

# **Discrete Mathematics**

**F. Oggier**

Division of Mathematical Sciences, Nanyang Technological University,  
Singapore

These notes were prepared for the course MH1812, Discrete Mathematics, a course given for the 1st year students of the school of computer engineering at Nanyang Technological University. It replaces the previous course CE/CZ1001, Discrete Mathematics.

*Disclaimer. Since this is the first version of the notes, it is highly likely that typos and other glitches may still be found...this may be fixed during the next iteration of this course...*

These notes (and exercises) were made based on the material prepared by the previous lecturers of this course: I used the pdf slides of Leong Peng Chor for the topics ranging from complex number to graph theory, and the slides and exercises of Anwitaman Datta for the topics ranging from elementary number theory to linear recurrences.

I am particularly thankful to Anwitaman Datta who further shared with me the source slides for the topics mentioned, but also some source slides on relations, functions and graphs, coming from another course he taught. I have prepared my own slides based on his pptx slides.

I further used the following notes and slides found online: notes by Richard Hammack, who has written nice notes on proofs by contradiction, slides by Brody Dylan Johson, on The Hat Problem, very clear (and humorous) notes by Srini Devadas and Eric Lehman, on the Hanoi Tower Problem, slides by Niloufar Shafiei on Solving Linear Recurrence Relations, and notes by Alfred Brousseau on Second-Order Linear Recursion Relations.

*Copyright. I do not owe the copyright of most figures used in these notes, whenever available, copyright is acknowledged on the slides themselves. All the figures were used for a purely educational purpose.*

# Contents

1 Elementary Number Theory	5
2 Propositional Logic	33
3 Predicate Logic	77
4 Set Theory	107
5 Combinatorics	127
6 Linear Recurrences	145
7 Complex Numbers	161
8 Linear Algebra	175
9 Relations	209
10 Functions	231
11 Graph Theory	251
12 Solutions to the Exercises	267



# Chapter 1

## Elementary Number Theory

*“Without mathematics, there’s nothing you can do. Everything around you is mathematics. Everything around you is numbers”.*  
*(S. Devi)*

The topic of this first chapter is elementary number theory, that is the study of numbers and some of their basic properties.

**Definition 1.** **Natural numbers** are numbers used for counting, that is

$$1, 2, 3, 4, 5, 6, \dots$$

Whether 0 belongs to natural numbers is often a matter of convention. We will refer to **whole numbers** to indicate that the zero should be included:

$$0, 1, 2, 3, 4, 5, 6, \dots$$

The set of natural numbers is often denoted by  $\mathbb{N}$ :

$$\mathbb{N} = \{1, 2, 3, 4, 5, 6, \dots\}.$$

**Definition 2.** **Integer numbers** are natural numbers, with their opposites (natural numbers with a negative sign), and zero, that is

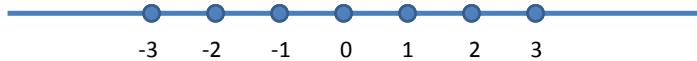
$$\dots, -6, -5, -4, -3, -2, -1, 0, 1, 2, 3, 4, 5, 6, \dots$$

The set of integer numbers is often denoted by  $\mathbb{Z}$ :

$$\mathbb{Z} = \{\dots, -6, -5, -4, -3, -2, -1, 0, 1, 2, 3, 4, 5, 6, \dots\}.$$

## Integer and Real Numbers

- **Natural numbers ( $\mathbb{N}$ )** are the counting numbers:  
1, 2, 3, ...  
– Sometimes 0 is also included (**whole numbers**)
- **Integers ( $\mathbb{Z}$ )** are the natural numbers, including zero, and their negatives:  
... -3, -2, -1, 0, 1, 2, 3, ...
- **Real numbers ( $\mathbb{R}$ )** are any value on the continuous line.  
– for example: 0.31, -4,  $\pi$ , 2



## Ir(rational) Numbers

- **Rational numbers ( $\mathbb{Q}$ )** are real numbers which can be represented in the form  $a/b$ , where  $a$  and  $b$  are integers  
– for example:  $3/7$ ,  $0.999 = 999/1000$ , ...
- **Irrational numbers ( $\mathbb{I}$ )** are real numbers which can NOT be represented in the form  $a/b$  for any integers  $a$  &  $b$   
– for example:  $\pi$ ,  $e$ ,  $2^{1/2}$

Real numbers are difficult to define formally. For us, we will say that **real numbers** are values on the continuous real line. Real numbers include for example  $\sqrt{2}$ , or  $\pi$ , but also  $-3$  or  $0.5$ .

**Definition 3.** **Rational numbers** are real numbers that can be written as the ratio of two integer numbers. The set  $\mathbb{Q}$  of rational numbers is formally defined as

