Discrete Mathematics

Haluk Bingol

Contents

Part 1

Preliminaries

CHAPTER 1

Preliminaries

1. Motivation

This text is not meant to be printed. It is designed to be read electronically. You will find many hyperlinks to sources in the web. Especially incredible wikipedia.com, which this book is dedicated to, gets many of them.

2. On definitions

Definitions are one of the starting points of mathematics. We should understand them well. By definition what we actually do is to give a "name" to "something". To start with, "that something" should be well-defined, that is, everybody understand the same without any ambiguity. What is in it, what is not in it should be clearly understood. Once we are all agree on "it", we give a "name".

The given name is not important. It could be some other name. Consider a text on geometry. Suppose we replace every occurrence of rectangle with triangle. The entire text would be still perfectly proper geometry text. This would be obvious if one considers the translation of the text to another language.

EXAMPLE 2.1. Suppose we all agree on parallelogram and right angle and try to define rectangle. A parallelogram is called rectangle if it has a right angle. Here we have an object which satisfies the conditions of both parallegram and right angle.

Note that in plain English we use the form "A is called B if A satisfies the followings ..." to define B. This may be falsely interpreted as one way implication such as "A satisfies the followings ... $\longrightarrow$ B". Actually what is intented is two-way implication such as "A is called B if and only if A satisfies the followings ...". More formally, it should be something like "A satisfies the followings ... $\longrightarrow$ B" and "$B \longrightarrow A$ satisfies the followings ..." at the same time. Instead of this long form, we write "A satisfies the followings ... $\longleftrightarrow$ B" in short.

In the language of mathematics, we use " $\longleftrightarrow$ " symbol in our definition. For example let n be a natural number. We want to define evenness of natural numbers.

$$n \text{ is even} \quad \longleftrightarrow \quad n \text{ is divisible by 2.}$$

Here the left hand side is not derived from the right hand side. It is just defined to be that way. In order to emphasize this we use the following notation:

$$n \text{ is even} \quad \overset{\Delta}{\longleftrightarrow} \quad n \text{ is divisible by 2.}$$

Unfortunately. this symbol is also used in different meaning. "$a \longleftrightarrow b$" means b can be obtained from a using some applications of rules, and similarly, a can also be obtained from b. This is the regular use of " $\longleftrightarrow$ ".

We feel that regular use of "=" should be differentiated from the usage of "=" in definitions. For example in

$$1 + (1 + 1) = 1 + 2 = 3$$

the usage of "=" is the regular usage meaning the right hand side of "=" is obtained from the left hand side by applying some rules. In the case of defining subtraction as

$$a - b = a + b^{-1}$$

where b^{-1} is the additive inverse of b, $a - b$ is defined in terms of known binary operation $+$ and unary operation of additive inverse. Therefore these will be written as

$$\begin{aligned} 1 + (1 + 1) &= 1 + 2 = 3 \\ a - b &\triangleq a + b^{-1} \end{aligned}$$

in this text.

EXAMPLE 2.2. Golden ration is the ratio of the sides of a rectangle which is presumable the aesthetically best. It is usually represented by ϕ. This can be given as:

$$\phi \triangleq \frac{1+\sqrt{5}}{2}.$$

As a summary, we exclusively use $\triangleq$ and $\overset{\Delta}{\longleftrightarrow}$ in the definitions. Therefore, it does not make sense trying to prove expressions such as $A \overset{\Delta}{\longleftrightarrow} B$ or $A \triangleq B$. On the other hand, in the expressions such as $A \longleftrightarrow B$ or $A = B$, the right hand side should be able to obtained from the left hand side. At the same time, the left hand side also should be able to obtained from the right hand side. That is, they are "provable".

In addition to this notation, the concept defined is presented in different color as in the case of *new concept*.

EXAMPLE 2.3. $\mathbb{Z}^+ \triangleq \{ z \in \mathbb{Z} \mid z > 0 \}$.

EXAMPLE 2.4. n is *even* $\overset{\Delta}{\longleftrightarrow}$ n is divisible by 2.

3. Similar statements

Sometimes two statements are very similar. They differ in a very few points. For example definition of evenness and oddness in natural numbers is given as follows:

DEFINITION 3.1. n is *even* $\overset{\Delta}{\longleftrightarrow}$ n is divisible by 2.

DEFINITION 3.2. n is *odd* $\overset{\Delta}{\longleftrightarrow}$ n is not divisible by 2.

In order to emphasize the differences of such cases the following notation is used.

DEFINITION 3.3. n is $\left|\begin{array}{l} \textit{even} \\ \textit{odd} \end{array}\right|$ $\overset{\Delta}{\longleftrightarrow}$ n is $\left|\begin{array}{l} \text{divisible} \\ \text{not divisible} \end{array}\right|$ by 2.

EXAMPLE 3.1. Let ρ be a relation on A, that is $\rho \subseteq A \times A$.

$$\rho \text{ is called } \left|\begin{array}{l} \textit{reflexive} \\ \textit{symmetric} \\ \textit{antisymmetric} \\ \textit{transitive} \end{array}\right| \overset{\Delta}{\longleftrightarrow} \left|\begin{array}{l} \forall a \in A \; [a \,\rho\, a] \\ \forall a, b \in A \; [a \,\rho\, b \longrightarrow b \,\rho\, a] \\ \forall a, b \in A \; [a \,\rho\, b \wedge b \,\rho\, a \longrightarrow a = b] \\ \forall a, b, c \in A \; [a \,\rho\, b \wedge b \,\rho\, c \longrightarrow a \,\rho\, c] \end{array}\right|.$$

4. Set of Numbers

We use the following symbols to represent the sets of various numbers.

$\mathbb{N}$	The set of natural numbers	$\mathbb{N} \triangleq \{0, 1, 2, \dots\}$
$\mathbb{Z}$	The set of integers	$\mathbb{Z} \triangleq \{\dots, -2, -1, 0, 1, 2, \dots\}$
$\mathbb{Z}^+$	The set of positive integers	$\mathbb{Z}^+ \triangleq \{1, 2, \dots\}$
$\mathbb{Z}^-$	The set of negative integers	$\mathbb{Z}^- \triangleq \{-1, -2, \dots\}$
$\mathbb{Z}_{\geq 0}$	The set of non negative integers	$\mathbb{Z}_{\geq 0} \triangleq \mathbb{Z}^+ \cup \{0\}$
$\mathbb{Z}_{\neq 0}$	The set of non zero integers	$\mathbb{Z}_{\neq 0} \triangleq \mathbb{Z} \setminus \{0\}$
$\mathbb{Z}_{\leq 0}$	The set of non positive integers	$\mathbb{Z}_{\leq 0} \triangleq \mathbb{Z}^- \cup \{0\}$
$\mathbb{Q}$	The set of rational numbers	$\mathbb{Q} \triangleq \{p/q \mid p, q \in \mathbb{Z} \text{ and } q \neq 0\}$
$\mathbb{Q}^+$	The set of positive rational numbers	$\mathbb{Q}^+ \triangleq \{p/q \mid p, q \in \mathbb{Z}^+\}$
$\mathbb{Q}^-$	The set of negative rational numbers	$\mathbb{Q}^- \triangleq \{-p/q \mid p, q \in \mathbb{Z}^+\}$
$\mathbb{Q}_{\geq 0}$	The set of non negative rational numbers	$\mathbb{Q}_{\geq 0} \triangleq \mathbb{Q}^+ \cup \{0\}$
$\mathbb{Q}_{\neq 0}$	The set of non zero rational numbers	$\mathbb{Q}_{\neq 0} \triangleq \mathbb{Q} \setminus \{0\}$
$\mathbb{Q}_{\leq 0}$	The set of non positive rational numbers	$\mathbb{Q}_{\leq 0} \triangleq \mathbb{Q}^- \cup \{0\}$
$\mathbb{R}$	The set of real numbers	$\mathbb{R}$
$\mathbb{R}^+$	The set of positive real numbers	$\mathbb{R}^+ \triangleq \{r \in \mathbb{R} \mid r > 0\}$
$\mathbb{R}^-$	The set of negative real numbers	$\mathbb{R}^- \triangleq \{r \in \mathbb{R} \mid r < 0\}$
$\mathbb{R}_{\geq 0}$	The set of non negative real numbers	$\mathbb{R}_{\geq 0} \triangleq \mathbb{R}^+ \cup \{0\}$
$\mathbb{R}_{\neq 0}$	The set of non zero real numbers	$\mathbb{R}_{\neq 0} \triangleq \mathbb{R} \setminus \{0\}$
$\mathbb{R}_{\leq 0}$	The set of non positive real numbers	$\mathbb{R}_{\leq 0} \triangleq \mathbb{R}^- \cup \{0\}$
$\mathbb{C}$	The set of complex numbers	$\mathbb{C} \triangleq \{a + \mathrm{i}\, b \mid a, b \in \mathbb{R}\}$
$\mathbb{C}_{\neq 0}$	The set of non zero complex numbers where $\mathrm{i}^2 = -1$	$\mathbb{C}_{\neq 0} \triangleq \mathbb{C} \setminus \{0\}$

Part 2

Basics

CHAPTER 2

Logic

1. Motivation

In the first half of 1900s mathematicians believed that entire mathematics can be constructed from a set of axioms, inference rules and symbolic logic. In 1910's, Bertrand Russell, now known due to his works in philosophy, and Alfred North Whitehead published Principia Mathematica which provided carefully designed construction of mathematics. They claim that every true mathematical statement can be proved by this way. Unfortunately in 1931 Kurt Gödel proved, in his incompleteness theorem, that there are some true statements that cannot be proven if the axiomatic system is consistent and sufficiently powerful to express the arithmetic of the natural numbers. The famous incompleteness theorem becomes one of the important milestones in Computer Science, too.

The logic is important for Computer Science in many ways. Search in the web is one of them. When we do search using a search engine in the Internet, we express ourselves using logic. For example typing

"Bertrand Russell Mathematica OR Kurt -philosopher"

to Google for search means we are looking pages which contains "Bertrand" and "Russell" words together, either of the words "Mathematica" or "Kurt" but we do not want word "philosopher" in our search. This can be rewritten in logic as "Bertrand" AND "Russell" AND *"Mathematica" OR "Kurt") AND NOT"philosopher".

2. Foundations

We use the notation of [**TZ82**] in defining well-formed logical formula. We use spaces in order to improve readability as in the case of $\forall x\ \phi(x)$ or $\phi \wedge \psi$ which should formally be written as $\forall x \phi(x)$ or $\phi \wedge \psi$.

2.1. Well-formed formula. It is not possible to evaluate an expression such as $2+3+\times 7$ since it is not properly formed.

Example 2.1. The following expressions are not meaningful. Try to interpret them.

i. $p\ \wedge\ q\ \wedge\ \ \wedge qq$
ii. $1\ +\ +\ +\ 1$
iii. $\times\ /\ /\ +\ -\ -\ +\ 2\ 3\ 4$
iv. $1\ 2\ +$
v. $1\ 2\ 3\ \times\ +$

It is correct that these expressions cannot be interpreted in the usual interpretation which is called infix notation. Actually the last two can be interpreted in postfix notation as $1+2$ and $1+(2\times 3)$, respectively. The postfix notation is sometimes called reverse polish notation since it is the reverse of prefix notation invented by Polish mathematician Jan Lukasiewicz around 1920's.

Properly formed expressions is the starting point of logic. Formally properly expression is called well-formed formula.

The language consists of:
Free variables: $a_0, a_1, \ldots$
Bound variables: $x_0, x_1, \ldots$
A predicate symbol: $\in$
Logical symbols: $\neg,\ \vee\ ,\ \wedge\ ,\ \longrightarrow\ ,\ \longleftrightarrow\ , \forall, \exists$.
Auxiliary symbols: $(,), [,]$.
ϕ, ψ, η are meta symbols.

Definition 2.1 (well-formed formula (wff)).

A formula is *well-formed formula (wff)* $\overset{\Delta}{\longleftrightarrow}$ it is deducible from the following rules:

i. If a and b are free variables, then $[a \in b]$ is a wff.
ii. If ϕ and ψ are wffs, then $\neg\phi, [\phi \vee \psi], [\phi \wedge \psi], [\phi \longrightarrow \psi]$, and $[\phi \longleftrightarrow \psi]$ are wff.
iii. If ϕ is a wff and x is a bound variable, then $\forall x\ \phi(x)$ and $\exists x\ \phi(x)$ are wff, where $\phi(x)$ is the formula obtained from the wff ϕ by replacing each occurrence of some free variable a by the bound variable x. We call $\forall x\phi(x)$ and $\exists x\phi(x)$ respectively, the formula obtained from ϕ by *universally*, or *existentially* qualifying on the variable a.

EXAMPLE 2.2. Examples of wffs are as follows where $p = x_0$ and $q = x_1$.
i. p and q are wffs due to Definition 2.1(i).
ii. $\neg p, p \vee q, p \wedge q, p \longrightarrow q, p \longleftrightarrow q$ are wffs due to Definition 2.1(ii).
iii. $[p \wedge q] \vee \neg p$, $[p \wedge q] \vee \neg[p \longrightarrow q]$ are wffs due to Definition 2.1(i) and (ii) .
iv. $\exists x\ [x \in a_1]$ is a wff. Since
$[a_0 \in a_1]$ by Definition 2.1(i)
$\exists x\ [x \in a_1]$ by existential qualifying on a_0.
v. $\exists x\ [x \in a_1] \wedge \forall x\ [x \in a_1]$ is a wff.

2.2. Logical Axioms.

AXIOM 1 (Logical Axioms).
i. $\phi \longrightarrow [\psi \longrightarrow \phi]$.
ii. $[\phi \longrightarrow [\psi \longrightarrow \eta]] \longrightarrow [[\phi \longrightarrow \psi] \longrightarrow [\phi \longrightarrow \eta]]$.
iii. $[\neg\phi \longrightarrow \neg\psi] \longrightarrow [\psi \longrightarrow \phi]$.
iv. $\forall x[\phi \longrightarrow \psi(x)] \longrightarrow [\phi \longrightarrow \forall x\psi(x)]$ *where free variable* a *on which we are quantifying does not occur in* ϕ.
v. $\forall x\phi(x) \longrightarrow \phi(a)$ *where* $\phi(a)$ *is the formula obtained by replacing each occurrence of the bound variable* x *in* $\phi(x)$ *by the free variable* a.

2.3. Rules of Inference.

AXIOM 2 (Rules of Inference).
i. From ϕ *and* $\phi \longrightarrow \psi$ *to infer* ψ.
ii. From ϕ *to infer* $\forall x\,\phi(x)$ *where* $\phi(x)$ *is obtained from* ϕ *by replacing each occurrence of some free variable by* x.

NOTATION.
i. ϕ and $\phi \longrightarrow \psi \implies \psi$.
ii. $\phi \implies \forall x\phi(x)$.

DEFINITION 2.2 (Logically Equivalence).
ϕ is *logically equivalent* to $\psi \overset{\Delta}{\longleftrightarrow} \phi$ is deducible using only the logical axioms. It is denoted by $\phi \longleftrightarrow \psi$.

2.4. Equality.

DEFINITION 2.3 (Equality).
$a{=}b \overset{\Delta}{\longleftrightarrow} \forall x\ [x \in a \longleftrightarrow x \in b]$.

PROPOSITION 2.1.
i. $a = a$.
ii. $a = b \longrightarrow b = a$.
iii. $a = b \wedge b = c \longrightarrow a = c$.

PROOF.
i. $\forall x[x \in a \longleftrightarrow x \in a]$.
ii. $\forall x[x \in a \longleftrightarrow x \in b] \implies \forall x[x \in b \longleftrightarrow x \in a]$.
iii. $\forall x[x \in a \longleftrightarrow x \in b] \wedge \forall x[x \in b \longleftrightarrow x \in c] \implies \forall x[x \in a \longleftrightarrow x \in c]$.

□

QUESTION 2.1. Do you think that
- NAN = NAN where NAN is "not a number" in programming.
- $$\left[\frac{\infty}{\infty}\right] = \left[\frac{\infty}{\infty}\right]$$
where $\left[\frac{\infty}{\infty}\right]$ is an indeterminate form of limits in Calculus.

p	$f_0 = F$	$f_1 = p$	$f_2 = \neg p$	$f_3 = T$
F	F	F	T	T
T	F	T	F	T

TABLE 1. Boolean functions of one variable

p	q	$f_0 = F$	$f_1 = p \wedge q$	$f_2 = \neg(p \longrightarrow q)$	$f_3 = p$	f_4	$f_5 = q$	$f_6 = p \oplus q$	$f_7 = p \vee q$	$f_8 = p$ NOR q	$f_9 = p \longleftrightarrow q$	$f_{10} = \neg q$	f_{11}	$f_{12} = \neg p$	$f_{13} = p \longrightarrow q$	$f_{14} = p$ NAND q	$f_{15} = T$
F	F	F	F	F	F	F	F	F	F	T	T	T	T	T	T	T	T
F	T	F	F	F	F	T	T	T	T	F	F	F	F	T	T	T	T
T	F	F	F	T	T	F	F	T	T	F	F	T	T	F	F	T	T
T	T	F	T	F	T	F	T	F	T	F	T	F	T	F	T	F	T

TABLE 2. Boolean functions of two variables

REMARK 2.1. If $a = b$ and a wff holds for a, then it must hold for b.

$$a = b \implies [\phi(a) \longleftrightarrow \phi(b)].$$

3. Propositional Logic

DEFINITION 3.1. A *proposition* is a statement that is either true or false but not both. The *truth value* of a true proposition is true, denoted by T and that of false proposition is false, denoted by F.

NOTATION. Propositions are represented by lower case letters such as p, r, q.

3.1. Compound Propositions. We generate new propositions from existing ones by means of well-formed-formulation. Any wff is a generated new proposition based on already established propositions.

DEFINITION 3.2. The set $\mathbb{B} \triangleq \{T, F\}$ is called *boolean domain* where T and F denotes true and false, respectively. An $n-$tuple $(p_1, p_2, \ldots, p_n)$ where $p_i \in \mathbb{B}$ is called a *boolean $n-$tuple*.

DEFINITION 3.3. An $n-$operand *truth table* is a table that assigns a boolean value to all boolean $n-$tuples. A *propositional operator* is a rule defined by a truth table.

DEFINITION 3.4. An operator is called | *monadic* / *dyadic* | if it has | one / two | operand(s).

REMARK 3.1.

i. A boolean $n-$tuple is an element of $\mathbb{B}^n$, that is, $(p_1, p_2, \ldots, p_n) \in \mathbb{B}^n$.
ii. There are 2^n different binary $n-$tuples.
iii. A truth table of a predicate p actually defines a function $f_p : \mathbb{B}^n \to \mathbb{B}$.
iv. There are 2^{2^n} different truth tables (functions) of binary $n-$tuples.
v. There are $2^{2^1} = 4$ monadic operators, *identity*, *negation*, *constant-True*, *constant-False*, as given in Table 3.1.
vi. There are $2^{2^2} = 16$ dyadic operators as given in Table 3.1.
vii. Note that the functions in the Table 3.1 have interesting properties: Firstly, notice the relation between f_i and the binary representation of i if F and T are represented as 0 and 1, respectively. For example f_{11} corresponds to $TFTT = 1011$. Secondly, $f_i = \neg f_{15-i}$ as in the case of $f_3 = \neg f_{12}$
viii. NAND and NOR are extensively used in logic design in Computer Engineering.
ix. p NAND $q \triangleq \neg(p \wedge q)$. That is, $f_{14}(p, q) = \neg f_1(p, q)$.
x. p NOR $q \triangleq \neg(p \vee q)$. That is, $f_8(p, q) = \neg f_7(p, q)$.

DEFINITION 3.5. Any wff is a *compound propositions*.

REMARK 3.2. In other words, propositions formed from existing propositions using logical operators are called *compound propositions*.

DEFINITION 3.6. Let p be a proposition. The *negation of* p, denoted by $\neg p$ or $\bar{p}$, is the statement "It is not the case that p".

DEFINITION 3.7. Let p and q be propositions. The *conjunction* of p and q, denoted by $p \wedge q$, is the proposition "p and q".

$$p \wedge q \triangleq \begin{cases} T, & \text{if both } p \text{ and } q \text{ are true,} \\ F, & \text{otherwise.} \end{cases}$$

DEFINITION 3.8. Let p and q be propositions. The *disjunction* of p and q, denoted by $p \vee q$, is the proposition "p or q".

$$p \vee q \triangleq \begin{cases} F, & \text{if both } p \text{ and } q \text{ are false,} \\ T, & \text{otherwise.} \end{cases}$$

DEFINITION 3.9. Let p and q be propositions.

The | *exclusive or* / *conditional statement* / *biconditional statement* |, denoted by | $p \oplus q$ / $p \longrightarrow q$ / $p \longleftrightarrow q$ |, is the function | f_6 / f_{13} / f_9 | in the Table 3.1.

REMARK 3.3.

i. Conditional $p \longrightarrow q$, sometimes called *implication*.
ii. Some other English usages are "if p, then q", "p implies q" and many more.
iii. p is called the *hypothesis*. q is called the *conclusion*.

REMARK 3.4.

i. Biconditional $p \longleftrightarrow q$, sometimes called *bi-implication* or *if-and-only-if*, *iff* in short.
ii. Some other English usages are "p is necessary and sufficient for q", "p iff q".
iii. Note that $p \longleftrightarrow q$ is equivalent to $(p \longrightarrow q) \wedge (q \longrightarrow p)$.

DEFINITION 3.10. Two compound propositions $\phi(x_1, x_2, \ldots, x_n)$ and $\psi(x_1, x_2, \ldots, x_n)$ of the same variables x_1, x_2, ..., x_n, are called *equivalent* $\overset{\Delta}{\longleftrightarrow}$ they have the same truth tables. It is denoted by $\phi(x_1, x_2, \ldots, x_n) \iff \psi(x_1, x_2, \ldots, x_n)$,

REMARK 3.5. The biconditional, $p \longleftrightarrow q$, is an operator. The equivalence of two compound propositions, $p \iff q$, is an equivalence relation on the set of all propositions.

DEFINITION 3.11. Let $p \longrightarrow q$.

The | *converse* / *contrapositive* / *inverse* | of $p \longrightarrow q$ is | $q \longrightarrow p$ / $\neg q \longrightarrow \neg p$ / $\neg p \longrightarrow \neg q$ |.

COROLLARY 3.1.

The | *implication,* $p \longrightarrow q$ / *converse,* $q \longrightarrow p$ | *is equivalent to* | *contrapositive,* $\neg q \longrightarrow \neg p$ / *inverse,* $\neg p \longrightarrow \neg q$ |.

3.2. Application. Logic descriptions is used in all branches of science and engineering. Unambiguous, precise, consistent reporting is a must.

3.2.1. *Translating English sentences.* Consider a detective story such as one from Sherlock Holmes. There are people $P_1, P_2, \ldots, P_n$. There are corresponding propositions $p_1, p_2, \ldots, p_n$ where p_i means person P_i is the murderer. Of course there is a description of the rest of the story which can be represented as $q(p_1, p_2, \ldots, p_n)$. In this formulization if person P_3 is the murderer then the truth assignment of $(F, F, T, F, \ldots, F)$ makes q true, that is, $q(F, F, T, F, \ldots, F) = T$. Here we assume that there is one murderer so there is only one entry $p_i = T$.

3.2.2. *System specifications.* In engineering precise, formal descriptions are needed. Software development is one of them. A typical software life cycle is as follows: A customer who needs a custom tailored software solution defines what she wants. This definition will be given to the contracting company. The developers start developing the software. At the end the software is delivered to the costumer. The costumer checks if the developed software meets the specification.

Equivalence	Name
$p \wedge T \equiv p$ $p \vee F \equiv p$	Identity laws
$p \wedge F \equiv F$ $p \vee T \equiv T$	Domination laws
$p \wedge p \equiv p$ $p \vee p \equiv p$	Idempotent laws
$p \wedge q \equiv q \wedge p$ $p \vee q \equiv q \vee p$	Commutativity laws
$p \wedge (q \wedge r) \equiv (p \wedge q) \wedge r$ $p \vee (q \vee r) \equiv (p \vee q) \vee r$	Associativity laws
$p \wedge (q \vee r) \equiv (p \wedge q) \vee (p \wedge r)$ $p \vee (q \wedge r) \equiv (p \vee q) \wedge (p \vee r)$	Distributivity laws
$p \wedge (p \vee q) \equiv p$ $p \vee (p \wedge q) \equiv p$	Absorption laws
$p \wedge \neg p \equiv F$ $p \vee \neg p \equiv T$	Negation laws
$\neg(p \wedge q) \equiv \neg p \vee \neg q$ $\neg(p \vee q) \equiv \neg p \wedge \neg q$	De Morgan's laws
$\neg(\neg p) \equiv p$	Double negation law

TABLE 3. The laws of logic

In such a scenario the definition should be as precise as possible. Think about the consequences if the definition is not precise, informal, possibly ambiguous. It is not that unusual that the definitions have some conflict or contradicting requirements.

3.2.3. *Boolean search. Search in web.* It is already mentioned in the motivation that search engines understand the language of predicates.

- Translating English sentences
- System specifications
- Boolean search. Search in web.
- Logic puzzles
- Logic and bit operations

4. Propositional Equivalence

We use wff for compound propositions.

DEFINITION 4.1. A wff is called a | *tautology* / *contradiction* | $\overset{\Delta}{\longleftrightarrow}$ it is always | true / false | independent of the truth values of its propositions. A wff that is neither tautology not contradiction is called a *contingency*.

EXAMPLE 4.1. Some simple forms are as follows:
Tautologies: T, $\neg F$, $p \vee T$, $p \vee \neg p$.
Contradictions: F, $\neg T$, $p \wedge F$, $p \wedge \neg p$.
Contingencies: p, $\neg p$, $p \vee F$, $p \wedge T$.

DEFINITION 4.2 (Logically Equivalence).
Two wffs p and q are called *logically equivalent* $\overset{\Delta}{\longleftrightarrow}$ The wff $p \longleftrightarrow q$ is a tautology. Logically equivalence of p and q is denoted by $p \equiv q$ or $p \iff q$.

REMARK 4.1. Note that logically equivalence is an equivalence relation on the set of all wff since:

i. Reflexivity: $\forall p \ [p \iff p]$.
ii. Symmetry: $\forall p, q \ [(p \iff q) \longrightarrow (q \iff p)]$.
iii. Transitivity: $\forall p, q, r \ [(p \iff q) \wedge (q \iff r) \longrightarrow (p \iff r)]$.

Laws of logically equivalence are given in Table 3.

EXAMPLE 4.2. $T \iff \neg F$, $\neg F \iff p \vee T$, $p \vee T \iff p \vee \neg p$.

5. Quantifiers

$\forall x\,[p(x) \vee q(x)] \Longleftarrow \forall x\,p(x) \vee \forall x\,q(x)$
$\forall x\,[p(x) \wedge q(x)] \Longleftrightarrow \forall x\,p(x) \wedge \forall x\,q(x)$
$\exists x\,[p(x) \vee q(x)] \Longleftrightarrow \exists x\,p(x) \vee \exists x\,q(x)$
$\exists x\,[p(x) \wedge q(x)] \Longrightarrow \exists x\,p(x) \wedge \exists x\,q(x)$

Acknowledgment. These notes are based on various books but especially [**Ros07, TZ82**].

Problems with Solutions

P 2.1. a) Express each of these statements using quantifiers and the following predicates where the domain consists of all people.

$S(x) : x$ is a student in this class. $M(x) : x$ is a mathematician
$L(x) : x$ likes discrete mathematics course. $C(x, y) : x$ and y are colleagues
$K(x, y) : x$ knows y

i) There are exactly two students in this class who like discrete mathematics course.
ii) Every student in this class knows Kurt Gödel or knows a mathematician who is a colleague of Kurt Gödel.
iii) There is no student in this class who knows everybody else in this class

b) Using rules of inference provide a formal proof for
If $\forall x\,[S(x) \vee Q(x)]$, and $\forall x\,[(\neg S(x) \wedge Q(x)) \to P(x)]$ are true then $\forall x\,[\neg P(x) \to S(x)]$ is also true where the domains of all quantifiers are the same.

Solution.

a)
i) $\exists x \exists y\,[x \neq y \wedge S(x) \wedge S(y) \wedge L(x) \wedge L(y) \wedge \forall z\,(S(z) \wedge L(z) \to z = x \vee z = y)]$
ii) $\forall x\,[S(x) \to [K(x, Godel) \vee \exists y\,(M(y) \wedge C(y, Godel) \wedge K(x, y))]]$
iii) $\neg\exists x \forall y\,[S(x) \wedge ((S(y) \wedge x \neq y) \to K(x, y))]$

b)

1.	$\forall x\,[S(x) \vee Q(x)]$	Premise
2.	$\forall x\,[(\neg S(x) \wedge Q(x)) \to P(x)]$	Premise
3.	$S(a) \vee Q(a)$	(1) universal generalization
4.	$(\neg S(a) \wedge Q(a)) \to P(a)$	(2) universal generalization
5.	$\neg(\neg S(a) \wedge Q(a)) \vee P(a)$	(4) logical equivalence $p \to q \equiv \neg p \vee q$
6.	$(S(a) \vee \neg Q(a)) \vee P(a)$	(5) De Morgan
7.	$(P(a) \vee S(a)) \vee \neg Q(a)$	(6) Commutativity and associativity of $\vee$
8.	$(P(a) \vee S(a)) \vee S(a)$	(7) and (3) resolution
9.	$P(a) \vee S(a)$	(8) Idempotent law
10.	$\neg P(a) \to S(a)$	(9) logical equivalence $p \to q \equiv \neg p \vee q$
11.	$\forall x\,(\neg P(x) \to S(x))$	Universal generalization (a was arbitrary)

CHAPTER 3

Sets, Relations, and Functions

1. Set

1.1. Sets.

DEFINITION 1.1. A *set* is unordered collection of objects.

REMARK 1.1. We do not define *set*, *element*, and *membership* properly. A set is a collection of elements. Sets are usually represented by capital letters A, B, ... Sets are defined either listing of the elements as in $A = \{ a_1, a_2, \dots \}$. or those elements that satisfy predicate $P(a)$ as in $A = \{ a \mid P(a) \}$. Note that the order of the elements is not important. Due to that *unordered n-tuple* is represented as $\{ a_1, a_2, \dots, a_n \}$. If a is an element of A, it is denoted as $a \in A$, otherwise as $a \notin A$.

EXAMPLE 1.1. The set of natural numbers $\mathbb{N} = \{ 0, 1, 2, \dots \}$. $8 \in \mathbb{N}$ but $-3 \notin \mathbb{N}$.

EXAMPLE 1.2. $E = \{ x \mid x \in \mathbb{N} \ \wedge \ x \text{ is even} \} = \{ x \in \mathbb{N} \mid x \text{ is even} \}$ where the second form is a short form of the first.

DEFINITION 1.2. The *empty set*, denoted $\emptyset$, has no elements in it.

REMARK 1.2. There is one and only one empty set. $\emptyset$ has interesting properties: Let $A = \emptyset$ and $B = \{ \emptyset \}$. Then $A \in B$ and $A \subseteq B$.

REMARK 1.3. Let $P(x)$ be a property. Then the following two propositions are true:

i. $\forall x \in \emptyset \ [P(x)]$
ii. $\neg \exists x \in \emptyset \ [P(x)]$

DEFINITION 1.3 (Equality of Sets).

$$A \text{ is } \left| \begin{array}{l} \textit{equal} \\ \textit{not equal} \end{array} \right| \text{ to } B, \text{ denoted by } \left| \begin{array}{l} A = B \\ A \neq B \end{array} \right|, \ \overset{\Delta}{\longleftrightarrow} \ \left| \begin{array}{l} \forall x \ [x \in A \ \longleftrightarrow \ x \in B] \\ \exists a \ (a \in A \ \wedge \ a \notin B) \ \vee \ \exists b \ (b \notin A \ \wedge \ b \in B) \end{array} \right|.$$

EXAMPLE 1.3. $\{ 1, 2, 3 \} = \{ 2, 1, 3 \}$. Order of elements is not important.

