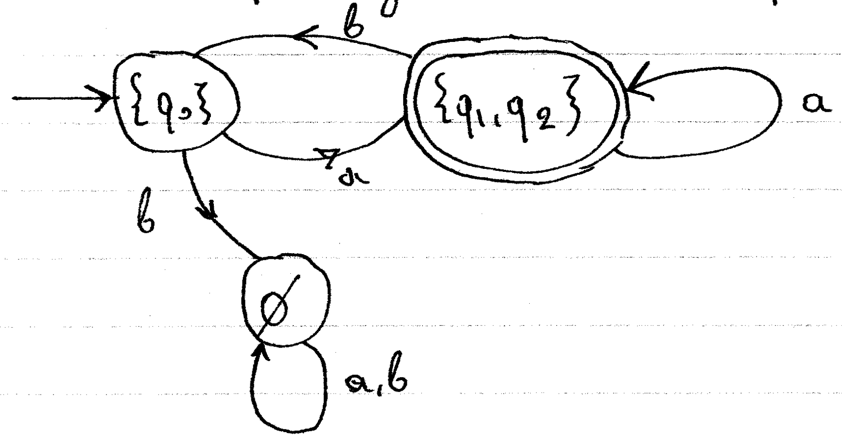
The set of final states is

$F_2 = \{q \in Q_2 \mid \exists \xi \in q : \xi \in F_1\}$

$= \{q \in \{\{q_0\}, \{q_1, q_2\}, \emptyset\} \mid \exists \xi \in q : \xi \in \{q_1\}\} =$

$= \{\{q_1, q_2\}\}$

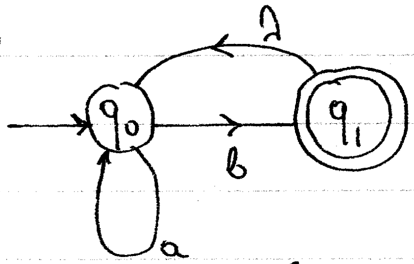
The corresponding dfa M_2 has representation



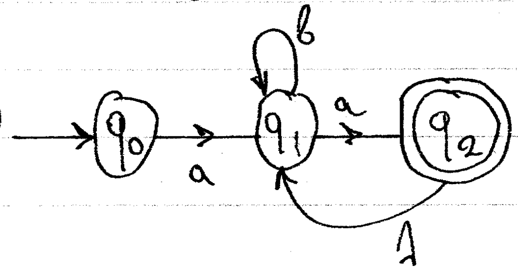
EXERCISES

(24) Define dfas that are equivalent to the following nfas and show the details of the constructions:

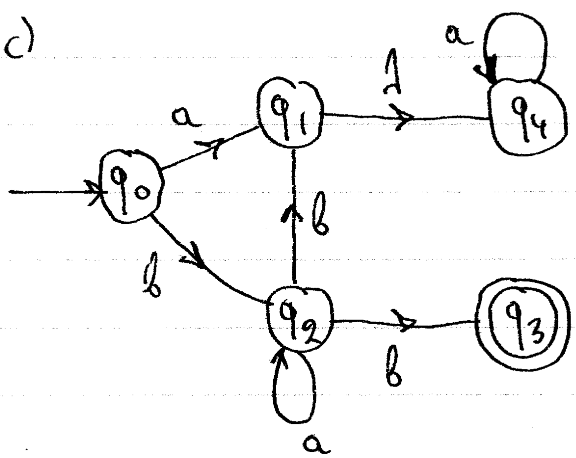
a)



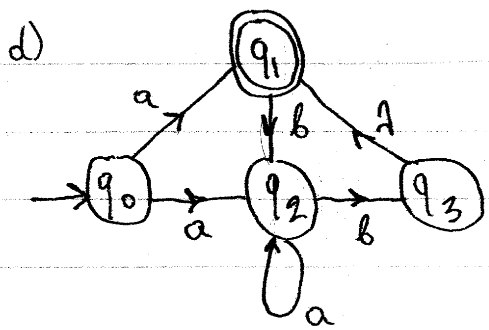
b)



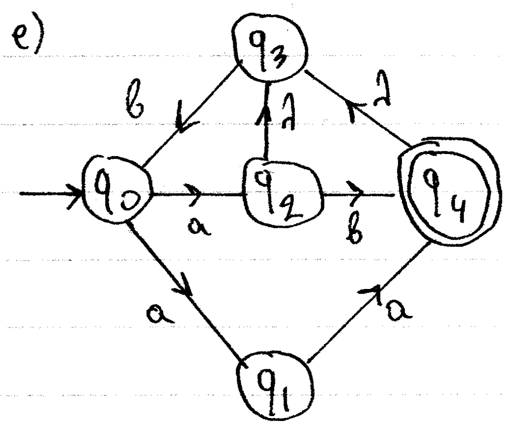
c)



d)



e)



Regular Expressions

Regular expressions can be used to provide a concise representation of regular languages. It is also simple to deduce a corresponding nfa from a regular expression and then convert it to a dfa. We use recursion to define regular expressions and the language induced by a regular expression, as follows:

<u>Def</u>: Let Σ be an alphabet. The set $Reg(\Sigma)$ of all regular expressions is defined as follows:
a) $\emptyset$, λ, and all $a \in \Sigma$ are regular expressions
b) Given $r_1, r_2 \in Reg(\Sigma)$, $r_1 \vee r_2$, $r_1 r_2$, $r_1 *$, (r_1) are also regular expressions
c) We build $Reg(\Sigma)$ by combining (a) and (b) in a finite number of steps.

A formal definition of the set $Reg(\Sigma)$ of all regular expressions can be given via a formal grammar:

<u>Def</u>: Let Σ be an alphabet and consider a grammar G with variables $V = \{S\}$, alphabet $\Sigma \cup \{\emptyset, (,), *\}$ and the following production rules:

$$\begin{cases} \forall a \in \Sigma: \quad S \to a \\ S \to \emptyset \mid \lambda \mid S \vee S \mid S S \mid S * \mid (S) \end{cases}$$

Then, we define $Reg(\Sigma) = \mathcal{L}(G)$.