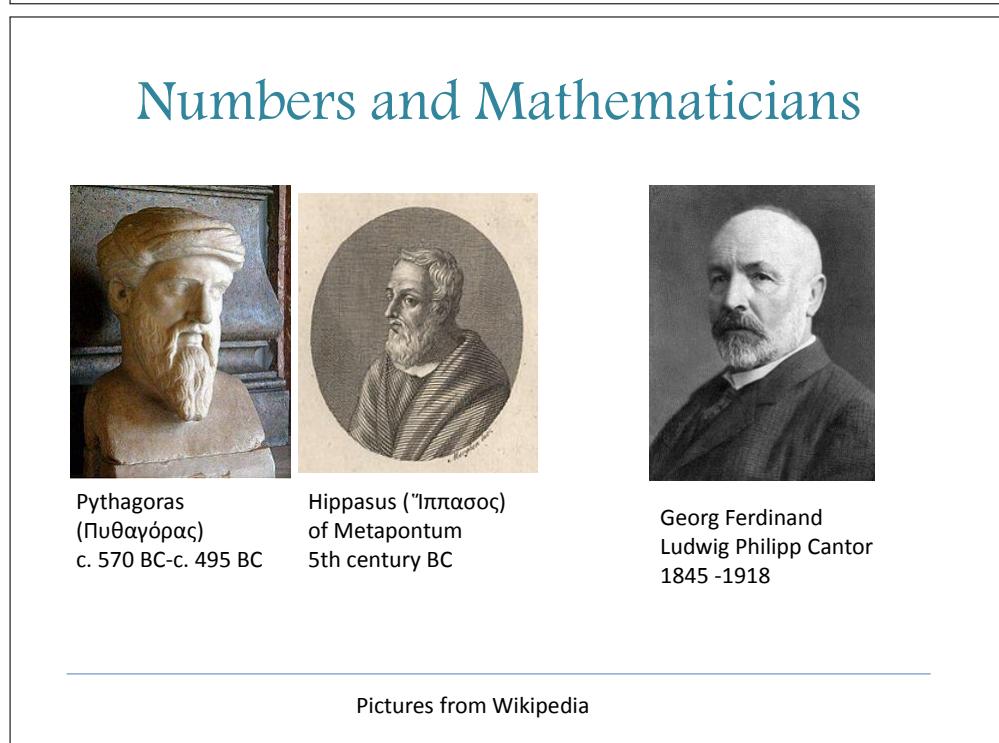
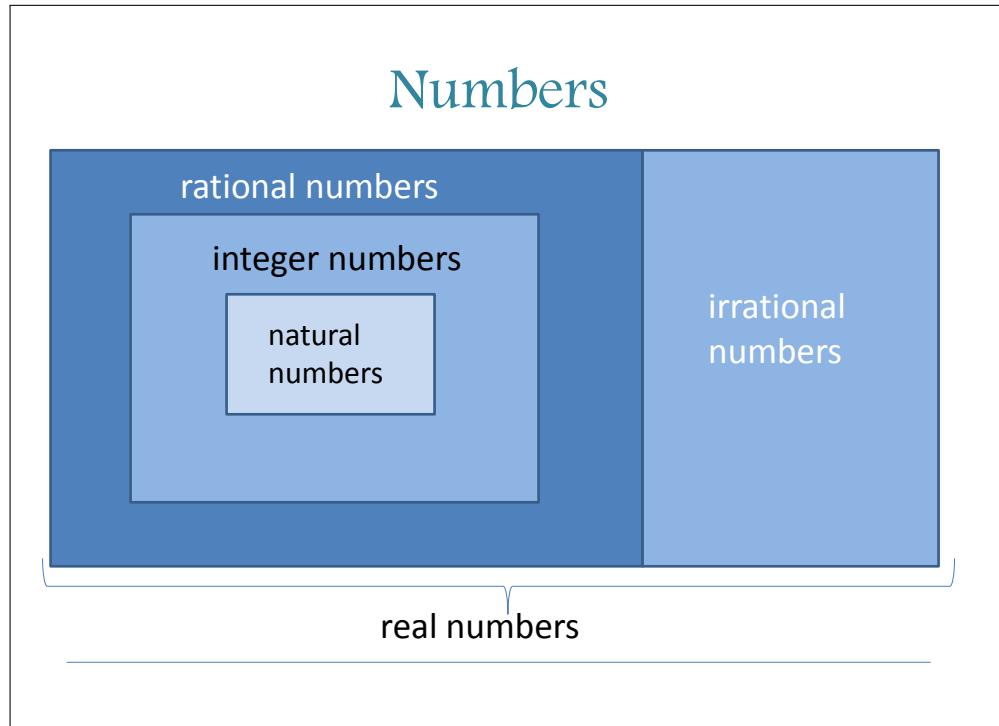
$$\mathbb{Q} = \left\{ \frac{a}{b}, a, b \in \mathbb{Z}, b \neq 0 \right\}.$$

**Definition 4.** **Irrational numbers** are real numbers that **cannot** be written as the ratio of two integer numbers.

It is easier to show that a number is rational than to show that it is irrational. Indeed, to show that a number is rational, it is enough to find two integers  $a$  and  $b$  such that  $a/b$  is your number. For example,  $0.5$  is rational, because if you take  $a = 1$  and  $b = 2$ , then  $a/b = 1/2 = 0.5$ . There are other choices of  $a$  and  $b$ . For example, you could have taken  $a = 2$  and  $b = 4$ , then  $a/b = 2/4 = 1/2 = 0.5$ . However, to show that a number is irrational, you need to show that no matter which choice of integers  $a$  and  $b$  you take, then  $a/b$  will never be your number (see Exercise 3 for an example).

We already saw 5 types of numbers: natural, integer, real, rational, and irrational. To remember them, it may be useful to think about how these types of numbers are connected.

- Real numbers form the largest set of numbers that we will encounter for many chapters. All 4 other sets of numbers are included in  $\mathbb{R}$ .
- Both rational numbers ( $\mathbb{Q}$ ) and irrational numbers are real numbers. These two sets are disjoint, that is there is no number which is both rational and irrational! Either a number is a ratio of two integers, or it is not. One way to help in remembering the term “rational” is to think of a “ratio” of two integers.
- The set  $\mathbb{Z}$  of integer numbers is included in  $\mathbb{Q}$ . Indeed, take any integer number  $a$  (positive, negative, or 0), then  $a = a/1$ , thus  $a$  is always the ratio of two integers.
- The set  $\mathbb{N}$  is included in  $\mathbb{Z}$ .



Mathematicians have been continuously studying numbers. The notion of irrationality is as old as 5th century BC. The school of Pythagoreans (Greek philosophers who followed the mathematician and philosopher Pythagoras) believed that all numbers were rationals. The legend has it that Hippasus, who is believed to have discovered the irrationality of some numbers, may have been drowned, either to be prevented to report this finding, or to punish him for it! From the Pythagoreans time, it took more than 20 centuries to mathematicians to come up with a formal definition for real numbers, which is due to the mathematician Cantor. Numbers are still studied by mathematicians nowadays, in a branch of mathematics called number theory.

Numbers have also been of great interest for computer scientists. While it is fairly easy for a computer to represent an integer number (not "too big") in its memory, it is much more complicated for real numbers in general, and in fact, many real numbers need to be approximated, leading to potential numerical problems. Rational numbers are much nicer, exactly because they can be represented by a pair of integers. Some programming languages (such as C), will require you to handle *types*, and thus to identify which numbers you are working with.

What we call now Euclidean division was discovered after irrational numbers.

**Definition 5.** The **Euclidean division** states that for any positive integer  $n$  and any integer  $m$ , there exist unique integers  $q$  (called **quotient**) and  $r$  (called **remainder**) such that

$$m = qn + r, \quad 0 \leq r < n.$$

Pay attention to the condition on  $r$ , it is important. It tells that  $r$  is strictly smaller than the number  $n$  with which we divide  $m$ . When the remainder  $r = 0$ , then we say that  **$n$  divides  $m$** , also written

$$n \mid m.$$

Alternatively, we may say that  **$n$  is divisible by  $m$** .