EXAMPLE 1.4. $\{ 1, 2, 3 \} = \{ 1, 1, 2, 3 \}$. Repetition of elements is not important.

EXAMPLE 1.5. $x \neq \{ x \}$, and $\{ x \} \neq \{ \{ x \} \}$.

DEFINITION 1.4 (Subset).

$$A \text{ is a } \left| \begin{array}{l} \textit{subset} \\ \textit{proper subset} \end{array} \right| \text{ of } B, \text{ denoted by } \left| \begin{array}{l} A \subseteq B \\ A \subset B \end{array} \right|, \ \overset{\Delta}{\longleftrightarrow} \ \left| \begin{array}{l} \forall a \ (a \in A \ \longrightarrow \ a \in B) \\ A \subseteq B \ \wedge \ \exists b \ (b \notin A \ \wedge \ b \in B) \end{array} \right|.$$

THEOREM 1.1.
Let A be a set.

i. $\forall A \ [\emptyset \subseteq A]$.
ii. $\forall A \ [A \subseteq A]$.

DEFINITION 1.5 (Cardinality, Finite Set, Infinite Set). A is *finite* and n is the *cardinality* of A $\overset{\Delta}{\longleftrightarrow}$ There are exactly n distinct elements in A. The cardinality of A is denoted by $| A |$. A is *infinite* $\overset{\Delta}{\longleftrightarrow}$ A is not *finite*.

REMARK 1.4. This definition of infinity needs elaboration.

EXAMPLE 1.6.

$0 = | \emptyset |$.
$1 = | \{ a \} | = | \{ a, a \} | = | \{ \emptyset \} | = | \{ \{ \emptyset \} \} | = | \{ \{ \{ \emptyset \} \} \} | = | \{ \{ \emptyset, \{ \emptyset \} \} \} |$.
$2 = | \{ a, b \} | = | \{ \emptyset, \{ \{ \emptyset \} \} \} | = | \{ \emptyset, \{ \emptyset \} \} |$.

DEFINITION 1.6 (Power Set). The *power set* of A: $2^A \triangleq \{ S \mid S \subseteq A \}$.

EXAMPLE 1.7. $2^{\{1,2,3\}} = \{\emptyset, \{1\}, \{2\}, \{3\}, \{1,2\}, \{1,3\}, \{2,3\}, \{1,2,3\}\}$.
$2^{\emptyset} = \{\emptyset\}$, $2^{\{\emptyset\}} = \{\emptyset, \{\emptyset\}\}$, $2^{2^{\{\emptyset\}}} = \{\emptyset, \{\emptyset\}, \{\{\emptyset\}\}, \{\emptyset, \{\emptyset\}\}\}$.

THEOREM 1.2. $A = \bigcup_{S \in 2^A} S$.

THEOREM 1.3. *If A is finite,* $\left| 2^A \right| = 2^{|A|}$.

THEOREM 1.4. $2^A = 2^B \longrightarrow A = B$.

PROOF. By Theorem 1.2, $2^A = 2^B \Longrightarrow \bigcup_{S \in 2^A} S = \bigcup_{S \in 2^B} S$. Therefore $A = B$. □

DEFINITION 1.7 (Ordered $n-$tuple). The *ordered $n-$tuple* $(a_1, a_2, \ldots, a_n)$ is the ordered collection that has a_i as its ith element. (a_1, a_2) is called *ordered pairs*. $(a_1, a_2, \ldots, a_n)$ is *equal* to $(b_1, b_2, \ldots, b_n)$, denoted by $(a_1, a_2, \ldots, a_n) = (b_1, b_2, \ldots, b_n)$, $\overset{\Delta}{\longleftrightarrow} \forall i \in \{1, \ldots, n\}\ a_i = b_i$.

REMARK 1.5. An unordered $n-$tuple is represented by a set. Sets can be used to represent ordered tuples, too. Ordered $n-$tuple can be represented as sets as:

$$(a_1, a_2) \triangleq \{a_1, \{a_2\}\}$$
$$(a_1, a_2, a_3) \triangleq \{a_1, \{a_2, \{a_3\}\}\}$$
$$\ldots$$

DEFINITION 1.8 (Cartesian Product). The *cartesian product* of A and B is defiend as $A \times B \triangleq \{ (a, b) \mid a \in A \ \wedge \ b \in B \}$.

REMARK 1.6. Note that $A \times B \neq B \times A$. As an example $A = \{1\}$ and $B = \{b\}$. Then $A \times B = \{(1, b)\}$ and $B \times A = \{(b, 1)\}$. Hence $A \times B \neq B \times A$.

THEOREM 1.5. $A \times \emptyset = \emptyset \times A = \emptyset$.

THEOREM 1.6. $A \times B = \emptyset \longrightarrow (A = \emptyset \ \vee \ B = \emptyset)$.

PROOF. Suppose $\neg(A = \emptyset \ \vee \ B = \emptyset)$.
$\Rightarrow A \neq \emptyset \ \wedge \ B \neq \emptyset$.
$\Rightarrow \exists a \in A \ \wedge \ \exists b \in B$.
$\Rightarrow (a, b) \in A \times B$.
$\Rightarrow A \times B \neq \emptyset$.
Hence $A = \emptyset \ \vee \ B = \emptyset$. □

DEFINITION 1.9. The Cartesian product of the sets $A_1, A_2, \ldots, A_n$ is defined as $A_1 \times A_2 \times \cdots \times A_n \triangleq \{ (a_1, a_2, \ldots, a_n) \mid \forall i \in \{1, \ldots, n\}\ a_i \in A_i \}$.

DEFINITION 1.10 (Power of a Set A^n). The nth *power of a set*, denoted by A^n, is defined as

$$A^1 = A$$
$$A^{n+1} = A^n \times A \text{ where } n \in \mathbb{Z}^+.$$

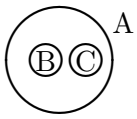
1.2. Set Operations.

DEFINITION 1.11 (Union, Intersection).

The $\left|\begin{array}{l} \textit{union} \\ \textit{intersection} \end{array}\right|$ of A and B defined as $\left|\begin{array}{ll} A \cup B & \triangleq \{ x \mid x \in A \ \vee \ x \in B \} \\ A \cap B & \triangleq \{ x \mid x \in A \ \wedge \ x \in B \} \end{array}\right|$.

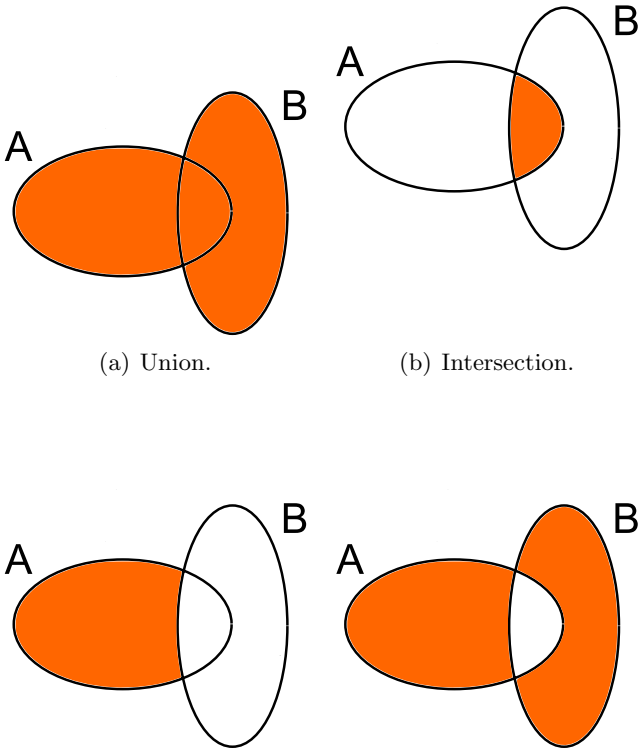
REMARK 1.7. Set operations can be visualized by Venn diagrams as in Fig. 1

EXAMPLE 1.8. For sets A, B, C in the figure, $A \cup B = A \cup C$ but $B \neq C$.



DEFINITION 1.12. Let $C = \{A_1, A_2, \ldots, A_n\}$ be a collection of sets.

The $\left|\begin{array}{l} \textit{union} \\ \textit{intersection} \end{array}\right|$ of collection C is defined as $\left|\begin{array}{ll} \bigcup_{i=1}^n A_i & \triangleq A_1 \cup A_2 \cup \cdots \cup A_n = \{ x \mid \exists i \in \{1, \ldots, n\}\ x \in A_i \} \\ \bigcap_{i=1}^n A_i & \triangleq A_1 \cap A_2 \cap \cdots \cap A_n = \{ x \mid \forall i \in \{1, \ldots, n\}\ x \in A_i \} \end{array}\right|$.



(a) Union. (b) Intersection.

(c) Set difference. (d) Symmetric set difference.

FIGURE 1. Set operations.

DEFINITION 1.13. A and B are *disjoint* $\overset{\Delta}{\longleftrightarrow}$ $A \cap B = \emptyset$.

THEOREM 1.7 (Principle of Inclusion-Exclusion).

$$|A \cup B| = \begin{cases} |A| + |B| & \text{if } A \cap B = \emptyset, \\ |A| + |B| - |A \cap B| & \text{if } A \cap B \neq \emptyset. \end{cases}$$

DEFINITION 1.14 (Set Difference). The *difference* of A and B is defined as $A \setminus B \triangleq \{ x \mid x \in A \ \wedge \ x \notin B \}$.

DEFINITION 1.15 (Symmetric Difference). The *symmetric difference* of A and B is defined as $A \oplus B \triangleq (A \cup B) \setminus (A \cap B)$.

DEFINITION 1.16 (Complement). The *complement* of A with respect to the universal set U: $\overline{A} \triangleq U \setminus A$.

THEOREM 1.8 (Set Identities).
Let A, B, C be sets and U be the universal set.

$A \cup \emptyset = A$	$A \cap U = A$	*Identity*
$A \cup U = U$	$A \cap \emptyset = \emptyset$	*Domination*
$A \cup A = A$	$A \cap A = A$	*Idempotent*
$\overline{(\overline{A})} = A$		*Complementation*
$A \cup B = B \cup A$	$A \cap B = B \cap A$	*(Commutativity)*
$A \cup (B \cup C) = (A \cup B) \cup C$	$A \cap (B \cap C) = (A \cap B) \cap C$	*(Associativity)*
$A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$	$A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$	*(Distributivity)*
$\overline{A \cup B} = \overline{A} \cap \overline{B}$	$\overline{A \cap B} = \overline{A} \cup \overline{B}$	*(De Morgan)*
$A \cup (A \cap B) = A$	$A \cap (A \cup B) = A$	*(Absorption)*
$A \cup \overline{A} = U$	$A \cap \overline{A} = \emptyset$	*(Complement)*

PROOF OF $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$.

$$
\begin{array}{ll}
A \cap (B \cup C) & \\
= \{\, x \mid x \in A \cap (B \cup C) \,\} & \text{definition of membership} \\
= \{\, x \mid x \in A \ \wedge\ (x \in B \cup C) \,\} & \text{definition of } \cap \\
= \{\, x \mid x \in A \ \wedge\ (x \in B \ \vee\ x \in C) \,\} & \text{definition of } \cup \\
= \{\, x \mid (x \in A \ \wedge\ x \in B) \ \vee\ (x \in A \ \wedge\ x \in C) \,\} & \text{distributivity of } \wedge \text{ over } \vee \\
= \{\, x \mid (x \in (A \cap B) \ \vee\ (x \in (A \cap C) \,\} & \text{definition of } \cap \\
= \{\, x \mid x \in (A \cap B) \cup (A \cap C) \,\} & \text{definition of } \cup \\
= (A \cap B) \cup (A \cap C) & \text{definition of membership}
\end{array}
$$

□

2. Relation

REMARK 2.1. $A \times B = \{\, (a, b) \mid a \in A \ \wedge\ b \in B \,\}$.

DEFINITION 2.1 (Matrix). An array of numbers with n rows and m columns is called an $n \times m$ *matrix*. The entry at the ith row and jth column of matrix M is denoted by $[M]_{ij}$. A matrix with entries 0 and 1 only is called a *binary matrix*. Binary matrices are also called $(0, 1)-$*matrices*.

DEFINITION 2.2. α is called a *binary relation* from A to $B \overset{\Delta}{\longleftrightarrow} \alpha \subseteq A \times B$. We use the infix notation of $a\, \alpha\, b$ whenever $(a, b) \in \alpha$.

REMARK 2.2. If sets A and B are finite with $|A| = n$ and $|B| = m$, the elements of A and B can be listed in an arbitrary order as $A = \{\, a_1, a_2, \ldots, a_n \,\}$ and $B = \{\, b_1, b_2, \ldots, b_m \,\}$. Then binary relation $\alpha \subseteq A \times B$ can be represented by an $n \times m$ $(0, 1)-$matrix, denoted by M_α, as

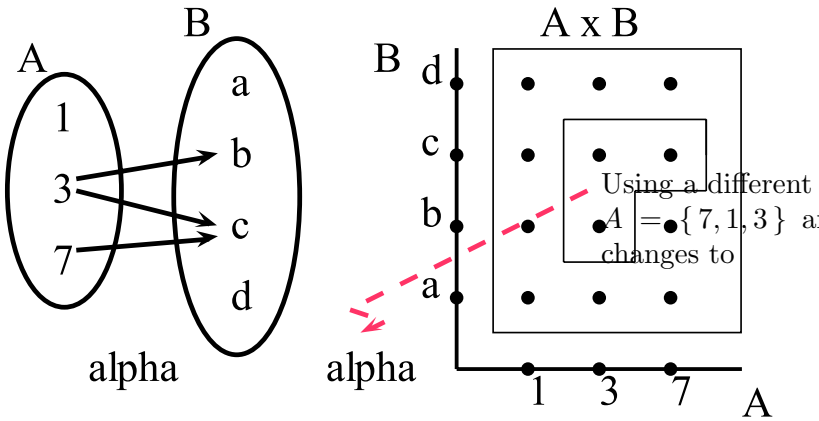
$$[M_\alpha]_{ij} \triangleq \begin{cases} 1, & a_i\, \alpha\, b_j \\ 0, & \textit{otherwise}. \end{cases}$$

Note that there are n rows correspond to the ordered elements of A, and m columns correspond to the ordered elements of B.

EXAMPLE 2.1.
Let $\alpha = \{\, (3, b), (3, c), (7, c) \,\} \subseteq A \times B$ where $A = \{\, 1, 3, 7 \,\}$ and $B = \{\, a, b, c, d \,\}$.

Using the orderings of $A = \{\, 1, 3, 7 \,\}$ and $B = \{\, a, b, c, d \,\}$ we have

$$M_\alpha = \begin{array}{c} \\ 1 \\ 3 \\ 7 \end{array} \begin{array}{c} \begin{array}{cccc} a & b & c & d \end{array} \\ \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \end{array}$$



Using a different orderings such as $A = \{\, 7, 1, 3 \,\}$ and $B = \{\, c, a, d, b \,\}$ the matrix changes to

$$M_\alpha = \begin{array}{c} \\ 7 \\ 1 \\ 3 \end{array} \begin{array}{c} \begin{array}{cccc} c & a & d & b \end{array} \\ \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 \end{bmatrix} \end{array}.$$

REMARK 2.3. The cartesian and the matrix representations are related. Rotate the cartesian representation by 90° clockwise, and compare with the matrix representation.

QUESTION 2.1. How many different binary relations from A to B can be defined?

2.1. Composition of Relations.

DEFINITION 2.3 (Composition of Relations).
Let $\alpha \subseteq A \times B, \beta \subseteq B \times C$. The *composition* of α and β, denoted by $\alpha \circ \beta$, is defined as $\alpha \circ \beta \triangleq \{\, (a,c) \in A \times C \mid \exists b \in B \;\; [a \,\alpha\, b \;\wedge\; b \,\beta\, c] \,\}$.

REMARK 2.4. Note that $\alpha \circ \beta \subseteq A \times C$. Note that this notation of composition is different that the notation of composition of functions which will be discussed at Sec. 3.

DEFINITION 2.4 (Boolean Matrix Multiplication).
Let M_α and M_β be $n \times m$ and $m \times p$ binary matrices. The *binary product* of M_β and M_α, denoted by $M_\alpha \odot M_\beta$, is an $n \times p$ binary matrix defined as

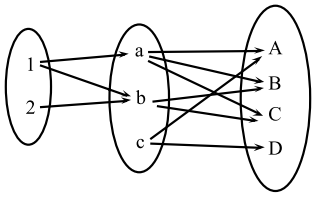
$$[M_\alpha \odot M_\beta]_{ij} \triangleq \begin{cases} 1, & \exists k \;\; [1 \le k \le m \;\wedge\; [M_\alpha]_{ik} = 1 \;\wedge\; [M_\beta]_{kj} = 1] \\ 0, & \textit{otherwise.} \end{cases}$$

REMARK 2.5. Binary matrix multiplication can be defined by means of logic functions.

$$[M_\alpha \odot M_\beta]_{ij} = \bigvee_{k=1}^{m} [M_\alpha]_{ik} \;\wedge\; [M_\beta]_{kj}$$

where $\wedge$ and $\vee$ are logical AND and OR functions. The notation $\bigvee_{k=1}^{m}$, is similar to $\sum_{k=1}^{m}$, is defined as $\bigvee_{k=1}^{n} \triangleq p_1 \vee p_2 \vee \cdots \vee p_n$.

EXAMPLE 2.2. Using orders $A = \{\, 1, 2 \,\}$, $B = \{\, a, b, c \,\}$ and $C = \{\, A, B, C, D \,\}$:



$$\begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \end{bmatrix} \odot \begin{bmatrix} 1 & 1 & 1 & 0 \\ 0 & 1 & 1 & 0 \\ 1 & 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 & 0 \\ 0 & 1 & 1 & 0 \end{bmatrix}$$
$$M_\alpha \quad \odot \quad M_\beta \quad = \quad M_{\alpha\circ\beta}$$

But regular matrix multiplication gives:

$$\begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \end{bmatrix} \times \begin{bmatrix} 1 & 1 & 1 & 0 \\ 0 & 1 & 1 & 0 \\ 1 & 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 2 & 2 & 0 \\ 0 & 1 & 1 & 0 \end{bmatrix}$$
$$M_\alpha \quad \times \quad M_\beta \quad = \quad M_\alpha \times M_\beta.$$

THEOREM 2.1. $M_{\alpha\circ\beta} = M_\alpha \odot M_\beta$.

THEOREM 2.2 (Associativity).
$(\alpha \circ \beta) \circ \gamma = \alpha \circ (\beta \circ \gamma)$ *whenever* $(\alpha \circ \beta) \circ \gamma$ *is defined.*

COROLLARY 2.3 (Associativity).
$M_\alpha \odot (M_\beta \odot M_\gamma) = (M_\alpha \odot M_\beta) \odot M_\gamma$ *whenever* $M_\alpha \odot (M_\beta \odot M_\gamma)$ *is defined.*

DEFINITION 2.5 (Inverse of a Binary Relation). The *inverse* of a binary relation, denoted by α^{-1}, is defined as $b\,\alpha^{-1}\,a \overset{\Delta}{\longleftrightarrow} a\,\alpha\,b$.

DEFINITION 2.6. The *transpose* of a matrix, denoted by $M^\top$, is defined as $[M^\top]_{ij} \triangleq [M]_{ji}$.

THEOREM 2.4. $M_{\alpha^{-1}} = (M_\alpha)^\top$.

THEOREM 2.5. $[\alpha \circ \beta]^{-1} = \beta^{-1} \circ \alpha^{-1}$.

DEFINITION 2.7. The *complement* of α, denoted as $\overline{\alpha}$, is defined as $a\,\overline{\alpha}\,b \overset{\Delta}{\longleftrightarrow} \neg\, a\,\alpha\,b$.

THEOREM 2.6. $M_{\bar{\alpha}} = \mathbf{1} - M_\alpha$ *where* $\mathbf{1}$ *is matrix of all* 1*s.*

EXAMPLE 2.3. Show that $(\overline{\alpha})^{-1} = \overline{(\alpha^{-1})}$
$(a,b) \in (\overline{\alpha})^{-1} \Leftrightarrow (b,a) \in \overline{\alpha} \Leftrightarrow (b,a) \notin \alpha \Leftrightarrow (a,b) \notin \alpha^{-1} \Leftrightarrow (a,b) \in \overline{(\alpha^{-1})}$.

3. Functions

REMARK 3.1. Let f be a relation from A to B. Pick an $a \in A$ and consider the corresponding set $B_a \subseteq B$ defined as $B_a \triangleq \{\, b \mid (a,b) \in f \,\}$. Note that $|\, B_a \,|$ could be 0, 1, 2, ...

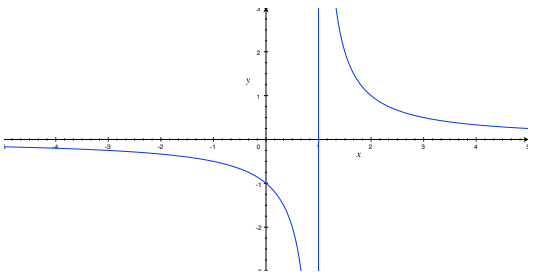
If $|\, B_a \,| = \left| \begin{array}{l} 0 \\ 1 \\ n \geq 2 \end{array} \right|$, then a is $\left| \begin{array}{l} \text{not mapped to any } b \in B \\ \text{mapped to exactly one } b \in B \\ \text{mapped to } n \text{ elements of } B \end{array} \right|$.

DEFINITION 3.1. A relation $f \subseteq A \times B$ is called *partial function*
$\overset{\Delta}{\longleftrightarrow} \forall a \in A \;\; |\, B_a \,| \leq 1$.

REMARK 3.2. Any computer program is actually a partial function from its input space to its output space. For some inputs in its domain it terminates and produces outputs. For some other inputs it does not terminate. Hence for those inputs there is no corresponding outputs. That is the reason that it is a partial function.

DEFINITION 3.2. A relation $f \subseteq A \times B$ is called *function*
$\overset{\Delta}{\longleftrightarrow} \forall a \in A \;\; |\, B_a \,| = 1$.

EXAMPLE 3.1. Remember function $y = f(x) = \frac{1}{x-1}$ from Calculus. It is considered to be a function from $\mathbb{R}$ to $\mathbb{R}$. Properly speaking this statement is not true since it is not defined at $x = 1 \in \mathbb{R}$.



Actually, it is a function from $\mathbb{R} \setminus \{\, 1 \,\}$ to $\mathbb{R}$. On the other hand, it is a partial function from $\mathbb{R}$ to $\mathbb{R}$.

DEFINITION 3.3. $\left| \begin{array}{l} A \\ B \\ f(A) \end{array} \right|$ is called the $\left| \begin{array}{l} \textit{domain} \\ \textit{codomain} \\ \textit{range} \end{array} \right|$ of f and written as $\left| \begin{array}{l} \text{dom } f \\ \text{cod } f \\ \text{ran } f \end{array} \right|$.

QUESTION 3.1. Consider the ceiling function $f(x) = \lceil\, x \,\rceil$. dom f = ?, cod f = ?, ran f = ?

NOTATION.
- $b = f(a) \overset{\Delta}{\longleftrightarrow} a\,f\,b$
- A function f from A to B is represented by $f : A \to B$.
- The set all functions from A to B is represented by B^A.
- $f(C) = \{\, f(c) \mid c \in C \,\}$ for $C \subseteq A$.

REMARK 3.3.
Let $\left| \begin{array}{l} B^A \\ \mathcal{P} \\ \mathcal{R} \end{array} \right|$ be the set of all $\left| \begin{array}{l} \text{functions} \\ \text{partial functions} \\ \text{relations} \end{array} \right|$ from A to B. Then $B^A \subseteq \mathcal{P} \subseteq \mathcal{R}$.

QUESTION 3.2.

- $\left|\, B^A \,\right| = ?$
- $|\, \mathcal{P} \,| = ?$
- $|\, \mathcal{R} \,| = ?$

THEOREM 3.1. *If A and B are finite sets, not both empty, then* $\left|\, B^A \,\right| = |\, B \,|^{|\, A \,|}$.

QUESTION 3.3. Let A be a nonempty set.

- $\left|\, \emptyset^A \,\right| = ?$
- $\left|\, A^\emptyset \,\right| = ?$
- $\left|\, \emptyset^\emptyset \,\right| = ?$

DEFINITION 3.4. Let $f : A \to B$ and $A_S \subseteq A \subseteq A_L$.
Function $f_S \subseteq A_S \times B$ is the *restriction* of f to A_S
$\overset{\Delta}{\longleftrightarrow}$ $[\forall a \in A, \forall b \in B\ ((a,b) \in f_S \iff (a,b) \in f)]$.
Then partial function $f_L \subseteq A_L \times B$ is an *extension* of f to A_L
$\overset{\Delta}{\longleftrightarrow}$ $[\forall a \in A_1, \forall b \in B\ ((a,b) \in f_L \iff (a,b) \in f)]$.

REMARK 3.4 (The inverse of a function). Note that for any relation α from A to B, there is a unique inverse relation from B to A. This inverse relation usually denoted by α^{-1}. Since a function f from A to B is also a relation, there is an inverse relation from B to A which is also denoted as f^{-1}. Note that f^{-1} is a relation but not necessarily a function. f^{-1} becomes a function if and only if f is a bijection.

DEFINITION 3.5. Let $f : A \to B$.

f is called		$\overset{\Delta}{\longleftrightarrow}$	
	the *identity*		$A = B \ \wedge\ \forall a \in A\ f(a) = a$
	a *surjection*		$f(A) = B$
	an *injection*		$\forall a_1, a_2 \in A\ [a_1 \neq a_2 \to f(a_1) \neq f(a_2)]$
	a *bijection*		surjection $\wedge$ injection
	a *permutation*		$f : A \to A$ and f is a bijection

THEOREM 3.2. *Let $f : A \to B$ and $g : B \to C$ be functions If f and g are* | *surjections* / *injections* / *bijections* |, *then* $g \circ f$ *is also* | *a surjection* / *an injection* / *a bijection* |.

REMARK 3.5 (Composition of functions). The notations of composition of relations and composition of functions are unfortunately inconsistent. Let $f : A \to B$ and $g : B \to C$. Then the composition of function f with function g is represented by $g \circ f$. On the other hand, the composition of relation f with relation g is $f \circ g$. Although we end up with two different expressions, they actually represent the same set

$$\{(a,c) \in A \times C \mid a \in A,\ c \in C,\ \exists b \in B\ ((a,b) \in A \times B \wedge (b,c) \in B \times C)\}.$$

When f and g are functions, we always use functional interpretation and notation.

Functional interpretations. In traditional notation for functions, we have $b = f(a)$ and $c = g(b)$. Hence $c = g(b) = g(f(a)) = (g \circ f)(a)$.

Relational interpretations. Since functions are special relations, we can apply composition of relations to these special relations. Then functions f and g are also two relations. So in relation notation, we have afb and bgc. Then $f \circ g$ is the composition of relation f with relation g. Note that $f \circ g$ is relation from A to C. Hence we have $a(f \circ g)c$.

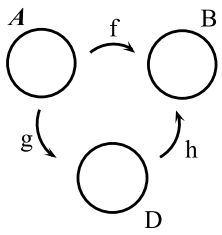
This inconsistency in notation is probably due to the convenience of matrix representations in the composition of relations. Remember that $M_{\alpha\circ\beta} = M_\alpha \odot M_\beta$ where $M_{\alpha\circ\beta}$ is the matrix of the composition of α with β. A formula such as $M_{\beta\circ\alpha} = M_\alpha \odot M_\beta$ would not be that convenient.

THEOREM 3.3. *Let $f : A \to B$ and $g : B \to C$ be functions. Then the composition of functions $g \circ f$ is a function: $g \circ f : A \to C$ where $(g \circ f)(a) = g(f(a))$.*

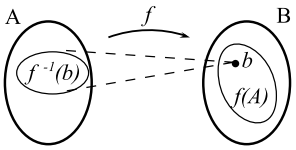
Remark 3.6.

- If functions $f : A \to B$ and $g : B \to C$ are invertible, then the function $g \circ f$ is also invertible and $(g \circ f)^{-1} = f^{-1} \circ g^{-1}$.
- $f^{-1}(B) = \{ a \in A \mid f(a) \in B \}$
- Let $f : \mathbb{R} \to \mathbb{R}$. Then f and f^{-1} are symmetric with respect to $y = x$ line.

Example 3.2.
Show that any function $f : A \to B$ can be represented as the composition of functions g and h, $f = h \circ g$, where g is a surjection, h is an injection.



Define $D = \{ D_i \subseteq A \mid d_1, d_2 \in D_i \Leftrightarrow f(d_1) = f(d_2) \}$.
Define $g : A \to D \ni g(a) = D_i = f^{-1}(f(a))$.
Then, clearly g is a surjection since $\forall D_i \in D \ \left[\exists a \in A \ \left[D_i = f^{-1}(f(a))\right]\right]$, hence $g(a) = D_i$.
Define $h : D \to B \ni h(D_i) = h(f^{-1}(f(a))) = f(a)$.
Let $D_i \neq D_j$. Then, for $d_i \in D_i$ and $d_j \in D_j$, $f(d_i) \neq f(d_j)$. So $h(D_i) \neq h(D_j)$. Therefore, h is an injection.
$(h \circ g)(a) = h(g(a)) = h(f^{-1}(f(a))) = f(a)$
So, $h \circ g = f$.



Question 3.4.

- Are D_i's disjoint?
- What is $f^{-1}(B)$?
- What is $f(f^{-1}(B))$?
- What kind of function is f if $B \backslash f(A) = \emptyset$?
- What kind of function is f if $\forall b \in B \ \left[\left| f^{-1}(b) \right| = 1\right]$?
- If $|A| = n$, $|B| = n$, what is the number different functions f ?

Remark 3.7. Let $f : A \to B$ be a function and A and B be finite sets. If f is $\left|\begin{array}{l}\text{an injection}\\ \text{a surjection}\\ \text{a bijection}\end{array}\right|$.
then $\left|\begin{array}{l}|A| \leq |B|\\ |A| \geq |B|\\ |A| = |B|\end{array}\right|$.

For functions from a set to itself has an interesting property that it can be applied repeatedly. Let $f : A \to A$ be a function and $a \in A$. Then $f(a)$, $f(f(a)), \ldots$, $f(f(f(a))), \ldots$ are all defined.

Definition 3.6 (Power of a function). Let $f : A \to A$ be a function. Power of a function is defined as
i. $f^1 \triangleq f$.
ii. $f^{n+1} \triangleq f \circ f^n$ for $n \in \mathbb{Z}^+$.

Definition 3.7 (fixed point). Let $f : A \to A$ be a function. $a \in A$ is called a *fixed point* of $f \overset{\Delta}{\longleftrightarrow} f(a) = a$,

REMARK 3.8. Fixed points are important in Computer Science. If a is a fixed point of f then $f^n(a) = a$ for all $n \in \mathbb{N}$.

QUESTION 3.5. Consider functions from $\mathbb{R}$ to $\mathbb{R}$. Let $a, b, c \in \mathbb{R}$.