<u>Example</u>: Given the alphabet $\Sigma = \{a, b\}$, the following strings are possible regular expressions:

$(ab)*$

$a \vee (b*)$

$(a*)(b*) \vee (ba)*$

<u>Def</u>: Let Σ be an alphabet and $r \in Reg(\Sigma)$ a regular expression. The language $\mathcal{L}(r)$ defined by the regular expression r is given recursively, according to the following rules:

(a) $\mathcal{L}(\emptyset) = \emptyset$

(b) $\mathcal{L}(\lambda) = \{\lambda\}$

(c) $\forall a \in \Sigma: \mathcal{L}(a) = \{a\}$

(d) $\forall r_1, r_2 \in Reg(\Sigma): \mathcal{L}(r_1 \vee r_2) = \mathcal{L}(r_1) \cup \mathcal{L}(r_2)$

(e) $\forall r_1, r_2 \in Reg(\Sigma): \mathcal{L}(r_1 r_2) = \mathcal{L}(r_1)\mathcal{L}(r_2)$

$$= \{uv \mid u \in \mathcal{L}(r_1) \wedge v \in \mathcal{L}(r_2)\}$$

(f) $\forall r \in Reg(\Sigma): \mathcal{L}((r)) = \mathcal{L}(r)$

(g) $\forall r \in Reg(\Sigma): \mathcal{L}(r*) = [\mathcal{L}(r)]^* = \bigcup_{n \in \mathbb{N}} [\mathcal{L}(r)]^n$

<u>Def</u>: (Equivalent regular expressions)

Let $u \in \Sigma$ be an alphabet and let $r_1, r_2 \in Reg(\Sigma)$.

We say that

$$r_1 \equiv r_2 \Leftrightarrow \mathcal{L}(r_1) = \mathcal{L}(r_2)$$

<u>Remark</u> : To minimize ambiguity in applying the above rules we adopt the following precedence rules:

(a) $*$ takes precedence over all operations
(e.g. $ab* \equiv a(b*)$, $a \vee b* \equiv a \vee (b*)$

(b) Concatenation takes precedence over disjunction
(e.g. $ab \vee c \equiv (ab) \vee c$).

EXAMPLE

a) Define the language defined by the regular expression $r = (ab)*$

Solution

$$\mathcal{L}(r) = \mathcal{L}((ab)*) = [\mathcal{L}(ab)]^* = [\mathcal{L}(a)\mathcal{L}(b)]^* =$$
$$= [\{a\}\{b\}]^* = \{ab\}^* = \bigcup_{n\in\mathbb{N}} \{ab\}^n =$$
$$= \{(ab)^n \mid n \in \mathbb{N}\}.$$

b) Define the language defined by the regular expression $r = a*(a \vee b)$

Solution

$$\mathcal{L}(r) = \mathcal{L}(a*(a\vee b)) = \mathcal{L}(a*)\mathcal{L}(a\vee b) = \mathcal{L}(a^*)[\mathcal{L}(a) \cup \mathcal{L}(b)]$$
$$= \{a\}^*[\{a\}\cup\{b\}] = \{a\}^*\{a,b\} =$$
$$= \{a^n \mid n \in \mathbb{N}\}\{a,b\} =$$
$$= \{a^{n+1}, a^n b \mid n \in \mathbb{N}\}.$$

c) Let $L = \{a^{2n} b^{2m+1} \mid n\in\mathbb{N} \wedge m\in\mathbb{N}\}$ be a language. Find a regular expression r that generates the language L.

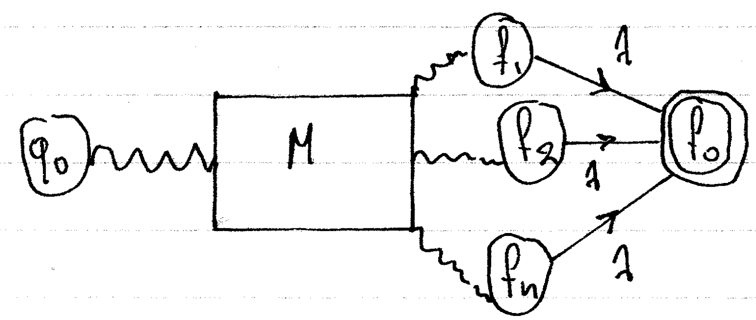
Solution

$$L = \{a^{2n} b^{2m+1} \mid n\in\mathbb{N} \wedge m \in\mathbb{N}\} =$$
$$= \{a^{2n} \mid n\in\mathbb{N}\}\{b^{2m} \mid n\in\mathbb{N}\}\{b\} =$$
$$= \{(aa)^n \mid n\in\mathbb{N}\}\{(bb)^m \mid m\in\mathbb{N}\}\{b\}$$
$$= \{aa\}^*\{bb\}^*\{b\} = [\mathcal{L}(aa)]^*[\mathcal{L}(bb)]^*\mathcal{L}(b)$$
$$= \mathcal{L}((aa)*)\mathcal{L}((bb)*)\mathcal{L}(b) = \mathcal{L}((aa)*(bb)*b) \Rightarrow$$
$$\Rightarrow r \equiv (aa)*(bb)*b.$$

➤ Constructing an nfa from a regular expression

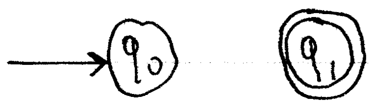
Given a regular expression $r \in Reg(\Sigma)$, the problem is to construct an nfa $M \in nfa(\Sigma)$ such that $\mathcal{L}(r) = \mathcal{L}(M)$.

Remark: With no loss of generality we can assume that our nfas have a unique final state. In any other nfa with multiple final states $f_1, f_2, \ldots, f_n \in F$, we can simply create a new final state f_0 and connect it with $f_1, f_2, \ldots, f_n$ using λ-edges:

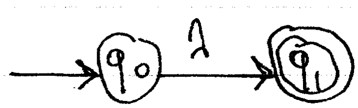


Algorithm: The needed nfa can be constructed recursively as follows:

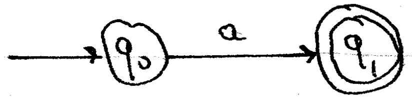
a) For $r = \emptyset$, the corresponding nfa is:



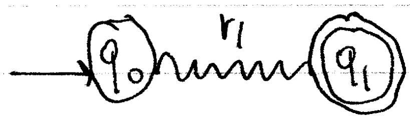
b) For $r = \lambda$, the corresponding nfa is



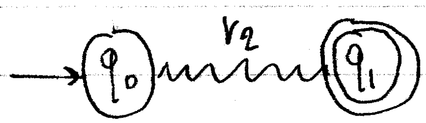
c) For $r = a$ with $a \in \Sigma$, the corresponding nfa is:



d) <u>Recursive cases</u>: Let us assume that we have already constructed the following nfas for the regular expressions $r_1, r_2 \in Reg(\Sigma)$:



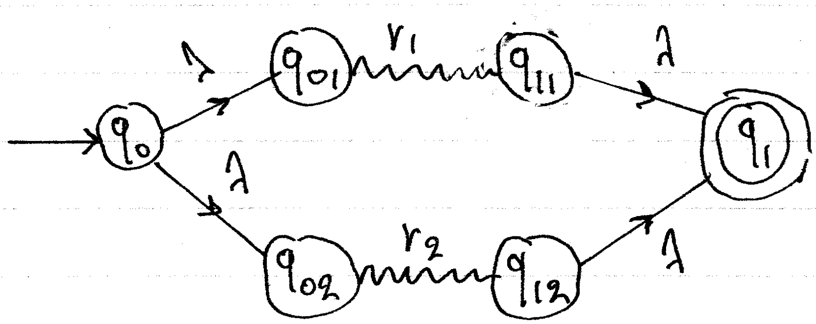
$\longleftarrow$ nfa for $\mathcal{L}(r_1)$