## Numbers and Computer Science

```
gap>
gap>
gap> 10/3;
10/3
gap> 10.0/3;
3.33333
gap>
```

- Real numbers need **approximation!**
- Rational numbers = pair of integers
- **Type** of numbers (e.g. in C)

```
frac.c ✘
#include <stdio.h>

void main()
{
    float a;
    a=10/3;
    printf("%f\n",a);
    a=(float)10/3;
    printf("%f\n",a);
}
```

```
frederique@frederique-desktop:~$ gcc frac.c -o frac
frederique@frederique-desktop:~$ ./frac
3.000000
3.333333
frederique@frederique-desktop:~$
```

## Euclidean Division

- Take any integer  $n$ ,  $n > 0$ , and any integer  $m$ . There exist unique integers  $q$  and  $r$  such that

$$m = qn + r, \quad 0 \leq r < n.$$

- $q$  = **quotient**,  $r$  = **remainder**
- When  $r=0$ , then
  - $n$  **divides**  $m$ , or
  - $m$  **is divisible by**  $n$ .
  - Notation:  $n \mid m$



Euclide (Εὐκλείδης)  
300 BC

Picture from Wikipedia

**Example 1.**

If  $n = 7$  and  $m = 17$ , then we want to write  $17 = 7q + r$ , with  $0 \leq r < 7$ . We get

$$17 = 7 \cdot 2 + 3, \quad q = 2, \quad r = 3.$$

If  $n = 7$  and  $m = 35$ , then we want to write  $35 = 7q + r$ , with  $0 \leq r < 7$ . We get

$$35 = 7 \cdot 5 + 0, \quad q = 5, \quad r = 0.$$

Thus  $7 \mid 35$ , in words, 7 divides 35.

What happens if  $m$  is a negative number? Here is an example. If  $n = 7$  and  $m = -5$ , then we want to write  $-5 = 7q + r$ , with  $0 \leq r < 7$ . We get

$$-5 = 7(-1) + 3, \quad q = -1, \quad r = 2.$$

Thanks to the notion of divisibility, we may define two new types of numbers.

**Definition 6.** A **prime number**, often denoted by  $p$ , is a natural number divisible by exactly two factors: 1 and itself.

We thus consider that 1 is not a prime number, because 1 has only one divisor, namely 1. The first prime numbers are

$$2, 3, 5, 7, \dots$$

But 9 is not a prime number, because  $1 \mid 9$ ,  $3 \mid 9$  and  $9 \mid 9$ , thus there are 3 factors, 1, 3, and 9 which divide 9.

**Definition 7.** An **even number** is an integer number divisible by 2. An **odd number** is an integer number which is not divisible by 2.

If  $a$  is an even number, we may formally write it as  $a = 2a'$  for  $a'$  some integer, to emphasize that 2 appears in its factorization.

For example, 4 is even. Among the prime numbers, only 2 is even! (see Exercise 1). Then 9 is odd, because we just saw that its factors are 1, 3 and 9, 2 is not a factor of 9.

## Euclidean Division: Examples

- Take any integer  $n$ ,  $n > 0$ , and any integer  $m$ . There exist unique integers  $q$  and  $r$  such that

$$m = qn + r, \quad 0 \leq r < n.$$

- E.g.  $n=7$ ,
  - If  $m = 17$ ,  $q = 2$ ,  $r = 3$
  - If  $m = 35$ ,  $q = 5$ ,  $r = 0$
  - If  $m = -5$ ,  $q = -1$ ,  $r = 2$

## Prime Numbers, Even Numbers

- **Prime numbers** are natural numbers  $p$ , which have only two factors:  $p$  and 1, i.e., *not divisible* by any other integer
  - For it to have two factors, it has to be larger than 1
  - 2, 3, 5, 7, 11, 13, ...
- **Even numbers** are integers divisible by 2. **Odd numbers** are integers not divisible by 2.

The notion of divisibility is useful to define a new notion, that of *modulo*.

**Definition 8.** For a positive integer  $n$ , two integers  $a$  and  $b$  are said to be **congruent  $(\text{mod } n)$**  if  $a - b$  is an integer multiple of  $n$ . We use the notation:

$$a \equiv b \pmod{n}.$$

The notation  $\equiv$  emphasizes that we do not have a “normal” equality here.

However,  $a \equiv b \pmod{n}$  can be turned into an equality as we see next. We write that  $a - b$  is an integer multiple of  $n$  as  $a - b = qn$  for some integer  $q$ , then

$$a = qn + b.$$

Let us see examples, where we will further comment on the link with the Euclidean division introduced above.

### Example 2.

Take  $a = -8$ , which we want to compute modulo  $n = 5$ . Then the Euclidean division tells us that we can write  $a = -8 = q5 + b$  with a  $b$  such that  $0 \leq b < 5$ . We get

$$-8 = 5(-2) + 2, \quad b = 2.$$

Thus  $-8 \equiv 2 \pmod{5}$ . Now in the definition of modulo  $n$ , we are allowed to take  $b$  bigger, so we could write

$$-8 = 5(-3) + 7, \quad b = 7$$

and further obtain that  $-8 \equiv 2 \equiv 7 \pmod{5}$ . Unlike for the Euclidean division, infinitely many choices of  $b$  are possible.

Suppose that we want next to compute  $17 \pmod{5}$ . “Informally”, we are allowed to add and subtract multiples of 5. Thus if one removes  $3 \cdot 5$  to 17, we get 2, if one removes  $1 \cdot 5$  to 17, we get 12, and if one adds  $1 \cdot 5$ , this gives 22. To summarize

$$17 \equiv 2 \equiv 12 \equiv 22 \pmod{5}.$$

Given an integer  $a$ , there are many (infinitely many)  $b$  such that  $a \equiv b \pmod{n}$ . However, using the Euclidean division, there is only one  $b$  between 0 and  $n - 1$ . This is why we can represent **integers  $(\text{mod } n)$**  by the set

$$\{0, 1, \dots, n - 1\}.$$

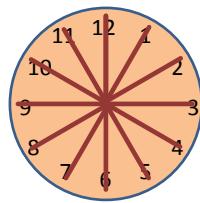
## Modulo n

- For a positive integer  $n$ , two integers  $a$  and  $b$  are said to be **congruent modulo  $n$** , if  $a - b$  is an integer multiple of  $n$ . We write  
$$a \equiv b \pmod{n}$$
- If  $a \equiv b \pmod{n}$ , then  $a - b = qn$  and  $a = qn + b$ .

## Modulo n: Examples

$$a \equiv b \pmod{n} \Leftrightarrow a = qn + b$$

- $-8 \equiv 2 \equiv 7 \pmod{5}$
- $17 \equiv 2 \equiv 12 \pmod{5}$
- Likewise  $17 \equiv 22 \pmod{5}$



Now that we have a new set of numbers, that of integers modulo  $n$ , one may wonder whether it is possible to compute with these numbers. And the answer is yes! It is possible to add two numbers modulo  $n$ , and to multiply them.

Let us try out with addition. Suppose that we have  $17 \pmod{5}$  and  $-8 \pmod{5}$ , and we want to add them up:

$$(7 \pmod{5}) + (-8 \pmod{5}) = ?$$

Shall we first compute  $7 \pmod{5} = 2$ , and  $-8 \pmod{5} = 2$ , and then add them to find that it is  $4 \pmod{5}$ ? Or should we first compute  $7 + (-8) = -1$ , and then compute  $-1 \pmod{5}$ ? It turns out that it does not matter (see Exercise 4), neither does it matter for multiplication, which is why computations are possible: they are consistent.