- What are the fixed points of $f(x) = ax + b$ where a and b are real parameters? Consider the cases where $a = 1$ and $b = 1$, $a = 2$ and $b = 1$, and $a = 1$ and $b = 0$.
- What are the fixed points of $f(x) = ax^2 + bx + c$ where a, b and c are real parameters? Consider the cases for different values of a, b and c.
- What are the fixed points of $f(x) = \sin x$?
- The function $f(x) = rx(1 - x)$ from $\mathbb{Z}$ to $\mathbb{Z}$ is called the *logistics map* where $r \in \mathbb{R}$ is a parameter [**Str94**]. Although it seems simple logistic map has unexpectedly rich properties if it is applied iteratively, i.e. $x_{n+1} = rx_n(1 - x_n)$. Try to plot logistic map for $r = 2.8$, $r = 3.3$, $r = 3.5$, $r = 3.857$.

Acknowledgment. These notes are based on various books but especially [**PY73, Ros07, TZ82, Gal89, Hol60, Nes09**].

Problems with Solutions

P 3.1. a) Prove or disprove that set difference distributes over union, that is,

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

b) Given a nonempty set A, let $f : A \to A$ and $g : A \to A$ where

$$\forall a \in A \;\; f(a) = g(f(f(a))) \text{ and } g(a) = f(g(f(a)))$$

Prove that $f = g$.

Solution.

a) $A - (B \cup C) = (A - B) \cup (A - C)$
Consider the following counter example which disproves the statement.
Let $A = \{1, 2, 4, 5\}$, $B = \{2, 3, 5, 6\}$ and $C = \{4, 5, 6, 7\}$.
Then $A - (B \cup C) = \{1, 2, 4, 5\} - \{2, 3, 4, 5, 6, 7\} = \{1\}$ and
$(A - B) \cup (A - C) = (\{1, 2, 4, 5\} - \{2, 3, 5, 6\}) \cup (\{1, 2, 4, 5\} - \{4, 5, 6, 7\}) = \{1, 2, 4\}$.
Hence, $A - (B \cup C) \neq (A - B) \cup (A - C)$.

b) Proof by contradiction: Let $f : A \to A$ and $g : A \to A$ and
$\forall a \in A \;\; f(a) = g(f(f(a)))$ (I), and $g(a) = f(g(f(a)))$ (II), but $f \neq g$.
Then $\exists s \in A \;\; f(s) \neq g(s)$.

$$\begin{array}{lrll}
 & \text{f(s)} & \neq g(s) & \\
\Leftrightarrow & g(f(f(s))) & \neq g(s) & \text{since } f = g(f(f)) \\
\Leftrightarrow & f(g(f(f(f(s))))) & \neq g(s) & \text{since } g = f(g(f)) \\
\Leftrightarrow & f(\underbrace{g(f(f}_{f}(f(s))))) & \neq g(s) & \\
\Leftrightarrow & f(f(f(s))) & \neq g(s) & \text{since } f = g(f(f)) \\
\Leftrightarrow & f(f(\underbrace{\quad f \quad}_{g(f(f(s)))})) & \neq g(s) & \\
\Leftrightarrow & f(f(g(f(f(s))))) & \neq g(s) & \text{since } f = g(f(f)) \\
\Leftrightarrow & f(\underbrace{f(g(f}_{g}(f(s))))) & \neq g(s) & \\
\Leftrightarrow & f(g(f(s))) & \neq g(s) & \text{since } g = f(g(f)) \\
\Leftrightarrow & g(s) & \neq g(s) & \text{since } g = f(g(f)). \text{ Contradiction!}
\end{array}$$

Hence, $f = g$.

CHAPTER 4

Relations on a Set

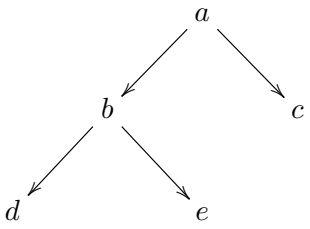
1. Relations on a Set

DEFINITION 1.1. Let ρ be a relation on A, that is $\rho \subseteq A \times A$.

$$\rho \text{ is called } \left| \begin{array}{l} \textit{reflexive} \\ \textit{symmetric} \\ \textit{antisymmetric} \\ \textit{transitive} \end{array} \right| \overset{\Delta}{\longleftrightarrow} \left| \begin{array}{l} \forall a \in A\ [a \rho a] \\ \forall a, b \in A\ [a \rho b \longrightarrow b \rho a] \\ \forall a, b \in A\ [a \rho b \wedge b \rho a \longrightarrow a = b] \\ \forall a, b, c \in A\ [a \rho b \wedge b \rho c \longrightarrow a \rho c] \end{array} \right|$$

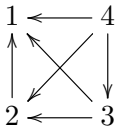
THEOREM 1.1. *If* ρ *is* $\left| \begin{array}{l} \textit{reflexive} \\ \textit{symmetric} \\ \textit{antisymmetric} \\ \textit{transitive} \end{array} \right|$ *then* ρ^{-1} *is* $\left| \begin{array}{l} \textit{reflexive} \\ \textit{symmetric} \\ \textit{antisymmetric} \\ \textit{transitive} \end{array} \right|$.

EXAMPLE 1.1.



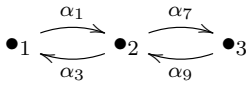
$$\begin{bmatrix} 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

A tree
$\overline{R}\overline{S}\overline{T}$



$$\begin{bmatrix} 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 1 & 1 & 0 \end{bmatrix}$$

greater then relation
$\overline{R}\overline{S}\overline{T}$



$$\begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}$$

$a \alpha b \overset{\Delta}{\longleftrightarrow} |a - b| = 1$
$\overline{R}\overline{S}\overline{T}$

2. Observations on the Matrix of a Relation

Let $\alpha \subseteq A \times A$.

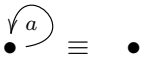
- α is ordinary $\longrightarrow$ No pattern in the matrix.

$$\begin{bmatrix} 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 \\ 1 & 0 & 1 & 0 \end{bmatrix}$$

Use directed graph.

- α is reflexive $\longleftrightarrow$ The main diagonal is all 1's.

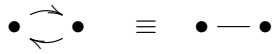
$$\begin{bmatrix} \boxed{1} & 1 & 0 & 0 \\ 0 & \boxed{1} & 1 & 0 \\ 1 & 0 & \boxed{1} & 1 \\ 1 & 0 & 1 & \boxed{1} \end{bmatrix}$$



Omit loops.

- α is symmetric $\longleftrightarrow$ The matrix is symmetric.

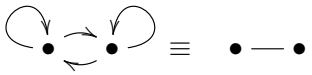
$$\begin{bmatrix} \boxed{1} & 1 & 0 & 0 \\ 1 & \boxed{0} & 1 & 0 \\ 0 & 1 & \boxed{1} & 0 \\ 0 & 0 & 0 & \boxed{1} \end{bmatrix} \begin{bmatrix} 1 & . & . & . \\ 1 & 0 & . & . \\ 0 & 1 & 1 & . \\ 0 & 0 & 0 & 1 \end{bmatrix}$$



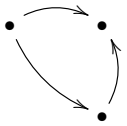
Use undirected graph.

- α is both reflexive and symmetric.

$$\begin{bmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 1 & 0 \\ 1 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 \end{bmatrix} \begin{bmatrix} . & . & . & . \\ 0 & . & . & . \\ 1 & 1 & . & . \\ 0 & 0 & 1 & . \end{bmatrix}$$



- α is transitive.



3. Closure of Relations

Example 3.1. Given a relation α_0 which is not reflexive, a new relation α_1 can be defined which is reflexive and $\alpha_0 \subseteq \alpha_1$. More than that, there are many reflexive relations α_j with $\alpha_0 \subseteq \alpha_j$. Note that α_1 is the smallest one satisfying this.

$$\alpha_0 = \begin{bmatrix} \boxed{1} & 1 & 0 & 0 \\ 0 & \boxed{1} & 1 & 0 \\ 1 & 0 & \boxed{1} & 1 \\ 1 & 0 & 1 & 0 \end{bmatrix} \quad \alpha_1 = \begin{bmatrix} \boxed{1} & 1 & 0 & 0 \\ 0 & \boxed{1} & 1 & 0 \\ 1 & 0 & \boxed{1} & 1 \\ 1 & 0 & 1 & \boxed{1} \end{bmatrix} \quad \alpha_2 = \begin{bmatrix} \boxed{1} & 1 & 1 & 0 \\ 0 & \boxed{1} & 1 & 0 \\ 1 & 0 & \boxed{1} & 1 \\ 1 & 0 & 1 & \boxed{1} \end{bmatrix}$$

Definition 3.1. Let α be a relation on a set A. Let P be a property such as reflexivity, symmetry, transitivity. β is called *closure* of α with respect to $P \overset{\Delta}{\longleftrightarrow}$ β is a relation with property P and $\alpha \subseteq \beta$ with $\forall\gamma \ [\alpha \subseteq \gamma]$ where γ is a relation with property P and $\beta \subseteq \gamma$

Remark 3.1. Note that β is the smallest relation satisfying this.

4. Compatibility Relation

Let $\gamma \subseteq A \times A$.

Definition 4.1 (Compatibility Relation).
A relation γ is a *compatibility relation* $\overset{\Delta}{\longleftrightarrow}$

i. γ is reflexive
ii. γ is symmetric

Remark 4.1. Note that equivalence relation has one more property, namely transitivity.
β is an equivalence relation $\longrightarrow$ β is a compatibility relation.

Definition 4.2. $C \subseteq A$ is called a *compatibility class* (*compatible*) $\overset{\Delta}{\longleftrightarrow}$
$\forall c_1, c_2 \in C \ [c_1 \, \gamma \, c_2]$ where γ is a compatibility relation.

Definition 4.3. A compatibility class which is not properly contained in any other compatibility class is called a *maximal compatibility class* (*maximal compatible*).

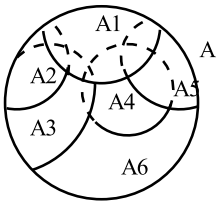
Definition 4.4. A *complete cover*, $C_\gamma(A)$, of A with respect to γ is a collection of all and only the maximal compatibles induced by γ.
$\overset{\Delta}{\longleftrightarrow}$ A collection, $C_\gamma(A)$, of all and only the maximal compatibles induced by γ on A is called a *complete cover* of A.

THEOREM 4.1. *If γ is compatibility relation on a finite set A and C is a compatibility class, then there is a maximal compatibility class C' such that $C \subseteq C'$.*

THEOREM 4.2. *There is a one-to-one correspondence between γ and $C_\gamma(A)$.*

THEOREM 4.3. *γ is a compatibility relation on $A \overset{\Delta}{\longleftrightarrow} \exists$ relation ρ from A to some $B \ni \gamma = \rho\rho^{-1}$ with $\forall a \in A \ [\exists b \in B \ [a \, \rho \, b]]$*

EXAMPLE 4.1.



Complete cover

$C_\gamma(A) = \{A_1, A_2, A_3, A_4, A_5, A_6\}$ is a complete cover.
$C'_\gamma(A) = \{A_1, A_2, A_3, A_4, A_5, A_7\}$ is not since $A_4 \subseteq A_7$.

4.1. Application of Compatibility Relation.

EXAMPLE 4.2 (Minimization of Incompletely Specified Finite State Machines). $S = \{a, b, c, d, e\}$

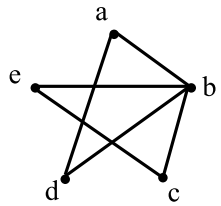
	I_1	I_2	I_3
a	c,0	e,1	-
b	c,0	e,-	-
c	b,-	c,0	a,-
d	b,0	c,-	e,-
e	-	e,0	a,-

DEFINITION 4.5. States a and b are *compatible*, $a\gamma b$, where $\gamma \subseteq S \times S \overset{\Delta}{\longleftrightarrow}$ If no applicable input sequence to both a and b produce conflicting outputs.

γ is a compatibility relation since $\forall a, b \in S$

i. $a\gamma a$.
ii. $a\gamma b \longrightarrow b\gamma a$.

a	$\surd$				
b	c,c e,e	$\surd$			
c	$\times$	b,c c,e	$\surd$		
d	b,c c,e	b,c c,e	$\times$	$\surd$	
e	$\times$	e,e	c,e a,a	$\times$	$\surd$
	a	b	c	d	e



$\{a, b\}$ is a compatible.
$\{a, b, d\}$ is a maximal compatible.
$\{\{a, b, d\}, \{b, c, e\}\}$ is a complete cover.
Let $A = \{a, b, d\}, B = \{b, c, e\}$

	I_1	I_2	I_3
A	B,0	B,1	B,-
B	B,0	B,0	A,-

5. Equivalence Relation

DEFINITION 5.1 (Equivalence Relation).

A relation γ is a *equivalence relation* $\overset{\Delta}{\longleftrightarrow}$

i. γ is reflexive

ii. γ is symmetric
iii. γ is transitive

THEOREM 5.1. *If γ is an equivalence relation $\longrightarrow$ γ is also a compatibility relation.*

DEFINITION 5.2 (Equivalence Class).
A maximal compatible of γ is called an *equivalence class* where γ is an equivalence relation.

THEOREM 5.2. *Let $\{ E_i \}$ be complete cover. $\forall E_1, E_2 \in C_\gamma(A)$ $[E_i \cap E_j = \emptyset]$*

THEOREM 5.3. *$\forall a \in A$ [a belongs to one and only one equivalence class].*

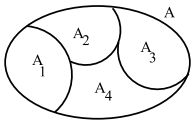
DEFINITION 5.3 (Partition).
A set $P = \{ A_i \neq \emptyset \mid A_i \subseteq A \}$ is called a *partition* of $A \overset{\Delta}{\longleftrightarrow}$
i. $\bigcup_i A_i = A$
ii. $A_i \cap A_j = \emptyset$ if $i \neq j$.
Each A_i is called a *block* of P.
P is called the *partition of singletons* $\overset{\Delta}{\longleftrightarrow}$ $\forall A_i \in P$ $[|\, A_i \,| = 1]$.
The partition of singletons and the partition $\{ A \}$ are called the *trivial partitions*.

EXAMPLE 5.1. Let $A = \{1,2,3,4,5\}$. Then sets $P_1 = \{\{1\},\{2\},\{3\},\{4\},\{5\}\}$, $P_2 = \{\{1\},\{2,3,4,5\}\}$, $P_3 = \{\{1\},\{2,4,5\},\{3\}\}$, $P_4 = \{\{1,2\},\{3,5\},\{4\}\}$, $P_5 = \{\{1,2,3,4,5\}\}$ are partitions of A.
Note that P_1 and P_5 are the trivial partitions.

THEOREM 5.4. *There is a one-to-one corresponding between equivalence relations on A and partitions on A.*

DEFINITION 5.4. *Dichotomy* is a partition with two blocks.

THEOREM 5.5. *γ is an equivalence relation on A* $\longleftrightarrow$ $\exists B$ $\left[\exists f : A \to B \ \left[\gamma = f f^{-1}\right]\right]$.



$P = \{ A_1, A_2, A_3, A_4 \}$.

THEOREM 5.6. *Let α and β be equivalence relations on A.*
α β is an equivalence relation $\longleftrightarrow$ $\alpha\beta = \beta\alpha$

PROOF.
($\Rightarrow$ part:) $\alpha\beta$ is an equivalence relation $\longrightarrow$ $\alpha\beta$ is symmetric $\longrightarrow$ $\alpha\beta = (\alpha\beta)^{-1} = \beta^{-1}\alpha^{-1} = \beta\alpha$.
Since α and β are symmetric.
($\Leftarrow$ part:)
i. $\forall a \in A$ $[a\alpha a \ \wedge \ a\beta a]$ $\longrightarrow$ $a\alpha\beta a$. reflexivity.
ii. $\forall a, b \in A$ $a(\alpha\beta)b$ $\longrightarrow$ $a(\beta\alpha)b$ $\longrightarrow$ $\exists d \in A$ $[a\beta d \ \wedge \ d\alpha b]$ $\longrightarrow$ $d\beta a \ \wedge \ b\alpha d$ $\longrightarrow$ $b(\alpha\beta)a$. symmetry.
iii. Left as an exercise. □

5.1. Applications of Equivalence Relations. An application of equivalence relations is state reduction of a completely specified FSM.

DEFINITION 5.5. A *finite state machine* (*FSM*) is a system $M = [Q, S, R, \alpha, \beta]$ where
Q is a finite set of states
S is a finite set of input symbols (input of alphabet, stimulus)
R is a finite set of output symbols (output alphabet, response)
$\alpha : Q \times S \to Q$ is the state function
$\beta : Q \times S \to R$ is the output function

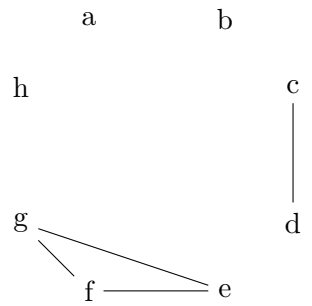
EXAMPLE 5.2 (State Reduction of a Completely Specified Finite State Machine).
Let FSM be defined as: $Q = \{ a, b, c, d, e, f, g, h \}$ $S = \{ I_1, I_2 \}$ $R = \{ 0, 1 \}$ and α and β are defined in the following table:

α, β	I_1	I_2
a	$b, 1$	$h, 1$
b	$f, 1$	$d, 1$
c	$d, 0$	$e, 1$
d	$c, 0$	$f, 1$
e	$d, 1$	$c, 1$
f	$c, 1$	$c, 1$
g	$c, 1$	$d, 1$
h	$c, 0$	$a, 1$

Define $\equiv \subseteq Q \times Q$ as follows:
a is *equivalent* to $b \longleftrightarrow a \equiv b$
$\overset{\Delta}{\longleftrightarrow}$ no input sequence can distinguish a from b.

$\equiv$ is an equivalence relation
since $\forall a, b, c \in S$
i. $a \equiv a$
ii. $a \equiv b \longrightarrow b \equiv a$
iii. $a \equiv b, b \equiv c \longrightarrow a \equiv c$

	a	b	c	d	e	f	g	h
a								
b	~~b,f~~ ~~d,h~~							
c	X	X						
d	X	X	c,d e,f					
e	~~b,d~~ ~~c,h~~	~~d,f~~ ~~c,d~~	X	X				
f	~~b,c~~ ~~c,h~~	~~c,f~~ ~~c,d~~	X	X	c,d c,c			
g	~~b,c~~ ~~d,h~~	~~c,f~~ ~~d,d~~	X	X	c,d c,d	c,c c,d		
h	X	X	~~c,d~~ ~~a,e~~	~~c,c~~ ~~a,f~~	X	X	X	



complete cover:
$\{ \{ a \}, \{ b \}, \{ c, d \}, \{ e, f, g \}, \{ h \} \}$

Let $A = \{ a \}, B = \{ b \}, C = \{ c, d \}, E = \{ e, f, g \}, H = \{ h \}$

	I_1	I_2
A	B,1	H,1
B	E,1	C,1
C	C,0	E,1
E	C,0	C,1
H	C,1	A,1

Acknowledgment. These notes are based on various books but especially [**PY73, Ros07, TZ82, Gal89**]. Class of CMPE220 of Fall 2008 did the initial LaTeX draft of hand written notes.

Problems with Solutions

P 4.1. Let $\rho \subseteq A \times A$, ρ^{-1} be the inverse relation of ρ, and i_A be the identity relation of A. What kind of relation is ρ if

i) $i_A \subseteq \rho$.
ii) $i_A \cap \rho = \emptyset$.
iii) $\rho^{-1} = \rho$.
iv) $\rho \cap \rho^{-1} \subseteq i_A$.
v) $\rho \cap \rho^{-1} = \emptyset$.
vi) $\rho = \cup_{i \in \mathbb{N}} \rho^i$.
Justify your answers.

Solution.

If $\left| \begin{array}{l} i_A \subseteq \rho \\ i_A \cap \rho = \emptyset \\ \rho^{-1} = \rho. \\ \rho \cap \rho^{-1} \subseteq i_A. \\ \rho \cap \rho^{-1} = \emptyset. \\ \rho = \cup_{i \in \mathbb{N}} \rho^i. \end{array} \right|$, then ρ is $\left| \begin{array}{l} \text{reflexive} \\ \text{irreflexive} \\ \text{symmetric} \\ \text{antisymetric} \\ \text{asymmetric} \\ \text{transitive} \end{array} \right|$.

Justification is left as exercise.

CHAPTER 5

Partial Ordering, Lattice

1. Partial Ordering

DEFINITION 1.1. Let $\leq$ be a relation on A, $\leq \subseteq A \times A$.

$\leq$ is a *Partial Ordering* $\overset{\Delta}{\longleftrightarrow}$

i. reflexive
ii. antisymmetric
iii. transitive.

$a < b \overset{\Delta}{\longleftrightarrow} a \leq b \wedge a \neq b$.

THEOREM 1.1. *The inverse relation $\leq^{-1}$ of a partial ordering $\leq$ is also a partial ordering, denoted by $\geq$.*

THEOREM 1.2. *The directed graph of a partial ordering relation contains no circuits of length greater than 1.*

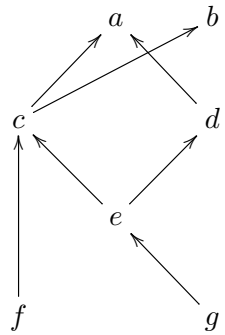
DEFINITION 1.2 (Poset).
A *partly ordered set* (*poset*), denoted $[A, \leq]$, consists of a set A and a partial ordering relation $\leq$ on A.

DEFINITION 1.3. Let $[A, \leq]$ be a poset. $a, b \in A$ are said to be *comparable* $\overset{\Delta}{\longleftrightarrow} a \leq b \vee b \leq a$.
$a, b \in A$ are *incomparable* $\overset{\Delta}{\longleftrightarrow}$ $a, b \in A$ are not comparable.

DEFINITION 1.4 (Linearly ordered set). $[A, \leq]$ is called a *linearly ordered set* $\overset{\Delta}{\longleftrightarrow}$ $\forall a, b \in A \ [a \leq b \vee b \leq a]$.

REMARK 1.1. Note that some elements of a poset are incomparable but every two elements in a totally ordered set should be comparable.

EXAMPLE 1.1.



Let A $= \{a, b, c, d, e, f, g\}$ and $b \leq a$ represented by $b \to a$.

Note that $c \leq a$, $f \leq a$. There should be an arc from f to a, too. In order to simplify the figure this kind of arcs are omitted.

a and b are not comparable. So there is no element that is larger than all the other elements. Similarly there is no element that is smaller than all the other elements.

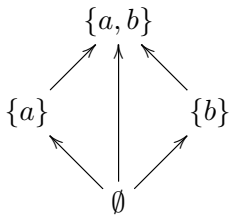
DEFINITION 1.5 (Consistent Enumeration).
A *consistent enumeration* of a finite poset A is a function $i : A \to \mathbb{N}$ such that
$\forall a_p, a_q \in A \ [a_p \leq a_q \longrightarrow i(a_p) \leq i(a_q)]$.

THEOREM 1.3. *Every finite poset admits of a consistent enumeration.*

EXAMPLE 1.2.

$[P(A), \subseteq]$ is a poset.

A $= \{a, b\}$
$2^A = \{\emptyset, \{a\}, \{b\}, \{a, b\}\}$


$\{a, b\}$
$\{a\}$
$\{b\}$
$\emptyset$

2. Hasse Diagram

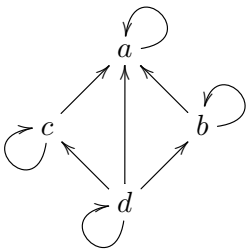
Definition 2.1. Let $[\,A, \leq\,]$ be a poset and $a, b \in A$ with $a \neq b$. a is an *immediate predecessor* of b, denoted by $a \prec b$, $\overset{\Delta}{\longleftrightarrow}$ $a < b$ and $\nexists c \in A\ [a < c < b]$.
b is an *immediate successor* of a, denoted by $b \succ a$, $\overset{\Delta}{\longleftrightarrow}$ $a < b$ and $\nexists c \in A\ [a < c < b]$.

Remark 2.1.

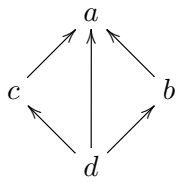
i. As convention an upper element is larger than a lower element.
ii. Immediate predecessor relation is
   - not reflexive
   - not symmetric
   - not transitive
iii. Given an immediate predecessor relation one can obtain the corresponding partial ordering.
iv. $\leq$ covers $\prec$.
v. Immediate relation simplifies the graph of partial ordering relation.

Definition 2.2. The graph of $\prec$ is called *Hasse Diagram*.

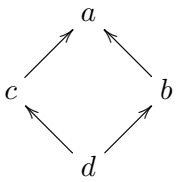
Example 2.1. Reflexive+Symmetric+Transitive



$R\overline{S}T$

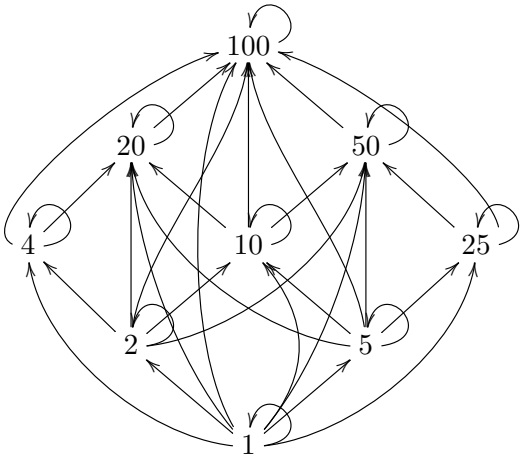


$\overline{RS}T$

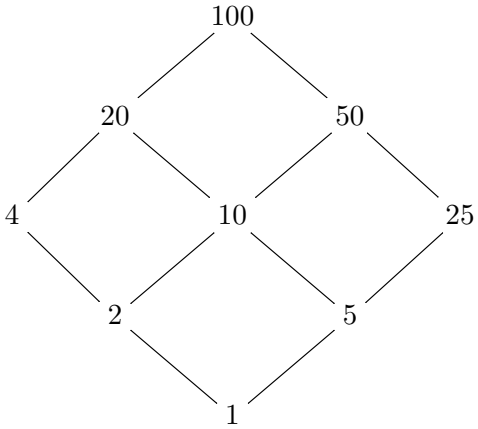


$\overline{RST}$

Example 2.2.
$a \mid b \overset{\Delta}{\longleftrightarrow} \exists c \in \mathbb{Z}\ [b = ca]$
Let $A = \{\, a \in \mathbb{N} \mid a \mid 100 \,\}$ and define a relation $\leq$ on A as $a \leq b \overset{\Delta}{\longleftrightarrow} a \mid b$.



Graph of $\leq$



Graph of $\prec$
- not reflexive
- not symmetric
- not transitive

Hasse diagram.

3. Lattice

DEFINITION 3.1. Let $[A, \leq]$ be a poset.

$m \in A$ is $\left| \begin{array}{l} \textit{maximal} \\ \textit{minimal} \end{array} \right| \overset{\Delta}{\longleftrightarrow} \left| \begin{array}{l} \nexists a \in A \ [m < a] \\ \nexists a \in A \ [a < m] \end{array} \right|$.

DEFINITION 3.2. Let $[A, \leq]$ be a poset and $B \subseteq A$.

$s \in A$ is called a *supremum* of set $B \overset{\Delta}{\longleftrightarrow}$

i. $\forall b \in B \ b \leq s$.

ii. $\nexists a \in A \, \forall b \in B \ b \leq a \longrightarrow a < s$.

QUESTION 3.1. Define *infimum* of B.

REMARK 3.1. Consider intervals $X = [0, 1]$ and $Y = (0, 1)$ of the real numbers $\mathbb{R}$. Then 0 and 1 are minimal and maximal of X, respectively. Since $0 \notin Y$, 0 cannot be a minimal of Y. 0 is a infimum of Y. 1 is a supremum of Y. Since $\mathbb{R}$ is totally ordered, there is no other infimum than 0. So we can say that 0 is the infimum of Y. Similarly 1 is the supremum of Y.

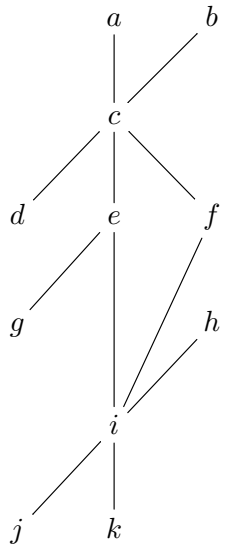
DEFINITION 3.3. Let $[A, \leq]$ be a poset and $I, O \in A$.

$\left| \begin{array}{l} I \\ O \end{array} \right|$ is the $\left| \begin{array}{l} \textit{greatest} \\ \textit{least} \end{array} \right| \overset{\Delta}{\longleftrightarrow} \left| \begin{array}{l} \forall a \in A \ [a \leq I] \\ \forall a \in A \ [O \leq a] \end{array} \right|$. I and O are called *universal upper bound* and *universal lower bound*, respectively.

REMARK 3.2.

- From now on all posets are finite.
- If poset is finite, there are minimal and maximal elements but there may not be universal upper and lower bounds.

EXAMPLE 3.1.



maximals: a, b, h.
minimals: d,g,j,k.
greatest: none.
least: none.

DEFINITION 3.4. Let $[A, \leq]$ be a poset and $a, b \in A$.
A *least upper bound* (*lub*) of a and b is $c \in A$

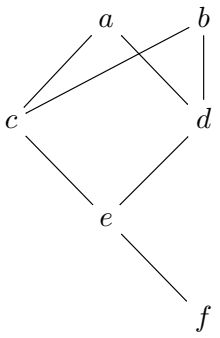
i. $a \leq c$ and $b \leq c$
ii. $\nexists x \in A \ [a \leq x \ \wedge \ b \leq x \ \wedge \ x < c]$

REMARK 3.3. Let $a, b \in A$.

i. Least upper bound of a and b may not exist.
ii. There may be more than one lub.
iii. It may be unique. If lub of a and b is unique, then it is denoted as $a + b$.

THEOREM 3.1. $\forall a, b \in A \ [$ *lub exists*$] \longrightarrow [A, \leq]$ *has the universal upper bound* I.

EXAMPLE 3.2.



- There is no universal upper bound.
- a and b are maximal elements.
- a and b are lub of c and d.
- e and f are lower bounds of c and d.
- e is unique glb of c and d.
- f is the universal lower bound O.

Definition 3.5. A *greatest lower bound* (*glb*) of a and b is $l \in A$ where
i. $l \leq a$ and $l \leq b$
ii. $\nexists x \in A \ [x \leq a \ \wedge \ x \leq b \ \wedge \ l < x]$

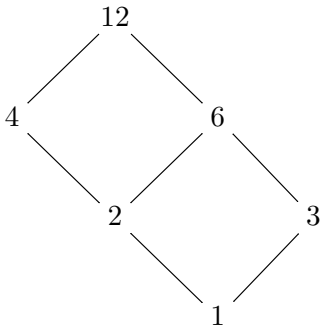
Remark 3.4. If glb of a and b is unique, then it is denoted as $a \cdot b$

Remark 3.5 (Duality Principle).
Let U be the Hasse diagram of poset $[A, \leq]$. Upside down version of U,call it D is the Hasse diagram of $[A, \geq]$.

Any property for U holds in D if the following substitutions are made:
- $+ \leftrightarrow \cdot$
- lub $\leftrightarrow$ glb
- $\leq \leftrightarrow \geq$
- $I \leftrightarrow O$

Definition 3.6 (Lattice). A *lattice* is a poset $[A, \leq]$ such that any two elements have a unique lub and glb. It is denoted as $[A, +, \cdot]$.