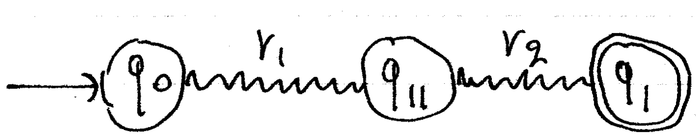
$\longleftarrow$ nfa for $\mathcal{L}(r_2)$

We may then construct rules for $r_1 \vee r_2$, $r_1 r_2$, r_1^* as follows:

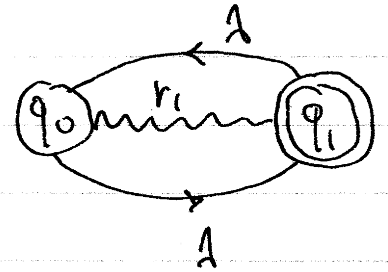
- 1) For $r_1 \vee r_2$, the corresponding nfa is:



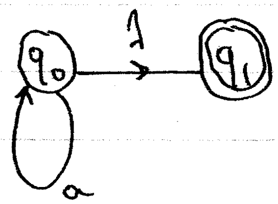
- 2) For $r_1 r_2$ the corresponding nfa is:



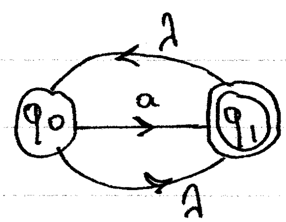
- For r_1^*, the corresponding nfa is:



↳ <u>Special case:</u> For $r = a^*$ with $a \in \Sigma$, the corresponding nfa can be written as:



as an alternative to



EXAMPLES

Construct nfas that accept the following regular expressions:

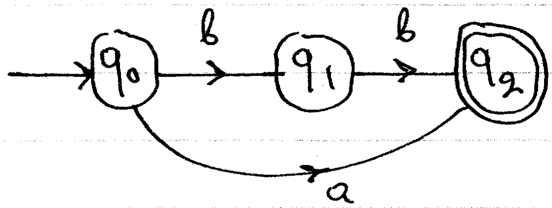
a) $r = a \vee (bb)$

b) $r = (ab^*a) \vee \lambda$

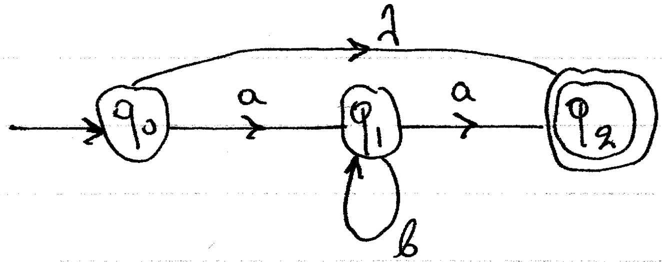
c) $r = (a \vee b)^* (ba^*)$

Solution

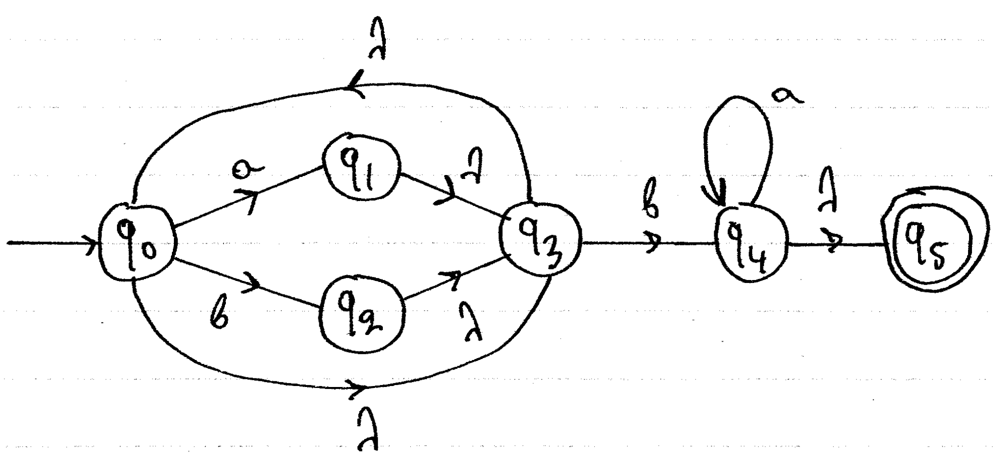
a) For $r = a \vee (bb)$



b) For $r = (ab^*a) \vee \lambda$



c) For $r = (a \vee b)^* (ba^*)$



EXAMPLE

Construct an nfa that accepts the language
$L=\{ab^n a^m b \mid n\in\mathbb{N} \wedge m\in\mathbb{N}\}$

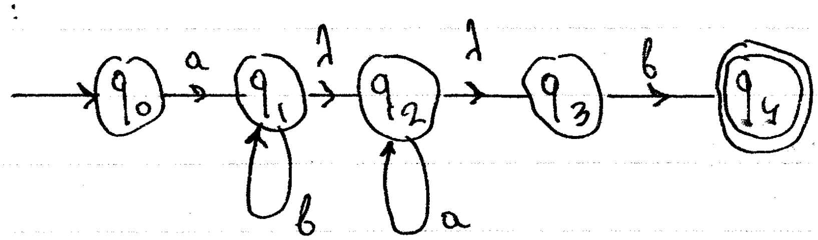
Solution

We note that

$$L=\{ab^n a^m b \mid n\in\mathbb{N} \wedge m\in\mathbb{N}\}=$$
$$=\{a\}\{b^n \mid n\in\mathbb{N}\}\{a^m \mid m\in\mathbb{N}\}\{b\}$$
$$=\{a\}\{b\}^*\{a\}^*\{b\}=$$
$$=\mathcal{L}(a)\mathcal{L}(b*)\mathcal{L}(a*)\mathcal{L}(b)=$$
$$=\mathcal{L}(ab*a*b)$$

and therefore the corresponding nfa is:

M:



EXERCISES

(25) Write the language accepted by the following expressions in set builder notation and construct an non-deterministic finite accepter that accepts that language

a) $r = (aba)*$

b) $r = ab(ba)*$

c) $r = b*a*\,b^2 a*$

d) $r = (ab)* \vee a*$

e) $r = (aa*) \vee (ba*b*)$

f) $r = (bab)* \vee (a \vee b*)*$

(26) Construct non-deterministic finite accepters that accept the following regular expressions and convert them to dfas.