We can summarize how to compute modulo  $n$  using an [addition table](#) or a [multiplication table](#). Let us try out with integers modulo 2, and addition:

$+$	0	1
0	0	1
1	1	0

We notice that this table is “almost normal”, but for the fact that  $1+1=0$ ! Similarly, we get for integers modulo 2 and multiplication:

*	0	1
0	0	0
1	0	1

## Modular Arithmetic

$$a \equiv b \pmod{n} \leftrightarrow a = qn + b$$

- Integers mod  $n$  can be represented as elements between 0 and  $n-1$ :  $\{0, 1, 2, \dots, n-1\}$
- Define **addition mod  $n$**  and **multiplication mod  $n$** :
 
$$(a \pmod{n}) + (b \pmod{n}) \equiv (a+b) \pmod{n}$$

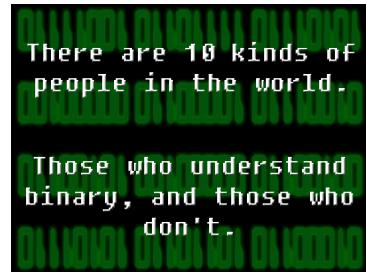
$$(a \pmod{n}) * (b \pmod{n}) \equiv (a*b) \pmod{n}$$
- Examples
  - $(17 \pmod{5}) + (-8 \pmod{5}) \equiv 4 \pmod{5}$
  - $(12 \pmod{5}) * (-3 \pmod{5}) \equiv 4 \pmod{5}$

## Integers mod 2

- Bits are integer modulo 2

+	0	1
0	0	1
1	1	0

*	0	1
0	0	0
1	0	1



We have seen many sets of numbers, all of them coming with “a way to compute”, that is to perform either addition, or multiplication, sometimes division. Consider more generally any set  $S$ , with any operator  $\Delta$ .

**Definition 9.** Consider a set  $S$  with an operator  $\Delta$ . Then  $S$  is [closed under  \$\Delta\$](#)  if the result of the operation  $\Delta$  on any two elements of  $S$  results in an element of  $S$ . This property is known as [closure](#).

**Example 3.** • If  $S = \mathbb{R}$ , the set of real numbers, and  $\Delta = +$  is the addition, then the addition of two real numbers results in another real number. Thus  $S$  is closed under addition.

- If  $S = \mathbb{Z}$ , the set of integer numbers, and  $\Delta$  is the division, then  $S$  is not closed under  $\Delta$ . Indeed, divide 1 by 2, both 1 and 2 are integer numbers, but  $1/2$  is not!
- The above definition can be applied to integers modulo  $n$  as well. Take  $S$  to be the set of integers modulo  $n$ , that is  $S = \{0, 1, \dots, n - 1\}$ , and  $\Delta$  is the addition modulo  $n$ , then  $S$  is closed under  $\Delta$ .

Why do we care about the closure property? Well, we usually care, because we like structure! A set as such is very generic, once we start adding an operator on it, and this set turns out to be closed under this operator, then we are gaining more structure! This becomes handy for example when one wants to write an algorithm on a set.

## Operator Closure

- Consider a set  $S$  with an operator  $\Delta$ . Then  $S$  is **closed under  $\Delta$**  if the result of the operation  $\Delta$  on any two elements of  $S$  results in an element of  $S$ .
  - This is known as the *closure* property.
- Examples:
  - $S=R=\{\text{real numbers}\}$  is closed under  $\Delta=+$  (and  $\Delta=*$ ).
  - $S=Z=\{\text{integer numbers}\}$  is not closed under  $\Delta=\text{division}$ .
  - $S=\{\text{integers mod } n\}$  is closed under  $\Delta=\text{addition mod } n$ .

---

## Summary

- Recognize different types of numbers (**natural, integer, real, rational, irrational, prime, even, modulo n**)
- Decide whether these sets of numbers are **closed** under a given operator.




---

Calvin & Hobbes belong to Bill Watterson

## Exercises for Chapter 1

**Exercise 1.** Show that 2 is the only prime number which is even.

**Exercise 2.** Show that if  $n^2$  is even, then  $n$  is even, for  $n$  an integer.

**Exercise 3.** The goal of this exercise is to show that  $\sqrt{2}$  is irrational. We provide a step by step way of doing so.

1. Suppose by contradiction that  $\sqrt{2}$  is rational, that is  $\sqrt{2} = \frac{m}{n}$ , for  $m$  and  $n$  integers with no common factor. Show that  $m$  has to be even, that is  $m = 2k$ .
2. Compute  $m^2$ , and deduce that  $n$  has to be even too, a contradiction.

**Exercise 4.** This exercise is optional, it requires to write things quite formally. Show the following two properties of integers modulo  $n$ :

1.  $(a \bmod n) + (b \bmod n) \equiv (a + b) \bmod n$ .
2.  $(a \bmod n)(b \bmod n) \equiv (a \cdot b) \bmod n$ .

**Exercise 5.** Compute the addition table and the multiplication tables for integers modulo 4.

**Exercise 6.** Show that  $\frac{p(p+1)}{2} \equiv 0 \pmod{p}$  for  $p$  an odd prime.

**Exercise 7.** Consider the following sets  $S$ , with respective operator  $\Delta$ .

- Let  $S$  be the set of rational numbers, and  $\Delta$  be the multiplication. Is  $S$  closed under  $\Delta$ ? Justify your answer.
- Let  $S$  be the set of natural numbers, and  $\Delta$  be the subtraction. Is  $S$  closed under  $\Delta$ ? Justify your answer.
- Let  $S$  be the set of irrational numbers, and  $\Delta$  be the addition. Is  $S$  closed under  $\Delta$ ? Justify your answer.

## Examples for Chapter 1

A first application of integers modulo 2 is counting in binary. In order to represent a natural number, we may write it in terms of power of 2:

$$2^0 = 1, \quad 2^1 = 2, \quad 2^2 = 4, \quad 2^3 = 8, \quad 2^4 = 16, \dots$$

Suppose that a number  $n$  may be written as

$$n = n_0 2^0 + n_1 2^1 + n_2 2^2 + \cdots + n_r 2^r$$

for some positive integer  $r$ . Then we may write  $n$  in binary as

$$n = (n_r \ n_{r-1} \ \cdots \ n_1 \ n_1 \ n_0).$$

Now all the  $n_i$ ,  $i = 0, \dots, n_r$  are either 0 or 1!