Example 3.3.
Divisibility relation and divisors of 12 makes a lattice.



$$A = \{ a \in \mathbb{N} \mid a \mid 12 \}$$
$$= \{ 1, 2, 3, 6, 4, 12 \}$$
$$\text{and } \forall a, b \in A \ \left[a \sim b \overset{\Delta}{\longleftrightarrow} a \mid b \right]$$

THEOREM 3.2. *Let* $[A, +, \cdot]$ *be a lattice and* $a, b, c \in A$.

i. $a + a = a$ *idempotency*
ii. $a + b = b + a$ *commutativity*
iii. $(a + b) + c = a + (b + c)$ *associativity*
iv. $a + (a \cdot b) = a$ *absorption*
v. $a + b = b \iff a \cdot b = a \iff a \leq b$ *consistency*

4. Applications

- PageRank of Google.
- Measure the similarity of two orderings (ranking) on a set, i.e. Pearson correlation.

Acknowledgment. These notes are based on various books but especially [**PY73, Ros07, Gal89**].

Problems with Solutions

P 5.1.

DEFINITION 4.1. Let $f, g : \mathbb{Z}^+ \to \mathbb{R}$. g *dominates* $f \overset{\Delta}{\iff} \exists m \in \mathbb{R}^+$ and $\exists k \in \mathbb{Z}^+$ such that $|f(n)| \leq m\ |g(n)|$ for all $n \in \mathbb{Z}^+$ where $n \geq k$

DEFINITION 4.2. For $f : \mathbb{Z}^+ \to \mathbb{R}$, f is *big Theta of* g, denoted by $f \in \Theta(g)$, $\overset{\Delta}{\iff}$ there exist constants $m_1, m_2 \in \mathbb{R}^+$ and $k \in \mathbb{Z}^+$ such that $m_1|g(n)| \leq |f(n)| \leq m_2|g(n)|$, for all $n \in \mathbb{Z}^+$, where $n \geq k$.

a) Let $\mathbb{R}^{\mathbb{Z}^+}$ be the set of all functions from $\mathbb{Z}^+$ to $\mathbb{R}$.
Define the relation β on $\mathbb{R}^{\mathbb{Z}^+}$ as

$$f \beta g \overset{\Delta}{\iff} f \in \Theta(g) \text{ for } f, g \in \mathbb{R}^{\mathbb{Z}^+}.$$

Prove that β is an equivalence relation on $\mathbb{R}^{\mathbb{Z}^+}$.

b) Let $[f]_\beta$ represent the equivalence class of $f \in \mathbb{R}^{\mathbb{Z}^+}$ for the relation β. Let E be the set of equivalence classes induced by β. Define the relation α on E by

$$[f]_\beta\ \alpha\ [g]_\beta, \text{ for } f, g \in \mathbb{R}^{\mathbb{Z}^+}, \overset{\Delta}{\iff} f \text{ is dominated by } g.$$

Show that α is a partial order.
Use shorthand notations F for $\mathbb{R}^{\mathbb{Z}^+}$ and $[f]$ for $[f]_\beta$.

Solution.

a) We need to show that β is reflexive, symmetric and transitive.
i. For each $f \in F$, $|f(n)| \leq 1\ |f(n)|$ for all $n \geq 1$. So $f \beta f$, and β is reflexive.
ii. For $f, g \in F$,

$$\begin{aligned}
f \beta g &\Rightarrow f \in \Theta(g) \\
&\Rightarrow m_f\ |g(n)| \leq |f(n)| \leq M_f\ |g(n)| \text{ for } n \geq k \text{ where } m_f, M_f \in \mathbb{R}^+ \text{ and } k \in \mathbb{Z}^+ \\
&\Rightarrow |g(n)| \leq 1/m_f\ |f(n)| \text{ and } 1/M_f\ |f(n)| \leq |g(n)| \\
&\Rightarrow m_g\ |f(n)| \leq |g(n)| \leq M_g\ |f(n)| \text{ for } n \geq k \text{ with } m_g = 1/M_f, M_g = 1/m_f \in \mathbb{R}^+ \\
&\Rightarrow g \in \Theta(f) \\
&\Rightarrow g \beta f.
\end{aligned}$$

So β is symmetric.
iii. Let $f, g, h \in F$ with $f \beta g, g \beta h$. Then, $f \in \Theta(g)$ and $g \in \Theta(h) \Rightarrow$ for all $n \in \mathbb{Z}^+$, there exist constants $m_f, M_f, m_g, M_g \in \mathbb{R}^+$ and $k_f, k_g \in \mathbb{Z}^+$ such that
$m_f\ |g(n)| \leq |f(n)| \leq M_f\ |g(n)|$ for $n \geq k_f$, and
$m_g\ |h(n)| \leq |g(n)| \leq M_g\ |h(n)|$ for $n \geq k_g$. Then for $n \geq \max\{k_f, k_g\}$,
$m_f m_g\ |h(n)| \leq m_f\ |g(n)| \leq |f(n)|$ and
$|f(n)| \leq M_f\ |g(n)| \leq M_f M_g\ |h(n)|$. Hence for $n \geq k$,
$m\ |h(n)| \leq |f(n)| \leq M\ |h(n)|$

where $m = m_f m_g, M = M_f M_g \in \mathbb{R}^+$ and $k = \max\{k_f, k_g\} \in \mathbb{Z}^+$. So $f\beta h$, that is, β transitive.

b) We need to show that α is reflexive, antisymmetric and transitive. Let $f, g, h \in \mathbb{R}^{\mathbb{Z}^+}$.

i. f is dominated by f since $|f(n)| \leq |f(n)|$ for $n \geq 1$. So $[f]\alpha[f]$, hence α is reflexive.

ii. Suppose $[f]\alpha[g]$ and $[g]\alpha[f]$. Then
$|f(n)| \leq m_f\ |g(n)|$ for $n \geq k_f$ for some m_f and k_f. Similarly,
$|g(n)| \leq m_g\ |f(n)|$ for $n \geq k_g$ for some m_g and k_g. Then for $n \geq \max\{k_f, k_g\}$
$1/m_f\ |f(n)| \leq |g(n)| \leq m_g\ |f(n)|$. That is, $g(n) \in \Theta(f(n))$. That means f and g are in the same equivalence class of β, i.e. $[f] = [g]$. So α is antisymmetric.

iii. Suppose $[f]\alpha[g]$ and $[g]\alpha[h]$. Then
$|f(n)| \leq m_f\ |g(n)|$ for $n \geq k_f$ for some m_f and k_f, and
$|g(n)| \leq m_g\ |h(n)|$ for $n \geq k_g$ for some m_g and k_g. Then for $n \geq \max\{k_f, k_g\}$
$|f(n)| \leq m_f m_g\ |h(n)|$. Therefore $[f]\alpha[h]$. Hence α is transitive.

P 5.2. Let F denote the set of all partial orderings on a set A. Define a relation $\leq$ on F such that for $\alpha, \beta \in F$, $\alpha \leq \beta \overset{\Delta}{\longleftrightarrow} \forall a, b \in A\ [a\alpha b \to a\beta b]$. Show that $\leq$ is a partial ordering on F.

Solution.

i. $\leq$ is reflexive. Since $\forall \alpha \in F\ \forall a, b \in A\ [a\alpha b \to a\alpha b] \Rightarrow \alpha \leq \alpha$.

ii. $\leq$ is antisymmetric. Suppose for $\alpha, \beta \in F$, $\alpha \leq \beta$ and $\beta \leq \alpha$. Then
$\alpha \leq \beta \Rightarrow \forall a, b \in A\ [a\alpha b \to a\beta b]$ and
$\beta \leq \alpha \Rightarrow \forall c, d \in A\ [c\beta d \to c\alpha d]$. That is, $\forall a, b \in A\ [a\alpha b \leftrightarrow a\beta b]$. Hence $\alpha = \beta$.

iii. $\leq$ is transitive. Suppose for $\alpha, \beta, \gamma \in F$, $\alpha \leq \beta$ and $\beta \leq \gamma$.
$\alpha \leq \beta \Rightarrow \forall a, b \in A\ [a\alpha b \to a\beta b]$. Similarly,
$\beta \leq \gamma \Rightarrow \forall a, b \in A\ [a\beta b \to a\gamma b]$. Hence
$\forall a, b \in A\ [a\alpha b \to a\gamma b]$. That is, $\alpha \leq \beta$.
Hence $\leq$ on F is a partial ordering.

Part 3

Algebra

CHAPTER 6

Algebraic Structures

1. Motivation

Suppose there are two research labs A and B. Lab A investigates gravitation. They do test on two masses m_1 and m_2. They discover that the attraction force F_g is given as

$$F_g = c_g \frac{m_1 m_2}{r^2}$$

where r is the distance between them and c_g is a constant.

Lab B investigates electrical charges. The force observed is attractive if the charges are opposite sign, repulsive otherwise. Yet, they measure that the force F_e between two spheres charged as q_1 and q_2 is given as

$$F_e = c_e \frac{q_1 q_2}{r^2}$$

where r is the distance between them and c_e is a constant.

Yet somewhere else, a theoretical physicist works on hypothetical forces. She assumes that the force between two bodies is proportional to some property of the body denoted by b. She also assumes that the force is inversely proportional to the square of the distance of the bodies. So she summarize her assumptions as

$$F_x = c_x \frac{b_1 b_2}{r^2}.$$

She did continue in her investigations. She figure out many properties of this hypothetical system.

Then in a conference somebody from Lab A happens to listen her presentation with amazement. This lady did all the work for them. All they have to do is to apply her findings with changing c_x with their constant c_g.

Mathematics is an abstraction. Yet, algebraic structures is just this kind of abstraction. It could be hard to find similarities between polynomials, integers and $N \times N$ square matrices. But actually they have very similar properties which will be call *ring* in this chapter.

EXAMPLE 1.1. Part a.
Let's solve $a + x = b$ for x where $a, b, x \in \mathbb{Z}$.

$$\begin{aligned} a + x &= b \\ (-a) + (a + x) &= (-a) + b \\ ((-a) + a) + x &= (-a) + b \\ 0 + x &= (-a) + b \\ x &= (-a) + b. \end{aligned}$$

Part b.
Let's solve $A + X = B$ for X where A, B, X are $N \times N$ real matrices.

$$\begin{aligned} A + X &= B \\ (-A) + (A + X) &= (-A) + B \\ ((-A) + A) + X &= (-A) + B \\ 0 + X &= (-A) + B \\ X &= (-A) + B. \end{aligned}$$

QUESTION 1.1. Compare Part a and Part b of Example 1.1. What are the differences and similarities?

QUESTION 1.2. Solve $A + X = B$ for X if

i. A, B, X are $N \times N$ rational matrices.
ii. A, B, X are $N \times N$ integer matrices.
iii. A, B, X are $N \times N$ natural number matrices.
iv. A, B, X are polynomials with complex coefficients in y.
v. A, B, X are polynomials with real coefficients in y.
vi. A, B, X are polynomials with rational coefficients in y.
vii. A, B, X are polynomials with integer coefficients in y.
viii. $A, B, X \in \mathbb{C}$.

ix. $A, B, X \in \mathbb{R}$.
x. $A, B, X \in \mathbb{Q}$.
xi. $A, B, X \in \mathbb{Z}$.
xii. $A, B, X \in \mathbb{N}$.
xiii. $A, B, X \in \mathbb{R} \setminus \mathbb{Q}$.
xiv. A, B, X are 2D vectors.
xv. A, B, X are 3D vectors.

QUESTION 1.3.

i. Some of the systems given in Question 1.2 have no solution. Can you find a pattern when there is a solution. What properties of what do you need in order to solve equation $A + X = B$?
ii. Reconsider Question 1.2 when addition is replaced by multiplication, that is, $A \times X = B$. Note that multiplication may not be defined in some concepts.

2. Algebraic Structures

Consider the equation $A + X = B$. In order to interpret the equation correctly we need to know couple of things: What is "+" represents? What are A, B and X? If A, B and X are of the same "type", what set do they member of? The only concept that does not need further explanation is the equality "=".

QUESTION 2.1. It should be an equivalence relation but do we really know what actually "=" means? We know that there are more than one equivalence relations can be defined on a set. So which one is this? Recall that *equivalence relations* 5.1 is covered in Chapter 4.

2.1. Binary Operations.

REMARK 2.1. At this point, you may want to refresh the definition of function 3.2.

DEFINITION 2.1. A *binary operation* $\star$ on set A is a function $\star : A \times A \to A$. A binary operation is represented by $a \star b$ instead of the traditional functional notation $\star((a, b))$ where $a, b \in A$.

QUESTION 2.2. What is the difference between $\star((a, b))$ and $\star(a, b)$?

DEFINITION 2.2. Let $A = \{a_1, \cdots, a_n\}$. An *operation table* represents the binary operation $\star$ in a table form where $a_i \star a_j = a_k$.

$\star$	a_1	$\cdots$	a_i	$\cdots$	a_j	$\cdots$	a_n
a_1	.	$\cdots$	.	$\cdots$	.	$\cdots$	.
$\vdots$	$\vdots$		$\vdots$		$\vdots$		$\vdots$
a_i	.	$\cdots$	.	$\cdots$	a_k	$\cdots$	.
$\vdots$	$\vdots$		$\vdots$		$\vdots$		$\vdots$
a_j	.	$\cdots$	a_m	$\cdots$	.	$\cdots$	.
$\vdots$	$\vdots$		$\vdots$		$\vdots$		$\vdots$
a_n	.	$\cdots$	.	$\cdots$	.	$\cdots$	.

REMARK 2.2. This representation is valid if the elements of A can be made into a list. Some sets have, some cannot have such a list. Making a list of elements is an importing concept which we will be looking at in Chapter 10 when we discuss finiteness and type of infinities in more detail. The elements of a finite set can always be made a list. If the set is not finite, there are two different cases. If the set is *countable infinite* such as $\mathbb{N}$, there is a natural list that can be used for operational table. Note that in this case the operational table would be an infinite table. If the set is *uncountable infinite* such as $\mathbb{R}$, then the elements cannot be put in a list. Concepts such as finiteness, infinity, countable infinity, uncountable infinity will be covered in Chapter 10.

REMARK 2.3. The order of operation is important. $a_i \star a_j = a_k \neq a_m = a_j \star a_i$.

DEFINITION 2.3. A binary operation $\star$ on A is called *associative* $\overset{\Delta}{\longleftrightarrow}$ $\forall a, b, c \in A \ [(a \star b) \star c = a \star (b \star c)]$

DEFINITION 2.4. A binary operation $\star$ on A is called *commutative* $\overset{\Delta}{\longleftrightarrow}$ $\forall a, b \in A \ [a \star b = b \star a]$.

2.2. Algebraic Structure.

DEFINITION 2.5. A nonempty set A and binary operations $f_1, f_2, \cdots, f_n$ defined on A together is called an *algebraic structure*, denoted by $[\, A, f_1, f_2, \cdots, f_n \,]$, where $n \in \mathbb{Z}^+$.

EXAMPLE 2.1. Let V be a vector space. Addition of two vectors is represented as $v_1 + v_2$ for $v_1, v_2 \in V$. Then $[V, +]$ is an algebraic structure. Note that $+$ is both associative and commutative.

QUESTION 2.3. In general, multiplication of two vectors is not defined. Only in 3D, *cross multiplication* of two vectors is defined, denoted as $v_1 \times v_2$. Let V be 3D vector space and $v_1, v_2, v_3 \in V$.

i. Is $\times$ associative?
ii. Is $\times$ commutative?
iii. Is it the case that $v_1 \times (v_2 + v_3) = (v_1 \times v_2) + (v_1 \times v_3)$?

QUESTION 2.4. In vector spaces, scalar multiplication is defined. How do you put that into algebraic structure notation?

2.3. Sub-Algebraic Structures. Two similar but not exactly the same systems can be investigated. A special case of similar structures is the case when one set is a subset of another.

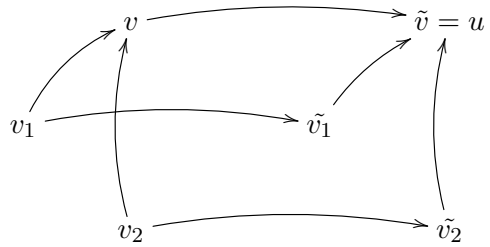
Suppose $B \subseteq A$ and a binary operation $\star$ on A is defined. Since a binary operation is a function from $A \times A$, and since $B \times B \subseteq A \times A$, we can try to restrict it to B.

One needs to be very careful at this point. The danger can be better seen in simpler case: What we have is a function $f : A \to A$. We have $B \subseteq A$ and we want to restrict function to B. It is perfectly possible that for some elements of B the image under the function may not be in B at all, that is, $f(b) \in A \setminus B$ for some $b \in B$. Whenever that happens, the function can not be a binary operation. This concern is called *closedness*.

EXAMPLE 2.2 (Subspace). Think about vectors in $X - Y$ plane. This makes a 2D vector space. Vectors in $X - Y - Z$ is a vector space in 3D. Let's denote 2D and 3D vector spaces by $\mathbb{R}^2$ and $\mathbb{R}^3$, respectively. Define addition of two vectors in the usual way in both $\mathbb{R}^2$ and $\mathbb{R}^3$. Then we obtained two independent algebraic structures $\left[\mathbb{R}^2, +\right]$ and $\left[\mathbb{R}^3, +\right]$.

Are they really independent? Actually $\left[\mathbb{R}^2, +\right]$ is a special case of $\left[\mathbb{R}^3, +\right]$. Any vector $v = [x\, y]^\top \in \mathbb{R}^2$ can be mapped to a unique vector, denoted by $\tilde{v} = [x\, y\, 0]^\top \in \mathbb{R}^3$. We say that $\mathbb{R}^2$ is a *subspace* of $\mathbb{R}^3$.

Note that for all $v_1, v_2 \in \mathbb{R}^2$ we have $v = v_1 + v_2 \in \mathbb{R}^2$. If we map v_1, v_2 into $\mathbb{R}^3$, we obtain $\tilde{v_1}, \tilde{v_2} \in \mathbb{R}^3$. This time use addition in $\mathbb{R}^3$ to obtain $u = \tilde{v_1} + \tilde{v_2} \in \mathbb{R}^3$. Is it the case that $\tilde{v} = u$? This can be visualized as: $\tilde{a}$



EXAMPLE 2.3 (Addition on Reals and Rationals). Consider the set of real numbers, $\mathbb{R}$. With ordinary addition, $+ : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$, it makes the algebraic system, $[\mathbb{R}, +]$. Now, consider the set of rational numbers. Depending how you look at it, a rational number is a real number or not. Here we assume that a rational number is also a real number. Hence $\mathbb{Q} \subseteq \mathbb{R}$. The addition $+$ in reals can be restricted to $\mathbb{Q}$. Let $+_Q : \mathbb{Q} \times \mathbb{Q} \to \mathbb{Q}$ be the restriction of addition in reals to rationals.

Fortunately addition of any two rational numbers is again a rational number. So we can safely restrict $+$ in $\mathbb{R}$ to $\mathbb{Q}$ and obtain binary operation $+_Q$ in $\mathbb{Q}$.

EXAMPLE 2.4 (Multiplication on Reals and Negative Integers). Multiplication on real numbers is a binary operation.

$$\times_{\mathbb{R}} : \mathbb{R} \times \mathbb{R} \to \mathbb{R}.$$

On the other hand restriction of it to negative integers

$$\times : \mathbb{Z}^- \times \mathbb{Z}^- \to \mathbb{Z}^+$$

is not a binary operation since the codomain not $\mathbb{Z}^-$ any more.

EXAMPLE 2.5 (Multiplication in Irrational Numbers). An irrational number is a real number that is not rational. Using this definition the set of irrational numbers can be represented as $A = \mathbb{R} \setminus \mathbb{Q}$. Note that we use A since there is no agreed symbol for the set of irrational numbers as we have $\mathbb{R}$ for reals.

Clearly $A \subseteq \mathbb{R}$. Then try to restrict multiplication $\times$ in reals to irrationals. The restricted function $\times_A$ would be

$$\times_A : A \times A \to \mathbb{R}.$$

Note that it is not the case that

$$\times_A : A \times A \to A$$

since $\sqrt{2} \in A$ but $\sqrt{2} \times_A \sqrt{2} = 2 \notin A$. That is, the restriction of multiplication in the set of reals to the set of irrationals is not a binary operation in irrationals. In other words, the restriction is not closed.

3. Algebraic Structures with One Binary Operation

3.1. Semigroup.

DEFINITION 3.1. An algebraic structure $G = [A, \star]$ is called *semigroup* $\overset{\Delta}{\longleftrightarrow}$

i. $\star$ is associative.

3.2. Monoid.

DEFINITION 3.2. Let $[A, \star]$ be an algebraic structure.

$$\left|\begin{array}{l}\ell\\ r\\ e\end{array}\right| \left|\begin{array}{l}\in A \text{ is called}\end{array}\right| \left|\begin{array}{l}\textit{left-identity}\\ \textit{right-identity}\\ \textit{(two-sided) identity}\end{array}\right| \overset{\Delta}{\longleftrightarrow} \left|\begin{array}{l}\forall a \in A \ [\ell \star a = a]\\ \forall a \in A \ [a \star r = a]\\ \forall a \in A \ [e \star a = a \star e = a]\end{array}\right|.$$

REMARK 3.1.

i. There may exist none, one or both of ℓ or r.
ii. There is no need to be semigroup in order to have left, right or two-sided identities.

THEOREM 3.1. *If ℓ and r are left and right identities of a semigroup G, then $\ell = r$.*

THEOREM 3.2. *If two-sided identity exists, then it is unique.*

QUESTION 3.1. The theorem says that there could not be two different identities. Is it possible that there are two different left-identities ℓ_1 and ℓ_2? The same question for right-identities?

DEFINITION 3.3. An algebraic structure $M = [A, \star]$ is called *monoid* $\overset{\Delta}{\longleftrightarrow}$

i. M is a semigroup.
ii. M has the identity, denoted by e.

QUESTION 3.2. Consider a row of the operation table of a monoid. What can be said about the number of identities?

DEFINITION 3.4 (Subsemigroup, Submonoid).

Let $\mathcal{A} = [A, \star]$ and $\mathcal{B} = [B, \circ]$ be $\left|\begin{array}{l}\text{semigroups}\\ \text{monoids}\end{array}\right|$.

$\mathcal{B}$ is said to be $\left|\begin{array}{l}\textit{subsemigroup}\\ \textit{submonoid}\end{array}\right|$ of $\mathcal{A}$ $\overset{\Delta}{\longleftrightarrow}$

i. $B \subseteq A$
ii. $\circ$ is the restriction of $\star$ to B.

REMARK 3.2. This is a typical definition of sub-structures. An equivalent but more compact definition would be as follows:

A $\left|\begin{array}{l}\text{semigroup}\\ \text{monoid}\end{array}\right|$ $\mathcal{B} = [B, \circ]$ is said to be a $\left|\begin{array}{l}\textit{subsemigroup}\\ \textit{submonoid}\end{array}\right|$ of another $\left|\begin{array}{l}\text{semigroup}\\ \text{monoid}\end{array}\right|$ $\mathcal{A} = [A, \star]$ $\overset{\Delta}{\longleftrightarrow}$

i. $B \subseteq A$
ii. $\circ$ is the restriction of $\star$ to B.

EXAMPLE 3.1. Let $A = \{1, 2, 3, 4\}$ and $B = \{3, 4\}$. Define binary operators as follows.

$*$	1	2	3	4
1	.	1	.	.
2	1	2	3	4
3	.	3	3	4
4	.	4	4	4

$\circ$	3	4
3	3	4
4	4	4

Then 2 and 3 are identities of $*$ and $\circ$ in sets A and B, respectively. Note that $\circ$ is the restriction of $*$ to B. Note also that 3 is not an identity in A.

QUESTION 3.3. Let $B = [B, \circ]$ be a submonoid of $A = [A, *]$ with identities e_B and e_A, respectively. Is it possible that $e_A \neq e_B$?

TABLE 1. default

$*$	1	2	3	4
1	.	1	.	.
2	1	2	3	4
3	.	3	3	4
4	.	3	4	4

3.3. Groups.

DEFINITION 3.5 (Inverse).
Let $G = [A, \star]$ be a monoid with the identity e. Let $a \in A$.

$$\left| \begin{matrix} \ell_a \\ r_a \\ b_a \end{matrix} \right| \in A \text{ is called } \left| \begin{matrix} \textit{left inverse} \\ \textit{right inverse} \\ \textit{(two-sided) inverse} \end{matrix} \right| \text{ of } a \overset{\Delta}{\longleftrightarrow} \left| \begin{matrix} \ell_a \star a = e \\ a \star r_a = e \\ b_a \star a = a \star b_a = e \end{matrix} \right| .$$

THEOREM 3.3. *If ℓ and r are left and right inverses of a, respectively, then $\ell = r$ in a monoid.*

QUESTION 3.4. Is it possible that a has two different left-inverses, ℓ_1 and ℓ_2?

NOTATION 3.1. The inverse of a is represented by a^{-1}.

DEFINITION 3.6 (Group).
An algebraic structure $G = [A, \star]$ is called *group* $\overset{\Delta}{\longleftrightarrow}$
i. G is a monoid.
ii. $\forall a \in A$, there is a unique inverse of a, denoted by a^{-1}.
If A is finite, G is said to be a *finite group* and $|A|$ is called the *order* of G.

THEOREM 3.4. *Let $G = [A, \star]$ be a group and $a, b \in A$. $a \star x = b$ and $y \star a = b$ have unique solutions, namely, $x = a^{-1} \star b$ and $y = b \star a^{-1}$.*

THEOREM 3.5 (Cancellation). *In a group,*
i. $a \star b = a \star c \Rightarrow b = c$.
ii. $b \star a = c \star a \Rightarrow b = c$.

REMARK 3.3. We know these two theorems in numbers since primary school. What the theorems say is that they are valid if the system satisfy the group axioms.

NOTATION 3.2. $a \star b$ is represented by ab when the binary operation $\star$ is clear in the context.

THEOREM 3.6 (Cayley). *Every finite group can be represented by a group of permutations.*

DEFINITION 3.7 (Permutation). Let A be a finite set. A bijection $\gamma : A \to A$ is called a *permutation*.

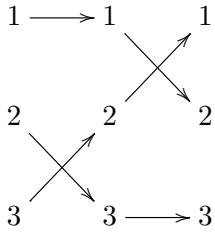
EXAMPLE 3.2 (Group of Permutations).

Let $A = \{ x_1, x_2, x_3 \}$ be a set with 3 elements.
A permutation on A can be represented as: $\begin{pmatrix} x_1 x_2 x_3 \\ x_2 x_1 x_3 \end{pmatrix} = (213)$.
There are $3! = 6$ permutations:

$a = (123)$
$b = (132)$
$c = (213)$
$d = (231)$
$e = (312)$
$f = (321)$

Let $S_3 \triangleq \{ a, b, c, d, e, f \}$ be the set of permutations on A.
Define a binary operation $\circledcirc$ on S_3 as the composition, that is, $\alpha \circledcirc \beta \triangleq \beta \circ \alpha$ where $\alpha, \beta \in S_3$. Hence $x \in A$ is mapped to $\beta(\alpha(x))$.

$$\begin{aligned} b \circledcirc c &= (132) \circledcirc (213) \\ &= (213) \circ (132) \\ &= (213)((132)) \\ &= (231) = d \end{aligned}$$



$\circledcirc$	a	b	c	d	e	f
a	a	b	c	d	e	f
b	b	?	d	?	?	?
c	c	?	?	?	?	?
d	d	?	?	?	?	?
e	e	?	?	?	?	?
f	f	?	?	?	?	?

Definition 3.8. $S_3 = [S_3, \circledcirc]$ is called the *symmetric group of order 3, S_3*. Symmetric groups S_n can be extended for any $n \in \mathbb{N}$.

Remark 3.4. $| S_n | = n!$

Definition 3.9. Let $G = [A, \star]$ and $H = [B, \circ]$ be two groups. Group G is said to be a *subgroup* of group $H \overset{\Delta}{\longleftrightarrow}$
i. $A \subseteq B$.
ii. $\star$ is the restriction of $\circ$ to A.

Remark 3.5.

i. $[G, \star]$ is a subgroup of itself.
ii. $[\{ e \}, \star]$ is a subgroup of G.

Definition 3.10. Subgroups G and $\{ e \}$ are called the *trivial subgroups*. Any subgroup that is not trivial is called *proper subgroup*.

Theorem 3.7. *$T \neq \emptyset$ is a subgroup of G* $\longleftrightarrow \forall a, b \in T \ \left[ab^{-1} \in T \right]$.

Theorem 3.8. *Let H be a subgroup of G. Then the order of H divides the order of G.*

Definition 3.11. A group with commutative binary operation is called *abelian group*.

4. Algebraic Structures with two Binary Operations

4.1. Ring.

Definition 4.1 (Ring). An algebraic structure $R = [A, \star, \circ]$ is called *ring* $\overset{\Delta}{\longleftrightarrow}$
i. $[A, \star]$ is an abelian group.
ii. $[A, \circ]$ is a semigroup.
iii. $\forall a, b, c \in A$
$a \circ (b \star c) = (a \circ b) \star (a \circ c)$
$(b \star c) \circ a = (b \circ a) \star (c \circ a)$

NOTATION 4.1. Usual notation for a ring is $R = [A, +, \cdot]$
Use ab for $a \cdot b$.
0 is the *additive identity*.
$-a$ is the *additive inverse* of a.
1 is the *multiplicative identity*.
a^{-1} is the *multiplicative inverse* of a.

REMARK 4.1. Note that since $[A, +]$ is a group, the additive identity 0 and additive inverse $-a$ for all a should be there. On the other hand, $[A, \cdot]$ is simple a semigroup. Therefore the multiplicative identity may not exist. Even if the multiplicative identity exists, multiplicative inverse may not.

DEFINITION 4.2 (Commutative Ring). A ring with commutative multiplication is called *commutative ring*.

THEOREM 4.1. *Let $R = [A, \star, \circ]$ be a ring. $\forall a, b \in A$*
i. $0 \circ a = a \circ 0 = 0$
ii. $(-a) \circ b = a \circ (-b) = -(a \circ b)$
iii. $(-a) \circ (-b) = a \circ b$.

PROOF.
Part (i)

$$\begin{aligned}(a \circ 0) \star 0 &= a \circ 0 && \text{//definition of additive identity}\\ &= a \circ (0 \star 0) && \text{//definition of additive identity}\\ &= (a \circ 0) \star (a \circ 0) && \text{//multiplication distributes over addition.}\end{aligned}$$

Using cancellation by $a \circ 0$ in group $[A, \star]$, we have $0 = a \circ 0$. The remaining part $0 = 0 \circ a$ can be proven similarly.

Part (ii)

$$\begin{aligned}(a \circ b) \star (-(a \circ b)) &= 0 && \text{//additive inverse}\\ &= 0 \circ b && \text{//from part (i)}\\ &= (a \star (-a)) \circ b && \text{//definition of additive identity}\\ &= (a \circ b) \star ((-a) \circ b) && \text{//distributivity of } \circ \text{ over } \star.\end{aligned}$$

Using cancellation by $a \circ 0$ in group $[A, \star]$, we have $-(a \circ b) = (-a) \circ b$. The remaining part $-(a \circ b) = a \circ (-b)$ can be proven similarly.

Part (iii)

$$\begin{aligned}(-a) \circ (-b) &= -(a \circ (-b)) && \text{//from part (ii)}\\ &= -(-(a \circ b)) && \text{//from part (ii)}\\ &= a \circ b && \text{//definition of additive inverse.}\end{aligned}$$

□

REMARK 4.2. We knew these properties of numbers since primary school. Now we learn that they are valid not only in numbers but any algebraic structure that satisfy conditions of ring such as the ring of polynomials or the ring of matrices as seen shortly.