a) $r = (ab)* \vee a$

b) $r = (a \vee b)*\,b*$

c) $r = (a* \vee (ba))*$

d) $r = [(ab)*] \vee a*$

e) $r = (a*bab*)*$

(27) Use regular expressions to construct non-deterministic finite accepters that accept the following languages, thereby establishing that they are regular.

a) $L = \{x^a y^b z^c \mid a, b, c \in \mathbb{N}\}$

b) $L = \{x^{a+2} y^b \mid a, b \in \mathbb{N}\}$

c) $L = \{xyx^a, x^b y^{a+2} \mid a, b \in \mathbb{N}\}$

d) $L = \{x^2 y x^{a+1} y^b \mid a, b \in \mathbb{N}\}$

e) $L = \{x^2 y^a, x^{a+1} y^a \mid a \in \mathbb{N}\}$

DST7: Turing machines

TURING MACHINES

Definition of Turing machine

Turing machines are believed to be the most powerful generalization of dfa/nfa automata that is possible. Modern computers are, in principle, reducable to Turing machines. The definition of the Turing machine is done in 3 steps:

- 1 We define the machine itself
- 2 We define the "tape", i.e. the input/output device used by the machine
- 3 We define the process by which the machine converts its input into output.

Def: A Turing machine M is defined as $M = (Q, \Sigma, \Gamma, \delta, q_0, B, F)$ where:

a) Q is a finite set of internal states

b) Σ is a finite set, the input alphabet

c) Γ is a finite set, the tape alphabet

d) δ is a transition function $\delta : Q \times \Gamma \to [Q \times \Gamma \times \{L, R\}] \cup \{H\}$ where L, R, H are fixed symbols.

e) $q_0 \in Q$ is an initial state.

f) $B \in \Gamma$ is the blank symbol

g) $F \subseteq Q$ is a set of final states

such that $\Sigma \subseteq \Gamma - \{B\}$ and $\forall q \in F : \forall a \in \Gamma : \delta(q, a) = H$

<u>notation</u> : $Tur(\Sigma)$ will denote the set of all Turing machines that can be defined on some input alphabet Σ.

<u>Remarks</u>

a) The symbols R, L are instructions to the attached tape device (to be defined below) that instruct the header reading the tape to move <u>right</u> or <u>left</u>. The symbol H corresponds to a command to halt the computation.

b) It is assumed that if the machine finds itself in a final state $q \in F$, then the transition mapping δ will halt the computation. However it is not necessary to reach a final state for the computation to halt.

The tape device

Attached to a Turing machine is an input/output device that we will call "tape". Informally, we envision the tape as follows:

a) The tape is a one-dimensional storage device of symbols from Γ with infinite length in either direction.

b) The Turing machine itself is envisioned as a <u>header</u> that points to some symbol somewhere on the tape. The header also has an internal state $q \in Q$.

c) If $q \in Q$ is the state of the Turing machine and $a \in \Gamma$ the symbol on the tape currently under the machine the

1) $\delta(q,a) = (p,b,L)$ means that the machine will replace a with b on the tape, transition from state q to p, then move the machine left.
2) $\delta(q,a) = (p,b,R)$ means that the machine will replace a with b on the tape, transition from state q to p, then move the machine right
3) $\delta(q,a) = H$ means that the machine terminates the algorithm

<u>Def</u>: Let $M = (Q,\Sigma,\Gamma,\delta,q_0,B,F) \in \text{Tur}(\Sigma)$ be a Turing machine

a) A string $u = XqaY$ with $X,Y \in \Gamma^*$ and $q \in Q$ and $a \in \Gamma$ is a <u>configuration</u> of the Turing machine where the tape content is XaY and the header is pointing at the character a while in internal state q.

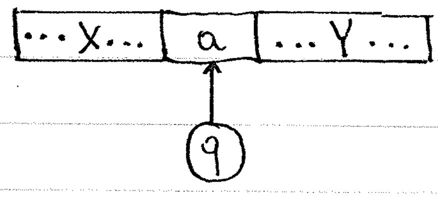
b) A string $u = XaY$ with $X,Y \in \Gamma^*$ and $a \in \Gamma$ is a configuration of the Turing machine, where the tape content is XaY and the machine is in a <u>halted</u> state (i.e. the header is unmounted from the tape).

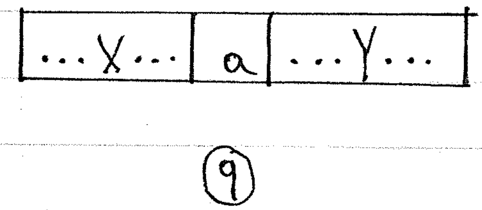
c) The set of all possible configurations is:

$$\text{config}(M) = \{XqaY, XaY \mid X,Y \in \Gamma^* \wedge q \in Q \wedge a \in \Gamma\}$$
$$= (\Gamma^* Q \Gamma^+) \cup (\Gamma^+)$$

<u>Remark</u>: The configurations described above can be visually represented as follows:

$\$ = XqaY$ $\qquad\qquad$ $\$ = XaY$





The operation of the Turing machine is defined formally via a deterministic transition function $\Delta : \text{config}(M) \to \text{config}(M)$ as follows:

<u>Def</u>: Let $M = (Q, \Sigma, \Gamma, \delta, q_0, B, F) \in \text{Tur}(\Sigma)$ be a Turing machine. We define:

the deterministic <u>configuration transition function</u> $\Delta : \text{config}(M) \to \text{config}(M)$ as follows:

$$\forall X, Y \in \Gamma^* : \forall q \in Q : \forall a, b \in \Gamma : \; : \begin{cases} \delta(q,a) = (p,c,R) \Rightarrow \Delta(XqabY) = XcpbY \\ \delta(q,a) = (p,c,L) \Rightarrow \Delta(XaqbY) = XpacY \end{cases}$$

$$\forall X \in \Gamma^* : \forall q \in Q : \forall a \in \Gamma : \; : \begin{cases} \delta(q,a) = (p,c,R) \Rightarrow \Delta(Xqa) = XcpB \\ \delta(q,a) = (p,c,L) \Rightarrow \Delta(qaX) = pBcX \end{cases}$$