**Example 4.** Take  $n = 18$ . Then

$$\begin{aligned} 18 &= 16 + 2 \\ &= 2^4 + 2^1 \end{aligned}$$

and

$$18 = (10010).$$

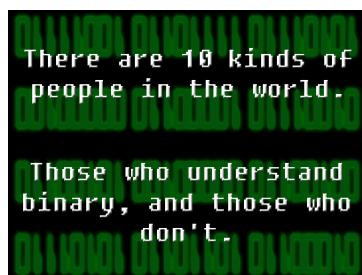
Our main application for binary arithmetic will be coming from storage. In order to explain it, we need the notion of binary vector. A vector in general is an array containing numbers (this array is either horizontal or vertical). For example, here are a row (horizontal) and a column (vertical) vector:

$$(a_1, a_2, \dots, a_n), \quad \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix}.$$

A binary vector contains only 0 and 1. Its length is the number of 0 and 1 it contains. For example  $(1, 0, 0, 1)$  is a binary vector of length 4.

## Counting in Binary

$$\begin{aligned}10010_2 &= 0*2^0 + 1*2^1 + 0*2^2 + 0*2^3 + 1*2^4 \\&= 18_{10} \\&= 8 \cdot 10^0 + 1 \cdot 10^1\end{aligned}$$



## Binary Vectors

- A **vector** is a row (or column) array containing numbers.
- $(1,0,0,1)$  is a binary row vector (of length 4)
- One can add two binary vectors component wise (vector addition).

One may add two vectors componentwise:

$$(a_1, a_2, \dots, a_n) + (b_1, b_2, \dots, b_n) = (a_1 + b_1, a_2 + b_2, \dots, a_n + b_n).$$

**Example 5.** We have

$$(1, 0, 0, 1) + (0, 0, 1, 1) = (1, 0, 1, 0)$$

since  $1 + 0 = 1$ ,  $0 + 0 = 0$ ,  $0 + 1 = 1$  and  $1 + 1 = 0$ .

We have seen in this chapter the notion of closure of a set  $S$  under an operator  $\Delta$  (see Definition 9). Let us see how to apply it here.

**Example 6.** Let  $S$  be the set of binary vectors of length  $n$ , and  $\Delta$  be the vector addition. Is  $S$  closed under  $\Delta$ ? To check this, we need to make sure that the addition of any two binary vectors will result in an element of  $S$ , that is a binary vector of length  $n$ . Suppose then that  $(a_1, a_2, \dots, a_n)$  and  $(b_1, b_2, \dots, b_n)$  are both in  $S$ . Then

$$(a_1, a_2, \dots, a_n) + (b_1, b_2, \dots, b_n) = (a_1 + b_1, a_2 + b_2, \dots, a_n + b_n),$$

and since  $a_i + b_i$  is either 0 or 1, then  $(a_1 + b_1, a_2 + b_2, \dots, a_n + b_n)$  is in  $S$ , thus  $S$  is closed under  $\Delta$ .

Note that binary vector addition is different from what we do to count in binary.

**Example 7.** We continue Example 4. We saw that

$$18 = (10010) = (n_4, n_3, n_2, n_1, n_0).$$

How do we count  $18 + 18$  in binary, using that  $18 = (10010)$ ? We start from  $n_0 = 0$ , and thus  $n_0 + n_0 = 0$ . Thus

$$18 + 18 = (*, 0).$$

We look at  $n_1$  next. We have  $n_1 + n_1 = 1 + 1 = 0$ , however here, it means that  $n_1 2^1 + n_1 2^1 = 2^1(n_1 + n_1)$  thus we do get a 0 in the second position from the right, but there is a 1 which is carried over!

$$18 + 18 = (*, 0, 0).$$

## Binary Vectors: Example

- For example  $(1,0,0,1)+(0,0,1,1)=(1,0,1,0)$
  - $S=\text{set of binary vectors of length } n, \Delta=\text{vector addition}$ . Is  $S$  closed under  $\Delta$ ?
  - Different from counting in binary!
- 

## Binary in the Real World

- Storage of data across multiple hard disks
- Data is in binary format.
- Data needs to be stored so as to tolerate disk failures.



---

Image belongs to Microsoft

To compute the next term, we have  $n_2 + n_2 = 0$ , but there is the 1 carried over, so we get

$$18 + 18 = (*, 1, 0, 0).$$

Then  $n_3 + n_3 = 0$ , and  $n_4 + n_4 = 0$  with again a 1 carried over, to finally get

$$18 + 18 = (1, 0, 0, 1, 0, 0) = 1 \cdot 2^5 + 1 \cdot 2^2 = 32 + 4$$

as it should be.

Now that we have seen how to perform vector additions (which is different from adding two numbers written in binary), let us go back to our storage application.

The engineering problem that we are facing is that of storing data across several hard disks, and in fact, data is stored in binary (so you can imagine the stored data as a binary vector). Now if your data is stored only on one disk, if this disk suffers from a failure, you are likely to lose your data. To prevent this to happen, typically you will store your data on two hard disks, each disk will contain one copy.

**Example 8.** Suppose you want to store 200GB of data, and the shop is selling disks of 100 GB each. Then you can buy 4 disks, store half of your data (let us call it  $D_1$ ) on disk 1, the other half (say  $D_2$ ) on disk 2, then copy the content of disk 1 to disk 3, and the content of disk 2 to disk 4. We get thus the following data allocation:

$$\text{disk 1 : } D_1, \text{ disk 2 : } D_2, \text{ disk 3 : } D_1, \text{ disk 4 : } D_2.$$

- **Good:** what is good now is that even if any of the 4 disks fails, your data is still safe. Note though that only one disk failure is ok, two failures is not, because if it happens to both disk 1 and disk 3 for example, then your data is lost.
- **Bad:** what is bad is that you did pay for 4 disks, while you only need 2 for the actual data...but then with only 2 disks, any failure will mean that your data is lost...

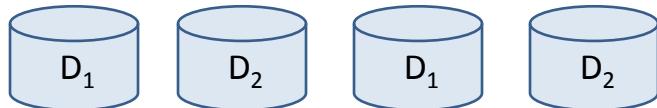
So there is a trade-off here: either you pay for 2 disks and tolerate no failure, or you pay for 4, but still can tolerate one failure. Is there any way to get a better trade-off?

Since 2 disks give no protection against failures, we need at least 3 disks. The question is then, can we have only 3 disks, yet tolerate any one of the 3 disks to fail?

## Example (I)

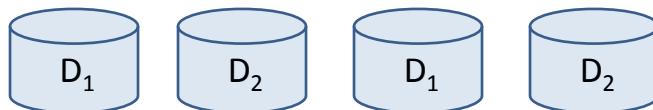
- Suppose you want to store 200GB of (binary) data
- Option 1: buy 4 disks of 100 GB each, store 2 copies of your data.

$$D = (D_1, D_2)$$



## Example (II)

$$D = (D_1, D_2)$$



- Good thing: if one hard disk fails, your data is safe.
- Bad thing: you paid for 4 hard disks instead of 2.
- Can we think of a better solution?