EXAMPLE 4.1.
i. $[\mathbb{N}, +, \cdot]$ is not a ring, since there is no additive inverse.
ii. $[\mathbb{Z}, +, \cdot]$ is a ring. So do $[\mathbb{Q}, +, \cdot]$, $[\mathbb{R}, +, \cdot]$ and $[\mathbb{C}, +, \cdot]$.
iii. Consider the set of polynomials with real coefficients in x, denoted by $\mathbb{R}[x]$. With regular addition and multiplication of polynomials, $[\mathbb{R}[x], +, \cdot]$ is a ring, called the *ring of polynomials*.

QUESTION 4.1.
i. What is the additive identity of group $[\mathbb{R}[x], +]$?
ii. What is the additive inverse of $5x^2 + 3x + 7$ in $[\mathbb{R}[x], +]$?
iii. Is there a multiplicative identity in semigroup $[\mathbb{R}[x], \cdot]$?
iv. What is the multiplicative inverse of $5x^2 + 3x + 7$ in $[\mathbb{R}[x], +, \cdot]$?

QUESTION 4.2. Consider the set of $N \times M$ matrices with real entries, denoted by $\mathbb{R}_{N\times M}$. With regular addition and multiplication of matrices, $[\mathbb{R}_{N\times M}, +, \cdot]$ is not a ring. Why? Can you make it a ring by additional constraints?

DEFINITION 4.3 (Subring). A ring $\mathcal{A} = [A, \oplus, \otimes]$ is said to be a *subring* of another ring $\mathcal{B} = [B, +, \times]$ $\overset{\Delta}{\longleftrightarrow}$
i. $A \subseteq B$.
ii. $\oplus$ is the restriction of $+$ to A.
iii. $\otimes$ is the restriction of $\times$ to A.

THEOREM 4.2. *Let T and R be rings. T is a subring of R* $\longleftrightarrow \forall a, b \in T\ [(a-b), ab \in T]$.

DEFINITION 4.4. Let T be a subring of ring R. If $\forall r \in T\ [\forall a \in R\ [ar, ra \in T]]$, then T is called an *ideal* of R.

4.2. Field.

DEFINITION 4.5. A *field*, $F = [A, +, \cdot]$, is a ring such that $[A \setminus \{0\}, \cdot]$ is an abelian group. If A is finite, F is called *finite field* (*Galois field*).

REMARK 4.3. Note that in 0 in $A \setminus \{0\}$ is the additive identity of group $[A, +]$.

DEFINITION 4.6. Let $\mathbb{Z}_n \triangleq \{0, \cdots, n-1\}$ where $n \in \mathbb{N}, n \geq 2$,
$a \oplus b \triangleq$ remainder of $\frac{a+b}{n}$,
$a \odot b \triangleq$ remainder of $\frac{ab}{n}$.

THEOREM 4.3. $[\mathbb{Z}_p, \oplus, \odot]$ *is a field if p is a prime number.*

REMARK 4.4. Take $n = 2$. Since 2 is prime, $[\mathbb{Z}_2, \oplus, \odot]$ is a field.
Actually, this is the field that Computer Engineering/Science is based on: $\mathbb{Z}_2 = \{0, 1\}$ where 0 and 1 are integers. Another interpretation of 0 and 1 would be "false" and "true", respectively. Then, one can interpret the binary operations $\oplus$ and $\odot$ as logical functions $f : \mathbb{B}\times\mathbb{B} \to \mathbb{B}$ and $g : \mathbb{B}\times\mathbb{B} \to \mathbb{B}$ where $\mathbb{B} = \{0, 1\}$ but this time in the logical meaning.

QUESTION 4.3. What logical functions do $\oplus$ and $\odot$ correspond? Can you express them in terms of $\wedge$, $\vee$ and $\neg$?

EXAMPLE 4.2.
i. $[\mathbb{Z}, +, \cdot]$ is not a field, since there is no multiplicative inverse.
ii. $[\mathbb{Q}, +, \cdot]$ is a field. So do $[\mathbb{R}, +, \cdot]$ and $[\mathbb{C}, +, \cdot]$.

QUESTION 4.4. $[\mathbb{R}[x], +, \cdot]$ is not a field, since no multiplicative inverse of $x + 1 \in \mathbb{R}[x] \setminus \{0\}$ exists. What is 0 in this context? Can you extend it into a field?

4.3. Lattice.

Lattice has two definitions: poset-wise and algebraic.

DEFINITION 4.7. A *(algebraic) lattice*, $L = [A, \sqcup, \sqcap]$ is a nonempty set A with binary operations $\sqcup, \sqcap$, called *join* and *meet*, iff

i)	$x \sqcap y = y \sqcap x$	$x \sqcup y = y \sqcup x$
ii)	$x \sqcap (y \sqcap z) = (x \sqcap y) \sqcap z$	$x \sqcup (y \sqcup z) = (x \sqcup y) \sqcup z$
iii)	$x \sqcap (x \sqcup y) = x$	$x \sqcup (x \sqcap y) = x$

REMARK 4.5. Let $[A, \sqcup, \sqcap]$ and $[A, +, \cdot]$ be lattices algebraic and poset sense, respectively. Then $a \sqcup b = a + b$ and $a \sqcap b = a \cdot b$.

4.4. Vector Spaces.

DEFINITION 4.8. Let $V = [V, +]$ be an additive abelian group. Let F be a field. Let $\cdot : F \times V \to V$ be a function. The group V is then called a *vector space over the field F*
$\overset{\Delta}{\longleftrightarrow}$ For $a, b \in F$, $\overrightarrow{v}, \overrightarrow{u} \in V$, the following conditions are satisfied:
i. $a \cdot (\overrightarrow{v} + \overrightarrow{u}) = a \cdot \overrightarrow{v} + a \cdot \overrightarrow{u}$
ii. $(a + b) \cdot \overrightarrow{v} = a \cdot \overrightarrow{v} + b \cdot \overrightarrow{v}$
iii. $a \cdot (b \cdot \overrightarrow{v}) = (ab) \cdot \overrightarrow{v}$

iv. $1 \cdot \overrightarrow{v} = \overrightarrow{v}$

where 1 is the multiplicative identity of F. Elements of V and F are called *vectors* and *scalars*, respectively. The function $\cdot$ is called *scalar multiplication*.

REMARK 4.6. Note that vector spaces are not algebraic structures.

EXAMPLE 4.3. In Physics we use vectors spaces over real numbers. For example 3D vectors are represented as $\mathbb{R}^3$. In our formal notation it should be $V = \mathbb{R}^3$ and $F = \mathbb{R}$. Then we have

$$\overrightarrow{v} \triangleq \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \in \mathbb{R}^3 \quad \text{and} \quad r \cdot \overrightarrow{v} = r \cdot \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \triangleq \begin{bmatrix} r \cdot x_1 \\ r \cdot x_2 \\ r \cdot x_3 \end{bmatrix}.$$

EXAMPLE 4.4. Let $\mathbb{Q}_{2\times 2}$ be the set of 2×2 matrices over rational numbers. Define the product of a rational number by a 2×2 matrix as a 2×2 matrix obtained by multiplying each of the entries by the rational number, i.e. for $a \in \mathbb{Q}$ and $M \in \mathbb{Q}_{2\times 2}$, $[a \cdot M]_{ij} = a[M_{ij}]$ where $[M]_{ij}$ is the i,j-th entry of the matrix. Then $\mathbb{Q}_{2\times 2}$ is a vector space over $\mathbb{Q}$.

EXAMPLE 4.5. Let $F[x]$ be the set of all polynomials in x with coefficients in F. Multiplication of a polynomial by a scalar is defined by multiplying each coefficient with that scalar. Then $F[x]$ is a vector space over F.

QUESTION 4.5. Prove that $0 \cdot \overrightarrow{v} = \overrightarrow{0}$.

QUESTION 4.6. What is wrong in the following?

$$\begin{aligned} 0 &= 0 \\ 0 \cdot 1 &= 0 \cdot 2 \\ 1 &= 2. \end{aligned}$$

5. Summary

Single Binary Operation, $[A, \oplus]$

- Binary operation
- Semigroup
  i. associativity
- Monoid
  i. semigroup
  ii. identity
- Group
  i. monoid
  ii. inverse
- Abelian Group
  i. group
  ii. commutativity

Two Binary Operations, $[A, \oplus, \otimes]$

- Ring $R = [A, \oplus, \otimes]$
  i. $[A, \oplus]$ is an abelian group
  ii. $[A, \otimes]$ is a semigroup
  iii. $\otimes$ right and left distributes over $\oplus$
- Field $F = [A, \oplus, \otimes]$
  i. $[A, \oplus, \otimes]$ is a ring
  ii. $[A \setminus \{0\}, \otimes]$ is an abelian group
- Lattice $L = [A, \oplus, \otimes]$

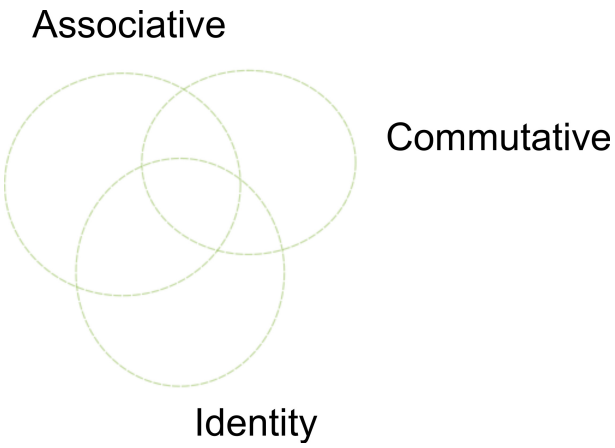


FIGURE 1. Note that an operation can be of any combinations of associativity, commutativity and identity.

Acknowledgment. These notes are based on various books but especially [**PY73, LP98, Ros07, Gal89**].

Problems with Solutions

P 6.1. Prove the following theorem.

THEOREM 5.1. *Let G be a group and $a_1, a_2, \ldots, a_n \in G$. Then $(a_1 a_2 \cdots a_n)^{-1} = a_n^{-1} a_{n-1}^{-1} \cdots a_1^{-1}$.*

Solution.

Use induction on n.
Induction Base. For $n = 1$, $(a)^{-1} = a^{-1}$.
Induction Hypotesis. Assume that $(a_1 a_2 \cdots a_n)^{-1} = a_n^{-1} a_{n-1}^{-1} \cdots a_1^{-1}$ for n.

$$\begin{aligned}
[a_1 a_2 \cdots a_n a_{n+1}][a_{n+1}^{-1} a_n^{-1} a_{n-1}^{-1} \cdots a_1^{-1}] &= [a_1 a_2 \cdots a_n][a_{n+1} a_{n+1}^{-1}][a_n^{-1} a_{n-1}^{-1} \cdots a_1^{-1}] \\
&= [a_1 a_2 \cdots a_n][e][a_n^{-1} a_{n-1}^{-1} \cdots a_1^{-1}] \text{ since } a_{n+1} a_{n+1}^{-1} = e \\
&= [a_1 a_2 \cdots a_n][a_n^{-1} a_{n-1}^{-1} \cdots a_1^{-1}] \\
&= e \text{ by the induction hypotesis.}
\end{aligned}$$

So $a_{n+1}^{-1}a_n^{-1}a_{n-1}^{-1}\cdots a_1^{-1}$ is a right inverse of $a_1a_2\cdots a_na_{n+1}$.

$$\begin{aligned}[a_{n+1}^{-1}a_n^{-1}a_{n-1}^{-1}\cdots a_1^{-1}][a_1a_2\cdots a_na_{n+1}] &= [a_{n+1}^{-1}a_n^{-1}a_{n-1}^{-1}\cdots a_2^{-1}][a_1^{-1}a_1][a_2\cdots a_na_{n+1}]\\ &= [a_{n+1}^{-1}a_n^{-1}a_{n-1}^{-1}\cdots a_2^{-1}][e][a_2\cdots a_na_{n+1}] \text{ since } a_1^{-1}a_1 = e\\ &= \ldots\\ &= a_{n+1}^{-1}a_{n+1} = e.\end{aligned}$$

So $a_{n+1}^{-1}a_n^{-1}a_{n-1}^{-1}\cdots a_1^{-1}$ is a left inverse of $a_1a_2\cdots a_na_{n+1}$.

Hence $a_{n+1}^{-1}a_n^{-1}\cdots a_1^{-1}$ is the inverse of $a_1a_2\cdots a_{n+1}$.

CHAPTER 7

Boolean Algebras

1. Reminders

i. Partial ordering = reflexive + antisysmmetic + transitive
ii. Poset $[S, \leq]$
iii. Immediate predecessor $\prec$
iv. Immediate successor $\succ$
v. Hasse diagram
vi. Maximal and minimal elements
vii. Universal upper bound (greatest element), 1
viii. Universal lower bound (least element), 0
ix. Least upper bound (lub, join), $a \oplus b$
x. Greatest lower bound (glb, meet), $a \odot b$
xi. Lattice, $L = [L, \oplus, \odot]$

2. Lattices

DEFINITION 2.1. $a \in L$ is called an *atom* $\overset{\Delta}{\longleftrightarrow}$ $0 \prec a$.

DEFINITION 2.2. Let $[L, \oplus, \odot]$ be a lattice.
$a \in L$ is said to be *join-irreducible* $\overset{\Delta}{\longleftrightarrow}$ $\forall x, y \in L$ $[x \oplus y = a \longrightarrow x = a \vee y = a]$.

REMARK 2.1.
i. The universal lower bound 0 is join-irreducible.
ii. All the atoms are join-irreducible.

DEFINITION 2.3. A lattice is called *finite length* $\overset{\Delta}{\longleftrightarrow}$ All chains in L are finite.

THEOREM 2.1. *If L is a finite length lattice then every element $a \in L$ can be represented as a joint of a finite number of join-irreducible elements of L.*

DEFINITION 2.4. Expression $x \oplus y = a$ is called an *irredundant join* of a $\overset{\Delta}{\longleftrightarrow}$ Any subset of $\{x, y\}$ no longer represents a.

THEOREM 2.2. *If L is a distributive, finite length lattice, then $\forall a \in L$ there is a unique representation as the join of irredundant set of join-irreducible elements.*

EXAMPLE 2.1. $[\mathbb{Z}^+, |]$ is a lattice where $|$ is divisibility. In this lattice prime numbers are atoms. The powers of primes are the join-irreducible elements.

3. Boolean Algebras

DEFINITION 3.1. A *lattice* is a poset $[L, \leq]$, any two elements of which have unique join and meet, denoted by $[L, \oplus, \odot]$.

3.1. Distributive Lattice.

DEFINITION 3.2. A lattice $[L, \oplus, \odot]$ is said to be *distributive* $\overset{\Delta}{\longleftrightarrow}$ $\forall a, b, c \in L$
i. $a \odot (b \oplus c) = (a \odot b) \oplus (a \odot c)$
ii. $a \oplus (b \odot c) = (a \oplus b) \odot (a \oplus c)$.

THEOREM 3.1. *Let A be a set.*
i. $[2^A, \cup, \cap]$ *is a lattice.*
ii. $\emptyset \in 2^A$ *is the universal lower bound.*
iii. $A \in 2^A$ *is the universal upper bound.*

iv. $\left[2^A, \cup, \cap\right]$ *is a distributive lattice.*

EXAMPLE 3.1. Let A be a finite set. Then $A = \{a_1, a_2, \ldots, a_n\}$ be a listing of elements of A. Let $A_j \in 2^A$. Define $b^{A_j} = (b_1^{A_j}, b_2^{A_j}, \ldots, b_n^{A_j}) \in \mathbb{B}^n$ such that

$$b_i^{A_j} = \begin{cases} 1, & a_i \in A_j, \\ 0, & a_i \notin A_j. \end{cases}$$

Notice that

$$\begin{aligned} f \ : \ 2^A &\to \mathbb{B}^n \\ A_j &\mapsto (b_1^{A_j}, b_2^{A_j}, \ldots, b_n^{A_j}) \end{aligned}$$

is a bijection.

EXAMPLE 3.2. Let $A = \{a, b, c\}$

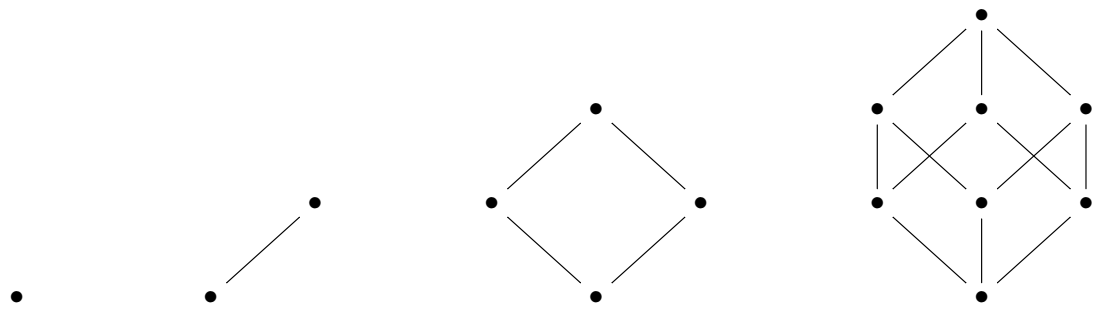
$$\begin{aligned} b^{\emptyset} &\to (0,0,0) \\ b^{\{a\}} &\to (1,0,0) \\ b^{\{a,c\}} &\to (1,0,1) \\ b^{\{a,b,c\}} &\to (1,1,1) \end{aligned}$$

QUESTION 3.1. Let $A_1, A_2 \in 2^A$. Then $A_1 \cup A_2, A_1 \cap A_2, \overline{A_1} \in 2^A$. What can you say about $b_j^{A_1 \cup A_2}, b_j^{A_1 \cap A_2}, b_j^{\overline{A_1}}$?

3.2. n-cube.

DEFINITION 3.3. $\mathbb{B}^n$ is called *n-cube* where $n \in \mathbb{N}$.

EXAMPLE 3.3.



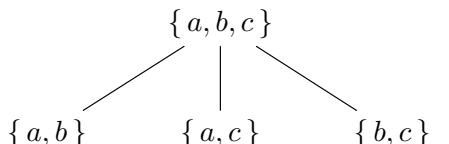
$\mathbb{B}^0$ $\mathbb{B}^1$ $\mathbb{B}^2$ $\mathbb{B}^3$

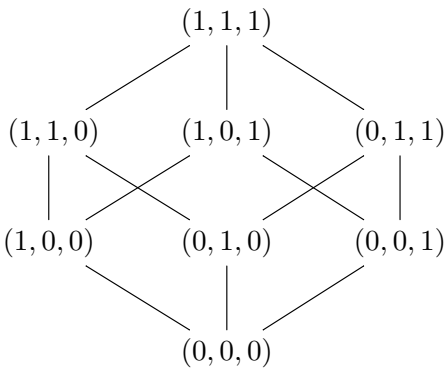
Note that

i. The diagram of $\mathbb{B}^n$ can be generated by two of diagrams of $\mathbb{B}^{n-1}$.
ii. The diagram of $\mathbb{B}^n$ is a undirected graph where vertices are $b \in \mathbb{B}^n$, and $b_1, b_2 \in \mathbb{B}^n$ are adjacent $\overset{\Delta}{\longleftrightarrow}$ they differ in exactly one coordinate.

THEOREM 3.2. *Let A be a finite set with $|A| = n$. The Hasse diagram of $\left[2^A, \cup, \cap\right]$ is the n-cube.*

EXAMPLE 3.4. Let $A = \{a, b, c\}$.




(1, 1, 1)
(1, 1, 0)
(1, 0, 1)
(0, 1, 1)
(1, 0, 0)
(0, 1, 0)
(0, 0, 1)
(0, 0, 0)

3.3. Bounded Lattice.

DEFINITION 3.4. A lattice L is called *bounded* $\overset{\Delta}{\longleftrightarrow}$ L has universal upper and lower bounds, 1 and 0, respectively.

3.4. Complemented Lattice.

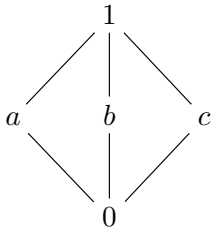
DEFINITION 3.5. Let $L = [L, \oplus, \odot]$ be a bounded lattice. $b \in L$ is called a *complement* of $a \in L$ $\overset{\Delta}{\longleftrightarrow}$ $a \odot b = 0 \ \wedge \ a \oplus b = 1$.

REMARK 3.1.

i. In general, complement may not exist.
ii. If it exists, it may not be unique. So complement is a relation rather then a function.
iii. $\begin{vmatrix} 0 \\ 1 \end{vmatrix}$ is a complement of $\begin{vmatrix} 1 \\ 0 \end{vmatrix}$.

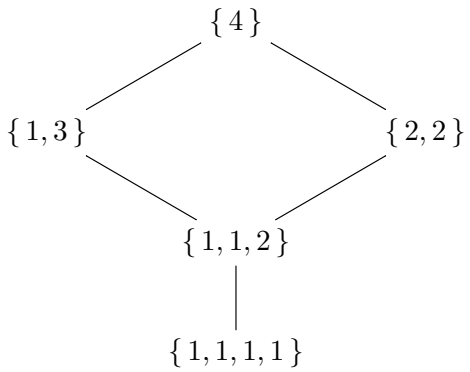
EXAMPLE 3.5.

$\begin{vmatrix} b \\ c \end{vmatrix}$ is a complement of a



Example 3.6.

The lattice of partitions of 4 is not complemented.



THEOREM 3.3 (Uniqueness of complement). *Let $L = [\, L, \oplus, \odot \,]$ be a bounded, distributive lattice. If b and c are complements of a, then $b = c$.*

REMARK 3.2. Note that complement of a may not exist. If it exists, then it is unique.

THEOREM 3.4 (Involution). *Let $L = [\, L, \oplus, \odot \,]$ be a bounded, distributive lattice. If $a \in L$ has the complement $\overline{a} \in L$, then $\overline{a}$ has its complement which is a. That is $\overline{\overline{a}} = a$.*

THEOREM 3.5 (De Morgan). *Let $L = [\, L, \oplus, \odot \,]$ be a bounded, distributive lattice. If complements of a and b exist, then*

$$\overline{a \oplus b} = \overline{a} \odot \overline{b} \text{ and } \overline{a \odot b} = \overline{a} \oplus \overline{b}$$

DEFINITION 3.6. A bounded lattice $L = [\, L, \oplus, \odot \,]$ is said to be *complemented* $\overset{\Delta}{\longleftrightarrow}$ $\forall a \in L \; \exists b \in L$ b is a complement of a.

4. Boolean Algebra

DEFINITION 4.1. A bounded, distributive, complemented lattice is called a *boolean algebra*. $\mathcal{B} = [\, B, \oplus, \odot, -, 0, 1 \,]$ denotes a boolean algebra with $\overline{a}$ is the complement of a, 0 and 1 are the universal lower and upper bounds, respectively.

EXAMPLE 4.1. Let A be a finite set. $\left[\, 2^A, \cup, \cap \,\right]$ is a booean algebra $\left[\, 2^A, \cup, \cap, -, 0, 1 \,\right]$.

5. Canonical Expressions in Boolean Algebras

THEOREM 5.1. *Let $\mathcal{B}$ be a boolean algebra and $x \in \mathcal{B}$. x is join-irreducible $\longleftrightarrow$ x is an atom.*

REMARK 5.1. Let $B = \{\, b_1, b_2, \ldots, b_n \,\}$ be the set of all atoms of boolean algebra $\mathcal{B}$.

i. No two elements of B is comparable.
ii. The join of any subset of B is irredundant.
iii. Any such join represents a unique element of $\mathcal{B}$.
iv. Therefore there is a bijection between 2^B and $\mathcal{B}$.

$$\varphi : \mathcal{B} \to 2^B$$
$$a \mapsto \text{ the subset whose join represents } a.$$

v. φ preserves $\oplus$ and $\odot$

$$\varphi(a \oplus b) = \varphi(a) \cup \varphi(b)$$
$$\varphi(a \odot b) = \varphi(a) \cap \varphi(b)$$
$$\varphi(\overline{a}) = \overline{\varphi(a)}.$$

THEOREM 5.2 (Stone representation). *A boolean algebra $\mathcal{B} = [\, B, \oplus, \odot, -, 0, 1 \,]$ of finite length is isomorphic to 2^B.*

THEOREM 5.3. *The Hasse diagram of a boolean algebra with n atoms is the n-cube.*

THEOREM 5.4. *A boolean algebra with n atoms has 2^n elements.*

REMARK 5.2. A boolean algebra $\mathcal{B}$ is entirely represented by n where n is the number of atoms. $\mathcal{B}_n$ denotes one such algebra.

Part 4

Number Systems

CHAPTER 8

Number Systems

1. Natural Numbers

DEFINITION 1.1 (Natural Numbers, (Peano Axioms)). The *set of natural numbers* is a set $\mathbb{N}$ that satisfies the following five axioms:
P1. $\exists 0 \in \mathbb{N}$.
P2. $\forall n \in \mathbb{N}\ \exists s(n) \in \mathbb{N}$. ($s(n)$ is called the *successor* of n).
P3. $\forall n \in \mathbb{N}$, $s(n) \neq 0$.
P4. $\forall m, n \in \mathbb{N}\,[s(m) = s(n) \longrightarrow m = n]$.
P5. $\forall A \subseteq \mathbb{N}\,[0 \in A \wedge \forall n\,(n \in A \longrightarrow s(n) \in A) \longrightarrow A = \mathbb{N}]$.

REMARK 1.1. These axioms are called Peano axioms [**Gal89**]. Note that s is a function given as $s : \mathbb{N} \to \mathbb{N}$.

REMARK 1.2. $\left|\begin{matrix} 0 \\ s(0) \\ s(s(0)) \\ \cdots \end{matrix}\right|$ is called $\left|\begin{matrix} \textit{zero} \\ \textit{one} \\ \textit{two} \\ \cdots \end{matrix}\right|$ and represented by $\left|\begin{matrix} \textit{0} \\ \textit{1} \\ \textit{2} \\ \cdots \end{matrix}\right|$

QUESTION 1.1. Does there exist such a set?

QUESTION 1.2. Is it unique? That is, if $\mathbb{N}_1$ and $\mathbb{N}_2$ are sets satisfying the Peano axioms, then are $\mathbb{N}_1$ and $\mathbb{N}_2$ isomorphic?

DEFINITION 1.2. The *predecessor* of n, $p(n)$, is $p(n) = m \overset{\Delta}{\longleftrightarrow} s(m) = n$.

DEFINITION 1.3 (Addition).
$+ : \mathbb{N} \times \mathbb{N} \to \mathbb{N}$ defined as:

$$n + m = \begin{cases} n, & m = 0, \\ s(n) + p(m), & \text{otherwise.} \end{cases}$$

REMARK 1.3. $n + m$ is called the *sum* of n and m.

DEFINITION 1.4 (Multiplication).
$\times : \mathbb{N} \times \mathbb{N} \to \mathbb{N}$ defined as:

$$n \times m = \begin{cases} 0, & m = 0, \\ n + n \times p(m), & \text{otherwise.} \end{cases}$$

REMARK 1.4. $n \times m$ is called the *product* of n and m.

DEFINITION 1.5 (Ordering).
$\leq : \mathbb{N} \times \mathbb{N} \to \{T, F\}$ defined as:

$$m \leq n \text{ means } \begin{cases} F, & \text{if } m \neq 0 \text{ and } n = 0, \\ T, & \text{if } m = 0, \\ p(m) \leq p(n), & \text{otherwise.} \end{cases}$$

QUESTION 1.3. Define exponentiation n^m.

QUESTION 1.4. What kind of algebraic structure is $\left|\begin{matrix} [\mathbb{N}, +] \\ [\mathbb{N}, \times] \\ [\mathbb{N}, +, \times] \end{matrix}\right|$?

2. Integers

Definition 2.1. Consider the relation $\sim$ on $\mathbb{N}\times\mathbb{N}$ defined as $\forall(a,b),(c,d)\in\mathbb{N}\times\mathbb{N}\,[(a,b)\sim(c,d) \iff a+d=b$

Example 2.1. Some elements of the relation $\sim$ are the followings:

$\cdots$	$(0,2)\sim(1,3)$	$(0,1)\sim(1,2)$	$(0,0)\sim(1,1)$	$(1,0)\sim(2,1)$	$(2,0)\sim(3,1)$	$\cdots$
$\cdots$	$(0,2)\sim(2,4)$	$(0,1)\sim(2,3)$	$(0,0)\sim(2,2)$	$(1,0)\sim(3,2)$	$(2,0)\sim(4,2)$	$\cdots$
$\cdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$

Theorem 2.1. *The relation $\sim$ is an equivalence relation on $\mathbb{N}\times\mathbb{N}$. Moreover, the set of equivalence classes, $\mathbb{N}\times\mathbb{N}/\sim$, is the set*

$$\mathbb{N}\times\mathbb{N}/\sim=\{[(n,0)]\mid n\in\mathbb{N}\}\cup\{[(0,n)]\mid n\in\mathbb{N}\wedge n\neq 0\}$$

Proof. Let $[(a,b)],[(c,d)],[(e,f)]\in\mathbb{N}\times\mathbb{N}/\sim$.

(1) $[(a,b)]\sim[(a,b)]$, since $a+b=b+a$. So $\sim$ is reflexive.
(2) $[(a,b)]\sim[(c,d)] \longrightarrow [(c,d)]\sim[(a,b)]$, since $a+d=b+c$. So $\sim$ is symmetric.
(3) If $[(a,b)]\sim[(c,d)]\wedge[(c,d)]\sim[(e,f)]$, then $a+d=b+c$ and $c+f=d+e$. Adding the two $a+d+c+f=b+c+d+e$, then $a+f=b+e$. Hence $[(a,b)]\sim[(e,f)]$. So $\sim$ is transitive.

Therefore $\sim$ is an equivalence relation on $\mathbb{N}\times\mathbb{N}$. □

Example 2.2. Some equivalence classes of $\mathbb{N}\times\mathbb{N}/\sim$ are:

$\cdots$
$[(2,0)]=\{(n+2,n)\mid n\in\mathbb{N}\}$
$[(1,0)]=\{(n+1,n)\mid n\in\mathbb{N}\}$
$[(0,0)]=\{(n,n)\mid n\in\mathbb{N}\}$
$[(0,1)]=\{(n,n+1)\mid n\in\mathbb{N}\}$
$[(0,2)]=\{(n,n+2)\mid n\in\mathbb{N}\}$
$\cdots$

Definition 2.2 (Integers). The set $\mathbb{N}\times\mathbb{N}/\sim$ is the *set of integers* and is denoted by $\mathbb{Z}$.

Definition 2.3 (Addition and Multiplication in $\mathbb{Z}$). Let $[(a,b)],[(c,d)]\in\mathbb{Z}$. Define $+:\mathbb{Z}\times\mathbb{Z}\to\mathbb{Z}$ and $\times:\mathbb{Z}\times\mathbb{Z}\to\mathbb{Z}$

$$[(a,b)]+[(c,d)]=[(a+c,b+d)]$$
$$[(a,b)]\times[(c,d)]=[(ac+bd,ad+bc)]$$

Theorem 2.2. $\forall x\in\mathbb{Z}\;\exists x'\in\mathbb{Z}\;[x+x'=x'+x=0]$. *Denote x' by $-x$.*

Remark 2.1. $x=[(a,b)] \longrightarrow -x=[(b,a)]$

Question 2.1. What kind of algebraic structure is $\left|\begin{array}{l}[\mathbb{Z},+]\\ [\mathbb{Z},\times]\\ [\mathbb{Z},+,\times]\end{array}\right|$?

Definition 2.4. $\mathbb{Z}^+=\{[(n,0)]\mid n\in\mathbb{N}\setminus\{0\}\}$ is called the *set of positive integers*. $\mathbb{Z}^-=\{[(0,n)]\mid n\in\mathbb{N}\setminus\{0\}\}$ is called the *set of negative integers*.