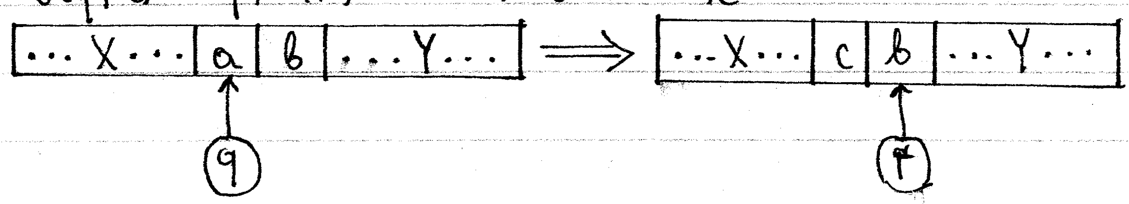
$$\forall X, Y \in \Gamma^* : \forall q \in Q : \forall a \in \Gamma : \delta(q,a) = H \Rightarrow \Delta(XqaY) = XaY$$

$$\forall P \in \Gamma^+ : \Delta(P) = P$$

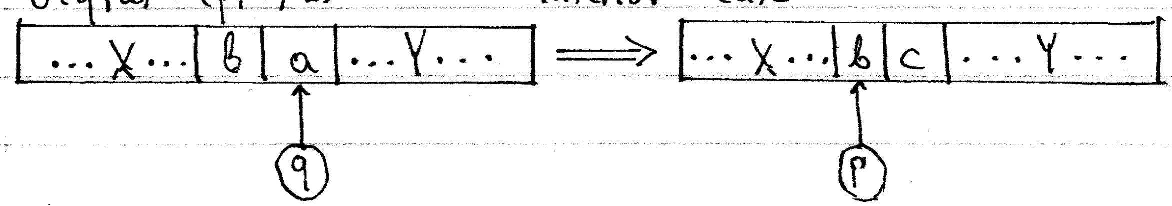
<u>Remark</u>: Note that we distinguish between configurations where the Turing machine is scanning the beginning or end of the tape string vs configurations where the Turing machine is scanning the interior of the tape string. In the first case it becomes necessary to utilize the "block" character B. Once the machine is halted, the configuration transition function Δ merely returns the content of the tape.

<u>Remark</u>: A graphical representation of the transitions accounted by the previous definition is given below:

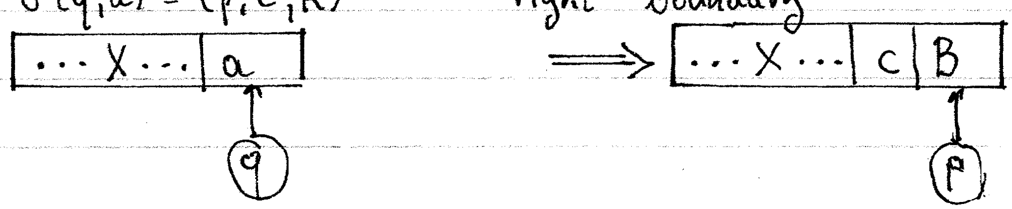
$\delta(q,a) = (p,c,R)$ — interior case



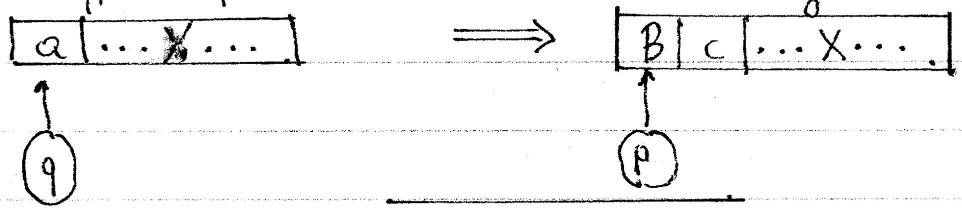
$\delta(q,a) = (p,c,L)$ — interior case



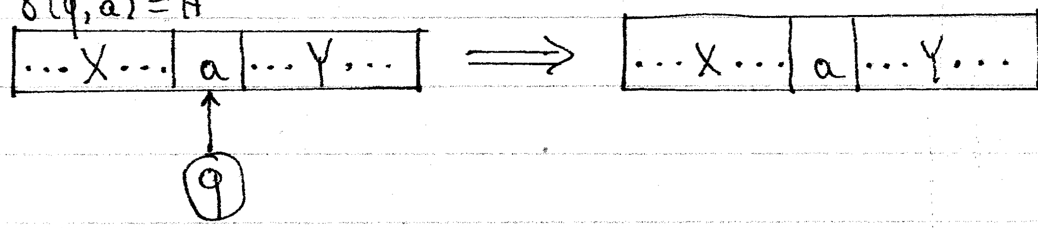
$\delta(q,a) = (p,c,R)$ — right boundary



$\delta(q,a) = (p,c,L)$ — left boundary case



$\delta(q,a) = H$



Configuration transitions and halting states

Turing machines are deterministic. Consequently, given an initial configuration, all subsequent configurations are predetermined. Eventually, the machine may reach a halting state. It is also possible that the machine never reaches a halting state or may even find itself in an infinite loop. We give the relevant definitions below:

Def: Let $M \in \mathrm{Tur}(\Sigma)$ be a Turing machine with configuration transition function $\Delta : \mathrm{config}(M) \to \mathrm{config}(M)$. Let $n \in \mathbb{N}^*$. We define the n-step configuration transition function $\Delta_n : \mathrm{config}(M) \to \mathrm{config}(M)$ recursively as:

$$\forall \$ \in \mathrm{config}(M) : \Delta_1(\$) = \Delta(\$)$$

$$\forall n \in \mathbb{N} - \{0,1\} : \forall \$ \in \mathrm{config}(M) : \Delta_n(\$) = \Delta(\Delta_{n-1}(\$))$$

Note that a halted configuration $\$ \in \text{config}(M)$ will satisfy $\$ \in \Gamma^+$, whereas a non-halted configuration will satisfy $\$ \in \Gamma^* Q \Gamma^+$. We may therefore give the following definitions:

Def: Let $M \in \text{Tur}(\Sigma)$ be a Turing machine and let $\$_1, \$_2 \in \text{config}(M)$ be two machine configurations. We say that

$$\$_1 \vdash \$_2 \iff \Delta(\$_1) = \$_2$$
$$\$_1 \vdash^* \$_2 \iff \exists n \in \mathbb{N}^* : \Delta_n(\$_1) = \$_2$$
$$\$_1 \vdash^* \infty \iff \forall n \in \mathbb{N}^* : \Delta_n(\$_1) \notin \Gamma^*$$
$$\$_1 \vdash^* \text{loop} \iff \exists n, m \in \mathbb{N}^* : (n \neq m \land \Delta_n(\$_1) = \Delta_m(\$_1))$$

interpretation

a) $\$_1 \vdash \$_2$ means that the machine will transition from configuration $\$_1$ to configuration $\$_2$ in one step.

b) $\$_1 \vdash^* \$_2$ means that the machine will transition from configuration $\$_1$ to configuration $\$_2$ in a finite number of steps.