If you put your data  $D_1$  and  $D_2$  in each of the first 2 disks

$$\text{disk 1 : } D_1, \text{ disk 2 : } D_2, \text{ disk 3 : ?}$$

then putting either  $D_1$  or  $D_2$  in the 3rd disk will not help. So what else could we do? We could use binary operations on  $D_1$  and  $D_2$ , namely compute the sum mod 2 (also called XOR) of  $D_1$  and  $D_2$ :

$$\text{disk 1 : } D_1, \text{ disk 2 : } D_2, \text{ disk 3 : } D_1 + D_2.$$

Let us make sure we understand what this means, by first looking at 3 bits,  $a, b, a + b$ , that is  $a, b, a + b$  are three integers mod 2. Now you have three cases:

1. suppose you lose  $a$ : then you have  $b$  and  $a + b$ , compute  $b + (a + b) = a$ .
2. suppose you lose  $b$ : then you have  $a$  and  $a + b$ , compute  $a + (a + b) = b$ .
3. suppose you lose  $a + b$ : then you have  $a$  and  $b$ .

In all cases, both  $a$  and  $b$  can be recovered from one missing bit.

We thus do the same thing with disks (drives), except that now, disks contain *binary vectors* instead of just one bit. Let us do an example, with  $D_1 = (0, 1, 1, 0, 1, 1, 0, 1)$ ,  $D_2 = (1, 1, 0, 1, 0, 1, 0, 0)$ :

$$\text{disk 1 : } D_1, \text{ disk 2 : } D_2, \text{ disk 3 : } D_1 + D_2 = ?.$$

We compute the binary vector addition of  $D_1$  and  $D_2$ :

$$D_3 = D_1 + D_2 = (1, 0, 1, 1, 1, 0, 0, 1).$$

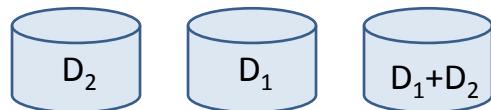
This strategy is thus better than the previous one: you are paying for one disk less (three hard disks instead of 4), and you still can tolerate any one disk failure.

### Example (III)

- Suppose  $a$  and  $b$  are bits, and take  
 $a, b, a+b$
- Do the same thing with disks

$a$	$b$	$a+b$
0	0	0
0	1	1
1	0	1
1	1	0

$$D = (D_1, D_2)$$



### Example (IV): Parity (RAID)

- Binary vector addition:

Drive 1: 01101101  
 Drive 2: 11010100

01101101
<b>XOR</b> 11010100
10111001

Store in Drive 3

You can still lose one disk, but paid for only 3.

**Example 9.** Let us try another storage example. You have 150 GB of data to store, and you are buying hard disks, each of 50GB storage capacity each. But unlike in the previous example where you were happy with one disk failure protection, this time you would like protection against any 2 failures!

Let us start, you have 150 GB of data, and disks of size 50 GB, so you split your data into 3 equal parts  $D_1, D_2, D_3$ :

$$\text{disk 1 : } D_1, \text{ disk 2 : } D_2, \text{ disk 3 : } D_3.$$

Here is now a *key observation*: to tolerate any 2 failures, each  $D_i, i = 1, 2, 3$  must appear at least 3 times! (if any appears only twice, when the corresponding 2 disks fail, then the data is lost...)

Can we manage something clever with only 4 disks?

$$\text{disk 1 : } D_1, \text{ disk 2 : } D_2, \text{ disk 3 : } D_3, \text{ disk 4 : ?}.$$

It is not possible to add only one piece of data in disk 4, and have  $D_1, D_2, D_3$  present at least 3 times. Let us try with 5 disks:

$$\text{disk 1 : } D_1, \text{ disk 2 : } D_2, \text{ disk 3 : } D_3, \text{ disk 4 : ?}, \text{ disk 5 : ?}.$$

The only way to have  $D_1, D_2, D_3$  present at least 3 times is to put  $D_1+D_2+D_3$  both in disk 4 and disk 5. Let us see whether you are protected against 2 failures. If  $D_1$  and  $D_2$  fail, you are left with

$$D_3, D_1 + D_2 + D_3.$$

It is only possible to recover  $D_1 + D_2$ ...so we have to move to 6 disks:

$$\text{disk 1 : } D_1, \text{ disk 2 : } D_2, \text{ disk 3 : } D_3, \text{ disk 4 : ?}, \text{ disk 5 : ?}, \text{ disk 6 : ?}.$$

Now we have space to put  $D_1, D_2, D_3$  such that they are each present at least 3 times. We are the possible choices of data to be put in your empty disks?

$$D_1 + D_2, D_1 + D_3, D_2 + D_3, D_1 + D_2 + D_3.$$

Let us try:

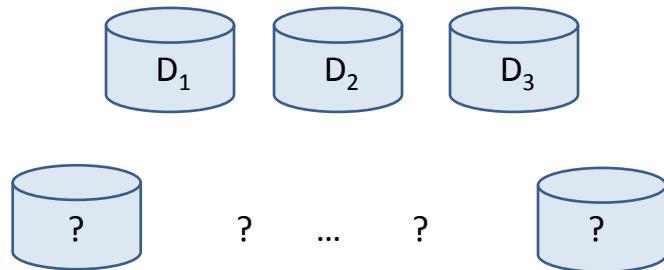
$$\begin{array}{lll} \text{disk 1 : } D_1, & \text{disk 2 : } D_2, & \text{disk 3 : } D_3, \\ \text{disk 4 : } D_1 + D_2, & \text{disk 5 : } D_1 + D_3, & \text{disk 6 : } D_2 + D_3. \end{array}$$

## Another Example (I)

- You want to store 150 GB data.
  - Now you are buying storage devices, each of 50GB capacity.
  - This time, even if any arbitrary two devices fail, you still want to recover all your data!
- 

## Another Example (II)

$$D = (D_1, D_2, D_3)$$



We can achieve what we want with a total of 9 disks. Can we do better?

---

Let us first look at the case where two of the  $D_i$  are gone:

1. If we lose  $D_1, D_2$ : we are left with only  $D_3$ , but  $D_3$  with disks 5 and 6 works.
2. If we lose  $D_1, D_3$ : we are left with only  $D_2$ , but  $D_2$  with disks 4 and 6 works.
3. If we lose  $D_2, D_3$ : we are left with only  $D_1$ , but  $D_1$  with disks 5 and 5 works.