Definition 2.5 (Ordering in $\mathbb{Z}$).
Let $x,y\in\mathbb{Z}$

- x is *less than* y, denoted by $x<y$, $\overset{\Delta}{\iff}$ $y-x=y+(-x)\in\mathbb{Z}^+$.
- x is *less than or equal to* y, denoted by $x\leq y$ $\overset{\Delta}{\iff}$ $(x<y\vee x=y)$.

Acknowledgment. These notes are based on various books but especially [**PY73, Ros07, Men08, TZ82, Gal89**].

CHAPTER 9

Division

1. Division

DEFINITION 1.1 (Division).
Let $d, n \in \mathbb{Z}$. d *divides* $n \overset{\Delta}{\longleftrightarrow} \exists c \in \mathbb{Z}\ n = cd$.
$\left|\begin{matrix} d \\ n \end{matrix}\right|$ is a $\left|\begin{matrix} \textit{factor} \text{ or } \textit{divisor} \\ \textit{multiple} \end{matrix}\right|$ of $\left|\begin{matrix} n \\ d \end{matrix}\right|$.

NOTATION. If d divides n, we write $d \mid n$. If d does not divide n, we write $d \nmid n$.

THEOREM 1.1. *Let* $d, n, m, a, b \in \mathbb{Z}$.
i. $n \mid n$ *(reflexivity)*
ii. $d \mid n \ \wedge\ n \mid m \longrightarrow d \mid m$ *(transitivity)*
iii. $d \mid n \ \wedge\ d \mid m \longrightarrow d \mid (an + bm)$ *(linearity)*
iv. $d \mid n \longrightarrow ad \mid an$ *(multiplication)*
v. $ad \mid an \ \wedge\ a \neq 0 \longrightarrow d \mid n$ *(cancellation)*
vi. $1 \mid n$ *(1 divides every integer)*
vii. $n \mid 0$ *(every integer divides 0)*
viii. $0 \mid n \longrightarrow n = 0$ *(0 divides only 0)*
ix. $d \mid n \ \wedge\ n \neq 0 \longrightarrow |d| \leq |n|$
x. $d \mid n \ \wedge\ n \mid d \longrightarrow |d| = |n|$
xi. $d \mid n \ \wedge\ d \neq 0 \longrightarrow (n/d) \mid n$

REMARK 1.1.
i. The expression $an + bm$ is called a *linear combination* of n and m.
ii. Every common multiple of n and m can be written as a linear combination of n and m.

Due to linearity $d \mid n \ \wedge\ d \mid m \longrightarrow d \mid (n+m) \wedge d \mid (nm)$.

QUESTION 1.1. Prove that $d \mid n \ \wedge\ d \mid m \longrightarrow d \mid (nm)$ using linearity.

THEOREM 1.2 (The Division Algorithm).
Let $n \in \mathbb{Z}, d \in \mathbb{Z}^+$. *Then there are unique* $q, r \in \mathbb{Z}$ *with* $0 \leq r < d$ *such that* $n = qd + r$.

DEFINITION 1.2.
Let $n = qd + r$ as in the division algorithm.
$\left|\begin{matrix} d \\ n \\ q \\ r \end{matrix}\right|$ is called $\left|\begin{matrix} \textit{divisor} \\ \textit{dividend} \\ \textit{quotient} \\ \textit{remainder} \end{matrix}\right|$ $\left|\begin{matrix} \\ \\ q = n \text{ div } d \\ r = n \bmod d \end{matrix}\right|$

The division algorithm is given as Algorithm 1. A trace of the algorithm for $n = 13, d = 3$ is given in Example 1.1.

EXAMPLE 1.1. 3 divides 13. That is, $n = 13, d = 3$.

n	d	q	r
13	3	0	
10		1	
7		2	
4		3	
1		4	
			1

Algorithm 1: The Division Algorithm

Input: Dividend $n > 0$ and divisor $d > 0$
Output: Quotient q and remainder r where $0 \leq r < d$

```
begin
    q ← 0
    while n ≥ d do
        q ← q + 1
        n ← n − d
    end
    r ← n
end
```

2. Prime Numbers

DEFINITION 2.1. Let $n \in \mathbb{Z}^+$. $n > 1$ is called *prime* $\overset{\Delta}{\longleftrightarrow}$ The only positive divisors of n are 1 and n. If n is not prime, then n is called *composite*.

NOTATION. Prime numbers are usually denoted by p, p', p_i, q, q', q_i.

EXAMPLE 2.1. There are 4 primes less than 10. There are 25 prime numbers less than 100 which are 2, 3, 5, 7, 11, 13, 17, 19, 23, 29, 31, 37, 41, 43, 47, 53, 59, 61, 67, 71, 73, 79, 83, 89, and 97. So for 40 % of the numbers between 1 and 10 are prime where as only 25 % of the numbers between 1 and 100 are prime. The primes becomes sparse as the numbers grow. There are 168, that is 17 %, prime numbers less than 1000.

THEOREM 2.1. *Every integer $n > 1$ is either a prime number or a product of prime numbers.*

PROOF. Use induction on n. As an induction base $n = 2$ is a prime. Assume that it is true for all the numbers m where $1 < m < n$, that is $\forall m \in \mathbb{Z}\ [1 < m < n]$ m is either a prime or a product of primes.
We want to prove that it is true for n, too. Consider n.

i. Case: n is prime. Then we are done.
ii. Case: n is not prime. Then there must be a positive divisor d that divides n, that is, $\exists d \in \mathbb{Z}(d > 0 \ \wedge \ d \neq 1 \ \wedge \ d \neq n) \exists c \in \mathbb{Z} n = cd$. Since both $1 < c < n$ and $1 < d < n$, by the induction hypothesis they are either prime or a product of primes. Therefore their product n should be a product of primes. □

THEOREM 2.2 (Euclid).
There are infinitely many prime numbers.

REMARK 2.1. The proof of the theorem is given in Elements by Euclid (300 BC).

NOTATION. Let $\mathbb{P}$ be the set of prime numbers.

THEOREM 2.3. $a \in \mathbb{Z} \ \wedge \ b \in \mathbb{Z} \ \wedge \ p \in \mathbb{P} \ \wedge \ p \mid ab \ \longrightarrow \ p \mid a \ \vee \ p \mid b$. *More generally,* $a_1, a_2, \ldots, a_i \in \mathbb{Z} \ \wedge \ p \in \mathbb{P} \ \wedge \ p \mid a_1 a_2 \cdots a_n \ \longrightarrow \ \exists i \in \{1, \ldots, n\}\ p \mid a_i$.

THEOREM 2.4 (The Fundamental Theorem of Arithmetic).
Every integer $n > 1$ can be written uniquely as a product of nondecreasing primes.

EXAMPLE 2.2.

$2 = 2$ $\quad 6 = 2 \cdot 3$ $\quad 30 = 2 \cdot 3 \cdot 5$
$4 = 2 \cdot 2 = 2^2$ $\quad 12 = 2 \cdot 2 \cdot 3 = 2^2 \cdot 3$ $\quad 720 = 2 \cdot 2 \cdot 2 \cdot 2 \cdot 3 \cdot 3 \cdot 5 = 2^4 \cdot 3^2 \cdot 5^1$

REMARK 2.2.

i. Given an integer $n > 1$, it is not easy to decide whether n is prime.
ii. There is no known formula that generates primes only.
iii. Prime numbers of the form $2^p - 1$, where $p \in \mathbb{P}$, are called *Mersenne primes*. As of 2009, there are 47 Mersenne primes known. The largest is $2^{43,112,609} - 1$ [**Wik09**].
iv. Numbers in the form of $F_n = 2^{2^n} + 1$ where $n \in \mathbb{N}$ are called *Fermat numbers*. $F_0 = 3, F_1 = 5, F_2 = 17, F_3 = 257, F_4 = 65,537$ are all prime but Euler (1732) found that $F_5 = 2^{32} + 1 = 641 \times 6,700,417$ is not prime. Beyond F_5 no Fermat primes have been found [**Apo**].

THEOREM 2.5. *Let $n \in \mathbb{Z}^+$ be composite. Then there is a prime divisor p of n and $p \leq \sqrt{n}$.*

DEFINITION 2.2. Let $\pi(x)$ be the number of primes p satisfying $2 \leq p \leq x$.

REMARK 2.3. The density of primes drops as the numbers grow.

n	10^1	10^2	10^3
Number of primes in $\{2, 3, \cdots, n\}$	4	25	168
Percentage of primes in $\{2, 3, \cdots, n\}$	40%	25%	17%

THEOREM 2.6 (The Prime Number Theorem (Hadamard + Vallee Poussin, 1896)).

$$\lim_{x \to \infty} \frac{\pi(x) \ln x}{x} = 1$$

COROLLARY 2.7 (Goldbach's Conjecture (1742)).
Every even integer $n > 2$ is the sum of two primes.

DEFINITION 2.3 (Twin Primes).
If p and $p + 2$ are primes, they are called *twin primes*.

THEOREM 2.8 (The Twin Prime Conjecture).
There are infinitely many twin primes.

3. Common Divisors and Multiples

DEFINITION 3.1 (Common Divisor).
Let $a, b \in \mathbb{Z}$. $d \in \mathbb{Z}^+$ is called a *common divisor* of a and b. $\overset{\Delta}{\longleftrightarrow}$ $d \mid a \wedge d \mid b$. Let $\mathrm{cd}(a, b)$ be the *set of all common divisors* of a and b.

REMARK 3.1. $\forall a, b \in \mathbb{Z}\ [1 \in \mathrm{cd}(a, b)]$. So, $\mathrm{cd}(a, b) \neq \emptyset$.

THEOREM 3.1. $\forall a, b \in \mathbb{Z}\ \exists d \in \mathbb{Z}^+$ *such that* $d \in \mathrm{cd}(a, b) \wedge \exists x, y \in \mathbb{Z}\ (d = ax + by)$. *Moreover,* $\forall k \in \mathrm{cd}(a, b)\ [k \mid d]$.

THEOREM 3.2. $\forall a, b \in \mathbb{Z}\ \exists! d \in \mathbb{Z}$ *with the following properties:*

i. $d \geq 0$
ii. $d \in \mathrm{cd}(a, b)$
iii. $\forall e \in \mathrm{cd}(a, b) \longrightarrow e \mid d$.

PROOF. By Theorem 3.1 there is at least one d satisfying (ii) and (iii). Note that $-d$ also satisfies the condition (iii). Suppose d' also satisfies (ii) and (iii), then $d \mid d'$ and $d' \mid d$, so $|d| = |d'|$. Hence there is exactly one $d \geq 0$ satisfying both (ii) and (iii). □

DEFINITION 3.2 (Greatest Common Divisor). The number d in Theorem 3.2 is called the *greatest common divisor* (gcd) of a and b and denoted by $a \mathbin{\mathrm{D}} b$ or $\gcd(a, b)$.

REMARK 3.2. $a \mathbin{\mathrm{D}} b$ is the operator notation, $\gcd(a, b)$ is the function notation of the greatest common divisor.

THEOREM 3.3. *The gcd has the following properties:*

i. $a \mathbin{\mathrm{D}} b = b \mathbin{\mathrm{D}} a$ *(commutativity)*
ii. $a \mathbin{\mathrm{D}} (b \mathbin{\mathrm{D}} c) = (a \mathbin{\mathrm{D}} b) \mathbin{\mathrm{D}} c$ *(associativity)*
iii. $|a| (b \mathbin{\mathrm{D}} c) = (ab) \mathbin{\mathrm{D}} (ac)$ *(distributivity)*
iv. $a \mathbin{\mathrm{D}} 1 = 1 \mathbin{\mathrm{D}} a = 1$
v. $a \mathbin{\mathrm{D}} 0 = 0 \mathbin{\mathrm{D}} a = |a|$

DEFINITION 3.3. If $a \mathbin{\mathrm{D}} b = 1$, then a and b are said to be *relatively prime*, denoted by $a \perp b$.

REMARK 3.3. $a \perp b \longleftrightarrow \exists x, y \in \mathbb{Z}\ xa + yb = 1$.

THEOREM 3.4 (Euclid's lemma).
$a \mid bc \wedge a \perp b \longrightarrow a \mid c$.

PROOF. Since $a \perp b$ we can write $1 = ax + by$. Therefore $c = cax + cby$. Since $a \mid cax$ and $a \mid bc$, $a \mid (cax + cby)$. Hence $a \mid c$. □

THEOREM 3.5. $p \in \mathbb{P} \wedge p \nmid a \longrightarrow p \mathbin{\mathrm{D}} a = 1$

THEOREM 3.6. *a and b are relatively prime.* $\longleftrightarrow \nexists p \in \mathbb{P}\ [p \mid a \ \wedge\ p \mid b]$.

EXAMPLE 3.1. Due to unique prime factorization, a positive integer can be represented as a vector where ith entry of the vector is the power of the ith prime number in the prime factorization of the number.

n	prime factors	vector
10	$2^1 \times 3^0 \times 5^1 \times 7^0 \times 11^0 \times \cdots$	$[10100\cdots]$
12	$2^2 \times 3^1 \times 5^0 \times 7^0 \times 11^0 \times \cdots$	$[21000\cdots]$
63	$2^0 \times 3^2 \times 5^0 \times 7^1 \times 11^0 \times \cdots$	$[02010\cdots]$

Consider the dot product of the corresponding vectors. The corresponding vectors are perpendicular if and only if the numbers are relatively prime:
$[10100\cdots] \cdot [02010\cdots] = 0 \ \longrightarrow 10 \perp 63$.
$[10100\cdots] \cdot [21000\cdots] = 2 \neq 0 \ \longrightarrow 10 \not\perp 12$.

QUESTION 3.1. What is the dimension of this vector space?

DEFINITION 3.4. Let $a, b \in \mathbb{Z}^+$. The smallest $m \in \mathbb{Z}^+$ with $a \mid m$ and $b \mid m$ is called the *least common multiple* of a and b, denoted by $\operatorname{lcm}(a, b)$.

REMARK 3.4. A more proper approach would be the approach of the greatest common divisor: First define the set of common multiples, denoted by $\operatorname{cm}(a, b)$. Then show that $\operatorname{cm}(a, b) \neq \emptyset$ since $ab \in \operatorname{cm}(a, b)$. Finally, show that there is the least element in $\operatorname{cm}(a, b)$.

THEOREM 3.7. *Let $[a_i]$ and $[b_i]$ be the vector representations of a and b. Then*

$$\gcd(a, b) = \prod_i p_i^{\min\{a_i, b_i\}} \text{ and } \operatorname{lcm}(a, b) = \prod_i p_i^{\max\{a_i, b_i\}}$$

where p_i is the ith prime number in the vector representation.

4. Modular Arithmetic

REMINDER 4.1. The division algorithm of Theorem 1.2 states that for $n \in \mathbb{Z}, d \in \mathbb{Z}^+$ there exist unique $q, r \in \mathbb{Z}$ with $n = qd + r$ and $0 \leq r < d$. Note that $r = n \bmod d$ and $q = n \operatorname{div} d$.

DEFINITION 4.1. Let $a, b \in \mathbb{Z}$ and $m \in \mathbb{Z}^+$. a is said to be *congruent* to b modulo m, denoted by $a \equiv b \pmod{m}$. $\overset{\Delta}{\longleftrightarrow}\ m \mid (a - b)$. If a and b are not congruent modulo m, then $a \not\equiv b \pmod{m}$.

THEOREM 4.1. $a \equiv b \pmod{m} \ \longleftrightarrow a \bmod m = b \bmod m$.

THEOREM 4.2. $a \equiv b \pmod{m} \ \longleftrightarrow \exists k \in \mathbb{Z}\ [a = b + km]$.

THEOREM 4.3. *Let $m \in \mathbb{Z}^+$ and $a, b, c, d \in \mathbb{Z}$.*
$a \equiv b \pmod{m} \ \wedge\ c \equiv d \pmod{m}$
$\longrightarrow a + c \equiv b + d \pmod{m} \ \wedge\ ac \equiv bd \pmod{m}$.

DEFINITION 4.2. The *congruence class* of a modulo m is defined as $[a]_m \triangleq \{\, n \in \mathbb{Z} \mid \exists k \in \mathbb{Z}\ n = a + km \,\}$

THEOREM 4.4. $\forall x, y \in [a]_m\ x \equiv y \pmod{m}$

THEOREM 4.5. *Let $a, b, c, d \in \mathbb{Z}$ and $m \in \mathbb{Z}^+$ $(m > 1)$. If $a \equiv b \pmod{m}$ and $c \equiv d \pmod{m}$ then*
a) $a + c \equiv b + d \pmod{m}$
b) $ac \equiv bd \pmod{m}$
c) $\forall x \in \mathbb{Z},\ ax \equiv bx \pmod{m}$
d) $\forall n \in \mathbb{Z}^+,\ a^n \equiv b^n \pmod{m}$.

———— @HB ————

APPLICATION 4.1.
i. Hashing Functions.
ii. Pseudorandom Numbers.
iii. Cryptology.

Acknowledgment. These notes are based on various books such as [**PY73, Ros07, Gal89**] but especially [**Apo**].

Part 5

Combinatorics

CHAPTER 10

Counting

1. Motivation

Counting the number of different ways of satisfying a condition is important in science including Probability, Statistical Physics and, of course, Computer Science. Theory of Algorithms in Computer Science deals with finding an algorithm which has the minimum number of steps. Then one needs to know what is the possible number of steps.

EXAMPLE 1.1.

i. How many vertices in a complete binary tree?
ii. How many steps are needed in order to traverse a binary tree?
iii. If you have n items, how many comparisons do you need to make in order to sort them?

QUESTION 1.1.

i. Suppose there are 3 balls looks like the same but one of them is different in weight only. How many comparisons in weight does it needed to figure out the different one?
ii. The same question for 12 balls?

EXAMPLE 1.2. Consider a room filled with air. The room divided into two halves. You are sitting in one of the halves. What is the possibility that all the molecules would be in the other half. Our every day observations says that that does not happen frequently. If that happens frequently enough, you would be suffocated. What is the probability that it happens, if the number of molecules in the room is n where

i. $n = 2$?
ii. $n = 8$?
iii. n is in the order of Avogadro number, that is $n \approx 10^{24}$?
iv. What is you estimate of n if the room is $4 \times 5 \times 3$ in meters?

Assume that air is an ideal gas where molecules are free to move without any interaction from each other.

This chapter deals with finite sets. Let A and B be finite sets. Let $|\,A\,| = \alpha$ and $|\,B\,| = \beta$. Since A is finite, the elements of A can be listed as an α-tuple $(a_1, a_2, \ldots, a_\alpha)$. This ordering is used for the following proofs.

DEFINITION 1.1 (Factorial).
Let $n \in \mathbb{N}$.

i. $0! = 1$
ii. $n! = n \cdot (n-1)!$ for $n > 0$.

REMARK 1.1. Use of factorial is quite old. One of the earliest use of factorials is in the proof of prime numbers are infinite given by Euclid around 300BC. The proof is based on the idea that there must be a prime between n and $n! + 1$.

DEFINITION 1.2 (Factorial Power).
Let $n, r \in \mathbb{Z}^+$.

$$\begin{array}{lll} n^{\overline{r}} & \triangleq & (n+0) \quad (n+1)(n+2)\cdots(n+(r-1)), \\ n^{\underline{r}} & \triangleq & (n-(r-1))\cdots(n-2)(n-1) \quad (n-0). \end{array}$$

$n^{\overline{r}}$ and $n^{\underline{r}}$ are called *rising factorial power* and *falling factorial power* and are read as *n to the r rising* and *n to the r falling*, respectively.

REMARK 1.2. Notice that

$$n^{\underline{r}} = \frac{n!}{(n-r)!}.$$

The notation of the rising and falling factorial power is due to [**GKP98**]

DEFINITION 1.3 (The set of bits).
$\mathbb{B} \triangleq \{0, 1\}$.

2. Cardinality: Finite and Infinite Sets

Numbers such as integers or reals are infinitely many. Then how do you represent them in computers? The number of bits in a register of a typical computer is usually 32 bits. Therefore the number of different bit patterns that can be obtained using 32-bit is 2^{32}. This is a quite big number but clearly not big enough to represent integers.

The real numbers are represented in computer by means of floating point arithmetic but we have the same representation problem since the set of real numbers is also infinite.

Do we really have the same problem?

DEFINITION 2.1. Let A and B be sets. A and B are said to be of the same *cardinality*, denoted by $|A| = |B|$, $\overset{\Delta}{\longleftrightarrow}$ There is a bijection from A to B.

DEFINITION 2.2. $I_m^n \triangleq \{m, m+1, \ldots, n\}$ where $m, n \in \mathbb{Z}$ and $m \leq n$.

DEFINITION 2.3. A set A is called *finite* $\overset{\Delta}{\longleftrightarrow}$ A has the same cardinality of I_1^n for some $n \in \mathbb{N}$, or $A = \emptyset$.

NOTATION 2.1. The cardinality of $\mathbb{Z}^+$ is denoted by $\aleph_0$, read as aleph null. That is, $\aleph_0 \triangleq |\mathbb{Z}^+|$.

DEFINITION 2.4. A set A is called *countable* $\overset{\Delta}{\longleftrightarrow}$ A is finite or A has the same cardinality of $\mathbb{Z}^+$.

DEFINITION 2.5. A set A is called *uncountable* $\overset{\Delta}{\longleftrightarrow}$ A is not countable.

EXAMPLE 2.1 (Hilbert's Hotel). Infinite sets have unintuitive properties. Hilbert provide a very nice story about a hotel with countably infinite rooms. Suppose the hotel is completely full and the officer at the reception is good in mathematics.

One person arrives. The receptionist asks every person in room number k to move the the room number $k+1$. By doing that the room number 1 becomes empty and the new comer gets it.

This time countably infinite group of people arrives. The receptionist asks every one in room number k to move the the room number $2k$. By doing so all the rooms with odd numbers become empty. So the kth person of the new group gets the room with number $2k+1$.

THEOREM 2.1. $|\mathbb{N}| = \aleph_0$.

PROOF. Define $f : \mathbb{N} \to \mathbb{Z}^+$ such that $f(n) = n + 1$. Since f is a bijection (left as exercise), $|\mathbb{N}| = |\mathbb{Z}^+|$. □

THEOREM 2.2. $|\mathbb{E}| = \aleph_0$ *and* $|\mathbb{O}| = \aleph_0$ *where* $\mathbb{E}, \mathbb{O}$ *are even and odd natural numbers, respectively.*

PROOF. Define $f : \mathbb{N} \to \mathbb{E}$ such that $f(n) = 2n$ and $g : \mathbb{N} \to \mathbb{O}$ such that $g(n) = 2n + 1$. Since f and g are bijections (left as exercise), $|\mathbb{E}| = |\mathbb{N}|$ and $|\mathbb{O}| = |\mathbb{N}|$. □

THEOREM 2.3. $|\mathbb{Z}^-| = \aleph_0$.

PROOF. Define $f : \mathbb{Z}^+ \to \mathbb{Z}^-$ such that $f(n) = -n$. Since f is a bijection (left as exercise), $|\mathbb{Z}^-| = |\mathbb{Z}^+|$. □

THEOREM 2.4. $|\mathbb{Z}| = \aleph_0$.

PROOF. Define $f : \mathbb{N} \to \mathbb{Z}$ such that

$$f(n) = \begin{cases} -k, & n = 2k, k \neq 0, k \in \mathbb{N} \\ 0, & n = 0 \\ k, & n = 2k+1, k \in \mathbb{N}. \end{cases}$$

Since f is a bijection, $|\mathbb{Z}| = |\mathbb{N}|$. □

THEOREM 2.5. $|\mathbb{Q}^+| = \aleph_0$.

PROOF.

$\downarrow q \to p$	1	2	3	4	5	6	$\cdots$
1	$\frac{1}{1}$	$\frac{1}{2}$	$\frac{1}{3}$	$\frac{1}{4}$	$\frac{1}{5}$	$\frac{1}{6}$	$\cdots$
2	$\frac{2}{1}$	$\frac{2}{2}$	$\frac{2}{3}$	$\frac{2}{4}$	$\frac{2}{5}$	$\frac{2}{6}$	$\cdots$
3	$\frac{3}{1}$	$\frac{3}{2}$	$\frac{3}{3}$	$\frac{3}{4}$	$\frac{3}{5}$	$\frac{3}{6}$	$\cdots$
4	$\frac{4}{1}$	$\frac{4}{2}$	$\frac{4}{3}$	$\frac{4}{4}$	$\frac{4}{5}$	$\frac{4}{6}$	$\cdots$
5	$\frac{5}{1}$	$\frac{5}{2}$	$\frac{5}{3}$	$\frac{5}{4}$	$\frac{5}{5}$	$\frac{5}{6}$	$\cdots$
6	$\frac{6}{1}$	$\frac{6}{2}$	$\frac{6}{3}$	$\frac{6}{4}$	$\frac{6}{5}$	$\frac{6}{6}$	$\cdots$
$\cdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$

Define $f : \mathbb{N} \to \mathbb{Q}^+$ such that $f(n) =?$[1]. Since f is a bijection, $|\mathbb{Q}^+| = |\mathbb{N}|$. □

THEOREM 2.6. $|\mathbb{Q}| = \aleph_0$.

PROOF. Let $f : \mathbb{Z}^+ \to \mathbb{Q}^+$ be a bijection. There is such a bijection since $|\mathbb{Z}^+| = |\mathbb{Q}^+|$. Define $q : \mathbb{Z} \to \mathbb{Q}$ using $f : \mathbb{Z}^+ \to \mathbb{Q}^+$ as

$$q(z) = \begin{cases} -f(z), & -z \in \mathbb{Z}^+ \\ 0, & z = 0 \\ f(z), & z \in \mathbb{Z}^+. \end{cases}$$

Since q is a bijection, $|\mathbb{Q}| = |\mathbb{Z}|$. □

SUMMARY 2.1. $\aleph_0 = |\mathbb{Z}^+| = |\mathbb{N}| = |\mathbb{E}| = |\mathbb{O}| = |\mathbb{Z}| = |\mathbb{Z}^-| = |\mathbb{Q}|$.

REMARK 2.1. So far it seems that there is only one kind of infinity, that is $\aleph_0$. Note that these sets are countable, that is, there is a bijection from $\mathbb{N}$ to them.

THEOREM 2.7 (Cantor's diagonalization).
$|[0,1)| \neq \aleph_0$ *where* $[0,1) = \{ x \in \mathbb{R} \mid 0 \leq x < 1 \}$.

PROOF. Assume that $|[0,1)| = \aleph_0$. Then we can make a list of elements of $[0,1)$. All the real numbers in $[0,1)$ have the decimal expansion of the form $0.d_1d_2d_3\ldots$. Let $x_k \in [0,1)$ be the kth real number in the list. Construct a new real number y in such a way that the kth digit of y would be different that the kth digit of the x_k. y is in $[0,1)$ since its decimal expansion is in the proper form. Yet, y is not in the list because for every k, y is different than x_k in the kth digit since its kth digit is different than d_k. This contradicts the assumption that such a list can be made. If such a list cannot be made that the set $[0,1)$ is not countable. □

REMARK 2.2. The set $[0,1)$ is uncountable. Note that the interval $[0,1)$ contains rational numbers such as $1/2$ as well as irrational numbers such as $\pi/4$ or $\sqrt{2}/2$. The rational numbers have repeating patterns in the decimal expansion where as irrational numbers do not.

2.1. Hierarchy of Infinities. It seems that there is one type of infinity which is $\aleph_0$. Cantor showed that there are actually infinitely many infinities.

THEOREM 2.8 (Cantor's Theorem). *Let A be a set. Then $|A| \neq |2^A|$.*

In other words "No set is the same size as its power set".

We show that mapping f from A to 2^A is not a surjection, therefore f is not a bijection.
Suppose $f : A \longrightarrow 2^A$. Define $B = \{x \in A \mid x \notin f(x)\}$. Clearly $B \subseteq A$. This means $B \in 2^A$.

We claim that $\forall x \in A\ f(x) \neq B$, that is $B \notin f(A)$. If the claim is correct, that means f cannot map into B. Hence f is not a surjection. Therefore f cannot be a bijection.
There is no bijection from A to 2^A. Therefore $|A| \neq |2^A|$.

THEOREM 2.9. $\forall x \in A\ f(x) \neq B$, *that is* $B \notin f(A)$.

PROOF. Suppose $\exists b \in A\ f(b) = B$. Ask if $b \in B$?

i) $b \in B$ case: By definition of B, $b \notin f(b)$. Since $f(b) = B$, we have $b \notin B$. Contradiction.
ii) $b \notin B$ case: Since $f(b) = B$, this means $b \notin f(b)$. This means $b \in B$ since B is defined so. Contradiction.

So we obtain $b \in B \Leftrightarrow b \notin B$. This contradiction means $\neg\exists b \in A\ f(b) = B$. That is $\forall b \in A\ f(b) \neq B$. The theorem is proved. □

[1]How to define this function?

REMARK 2.3. Using the Theorem 2.8 we can obtain infinitely many infinities based on $\mathbb{N}$. Let $N_0 = \mathbb{N}$. Then $|2^{\mathbb{N}}| \neq |\mathbb{N}|$. We can extend this as follows:

$$\text{Define} \left| \begin{array}{l} N_1 = 2^{N_0} \\ N_2 = 2^{N_1} \\ \dots \end{array} \right| . \text{ Then } \left| \begin{array}{l} |N_1| \neq |N_0| \\ |N_2| \neq |N_1| \\ \dots \end{array} \right| .$$

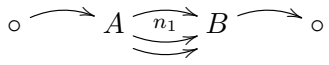
REMARK 2.4. The set of real numbers also produces an other chain of infinities as follows: Let $R_0 = \mathbb{R}$. Then $|2^{\mathbb{R}}| \neq |\mathbb{R}|$. We can extend this as follows:

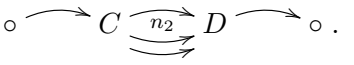
$$\text{Define} \left| \begin{array}{l} R_1 = 2^{R_0} \\ R_2 = 2^{R_1} \\ \dots \end{array} \right| . \text{ Then } \left| \begin{array}{l} |R_1| \neq |R_0| \\ |R_2| \neq |R_1| \\ \dots \end{array} \right| .$$

QUESTION 2.1. Are the hierarchies of $N_0, N_1, \dots$ and $R_0, R_1, \dots$ related or different?

3. The Number of Ways

Suppose there are n_1 ways to go from A to B and n_2 ways to go from C to D.

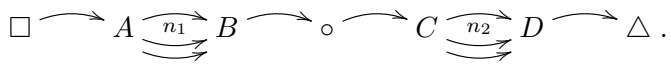




3.1. The Product Rule.

REMARK 3.1. Let A and B be finite sets where $|\,A\,| = \alpha$ and $|\,B\,| = \beta$.

THEOREM 3.1 (The Product Rule).
Suppose $A - B$ and $C - D$ are connected in serial. The number of different ways to go from $\square$ to $\triangle$ is $n_1 \times n_2$.



THEOREM 3.2 (The number of elements of cartesian product).
Let A be finite set. Then $|\,A^n\,| = |\,A\,|^n$.

PROOF. By induction on n using the product rule. □

COROLLARY 3.3 (The number of bit strings of length n).
Let $\mathbb{B}^n$ be the the set of bit strings of length n where $\mathbb{B} = \{\,0, 1\,\}$. Then $|\,\mathbb{B}^n\,| = 2^n$.

THEOREM 3.4 (The number of truth tables).
There are 2^{2^n} different truth tables for propositions in n variables.

PROOF. There are 2^n rows in a truth table of n variables. For each row, one can assign two choices, namely F or T. Hence there are 2^{2^n} different assignments. □

The following theorem will be proved in steps.