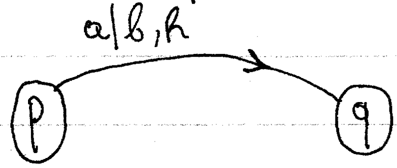
c) $\$_1 \vdash^* \infty$ means that once initialized with the configuration $\$_1$, the machine will never reach a halt state in subsequent steps.

d) $\$_1 \vdash^* \text{loop}$ means that once initialized with the configuration $\$_1$, the machine will enter an infinite loop where it cycles through a finite sequence of configurations infinite times.

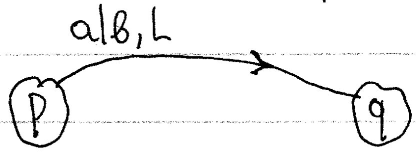
Graph representation of Turing machines

Turing machines can be represented as directed graphs according to the following conventions:

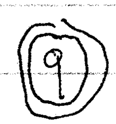
1) For every transition rule $\delta(p,a)=(q,b,R)$ we have the representation



Likewise, for the transition rule $\delta(p,a)=(q,b,L)$ we have the representation



2) Final internal states are denoted via a double oval as:



and according to the Turing machine definition, they have no outgoing arrows

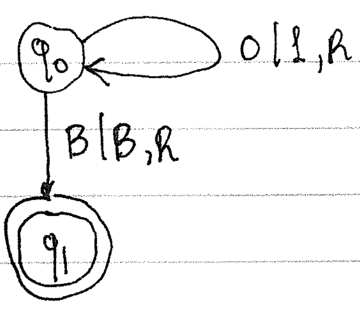
3) Transitions of the form $\delta(p,a)=H$ are not shown graphically. The absense of a needed outgoing arrow indicates that the machine will halt.

EXAMPLE

Consider a Turing machine $M = (Q, \Sigma, \Gamma, \delta, q_0, B, F)$ with $Q = \{q_0, q_1\}$ and $\Sigma = \{0, 1\}$ and $\Gamma = \{0, 1, B\}$ and $F = \{q_1\}$ and transition rules

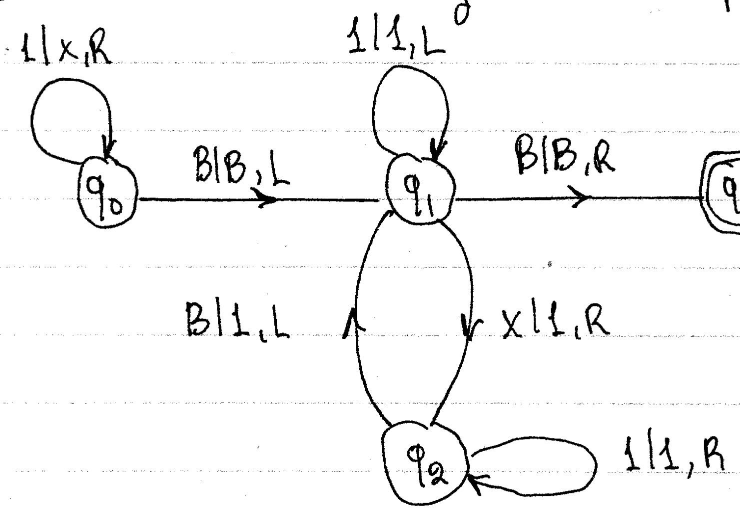
$$\delta(q_0, 0) = (q_0, 1, R)$$
$$\delta(q_0, B) = (q_1, B, R)$$

The corresponding graphical representation is



EXERCISES

(1) Consider the Turing machine represented by:



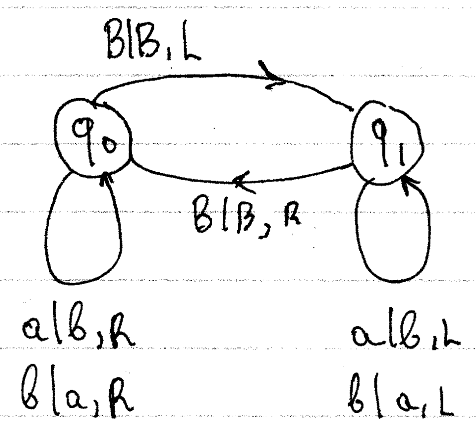
a) Give the corresponding set theoretic definition of this machine.

b) Consider the following initial configurations:

i) $\xi = q_0 11$ ii) $\xi = q_0 111$ iii) $\xi = q_0 1111$

Trace the calculation of the Turing machine from these initial configurations until it halts.

(2) Consider a Turing machine with graphical representation:



a) Give a set theoretic definition of this machine
b) Trace the first 20 steps of the machine configuration from initial state $\xi = q_0 aaba$
c) Explain why this machine never halts for any initial configuration $\xi = q_0 u$ with $u \in \{a, b\}^*$ any string.

(3) Let $M \in Tur(\Sigma)$ be a Turing machine with configuration transition function $\Delta_n : config(M) \to config(M)$ with $n \in \mathbb{N}^*$. Show that

a) $\forall n, m \in \mathbb{N}^* : \Delta_n(\Delta_m(\xi)) = \Delta_{n+m}(\xi)$
b) $\forall \xi_1, \xi_2 \in config(M) : (\xi_1 \vdash^* \xi_2 \land \xi_2 \vdash^* \xi_3 \Rightarrow \xi_1 \vdash^* \xi_3)$
c) $\forall \xi \in config(M) : (\xi \vdash^* loop \Rightarrow \xi \vdash^* \infty)$

▼ Turing machines as language accepters

Turing machines can be used as language accepters, analogously with dfas and nfas.

Def: Let $M = (Q, \Sigma, \Gamma, \delta, q_0, B, F)$ be a Turing machine $M \in \mathrm{Tur}(\Sigma)$. The language accepted by M is:
$\mathcal{L}(M) = \{u \in \Sigma^* \mid \exists n \in \mathbb{N}^* : \exists q \in F : \exists x, y \in \Gamma^* : \Delta_n(q_0 u) = xqy\}$

interpretation: We begin with a candidate string $u \in \Sigma^*$, and place the Turing machine on the leftmost character of u with q_0 as the initial internal state of the Turing machine. The string u is accepted if and only if after n steps the Turing computation terminates with the Turing machine in a final state. When $u \notin \mathcal{L}(M)$, the Turing machine could terminate in a non-final internal state or never terminate.

EXAMPLE

Given the alphabet $\Sigma = \{a, b\}$, implement a Turing machine that accepts the language
$L = \{a^n b^n \mid n \in \mathbb{N}^+\}$

Solution

▸ Strategy: Starting at leftmost a, we mark it by replacing it with x. Then we move to leftmost b and mark it with y. We move back left and look for the leftmost a and repeat. Eventually, when we cannot find any leftmost a, we should not be able to find a leftmost b either. To illustrate this process, the corresponding string modifications will look like:

$aaabbb \to xaabbb \to xaaybb \to xxaybb \to$
$\to xxayyb \to xxxyyb \to xxxyyy.$

▸ Implementation

- 1 We use the following tape alphabet $\Gamma = \{a, b, x, y, B\}$.
- 2 The following transitions replace the leftmost a with x and search for leftmost b by moving the machine right.