We recall the data allocation for convenience:

$$\begin{array}{lll} \text{disk 1 : } D_1, & \text{disk 2 : } D_2, & \text{disk 3 : } D_3, \\ \text{disk 4 : } D_1 + D_2, & \text{disk 5 : } D_1 + D_3, & \text{disk 6 : } D_2 + D_3. \end{array}$$

We look next at the case where one of the  $D_i$  is gone, and the other one is either disk 4, 5, or 6:

1. If we lose  $D_1$ : either disk 4 or 5 can be used.
2. If we lose  $D_2$ : either disk 4 or 6 can be used.
3. If we lose  $D_3$ : either disk 5 or 6 can be used.

So this strategy works! It needs 6 hard disks, which is less costly than making 3 copies of the data, which would have needed 9 disks...

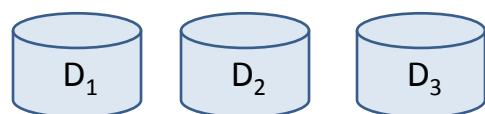
It is not the only strategy that works, you may come up with your solution, which is different than this one, and also works!

If you are wondering whether this way of thinking can be used for other examples, yes it can, as long as the numbers of disks are not too big, after which it becomes more tiresome...

If you are wondering whether there is a trick to make things faster, yes, there is one, it is called *the Hamming distance*, but it is beyond the scope of this course, so feel free to check it out if you are curious, but the notion of Hamming distance **is not needed** for this course.

### Another Example (III)

$$D = (D_1, D_2, D_3)$$



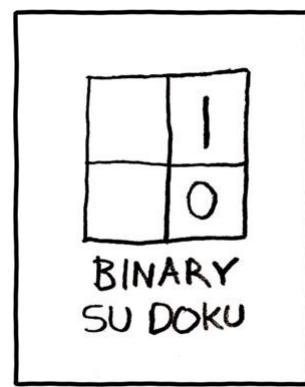
$$\begin{aligned} & D_1 + D_2 \\ & D_1 + D_3 \\ & D_2 + D_3 \\ & D_1 + D_2 + D_3 \end{aligned}$$



To tolerate two failures, we need each  $D_i$  to be present at least 3 times.

### Summary

- Modulo n
- Binary arithmetic
  - Binary vectors
  - Counting in binary
- Applications of binary arithmetic
  - storage





# Chapter 2

## Propositional Logic

*“Contrariwise,’ continued Tweedledee, ‘if it was so, it might be; and if it were so, it would be; but as it isn’t, it ain’t. That’s logic”.*  
*(Lewis Carroll, Alice’s Adventures in Wonderland and Through the Looking-Glass)*

This chapter is dedicated to one type of logic, called *propositional logic*. The word *logic* refers to the use and study of valid reasoning. Logic contains rules and techniques to formalize statements, to make them precise. Logic is studied by philosophers, mathematicians and computer scientists. Logic appears in different areas of computer science, such as programming, circuits, artificial intelligence and databases.

It is useful to represent knowledge precisely and to help extract information. This last sentence may not be clear at this point, but hopefully it will become once we progress.

The term *propositional logic* thus refers to a logic which relies on *propositions*, which is defined as follows:

**Definition 10.** A *proposition* is a statement either true or false, but not both.

**Example 10.**

1. The statement ” $1 + 1 > 3$ ” is false, while the statement ” $5 > 3$ ” is true. Both statements are propositions.
2. The statement ”What a great book!” is not a proposition. Someone is expressing an opinion.

## Logic

- Accepted rules for making **precise** statements.
  
  
  
  - Logic for computer science: programming, artificial intelligence, logic circuits, database.
  
  
  
  - Logic
    - Represents **knowledge precisely**
    - Helps to **extract information** (inference)
- 

## Proposition

A statement that is **either true or false** but not **both** is called **a proposition**.

- Examples of propositions
    - “ $1 + 1 = 2$ ” ⋯ True
    - “ $1 + 1 > 3$ ” ⋯ False
    - “Singapore is in Europe.” ⋯ False
  - Examples (which are not propositions)
    - “ $1 + 1 > x$ ” ⋯ **x**
    - “What a great book!” ⋯ **x**
    - “Is Singapore in Asia?” ⋯ **x**
- 

```
gap> (5>3);
true
gap> (1>3);
false
gap>
```

We just said that a proposition is a statement which is either true or false. We next give a definition for a statement which cannot be assigned a truth value.

**Definition 11.** A *paradox* is a statement that cannot be assigned a truth value.

Thus a paradox cannot be a proposition!

**Example 11.** Here is a paradox called the *liar paradox*: “This statement is false”. Suppose that “This statement is false” takes a true value, then it must be that the statement is false, but then if “this statement is false” is false, then the statement is true, and we can iterate this process which will never lead to any conclusion.

To formalize statements, we will use *symbols*, and these symbols will have the same truth values as the statements. For example, if we consider the proposition “ $1 + 1 > 5$ ”, we will denote it by  $p$ :  $p = “1 + 1 > 5”$ . Since “ $1 + 1 > 5$ ” is false,  $p$  will take the truth value false ( $F$ ).

Sometimes, we will use *logical operators* to combine statements. Here are three basic operators that we will discuss into details next:

- $\wedge$  conjunction (and)
- $\vee$  disjunction (or)
- $\neg$  negation (not)

The  $\neg$  operator is sometimes denoted  $\sim$  instead.

We are particularly interested in combining propositions (statements that either true or false).

**Definition 12.** A *compound proposition* is a statement obtained by combining propositions with logical operators.

Let us start with the negation operator:

$\neg$  negation (not) ].

We will use a *truth table* to describe how  $\neg$  operates on a proposition  $p$ :

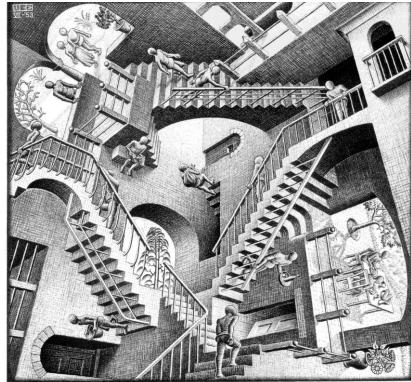
$p$	$\neg p$
$T$	$F$
$F$	$T$

The way the truth table is read is: on the first row, if  $p = T$ , then  $\neg p = F$ , on the second row, if  $p = F$ , then  $\neg p = T$ .

## Paradox

A **paradox** is a statement that *cannot be assigned a truth value*.