THEOREM 3.5 (The number of various type of functions).

$$\textit{The number of} \left| \begin{array}{l} \textit{functions} \\ \textit{injections} \\ \textit{surjections} \\ \textit{bijections} \\ \textit{partial functions} \end{array} \right| \textit{ from } A \textit{ to } B \textit{ is} \left| \begin{array}{l} \beta^\alpha \\ \beta^{\underline{\alpha}} \\ ? \\ \beta! \\ (\beta+1)^\alpha \end{array} \right| .$$

PROOF. The number of functions from A to B, that is $\left| B^A \right| = |B|^{|A|}$.
Since A is finite, the elements of A can be listed as $(a_1, a_2, \ldots, a_\alpha)$. Use this ordering of elements:
For $\left| \begin{array}{l} a_1 \in A \\ a_2 \in A \\ \ldots \\ a_\alpha \in A \end{array} \right|$, there are $\left| \begin{array}{l} \beta \\ \beta \\ \ldots \\ \beta \end{array} \right|$ different choices for $\left| \begin{array}{l} f(a_1) \in B \\ f(a_2) \in B \\ \ldots \\ f(a_\alpha) \in B \end{array} \right|$.
By product rule, there are $\beta\beta\ldots\beta = \beta^\alpha$ different ways. □

PROOF. The number of injections from A to B is $\beta^{\underline{\alpha}}$.
If $|A| > |B|$, then there is no injection from A to B. So consider the case of $|A| \leq |B|$. Since A is finite, the elements of A can be listed. Use this ordering of elements:
For $\left| \begin{array}{l} a_1 \in A \\ a_2 \in A \\ \ldots \\ a_\alpha \in A \end{array} \right|$, there are $\left| \begin{array}{l} \beta - 0 \\ \beta - 1 \\ \ldots \\ \beta - (\alpha - 1) \end{array} \right|$ different choices for $\left| \begin{array}{l} f(a_1) \in B \\ f(a_2) \in B \\ \ldots \\ f(a_\alpha) \in B \end{array} \right|$.
By product rule, there are $(\beta - 0)(\beta - 1)\ldots(\beta - (\alpha - 1)) = \beta^{\underline{\alpha}}$ different ways. □

EXAMPLE 3.1. A common approach to counting is to define a bijection to a set whose cardinality is already known.

We prove that the number of bit strings of length n is $|\mathbb{B}^n| = 2^n$ by defining a bijection to $\mathbb{B}^{\{1,2,\ldots,n\}}$ whose cardinality is 2^n. Consider the set $\mathbb{B}^{\{1,2,\ldots,n\}}$ of functions from $\{1, 2, \ldots, n\}$ to $\mathbb{B}$. Define a function

$$\begin{aligned} f : \mathbb{B}^{\{1,2,\ldots,n\}} &\to \{0, 1\}^n \\ g &\mapsto (b_1, b_2, \ldots, b_n) \end{aligned}$$

where $g \in \mathbb{B}^{\{1,2,\ldots,n\}}$ and $b_i = g(i), i \in \{1, 2, \ldots, n\}$.
f is a bijection (proof?). Therefore

$$|\mathbb{B}^n| = \left| \mathbb{B}^{\{1,2,\ldots,n\}} \right| = |\mathbb{B}|^{|\{1,2,\ldots,n\}|} = 2^n.$$

REMARK 3.2. IPv4 is the currently used protocol for the Internet in which every device on the Internet should have a unique ID. IPv4 uses 32-bit IDs which is usually written as four numbers separated by dots as in the case of 127.0.0.1. Therefore 2^{32} devices can be connected to the Internet at a given time. Although it is a very big number, it is not big enough for the demand of the future. IPv6 is planned to use 128-bit ID's.

THEOREM 3.6 (The number of subsets).
Let A be a finite set. Then $\left| 2^A \right| = 2^{|A|} = 2^\alpha$.

PROOF. Use α-tuple $(a_1, a_2, \ldots, a_\alpha)$ of A. Define a function

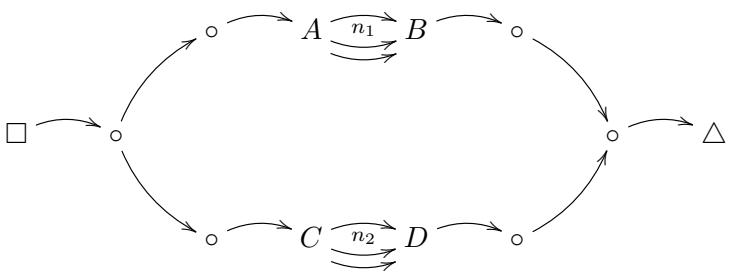
$$\begin{aligned} f : 2^A &\to \mathbb{B}^\alpha \\ \mathcal{A} &\mapsto (b_1, b_2, \ldots, b_\alpha) \end{aligned}$$

where $\mathcal{A} \in 2^A$ and $b_i = \begin{cases} 1, & a_i \in \mathcal{A}, \\ 0, & a_i \notin \mathcal{A}. \end{cases}$
f is a bijection (proof?). Therefore $\left| 2^A \right| = |\mathbb{B}^\alpha| = 2^\alpha = 2^{|A|}$. □

THEOREM 3.7. *Let $A_1, A_2, \ldots, A_n$ be finite sets. Then $|A_1 \times A_2 \times \ldots \times A_n| = |A_1||A_2|\cdots|A_n|$*

3.2. The Sum Rule.

THEOREM 3.8 (The Sum Rule).
Suppose $A - B$ and $C - D$ is connected in parallel. The number of different ways to go from □ to △ is $n_1 + n_2$.



THEOREM 3.9. *Let $A_1, A_2, \ldots, A_n$ be finite disjoint sets. Then*

$$|A_1 \cup A_2 \cup \ldots \cup A_n| = |A_1| + |A_2| + \ldots + |A_n|.$$

EXAMPLE 3.2. Suppose a computer system requires a password which is 6 to 8 characters long, where each character is an uppercase letter or a digit. Each password must contain at least one digit.

How many different passwords are there?

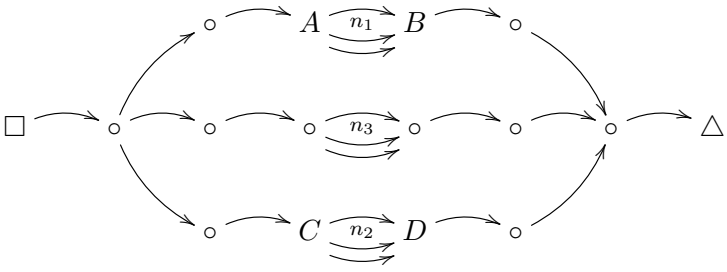
Ans. Let P denotes the total number of passwords. Let P_6, P_7, P_8 denote the number of passwords of length 6, 7, 8, respectively. By sum rule $P = P_6 + P_7 + P_8$.

$$\begin{aligned} P_6 &= |\{\text{ all strings of letters and digits of length } 6\}| - |\{\text{ all strings of letters of length } 6\}| \\ &= \left|\{A, B, \ldots, Z, 0, 1, \ldots, 9\}^6\right| - \left|\{A, B, \ldots, Z\}^6\right| \\ &= (26+10)^6 - 26^6 \\ &= 36^6 - 26^6. \end{aligned}$$

Similarly $P_7 = 36^7 - 26^7$ and $P_8 = 36^8 - 26^8$. Then

$$P = P_6 + P_7 + P_8 = 36^6 - 26^6 + 36^7 - 26^7 + 36^8 - 26^8 = 36^6(1 + 36 + 36^2) - 26^6(1 + 26 + 26^2).$$

3.3. The Inclusion-Exclusion Rule. Suppose $A - B$ and $C - D$ is connected in parellel but n_3 of ways from A to B are common ways from C to D. The number of different ways to go from $\square$ to $\triangle$ is $n_1 + n_2 - n_3$.



THEOREM 3.10. *Let A_1 and A_2 be finite sets. Then*

$$|A_1 \cup A_2| = |A_1| + |A_2| - |A_1 \cap A_2|.$$

EXAMPLE 3.3. What is the number of 8-bit strings starting with 1 or ending with 00. Let $\mathcal{A} = \{(b_1, b_2, \ldots, b_n) \in \mathbb{B}^8 \mid b_1 = 1 \ \vee \ (b_{n-1} = 0 \ \wedge \ b_n = 0)\}$. $\mathcal{A}$ is the set of 8-bit strings starting with 1 or ending with 00.

$|\mathcal{A}| = ?$

Ans. Let $\left|\begin{array}{l} A_{1\ldots} \\ A_{\ldots 00} \\ A_{1\ldots 00} \end{array}\right|$ be the 8-bit strings $\left|\begin{array}{l} \text{starting with 1} \\ \text{ending with 00} \\ \text{starting with 1 and ending with 00} \end{array}\right|$.

Then $|\mathcal{A}| = |A_{1\ldots} \cup A_{\ldots 00}| = |A_{1\ldots}| + |A_{\ldots 00}| - |A_{1\ldots} \cap A_{\ldots 00}|$ where $A_{1\ldots} \cap A_{\ldots 00} = A_{1\ldots 00}$.

$|A_{1\ldots}| = \left|\mathbb{B}^{8-1}\right| = 2^7$, $|A_{\ldots 00}| = \left|\mathbb{B}^{8-2}\right| = 2^6$ and $|A_{1\ldots 00}| = \left|\mathbb{B}^{8-3}\right| = 2^5$.

Therefore $|\mathcal{A}| = |A_{1\ldots}| + |A_{\ldots 00}| - |A_{1\ldots 00}| = 2^7 + 2^6 - 2^5 = 2^5(2^2 + 2^1 - 1)$.

TABLE 1. Summary of Counting Methods

	ordered	unordered
with repetition	n^r	$\binom{n+r-1}{r}$
without repetition	$\mathrm{P}(n,r) = \frac{n!}{(n-r)!} = n^{\underline{r}}$	$\binom{n}{r} = \frac{\mathrm{P}(n,r)}{r!} = \frac{n!}{(n-r)!r!} = \frac{n^{\underline{r}}}{r!}$

THEOREM 3.11. *Let $\mathcal{P}$ be the set of partial functions from A to B. Then $|\mathcal{P}| = (\beta+1)^\alpha$.*

PROOF. Let $\mathcal{B} = B \cup \{ b_0 \}$ where $b_0 \notin B$. Let $p \in \mathcal{P}$. Define set $A_p = \{ a \in A \mid p(a) \text{ is not defined} \}$. Then define function $p^* : A \to \mathcal{B}$ such that

$$p^*(a) = \begin{cases} p(a), & a \in A \backslash A_p, \\ b_0, & a \in A_p. \end{cases}$$

Note that p^* is a function. Then define function f as

$$\begin{aligned} f : \mathcal{P} &\to \mathcal{B}^A \\ p &\mapsto p^* \end{aligned}$$

f is a bijection (proof ?). Hence $|\mathcal{P}| = \left|\mathcal{B}^A\right| = |\mathcal{B}|^{|A|} = (\beta+1)^\alpha$ □

4. The Pigeonhole Principle

THEOREM 4.1 (The Pigeonhole Principle).
If $k+1$ or more objects are placed into k boxes, then there is at least one box containing 2 or more objects.

THEOREM 4.2. *Among any $n+1$ positive integers not exceeding $2n$, there must be an integer that divides one of the other integers.*

THEOREM 4.3 (The Generalized Pigeonhole Principle).
If n objects are placed into k boxes, then there is at least one box containing at least $\lceil n/k \rceil$ objects.

THEOREM 4.4. *Let A and B be finite sets and $f : A \to B$ be a function. If $|A| > |B|$, then f cannot be an injection.*

PROOF. Assume that f is an injection. Then f maps each $a \in A$ into different $b \in B$. Since $|A| > |B|$ by pigeonhole principle, $\exists a_1, a_2 \in A \ [f(a_1) = f(a_2) = b]$ for some $b \in B$. Hence f cannot be injective. □

EXAMPLE 4.1. Let $A_i = (x_i, y_i)$ $i = 1, 2, \ldots, 5$ be a set of five distinct points with integer coordinates in the xy-plane.

Show that the midpoints of the line joining at least one pair of these points have integer coordinates.

Ans. $\left(\frac{x_i+x_j}{2}, \frac{y_i+y_j}{2}\right)$ is the coordinate of the mid point of A_i and A_j. Notice that in order to have its coordinates integer, $x_i + x_j$ and $y_i + y_j$ should be both even. For (x_i, y_i) pair there are four boxes, these are (O,O), (O,E), (E,O), (E,E) where E and O represent even and odd numbers, respectively. Since there are five points, by the pigeonhole principle, two of them should be in the same box. Take these two points. Their mid point has integer coordinates since $x_i + x_j$ and $y_i + y_j$ are even numbers.

5. Counting Methods: Permutation, Combination and Others

Many counting problems can be put into the following form:

There are n objects. Select r of them ($r \leq n$). Then the result depends on the following questions as given in Table 1:

i. Can an item be reselected? That is, is repetition allowed?
ii. Is order of items important?

DEFINITION 5.1. Let A be any finite set. A *permutation* σ of A is a bijection from A to itself.

REMARK 5.1. If $|A| = n$, then it is represented as

$$\sigma = \begin{pmatrix} a_1 & a_2 & \cdots & a_n \\ a_{i_1} & a_{i_2} & \cdots & a_{i_n} \end{pmatrix} = \begin{pmatrix} a_{i_1} & a_{i_2} & \cdots & a_{i_n} \end{pmatrix}$$

where $f(a_j) = a_{i_j}$ with $i, j, i_j \in \{ 1, \ldots, n \}$.

DEFINITION 5.2. The number of distinct subsets with r elements that can be chosen from a set with n elements is called *binomial coefficient*, denoted by $\binom{n}{r}$, and is pronounced "n choose r".

THEOREM 5.1. *For $n, r \in \mathbb{N}$, the binomial coefficients satisfy the following recurrence relation:*

i.

$$\binom{n}{0} = \binom{n}{n} = 1$$

ii.

$$\binom{n}{r} = \binom{n-1}{r} + \binom{n-1}{r-1} \text{ for } 0 < r < n.$$

REMARK 5.2. Note that Theorem 5.1 leads to the famous *Pascal's triangle*.

	1	2	3	4	5	6	7	8	9	10
1	1	1	1	1	1	1	1	1	1	1
2	1	2	3	4	5	6	7	8	9	
3	1	3	6	10	15	21	28	36		
4	1	4	10	20	35	56	84			
5	1	5	15	35	70	126				
6	1	6	21	56	126					
7	1	7	28	84						
8	1	8	36							
9	1	9								
10	1									

The numbers in the first column are the natural numbers. The numbers in the second and the third columns are called the *triangular* and *tetrahedral* numbers, respectively. The nth triangular number is simply the sum of the first n integers. The tetrahedral numbers are the sums of the triangular numbers.

History of binomial coefficients and Pascal's triangle goes way back than Pascal. The Pythagoreans considered the triangular numbers around 540 BC. The Greek mathematicians investigated tetrahedral numbers at the beginning of 200 BC. The binomial numbers as coefficients of $(a+b)^n$ appeared in the works of mathematicians in China around 1100. Hindu mathematicians began to encounter the binomial coefficients in combinatorial problems in 1150s. Pascal puts his triangle in 1665.

THEOREM 5.2. *For $n, r \in \mathbb{N}$ and $r \leq n$,*

$$\binom{n}{r} = \frac{n!}{(n-r)!r!} = \frac{n^{\underline{r}}}{r!}.$$

THEOREM 5.3.

$$\binom{n}{r} = \binom{n}{n-r}.$$

EXAMPLE 5.1. Let A be a set with n elements. We want to count the number of distinct subsets of the set A that have exactly r elements.

EXAMPLE 5.2. Consider $A = \{a, b, c, d, e\}$.

i. How many different words of length 3?
Ans. $5 \times 5 \times 5 = 5^3 = 125$. Note that each $f \in A^{\{1,2,3\}}$ give a different word.

ii. How many different words of length 3 if no letter is allowed to repeat?
Ans. $5 \times 4 \times 3 = 5^{\underline{3}} = 60$. Note that each $f \in A^{\{1,2,3\}}$ where f is an injection gives a different word.

iii. How many different words of length 3 if the order of letters is not important?
Ans. There are $5^{\underline{3}}$ words with order. Each letter combination is counted $3! = 6$ times. So $\frac{5^{\underline{3}}}{3!} = \binom{5}{3} = 10$

iv. How many different words of length 3 if letters are allowed to repeat but the order of the letters is not important?
Ans.

aaa	bbb	ccc	ddd	eee	1234567
3	-	-	-	-	0001111
2	-	1	-	-	0011011
-	1	-	1	1	1011010
2	1	-	-	-	0010111

Notice that problem is equivalent to selecting 3 balls out of 5 different boxes. Hence $\binom{5+3-1}{3}$.

EXAMPLE 5.3. How many bit strings of length n contains exactly r 1's?

Ans. r positions out of n positions are selected and set to 1. The remaining $n-r$ positions are set to 0. The order of r positions is not important. So $\binom{n}{r}$.

EXAMPLE 5.4. How many ways are there for 8 men and 5 women to stand in a line so that no two women stand next to each other, given that all the women are identical, as well as the men?

Ans. There are 9 positions separated by men so that no two women stand next to each other. So $\binom{9}{5} = 126$.

men		1		2		3		4		5		6		7		8	
women	1		2		3		4		5		6		7		8		9

EXAMPLE 5.5.

i. How many solutions does the equation $x_1 + x_2 + x_3 = 11$ have, where $x_1, x_2, x_3 \in \mathbb{N}$?

Ans. Note that since $1+1+1+1+1+1+1+1+1+1+1 = 11$ the pattern 11/1111/11111 which corresponds to 2+4+5 is a solution. So the number of solutions is $\binom{11+3-1}{2}$.

ii. The same question with constraint $x_1 \geq 1, x_2 \geq 2, x_3 \geq 3$?

Ans. Pick 1 of type 1, 2 of type 2, and 3 of type 3. Then there are additional 5 selections out of 3 item types. So $\binom{5+3-1}{5}$.

EXAMPLE 5.6. How many ways are there to place 10 indistinguishable balls into 8 distinguishable bins? (Consider atoms. Electrons are indistinguishable. The energy levels of an atom are distinguishable.)

Ans. Rephrase the problem as 10 balls and 7 separators. So $\binom{10+8-1}{7}$.

EXAMPLE 5.7. Suppose that S is a set with n elements. How many ordered pairs (A, B) are there such that A and B are subsets with $A \subseteq B$?

Ans. $\{ A, B \setminus A, S \setminus B \}$ is a partition of S. So any $a \in S$ should belong to one of them. Since S is finite, one can list its elements as $(s_1, s_2, \ldots, s_n)$. Let

$$b_i = \begin{cases} 0, & s_i \in A, \\ 1, & s_i \in B \setminus A, \\ 2, & s_i \in S \setminus B. \end{cases}$$

Then for any $(b_1, b_2, \ldots, b_n) \in \{0,1,2\}^n$, define $A = \{ s_i \mid b_i = 0 \}$ and $B = \{ s_i \mid b_i = 0 \ \vee \ b_i = 1 \}$, hence $A \subseteq B$. Notice that we define a bijection between the set of ternary n-tuples $\{0,1,2\}^n$) and our set of (A, B) pairs. So the number is equal to the number of ternary n-tuples, that is 3^n.

6. Supplementary Materials

6.1. Some Useful Sequences. A great source of sequences is so called *The On-Line Encyclopedia of Integer Sequences* (OEIS) [**Slo09**]) available at https://oeis.org/. It is a catalog of more than 150000 sequences. It is a great source for combinatorics.

DEFINITION 6.1. The *Stirling number*, *S(n, k)* is the number of ways to partition a set of cardinality n into exactly k nonempty subsets.

DEFINITION 6.2. The nth *Bell number*, B_n is the number of partitions of a set with n members, or equivalently, the number of equivalence relations on it. (Sequence A000110 in [**Slo09**]). It is named in honor of Eric Temple Bell.

6.2. Approximations for $n!$ and $\binom{n}{r}$. There is no closed form of $n!$. Therefore calculating $n!$ for large n is not easy. It takes too much computation. Some approximations are usually used instead.

6.2.1. *Approximations for $n!$.* Stirling provided an approximation for $n!$ in 1730. Stirling's approximation is quite good for $n \gg 1$ [**Rei67, Mac03**]. Even as small values as $n = 10$, the error is less than 1%.

$$\begin{aligned} n! &= 1 \times 2 \times \cdots \times n \\ \ln n! &= \ln 1 + \ln 2 + \cdots + \ln n \\ &\approx \int_1^n \ln x \, \mathrm{d}x = [x \ln x - x]_1^n = n \ln n - n + 1 \\ &\approx n \ln n - n \\ n! &\approx e^{n \ln n - n} = e^{\ln n^n} e^{-n} = n^n e^{-n}. \end{aligned}$$

With the next order of correction, it is:

$$\begin{aligned} n! &\approx n^n e^{-n} \sqrt{2\pi n} \\ \ln n! &\approx n \ln n - n + \frac{1}{2} \ln(2\pi n). \end{aligned}$$

A better approximation is:

$$\sqrt{2\pi n} \left(\frac{n}{e}\right)^n e^{1/(12n+1)} < n! < \sqrt{2\pi n} \left(\frac{n}{e}\right)^n e^{1/12n}.$$

6.2.2. *Approximations for $\binom{n}{r}$.* Using approximations for $n!$, $\binom{n}{r}$ is obtained:

$$\begin{aligned} \binom{n}{r} &= \frac{n!}{(n-r)!r!} \\ \ln \binom{n}{r} &= \ln n! - \ln(n-r)! - \ln r! \\ &\approx n \ln n - n \\ &\quad - ((n-r)\ln(n-r) - (n-r)) \\ &\quad - (r \ln r - r) \\ &= n \ln n - (n-r)\ln(n-r) - r \ln r \\ &= \ln n^n - \ln(n-r)^{(n-r)} \ln r^r \\ &= \ln \frac{n^n}{(n-r)^{(n-r)} r^r}. \end{aligned}$$

$$\begin{aligned} \log_2 \binom{n}{r} &= \log_2 \frac{n^n}{(n-r)^{(n-r)} r^r} \\ &= \log_2 \frac{n^{(n-r)}}{(n-r)^{(n-r)}} \frac{n^r}{r^r} \\ &= \log_2 \left(\frac{n}{n-r}\right)^{n-r} + \log_2 \left(\frac{n}{r}\right)^r \\ &= (n-r) \log_2 \frac{n}{n-r} + r \log_2 \frac{n}{r} \\ &= n \left(\frac{(n-r)}{n} \log_2 \frac{n}{n-r} + \frac{r}{n} \log_2 \frac{n}{r} \right) \\ &= n \left(\frac{(n-r)}{n} \log_2 \frac{1}{\frac{(n-r)}{n}} + \frac{r}{n} \log_2 \frac{1}{\frac{r}{n}} \right) \\ &= n \left(\left(1 - \frac{r}{n}\right) \log_2 \frac{1}{\left(1 - \frac{r}{n}\right)} + \frac{r}{n} \log_2 \frac{1}{\frac{r}{n}} \right) \\ &= n \, \mathrm{H}_2 \left(\frac{r}{n}\right) . \end{aligned}$$

where the *binary entropy function*, $\mathrm{H}_2(x)$, is given as

$$\begin{aligned}\mathrm{H}_2(x) &= -x\log_2 x - (1-x)\log_2(1-x)\\ &= x\log_2\frac{1}{x} + (1-x)\log_2\frac{1}{(1-x)}.\end{aligned}$$

Acknowledgment. These notes are based on various books but especially [**Ros07**].

Problems with Solutions

P 10.1. Suppose that S is a set with n elements. How many ordered pairs (A, B) are there such that A and B are subsets of S with $A \subseteq B$? [Hint: Show that each element of S belongs to $A, B \setminus A$, or $S \setminus B$.]

Solution.
Let A and B subsets of S with $A \subseteq B$.

$$\begin{aligned}&\forall s \in S\ [s \in S] && \text{tautology}\\ \Leftrightarrow &\forall s \in S\ [(s \in B)\ \oplus\ (s \in S - B)] && \text{since } B \subseteq S\\ & && \Rightarrow S = B \cup (S - B)\\ & && \text{and } B \cap (S - B) = \emptyset\\ \Leftrightarrow &\forall s \in S\ [((s \in A)\ \oplus\ (s \in B - A))\ \oplus\ (s \in S - B)] && \text{since } A \subseteq B\\ & && \Rightarrow B = A \cup (B - A)\\ & && \text{and } A \cap (B - A) = \emptyset\end{aligned}$$

Hence, for every $s \in S$ there are three mutually disjoint alternatives: s belongs to either A or $B \setminus A$, or $S \setminus B$. Since the choice for different elements are independent from each other, the result can be found by the rule of product. The number of ordered pairs is

$$\underbrace{3 \cdot 3 \cdot \cdots \cdot 3}_{\text{n times}} = 3^n.$$

P 10.2. Show that a subset of a countable set is also countable.

Solution.

Let $A \subseteq B$. If A is finite by definition it is countable.

Suppose A is infinite. We define a bijection from $\mathbb{N}$ to A as follows: Since B is countable there is a bijection from $\mathbb{N}$ to B. Using this bijection we can make a list of elements of B. Use this list, drop all elements that are not in A. This will be a new list which contains only the elements of A. Define a function $f : \mathbb{N} \to A$ as $f(n) \mapsto a_n$ where a_n is the nth elements of A in the new list. This is a bijection since for any $n \in \mathbb{N}$ there is a unique $a_n \in A$. and vice versa.

P 10.3.

Definition 6.3. A function $f : A \to A$ is said to have a *fixed point* if there exists $x \in A$ such that $f(x) = x$.

Let $A = \{1, 2, \ldots, n\}$ for some $n \in \mathbb{N}$. How many one-to-one functions $f : A \to A$ have at least one fixed point?

Solution.

The number of one-to-one functions, which have at least one fixed point, can be computed by subtracting the number of one-to-one functions, which does not have any fixed points, from the number of all one-to-one functions.

The number of all one-to-one functions $f : A \to A$ is $n!$

A function which does not have any fixed point is a derangement. The number of derangements of a set with n element is D_n where:

$$D_n = n!\left[1 - \frac{1}{1!} + \frac{1}{2!} - \frac{1}{3!} + ... + (-1)^n\frac{1}{n!}\right].$$

Hence the answer is $n! - D_n$

P 10.4. Show that a subset of a countable set is also countable.

Solution.
Let A be a countable set and B be a subset of A.
i) If A is finite, then B should be finite too since $|B| \leq |A|$. Hence B is countable.
ii) If A is not finite, then there exists a bijection $f : A \longrightarrow \mathbb{N}$. If B is finite, it is countable. For infinite B we can list its elements as follows:

$$\begin{aligned} b_1 &= f^{-1}\left(min\{\ f(b) \mid b \in B\ \}\right) \\ b_2 &= f^{-1}\left(min\{\ f(b) \mid b \in B - \{b_1\}\ \}\right) \\ b_3 &= f^{-1}\left(min\{\ f(b) \mid b \in B - \{b_1, b_2\}\ \}\right) \\ ... &= ..... \\ b_n &= f^{-1}\left(min\{\ f(b) \mid b \in B - \{b_1, b_2, ..., b_{n-1}\}\ \}\right) \end{aligned}$$

This is equivalent to saying that we are listing B's elements in the same order as they are listed by f. Hence, B is countable.

CHAPTER 11

Recurrence

1. Motivation

2. Recurrence Equations

Problems with Solutions

P 11.1. Markov chains are powerful models that are often used in Computer Science. In one application of Markov models is the *population dynamics* where there are two types A and B in competition with populations n_A and n_B where the total population $n_A + n_B$ is constant. So if type A increases by one, type B should decrease by one. Then, the state i of the system can be represented by the population of A, that is, $i = n_A$. Hence there are $N + 1$ states represented by $0, 1, \cdots, N$. If the system is in state i, then it moves to state $i - 1$ and $i + 1$ with probabilities $p_{i,i-1}$ and $p_{i,i+1}$, respectively. Then with probability $1 - (p_{i,i-1} + p_{i,i+1})$ it stays in i. The states 0 and N are *absorbing states*. When the system gets in one of these, there is no way to leave them, i.e. $p_{0,1} = p_{N,N-1} = 0$ and $p_{0,0} = p_{N,N} = 1$.

Assuming $p_{i,i-1} = p_{i,i+1} = a$, one obtains the following recurrence equation:

$$\begin{aligned} &x_0 = 0, \\ &x_i = ax_{i+1} + (1 - 2a)x_i + ax_{i-1}, \ \forall i \ 0 < i < N, \\ &x_N = 1. \end{aligned}$$

Solve this recurrence relation.

Solution.
The characteristic equation of $x_i = ax_{i+1} + (1 - 2a)x_i + ax_{i-1}$ is $ar^2 - 2ar + a = 0$. $ar^2 - 2ar + a = a(r^2 - 2r + 1) = a(r - 1)^2$. Since $r = 1$ is the root with multiplicity 2, the solution would be $x_i = (\beta_0 + \beta_1 i)r^i = \beta_0 + \beta_1 i$.
Use boundry conditions to find the values of β_0 and β_1.
$x_0 = 0 = \beta_0 + \beta_1 0 = \beta_0$. So $\beta_0 = 0$.
$x_N = 1 = \beta_0 + \beta_1 N = \beta_1 N$. So $\beta_1 = 1/N$.
Finally, the solution is $x_i = i/N$ for $i = 0, 1, \ldots, N$.

P 11.2. Let $\{ a_n \}$ and $\{ b_n \}$ be two sequences whose terms are coupled as

$$\begin{aligned} a_{n+1} &= \alpha_1 a_n + \beta_1 b_n \\ b_{n+1} &= \alpha_2 a_n + \beta_2 b_n \end{aligned}$$

where $\alpha_1, \alpha_2, \beta_1, \beta_2 \in \mathbb{R}$ and $\Delta = \alpha_1\beta_2 - \alpha_2\beta_1 \neq 0$.

Solve a_n and b_n when $a_0 = C$ and $b_0 = D$.

Solution.
The degenerated case is when $\alpha_2 = 0$ and $\beta_1 = 0$. In this case the sequences are not coupled and they are solved separately. That is, $a_n = C\alpha_1^n$ and $b_n = D\beta_2^n$.

Consider the non-degenarated case. Assume $\beta_1 \neq 0$. Then

$$b_n = \frac{1}{\beta_1}(a_{n+1} - \alpha_1 a_n) \tag{1}$$

$$b_{n+1} = \frac{1}{\beta_1}(a_{n+2} - \alpha_1 a_{n+1}). \tag{2}$$

Substituting b_n and b_{n+1} into the second relation, we obtain

$$\begin{aligned} 0 &= b_{n+1} - \alpha_2 a_n - \beta_2 b_n \\ \Rightarrow 0 &= \frac{1}{\beta_1}(a_{n+2} - \alpha_1 a_{n+1}) - \alpha_2 a_n - \frac{\beta_2}{\beta_1}(a_{n+1} - \alpha_1 a_n) \\ \Rightarrow 0 &= (a_{n+2} - \alpha_1 a_{n+1}) - \beta_1\alpha_2 a_n - \beta_2(a_{n+1} - \alpha_1 a_n) \\ \Rightarrow 0 &= a_{n+2} + (-\alpha_1 - \beta_2)a_{n+1} + (-\alpha_2\beta_1 + \alpha_1\beta_2)a_n \end{aligned}$$

Let $A_1 = -\alpha_1 - \beta_2$ and $A_0 = -\alpha_2\beta_1 + \alpha_1\beta_2$. Then $a_{n+2} + A_1 a_{n+1} + A_0 a_n = 0$. By change of index $a_n + A_1 a_{n-1} + A_0 a_{n-2} = 0$. Use characteristic root technique to solve a_n. Once a_n is obtained as $a_n = f(a_{n-1}, a_{n-2})$, use Eq 1.