$\delta(q_0, a) = (q_1, x, R)$ // If 1st character is a, replace with x, move right. Go to state q_1 to find the leftmost b

// Skip all a, y characters while searching for the
// leftmost b

$\delta(q_1, a) = (q_1, a, R)$

$\delta(q_1, y) = (q_1, y, R)$

// When we find the leftmost b, mark it as y and
// go to state q_2 to search for next leftmost a

$\delta(q_1, b) = (q_2, y, L)$

• 3 The following transitions implement searching for the next "a" by moving the machine left.

// Moving left, skip all "y" and "a" characters

$\delta(q_2, a) = (q_2, a, L)$

$\delta(q_2, y) = (q_2, y, L)$

// When we find the first x, move right to return to the leftmost "a". Go to state q_0 in order to repeat the whole process

$\delta(q_2, x) = (q_0, x, R)$

// It is assumed that after the above transition is executed the machine is scanning the "a" character. However, if we have exhausted all "a", then it will be scanning a "y" character instead. In that case, we enter a new mode q_3 and move right to verify that we have exhausted all "b" characters.

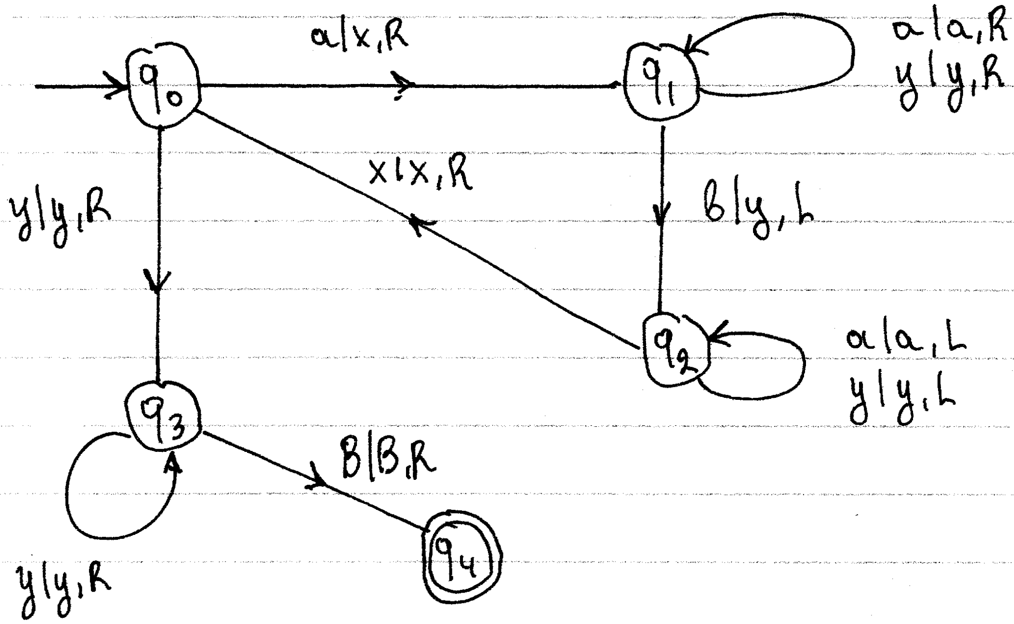
$\delta(q_0, y) = (q_3, y, R)$

$\delta(q_3, y) = (q_3, y, R)$

// Encountering a blank means that we exhausted all "b" characters so we go to final state q_4 and halt

$\delta(q_3, B) = (q_4, B, R)$

- It follows that
  $Q = \{q_0, q_1, q_2, q_3, q_4\}$
  $F = \{q_4\}$
  with q_4 being the halting state.
- A graphical representation of this machine is as follows:



Remarks

a) Note that q_3 has missing outgoing edges for a, b, x. If such characters are encountered, the machine halts in q_3 and since $q_3 \notin F$ the overall string is rejected.

b) q_1 and q_2 have two loops, but for convenience we show them as one loop with 2 labels. Implicitly each label corresponds with a different loop.

c) In describing a Turing machine, just like with any other programming language, you should COMMENT YOUR CODE.

EXERCISES

(4) Design and implement Turing machines that accept the following languages. Use both commented source code explicitly defining the Turing machine via set theory and also the corresponding graph representation.

a) $L = \{a, a^2b, ab^2\}$
b) $L = \{ab^n \mid n \in \mathbb{N}^*\}$
c) $L = \{w \in \{a,b\}^* \mid \exists k \in \mathbb{N}^* : |w| = 2k\}$
d) $L = \{w \in \{a,b\}^* \mid \exists k \in \mathbb{N}^* : |w| = 3k+1\}$
e) $L = \{a^n b^n c^n \mid n \in \mathbb{N}^*\}$
f) $L = \{a^n b^{2n} \mid n \in \mathbb{N}^*\}$
g) $L = \{a^n b^m a^{n+m} \mid n, m \in \mathbb{N}^*\}$
h) $L = \{u \in \{a,b\}^* \mid n_a(u) = n_b(u)\}$

Recursively enumerable and recursive languages

Def: Let Σ be an alphabet and $L \subseteq \Sigma^*$ a language. We say that

L recursively enumerable $\Leftrightarrow \exists M \in \mathrm{Tur}(\Sigma) : \mathcal{L}(M) = L$

The problem with this definition is that when a string $u \notin \mathcal{L}(M)$, it may throw the Turing machine into a computation that never terminates. This motivates the following stricter definition:

Def: Let Σ be an alphabet and let $L \subseteq \Sigma^*$ be a language. We say that

$$L \text{ recursive} \Leftrightarrow \exists M \in \mathrm{Tur}(\Sigma) : \begin{cases} \mathcal{L}(M) = L \\ \forall u \in \Sigma^+ : \overline{q_0 u \vdash \infty} \end{cases}$$

In a recursive language we make the demand that the Turing machine M that can accept the language L should halt in a finite number of steps when it is given non-empty strings that do not belong to L. Thus the machine M will be able to tell us for all $u \in \Sigma^+$, whether or not they belong to L with a finite number of steps.

Remarks

a) It is obvious that

$$\forall L \in \mathcal{P}(\Sigma^*) : (L \text{ recursive} \Rightarrow L \text{ recursively enumerable})$$

However there is a counterexample to the converse statement.

b) We will show below that there are languages in $\mathcal{P}(\Sigma^*)$ that are not recursively enumerable. This means that the Turing machine is not powerful enough to account for all languages in $\mathcal{P}(\Sigma^*)$.

c) On the other hand, according to the Church–Turing hypothesis, there are no possible modifications that can be made to the Turing machine definition to create a more powerful machine that can accept all recursively enumerable languages in addition to languages that are not recursively enumerable

d) The following modifications fail to make the machine more powerful:

i) Using a two-dimensional (or multi-dimensional) tape

ii) Adding a <u>stay</u> S operation, in addition to the <u>right</u> R and <u>left</u> L operations.

iii) Adding multiple tape devices, one-dimensional or multi-dimensional.

iv) Making the Turing machine non-deterministic (analogously to nfas vs. dfas)

or any combination of the above.

Limitations of Turing machines

- It takes a finite sequence of symbols to define every Turing machine. These symbols can be encoded as a binary string $u \in \{0,1\}^*$. We can thus construct a bijection from $Tur(\Sigma)$ to $\{0,1\}^*$ and conclude that

$$Tur(\Sigma) \sim \{0,1\}^* \qquad (1)$$

- We also know that

$$\{0,1\}^* = \bigcup_{n \in \mathbb{N}} \{0,1\}^n \Rightarrow \{0,1\}^* \text{ countably infinite}$$

$$\Rightarrow \{0,1\}^* \sim \mathbb{N} \qquad (2)$$

because $\{0,1\}^*$ is a countable union of finite sets.

- Likewise, for any finite alphabet Σ, we can show that

$$\Sigma^* \sim \mathbb{N} \qquad (3)$$

- From the above statements we can now prove the existence of languages that are not recursively enumerable.

> Thm: Given a finite alphabet Σ
> $\exists L \in \mathcal{P}(\Sigma^*)$: L not recursively enumerable

Proof

To show a contradiction, we assume the negation of the claim that:

$$\forall L \in \mathcal{P}(\Sigma^*) : L \text{ recursively enumerable}$$

$$\Rightarrow \forall L \in \mathcal{P}(\Sigma^*) : \exists M \in Tur(\Sigma) : L = \mathcal{L}(M)$$

This statement allows us to define a mapping $f: \mathcal{P}(\Sigma^*) \to Tur(\Sigma)$ such that for every language $L \in \mathcal{P}(\Sigma^*)$, $f(L)$ is a Turing machine such that $\mathcal{L}(f(L)) = L$. It is easy to show that f is one-to-one (i.e. the same machine cannot accept two different languages simultaneously). It follows that:

$$
\begin{aligned}
\Sigma^* &< \mathcal{P}(\Sigma^*) && [\text{Cantor's theorem}] \\
&\leqslant Tur(\Sigma) && [f \text{ one-to-one}] \\
&\sim \{0,1\}^* && [\text{Eq. } (1)] \\
&\sim \mathbb{N} && [\text{Eq. } (2)] \\
&\sim \Sigma^* && [\text{Eq } (3)]
\end{aligned}
$$

$\Rightarrow \Sigma^* < \Sigma^* \Rightarrow \Sigma^* \not\sim \Sigma^*$

which is a contradiction, since $\Sigma^* = \Sigma^* \Rightarrow \Sigma^* \sim \Sigma^*$.
It follows that

$\exists L \in \mathcal{P}(\Sigma^*) : L$ not recursively enumerable. $\square$

This theorem establishes the existence of languages that are not recursively enumerable and exposes the underlying problem: the set $\mathcal{P}(\Sigma^*)$ of all possible languages is uncountable whereas the set $Tur(\Sigma)$ of all possible Turing machines is countable. As a result, we cannot construct a distinct Turing machine for every language in $\mathcal{P}(\Sigma^*)$. Another way to show this is to construct an example of a specific language that is not recursively enumerable. This can be done as follows:

Construction: Let $\Sigma = \{x\}$ and consider the set $Tur(\Sigma)$ of all Turing machines. Since $Tur(\Sigma) \sim \mathbb{N}$, let $M_0, M_1, M_2, \ldots$ be an enumeration of all possible Turing machines. We define the language

$$L = \{ x^a \mid a \in \mathbb{N} \wedge x^a \notin \mathcal{L}(M_a) \}$$

Then L is NOT recursively enumerable.

Proof

To show that L is not recursively enumerable, we assume that L is recursively enumerable in order to derive a contradiction. Then:

$$\begin{aligned} L \text{ recursively enumerable} &\Rightarrow \exists M \in Tur(\Sigma) : L = \mathcal{L}(M) \\ &\Rightarrow \exists a \in \mathbb{N} : L = \mathcal{L}(M_a) \end{aligned}$$

Choose a $b \in \mathbb{N}$ such that $L = \mathcal{L}(M_b)$. We distinguish between the following cases:

Case 1: Assume that $x^b \in L$. Then:

$$\begin{aligned} x^b \in L &\Rightarrow x^b \in \{x^a \mid a \in \mathbb{N} \wedge x^a \notin \mathcal{L}(M_a)\} && [\text{Def of } L] \\ &\Rightarrow b \in \mathbb{N} \wedge x^b \notin \mathcal{L}(M_b) \\ &\Rightarrow x^b \notin \mathcal{L}(M_b) \\ &\Rightarrow x^b \notin L && [\text{via } L = \mathcal{L}(M_b)] \end{aligned}$$

which is a contradiction, therefore case 1 does not materialize.

Case 2: Assume that $x^b \notin L$. Then:

$$\begin{aligned} b \in \mathbb{N} \wedge x^b \notin L &\Rightarrow b \in \mathbb{N} \wedge x^b \notin \mathcal{L}(M_b) && [\text{via } L = \mathcal{L}(M_b)] \\ &\Rightarrow x^b \in \{x^a \mid a \in \mathbb{N} \wedge x^a \notin \mathcal{L}(M_a)\} \\ &\Rightarrow x^b \in L \end{aligned}$$

which is a contradiction. We conclude that case 2 does not materialize.
From the above argument, it follows that L is NOT recursively enumerable $\square$

EXERCISES

5. Let $\Sigma = \{a, b\}$ be an alphabet and consider the set

$$A = \{ L \in \mathcal{P}(\Sigma^*) \mid L \text{ not recursively enumerable} \}$$

Show that A is uncountalle.

6. Let $\Sigma = \{a, b\}$ and consider two languages $L_1, L_2 \in \mathcal{P}(\Sigma^*)$. Show that:

a) $\left\{ \begin{array}{l} L_1 \text{ recursively enumerable} \\ L_2 \text{ recursively enumerable} \end{array} \right. \Rightarrow L_1 \cap L_2$ recursively enumerable

b) $\left\{ \begin{array}{l} L_1 \text{ recursive} \\ L_2 \text{ recursive} \end{array} \right. \Rightarrow L_1 \cap L_2$ recursive