Part 6

Graphs

CHAPTER 12

Graphs

1. Introduction

Graph is a very powerful representation used in many disciplines including Computer Science, Management, Physics and of course Mathematics. Interactions, relations of objects are usually represented by a graph.

On the other hand, graphs are used for social entertainment. Remember questions such as "can you draw this without removing your pencil from the paper" as in Fig. 1.

EXAMPLE 1.1. Consider www, web pages and links. A web page has links to other web pages. A web page a can have many links to page b but there may be no link from b to a. Your web page probably have a link to google.com but you would be very lucky if home page of google.com has a link to your web page.

A web page can have a link to an item in itself. Long web pages have internal links to headings in it.

2. Graphs

DEFINITION 2.1 (Multigraph).
A *Multigraph* $G = (V, A, \varphi)$ consists of a nonempty set V of vertices, a set A of arcs and a function $\varphi : A \to V \times V$

REMARK 2.1.
- V is nonempty but A could be empty. So, a single vertex with no arcs is the most simple graph.
- arcs will be denoted by Greek letters.
- arc $\overrightarrow{\alpha} = \overrightarrow{(u, v)}$ if $\varphi(\overrightarrow{\alpha}) = (u, v)$.
- Function φ is in general not an injection. So it is possible that $\varphi(\overrightarrow{\alpha_i}) = \varphi(\overrightarrow{\alpha_j}) = (u, v)$ for $\overrightarrow{\alpha_i} \neq \overrightarrow{\alpha_j}$. Hence, two vertices can be connected more than once, and therefore $\overrightarrow{(u, v)}$ representation of arcs becomes ambiguous.

DEFINITION 2.2 (Multiplicity).
Multiplicity of $(u, v) \triangleq \left|\varphi^{-1}((u, v))\right|$.

DEFINITION 2.3 (Simple Graph).
A *simple graph* is a multigraph $G = (V, A, \varphi)$ such that φ is an injection.

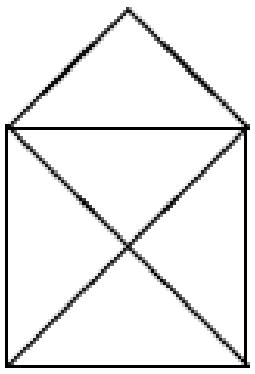


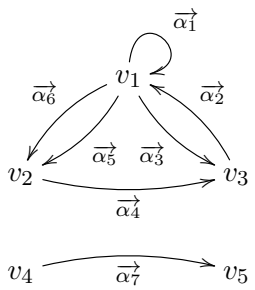
FIGURE 1. Envelope

REMARK 2.2. Multiplicity of any ordered pair (u, v) is at most 1. Therefore (u, v) is sufficient to denote the arc. Hence a simple graph can be represented as $G = (V, \varphi)$ where φ is a relation on V.

REMARK 2.3. A multigraph G can be represented as $G = (V, \varphi)$ where V is the vertex set, $\varphi : V \times V \to \mathbb{N}$ is the function expressing the multiplicity of (v_1, v_2).

EXAMPLE 2.1.

path: $\overrightarrow{\alpha_1\alpha_3\alpha_2}$
simple path: $\overrightarrow{\alpha_4\alpha_2\alpha_1\alpha_5}$
elementary path: $\overrightarrow{\alpha_6\alpha_4}$
circuit: $\overrightarrow{\alpha_3\alpha_2}$
loop: $\overrightarrow{\alpha_1}$



$$\begin{bmatrix} 1 & 2 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

Definition 2.4 (Path).
Let $G = (V, A, \varphi)$ be a multigraph. Given an arc $\overrightarrow{(u,v)} \in A$, u is called the *origin*, v is called *terminus*. A *path* P of G is a sequence of arcs $\overrightarrow{\alpha_0}\overrightarrow{\alpha_1}\dots$ such that for every pair, $\overrightarrow{\alpha_i}, \overrightarrow{\alpha_{i+1}}$, the origin of $\overrightarrow{\alpha_{i+1}}$ is the terminus of $\overrightarrow{\alpha_i}$.

Definition 2.5.
A path which does not traverse the same | arc / vertex | twice is called | a *simple path*. / an *elementary path*. |

Definition 2.6. *Circuit* is a finite path such that the origin of the first one coincides with the terminus of the last.

Definition 2.7. *Simple circuit* is a circuit which is a simple path.

Question 2.1. Define elementary circuit.

Definition 2.8. The number of arcs in a finite path is called the *order* of the path.

Definition 2.9. A circuit of order 1 is called *loop*.

2.1. Reachability and Strong-Connectedness.

Definition 2.10. v is said to be *reachable* from $u \in G = (V, A, \varphi) \overset{\Delta}{\longleftrightarrow} u = v$ or there is a path from u to v.

Definition 2.11. *Simple graph G^* associated with G* is the simple graph obtained by eliminating all but one of the arcs if the multiplicity is more than 1.

Remark 2.4.
i. Connection array M_{G^*} of G^* has only 0s and 1s.
ii. $[M_{G^*}]_{ij} = 1$ means there is an arc from v_i to v_j.

Definition 2.12. *Power of a relation on A.*

$$\rho^k = \begin{cases} \rho\rho^{k-1} & k > 1, \\ \rho & k = 1. \end{cases}$$

Remark 2.5.
- $M_{\rho^{k+1}} = M_\rho M_{\rho^k} = (M_\rho)^{k+1}$
- $\rho\rho^k = \rho^k\rho$

Theorem 2.1. *Let $G = (V, \rho)$ be a simple graph describing ρ.*
There is a path of order n from u to $v \longleftrightarrow (u, v) \in \rho^n$.

Remark 2.6. Given a simple graph $G = (V, \rho)$, the array $(M_\rho)^k$ describes the relation on V: "there is at least one path of order exactly k". In other words, $[(M_\rho)^k]_{ij} = 1 \longleftrightarrow$ there is at least one path of order k from v_i to v_j.

Reachability for $G^* = (V, \rho)$ can be obtained by Algorithm 2 where $|V| = n$.

Algorithm 2: Reachability

```
/* reachability for G* = (V, ρ)                    */
/* where |V| = n                                   */
1 begin
2 |   M(G) ← M_ρ + I
3 |   for k = 2 to n − 1 do
4 |   |   M(G) ← M(G) + M_{ρ^k}
5 |   end
6 end
```

DEFINITION 2.13. G is a multigraph and G^* is the associated simple graph.
The *reachability array* $M(G)$ of G is defined as

$$[M(G)]_{ij} = \begin{cases} 1, & \exists k \text{ for which } [M_{\rho^k}]_{ij} = 1 \text{ where } 1 \le k \le n-1, \\ 0, & \text{otherwise.} \end{cases}$$

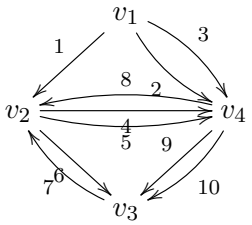
REMARK 2.7. $[M(G)]_{ij} = 1 \iff$ there is a path from v_i to v_j in G.

DEFINITION 2.14. A multigraph is said to be *strongly connected* $\overset{\Delta}{\iff}$ there is a path from v_i to v_j for all $v_i, v_j \in V$.

DEFINITION 2.15. $G = [V, A, \varphi]$ is a multigraph, $\{ V_1, V_2 \}$ dichotomy of V. The set of arcs from vertices of V_1 to vertices of V_2 is called the *cut-set* of G relative to the dichotomy $\{ V_1, V_2 \}$.

EXAMPLE 2.2.

a)Is it strongly connected? No, there is no way to v_1.



$$M_{G^*} = \begin{bmatrix} 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 \end{bmatrix}$$

$$M(G) = M_{G^*}$$

$$M_{G^2} = \begin{bmatrix} 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 \end{bmatrix} \vee \begin{bmatrix} 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 \end{bmatrix} \begin{bmatrix} 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 \end{bmatrix}$$

$$= \begin{bmatrix} 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 \end{bmatrix} \vee \begin{bmatrix} 0 & 1 & 1 & 1 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 0 & 1 & 1 & 1 \\ 0 & 1 & 1 & 1 \\ 0 & 1 & 0 & 1 \\ 0 & 1 & 1 & 1 \end{bmatrix}$$

$$M_{G^3} = \begin{bmatrix} 0 & 1 & 1 & 1 \\ 0 & 1 & 1 & 1 \\ 0 & 1 & 0 & 1 \\ 0 & 1 & 1 & 1 \end{bmatrix} \vee \begin{bmatrix} 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 \end{bmatrix} \begin{bmatrix} 0 & 1 & 1 & 1 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 1 \end{bmatrix}$$

$$= \begin{bmatrix} 0 & 1 & 1 & 1 \\ 0 & 1 & 1 & 1 \\ 0 & 1 & 0 & 1 \\ 0 & 1 & 1 & 1 \end{bmatrix} \vee \begin{bmatrix} 0 & 1 & 1 & 1 \\ 0 & 1 & 0 & 1 \\ 0 & 1 & 1 & 0 \\ 0 & 1 & 1 & 1 \end{bmatrix} = \begin{bmatrix} 0 & 1 & 1 & 1 \\ 0 & 1 & 1 & 1 \\ 0 & 1 & 1 & 1 \\ 0 & 1 & 1 & 1 \end{bmatrix}$$

$$M_{G^*} = \ldots = \begin{bmatrix} 0 & 1 & 1 & 1 \\ 0 & 1 & 1 & 1 \\ 0 & 1 & 1 & 1 \\ 0 & 1 & 1 & 1 \end{bmatrix} \text{ not strongly connected.}$$

b) use $V_0 = \{ \{ v_1, v_4 \}, \{ v_2, v_3 \} \}$ as an example.
Then the cut-set relative to V_0 is $\{ \alpha_1, \alpha_8, \alpha_9, \alpha_{10} \}$.

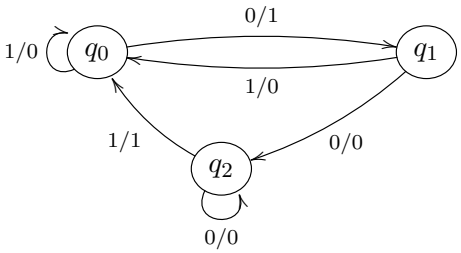
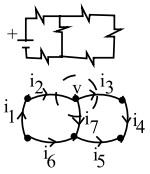


FIGURE 2. State transition diagram

EXAMPLE 2.3.



cut-set at V:
i_2- i_3- i_7= 0

THEOREM 2.2. *$G = [V, A, \varphi]$ is strongly connected $\longleftrightarrow$ for every dichotomy $\{V_1, V_2\}$, the cut-set is non empty.*

1

2.2. Application of Multigraphs.

- Graphs are used to represent state transition of finite state machines Fig. 2.
- Many data structures are represented as graphs Fig. 3, Fig. 4, Fig. 5.
- Some other systems are also represented by graphs Fig. 6.
- Computer networks
- Interconnecting networks
- Parallel architectures
- Graph algorithms

3. Undirected Graphs

If the relation is symmetric:  $\equiv$

DEFINITION 3.1 (Multiple undirected graph).
A *multiple undirected graph* $G = (V, E, \varphi)$ consists of a set V of vertices, a set E of edges and a function φ from E to the set of unordered pairs in V.

NOTATION. If $\varphi(\alpha) =< u, v >$ we write (u, v).

DEFINITION 3.2. For $(u, v) \in E$, u and v are called *terminals*.

DEFINITION 3.3. A *chain* is a sequence of edges $(v_0, v_1)(v_1, v_2)(v_2, v_3) \cdots (v_{n-1}, v_n)$.
A *simple chain* if no edge is repeated.
An *elementary chain* if no vertices is repeated.
A *cycle* if $v_0 = v_n$
A *simple cycle* $\equiv$ simple chain + cycle.

DEFINITION 3.4. $G = (V, E, \varphi)$ is *connected* $\overset{\Delta}{\longleftrightarrow}$ $\forall v_1, v_2 \in V$ [there is a chain between v_1 and v_2].

[1] @HB check

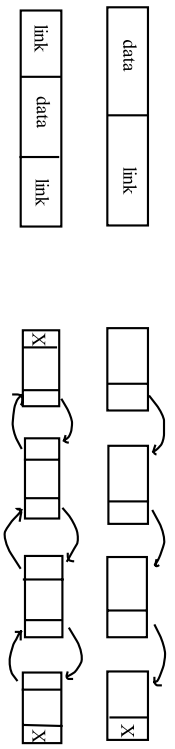


FIGURE 3. Linked list

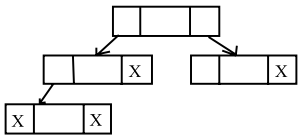


FIGURE 4. B-tree

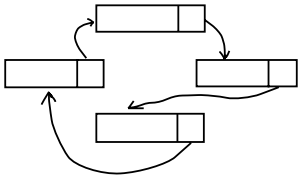


FIGURE 5. Circular queue

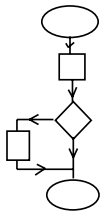


FIGURE 6. Flow chart

DEFINITION 3.5. Let G be an undirected graph. The corresponding *adjoint multigraph*, G^A, is obtained by replacing each edge by a pair of opposite arcs.

DEFINITION 3.6 (Subgraph).
$G' = (V', E', \varphi')$ is called a *subgraph* of $G = (V, E, \varphi)$ if $V' \subseteq V$, $E' \subseteq E$ and E' consists of all the edges in E joining vertices in V'. If E' is a subset of all the edges in E joining vertices in V', then G' is called *partial subgraph*.

DEFINITION 3.7. A *component* G' of G is a connected subgraph such that no vertex in V' is connected to a vertex in $V \setminus V'$ in G.

THEOREM 3.1. *Let G be an undirected multigraph.*
G is connected
$\longleftrightarrow$ G^A is strongly connected.
$\longleftrightarrow$ "something similar to cut-set theorem".
$\longleftrightarrow$ G has only one component.

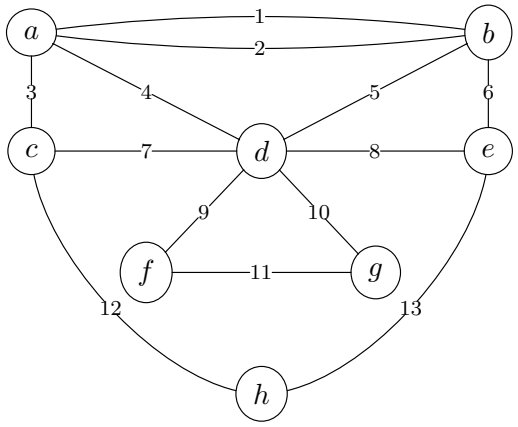
THEOREM 3.2. *A connected undirected simple graph with $|V| = n \geq 1$, must have at least $n - 1$ edges.*

DEFINITION 3.8. Let $G = (V, E, \varphi)$ be connected. A vertex v is called a *cut-point* if the subgraph obtained by deleting it is not connected.

THEOREM 3.3. *v is a cut-point $\longleftrightarrow$ there exist vertices u and w such that every chain connecting u and w passes through v.*

EXAMPLE 3.1.

i. A simple chain which is not an elementary chain: 4, 9, 11, 10, 5.
ii. A simple cycle: 9, 10, 11 or 1, 6, 13, 12, 3.
iii. Is G connected? Yes.
iv. How many components does G have? 1.



EXAMPLE 3.2. Show that a finite graph with n vertices is connected $\longleftrightarrow$ every pair of vertices is connected by a chain of order $\leq n-1$.

$\Rightarrow$ part: $G=[V,E,\phi]$ is connected, then there is a chain between u and v. We need to show that the order of the chain $\leq n-1$. Suppose the order $\geq n$. Then a vertex is repeated in the chain, hence the chain contains a cycle. Remove the cycle. This process can be repeated until the order $\leq n-1$.

$\Leftarrow$ part: If every vertex pair is connected, then by definition G is connected.

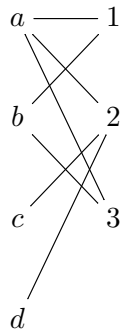
DEFINITION 3.9. A simple undirected graph is said to be *bipartite* $\overset{\Delta}{\longleftrightarrow}$ its vertices can be partitioned into two disjoint sets V_1 and V_2 so that every edge has one terminal vertex in each.

THEOREM 3.4. *G is bipartite $\longleftrightarrow$ there is a cut-set which contains all the edges.*

PROOF. $\Rightarrow$ part: G is a bipartite, then use V_1 and V_2, each edge has one terminal in V_1 and other in V_2. $\Rightarrow$ the cut-set of V_1, V_2 contains all the edges.
$\Leftarrow$ part: All the edges in the cut-set of $V_1, V_2 \Rightarrow$ each edge has one terminal vertex in V_1, the other in V_2. $\Rightarrow$ G is bipartite.

□

REMARK 3.1.



Applications of bipartite graphs
- classical 3 house, 3 utility problem
- optimum matching problems
- heterosexual relations

4. Path Problems

DEFINITION 4.1. For a multigraph G:
| *indegree* / *outdegree* | of a vertex v is the number of axes | terminate / originated | on v. |

DEFINITION 4.2. For an undirected multigraph G, *degree* of a vertex v is the number of edges incident on v.

DEFINITION 4.3. *isolated vertex* is a vertex of degree 0.

DEFINITION 4.4. An | *Eulerian chain* / *Eulerian cycle* | in an undirected multigraph is a | chain / cycle | that uses every edge once and only once.

THEOREM 4.1. *An undirected multigraph without isolated vertices has an Eulerian cycle* $\longleftrightarrow$ *it is connected and contains no vertices of odd degree.*

THEOREM 4.2. $G = [V, E, \varphi]$ *without isolated vertices has an Eulerian chain* $\longleftrightarrow$ *it is connected and contains exactly two vertices of odd degrees.*

DEFINITION 4.5. An | *Hamilton path* / *Hamilton circuit* | in a multigraph is a | path / circuit | which passes through each of the vertices exactly once.

DEFINITION 4.6. A multigraph is *complete* if every pair of vertices is joined by at least one arc.

THEOREM 4.3. *Every complete multigraph contains a Hamiltonian path.*

shortest path algorithm pp77. Read 2.4.1, 2.4.2, 2.4.3

QUESTION 4.1. $G = [V, A, \alpha]$. $d^+(v) = indegree(v), d^-(v) = outdegree(v)$.
Show $\sum_{v \in V} d^+(v) = \sum_{v \in V} d^-(v) = |\, A \,|$.

THEOREM 4.4. *There is even number of vertices which have odd degrees in an undirected multigraph.*

5. Planarity and Coloration

Application. Consider printed circuit boards (PCB) used in electronic devices. Legs of integrated circuits (IC) are electrically connected by means of channels of the PCB. While channel run on PCB, they should not cross each other. If they do, short circuits occur. So drawing channels without crossing each other becomes an importing issue. Very same problem occurs in the integrated circuit production. In this case electronic components such as resistors, transistors are structures on a silicon wafer. There is channels on silicon to connect them electrically. This is problem can be converted to the problem of drawing a graph on a plane or planarity of graphs.

DEFINITION 5.1. A finite undirected multigraph is *planar* if it can be drawn on a plane in such a way that no two of its edges intersect except, possibly, at vertices.

DEFINITION 5.2. An undirected multigraph $G = (V, E, \varphi)$ is said to be *n-colorable* $\overset{\Delta}{\longleftrightarrow} \exists f, f : V \to \{1, 2, \ldots, n\}$ such that if $(u, v) \in E$ then $f(u) \neq f(v)$.

THEOREM 5.1 (The Four-color Theorem).
Every planar graph is 4-colorable.

DEFINITION 5.3. Minimum number n for which an undirected multigraph is n-colorable is called the *chromatic number* of the graph.

THEOREM 5.2. *If the maximum degree of vertices is n, then the chromatic number is less than or equal to* $n + 1$.

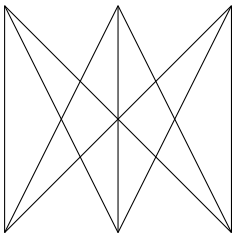
THEOREM 5.3. *An undirected multigraph is 2-colorable* $\longleftrightarrow$ *it contains no cycles of odd length.*

THEOREM 5.4. *A tree is 2-colorable.*

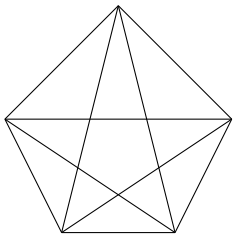
DEFINITION 5.4. Let G be a connected planar graph. A *region* of G is a domain of the plane surrounded by edges of the graph such that any two points in it can be joined by a line not crossing any edge. The edges touching a region contain a simple cycle called the *contour* of the region. Two regions are said to be *adjacent* if the contours of the two regions have at least one edge in common.

THEOREM 5.5 (Euler Formula).
If a connected planar graph has ν vertices, e edges and r regions then $\nu - e + r = 2$.

THEOREM 5.6. *If G is a connected simple graph without loops, and has ν vertices, e edges and r regions,*
then $\frac{3}{2}r \leq e \leq 3\nu - 6$.



$K_{3,3}$



K_5

THEOREM 5.7 (Kuratowski).
An undirected multigraph is planar
$\longleftrightarrow$ it contains no partial subgraphs of either $K_{3,3}$ or K_5.

6. Tree

DEFINITION 6.1. A *tree* is a connected undirected graph with no cycles. A tree of an isolated vertex is called *degenerated tree*.

THEOREM 6.1. *Let $G = [V, E, \alpha]$ and G is a nondegenerated tree.*
$\longleftrightarrow$ every pair of vertices is connected by one and only one chain
$\longleftrightarrow$ G is connected but deletion of an edge makes it disconnected
$\longleftrightarrow$ G has no cycles and if an edge is added, one and only one cycle is formed.

THEOREM 6.2. *A nondegenerated tree contains at least two vertices of degree 1.*

THEOREM 6.3. *$G = [V, E, \alpha]$ with $|\, V \,| = n \geq 1$*
G is a tree
$\longleftrightarrow$ G contains no cycle and has n-1 edges
$\longleftrightarrow$ G is connected and has n-1 edges

DEFINITION 6.2 (Spanning Tree).
A *spanning tree* of a connected undirected graph $G = [V, E, \alpha]$ is a tree $T = [V, E', \alpha]$ where $E' \subseteq E$ and α' is the restriction of α to E'.

REMARK 6.1. There may be many spanning trees for G.

DEFINITION 6.3. A *minimal spanning tree* is a spanning tree which has minimum number of edges.[2]

Algorithm 3: The Minimal Spanning Tree Algorithm

```
1 begin
  |  /* minimal spanning tree algorithm here                    */
2 end
```

6.1. Minimal Spanning Tree Algorithm.

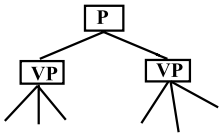
[2]@HB correct this

6.2. Rooted Tree.

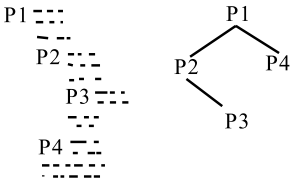
DEFINITION 6.4. A *rooted tree* R is a directed graph obtained by specifying as the *root* a special vertex v and each chain between v and some u is replaced by a path from v to u. The order of the path from v to u is called the *level* of u. For every arc $(\overrightarrow{u,w})$, u and w are a predecessor-successor pair. Any vertex whose outdegree is 0 is called a *leaf*.

REMARK 6.2. Rooted trees are classical representations for hierarchical structures.

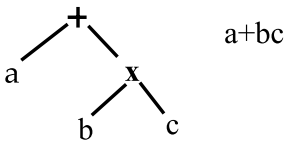
- organization charts



- procedures in programming languages



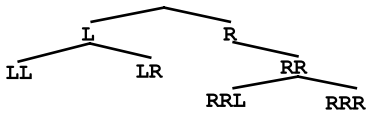
- algebra of commutative operations



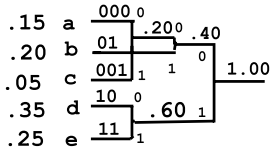
- scope of variables in procedural programming languages

DEFINITION 6.5. An *oriented rooted tree* is a rooted tree such that the set of arcs issuing from any vertex is an ordered set.

Representation



EXAMPLE 6.1. *Huffmann coding* is a coding technique which is optimum in mean coding length.



Acknowledgment. These notes are based on various books but especially [**PY73, Ros07, TZ82, Gal89**].

Problems with Solutions

P 12.1. Let $G = [V, E]$ be a simple, undirected and connected graph which does not contain any self-loops. Prove that G is a bipartite graph if and only if G does not contain any cycle of odd length.

Solution.
Let $G = [V, E]$ be a loop-free simple, undirected, connected and bipartite graph with $V = V_1 \cup V_2$.

Let $C = \{(v_1, v_2), (v_2, v_3), (v_3, v_4), \ldots, (v_n, v_1), \}$ be a cycle in G. Let $v_1 \in V_1$. (The proof for $v_1 \in V_2$ is similar.) Since G is a bipartite graph v_i and v_{i+1} must belong to different sets V_1 and V_2. Hence, the sequence $v_1 - v_2, \ldots, v_n - v_1$ is an alternating sequence between the edges of V_1 and V_2. For this sequence to start and end with the same vertex there must be odd number of vertices in this sequence. Hence, the number of edges on C must be even.

($\Leftarrow$ Part) Let $G = [V, E]$ be a loop-free simple, undirected and connected graph with no cycles of odd length. Let $x \in V$, and
$V_1 = \{v \in V \mid$ the length of a shortest path between x and v is odd$\}$ and
$V_2 = \{w \in V \mid$ the length of a shortest path between x and w is even$\}$.

Note that
i) $x \in V_2$
ii) $V_1 \cap V_2 = \emptyset$
iii) $V_1 \cup V_2 = V$.

Claim: each edge $\{a, b\} \in E$ has one vertex in V_1 and the other vertex in V_2.

To prove this claim suppose that there exists an edge $e = \{a, b\} \in E$ with $a \neq b$ and $a, b \in V_1$. (The proof for $a, b \in V_2$ is similar) Let $E_a = \{\{a, v_1\}, \{v_1, v_2\}, \ldots, \{v_{m-1}, x\}\}$ be the m edges in a shortest path from a to x and let $E_b = \{\{b, {v_1}'\}, \{{v_1}', {v_2}'\}, \ldots, \{{v_{n-1}}', x\}\}$ be the n edges in a shortest path from b to x. m and n are both odd since $a, b \in V_1$.
If $\{v_1, v_2, \ldots, v_{m-1}\} \cap \{{v_1}', {v_2}', \ldots, {v_{m-1}}'\} = \emptyset$, then the set of edges $C_1 = \{\{a, b\}\} \cup E_a \cup E_b$ is a cycle of odd length in G. Otherwise, let $w(\neq x)$ be the first vertex where the paths come together and let

$$C_2 = \{\{a, b\}\} \cup \{\{a, v_1\}, \{v_1, v_2\}, \ldots, \{v_i, w\}\} \cup \{\{b, {v_1}'\}, \{{v_1}', {v_2}'\}, \ldots, \{{v_j}', w\}\}$$

for some $1 \leq i \leq m-1$ and $1 \leq j \leq n-1$. Then either C_2 or $C_1 - C_2$ is a cycle of odd-length in G.

P 12.2. Suppose $d_1, d_2, \ldots, d_n \in \mathbb{Z}^+$ with $\sum_{i=1}^n d_i = 2n - 2$. Show that there is a tree that has n vertices with degree sequence $d_1, d_2, \ldots, d_n$.

Solution.
Note that $\forall i \ [d_i > 0]$. Use induction on the number of vertices n.

Case $n = 1$. Since $2n - 2 = 0$, the degree sequence is $d_1 = 0$ only. A degenerated tree with one vertex satisfies the proposition. Note that $d_1 = 0$ does not actually satisfy the rule that $d_i > 0$. So the question should have one more condition such as $n > 1$. Hence induction should start from $n = 2$.

Induction Base. $n = 2$. Since $2n - 2 = 2$, the degree sequence can only be $d_1 = 1, d_2 = 1$. A tree with 2 vertices connected by an edge satisfies it.

Induction. Assume that it is true for $n \geq 2$. Then show that is must be true for $n + 1$.

Consider a degree sequence $d_1, d_2, \ldots, , d_n, d_{n+1}$. If we can show that there exist vertices v_k of degree $d_k = 1$ and v_ℓ of degree $d_\ell > 1$, then we can construct such tree as follows: Reindex $d_1, d_2, \ldots, d_{n-1}, d_n, d_{n+1}$ so that we have $d_n > 1$ and $d_{n+1} = 1$. Hence, $d_n - 1 > 0$. By induction hypotesis, there exists a corresponding tree T with n vertices for the degree sequence $d_1, d_2, \ldots, d_{n-1}, d_n - 1$. Obtain a new tree T' by add vertex v_{n+1} to T by connecting v_{n+1} to v_n. The degree sequence of T' becomes $d_1, d_2, \ldots, , d_n - 1 + 1, d_{n+1} = 1$.

Existence of $d_k = 1$. Suppose $\forall i \in \{1, 2, \ldots, n\} \ [d_i > 1]$. Then $d_i \geq 2$, so $\sum_{i=1}^n d_i \geq \sum_{i=1}^n 2 = 2n > 2n - 2$. Since $\sum_{i=1}^n d_i = 2n - 2$, there must be at lease one vertex with $d_k < 2$, that is $d_k = 1$.

Existence of $d_\ell > 1$. Suppose $\forall i \in \{1, 2, \ldots, n\} \ [d_i = 1]$. Then $\sum_{i=1}^n d_i = \sum_{i=1}^n 1 = n < 2n - 2$. Since $\sum_{i=1}^n d_i = 2n - 2$, There must be at lease one vertex with $d_k > 1$.

Bibliography

[Apo] Tom M. Apostol. *Introduction to Analytic Number Theory*. Springer-Verlag.

[Gal89] Steven Galovich. *Introduction to Mathematical Structures*. Harcourt Brace Jovanovich, Inc., 1989.

[GKP98] Ronald L. Graham, Donald E. Knuth, and Oren Patashnik. *Concrete Mathematics: A Foundation for Computer Science*. Addison-Wesley, 1998.

[Hol60] Paul Holmos. *Naive Set Theory*. Van Nastrand, 1960.

[LP98] Rudolf Lidl and Gunter Pilz. *Applied Abstract Algebra*. Springer, 1998.

[Mac03] David MacKay. *Information Theory,Inference, and Learning*. Cambridge University Press, 2003.

[Men08] Eliott Mendelson. *Number Systems and the Foundation of Analysis*. Dover, 2008.

[Nes09] Ali Nesin. *Mathematige Giris: III. Sayma*. Nesin Yayincilik, 2009.

[PY73] Franco P. Preparata and Raymond T. Yeh. *Introduction to Discrete Structures*. Addison-Wesley, 1973.

[Rei67] Frederick Reif. *Berkeley Physics: Statistical Physics*. McGraw-Hill, 1967.

[Ros07] Kenneth H. Rosen. *Discrete Mathematics and its Applications*. McGraw-Hill, 2007.

[Slo09] N. J. A. Sloane. The On-Line Encyclopedia of Integer Sequences (OEIS), 2009.

[Str94] Steven H. Strogatz. *Nonlinear Dynamics and Chaos*. Perseus Books Publishing, 1994.

[TZ82] Gaisi Takeuti and Wilson M. Zaring. *Introduction to Axiomatic Set Theory*. Springer-Verlag, 1982.

[Wik09] Wikipedia. Mersenne Prime, 2009.

.

The Notation Index

The Concepts Index