- A paradox is not a proposition.
- Example: the liar paradox  
“This statement is false”



Art Work by Escher ("Relativity")

## Symbolic Logic

- Use **symbols** to represent statements (both have *the same truth values*)
- Use **logical operators** to combine statements
  - Compound propositions = propositions combined with logical operator(s)
- Three basic operators
  - $\wedge$ : conjunction (**and**)
  - $\vee$ : disjunction (**or**)
  - $\neg$ : negation (**not**, alternatively  $\sim$ )

Here is an example: if  $p$  = “you shall pass”,  $\neg p$  = “you shall not pass”. Next we have the disjunction operator:

 $\vee$  disjunction (or)

It is an operator that combines two propositions  $p$  and  $q$  as described in the following truth table:

$p$	$q$	$p \vee q$	$q \vee p$
$T$	$T$	$T$	$T$
$T$	$F$	$T$	$T$
$F$	$T$	$T$	$T$
$F$	$F$	$F$	$F$

The disjunction operator returns  $T$  when at least one of the two propositions  $p, q$  is true. This is why whenever there is at least one  $T$  on a row,  $p \vee q$  is true. We notice that  $p \vee q$  and  $q \vee p$  have the same truth table. In this case, we will say that the compound propositions  $p \vee q$  and  $q \vee p$  are [equivalent propositions](#). We also say that the operator  $\vee$  is [commutative](#).

Finally we have the conjunction operator:

 $\wedge$  conjunction (and)

It is an operator that combines two propositions  $p$  and  $q$  as described in the following truth table:

$p$	$q$	$p \wedge q$	$q \wedge p$
$T$	$T$	$T$	$T$
$T$	$F$	$F$	$F$
$F$	$T$	$F$	$F$
$F$	$F$	$F$	$F$

This time, for  $p \wedge q$  to be true, we need both  $p$  and  $q$  to be true, which is why there is only one row for which  $p \wedge q$  is true. Note that  $\wedge$  is also commutative.

We now have three operators  $\neg, \vee, \wedge$ , that we can apply on propositions  $p, q$ . Our first big result of this chapter is called [De Morgan’s Law](#), and states connections between these three operators.

**Theorem 1** (De Morgan’s Law.). *We have*

$$\neg(p \wedge q) \equiv \neg p \vee \neg q, \quad \neg(p \vee q) \equiv \neg p \wedge \neg q.$$

## Negation $\neg$

- Negation (not) of p:  $\neg p$  ( $\sim p$  is also used)

p	$\neg p$
T	F
F	T

← Truth Table

- p: You shall pass
- $\neg p$ : You shall not pass



Picture from the movie Lord of the Rings

## Disjunction $\vee$

- Disjunction (or) of p with q:  $p \vee q$

p	q	$p \vee q$	q $\vee$ p
T	T	T	T
T	F	T	T
F	T	T	T
F	F	F	F

True when  
‘at least one’  
of them is  
true

- $p \vee q \equiv q \vee p$ , i.e. operator  $\vee$  commutes

↑ means “equivalent”

```
gap>
gap> (5>3) or (1>5);
true
gap>
```

## Conjunction $\wedge$

- Conjunction (and) of p with q:  $p \wedge q$

p	q	$p \wedge q$
T	T	T
T	F	F
F	T	F
F	F	F

True only when  
'both' of them  
are true

- $\wedge$  is also commutative:  $p \wedge q \equiv q \wedge p$

```
gap> (5>3) and (7>5);
true
gap>
gap>
gap> (5>3) and (1>5);
false
```

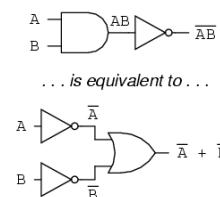
## De Morgan's Law

$$\begin{aligned}\neg(p \wedge q) &\equiv \neg p \vee \neg q \\ \neg(p \vee q) &\equiv \neg p \wedge \neg q\end{aligned}$$



Augustus De Morgan  
(1806-1871)

p	q	$\neg p$	$\neg q$	$p \wedge q$	$\neg(p \wedge q)$	$\neg p \vee \neg q$
T	T	F	F	T	F	F
T	F	F	T	F	T	T
F	T	T	F	F	T	T
F	F	T	T	F	T	T



$$\overline{AB} = \overline{A} + \overline{B}$$

Picture from Wikipedia

## Contradiction

A statement that is always false is called a **contradiction**.

Example: This course is easy  
 'and' this course is not easy  
 $p \wedge (\neg p) \equiv F$



© Mike Baldwin, www.cartoonstock.com

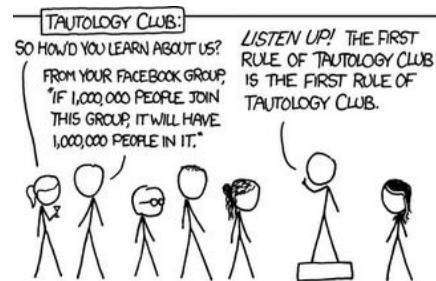
## Tautology

An expression that always gives a true value is called a **tautology**.

Example:  $p \vee (\neg p) \equiv T$

$p$	$\neg p$	$p \vee \neg p$
T	F	T
F	T	T

Always  
true!



Somewhat similar to “Head I win, Tail you lose”

© xkcd

*Proof.* The proof consists of computing the truth table. We need to show that the truth table for  $\neg(p \wedge q)$  and  $\neg p \vee \neg q$  are the same, which will say that both compound propositions are equivalent. Let us compute in details the first row of the truth table for  $\neg(p \wedge q)$  and  $\neg p \vee \neg q$ : if both  $p$  and  $q$  are true, then  $\neg p$  and  $\neg q$  are false, and  $p \wedge q$  is true, since it is true exactly when both  $p$  and  $q$  are true. Then  $\neg(p \wedge q)$  is the negation of  $p \wedge q$ , that is  $\neg(p \wedge q)$  is false. We only need to show that  $\neg p \vee \neg q$  is false too, which is the case, since both  $\neg p$  and  $\neg q$  are false. The other rows are filled in similarly. The proof of  $\neg(p \vee q) \equiv \neg p \wedge \neg q$  is done in Exercise 9.  $\square$

So far, we have seen two types of statements: (1) a proposition, which is a statement either always true, or always false, and (2) a paradox, which is a statement whose truth value cannot be assigned. Here are two new types of statements:

**Definition 13.** A contradiction is a statement that is always false.

**Example 12.** A typical example of a contradiction is  $p \wedge (\neg p)$ . Why? recall that  $p \wedge q$  takes the truth value true only when both  $p$  and  $q$  are true at the same time...but it is not possible for both  $p$  and  $\neg p$  to be true both at the same time.

If a statement which is always false is called a contradiction, you may wonder whether there is a name for a statement which is always true?

**Definition 14.** A tautology is a statement that always gives a true value.

**Example 13.** For an example of contradiction, we choose  $\wedge$ , because  $p \wedge q$  is true only when both  $p$  and  $q$  is true. Now we may similarly get an example of tautology. Pick  $\vee$  instead. For  $p \vee q$  to be true, it is enough that either  $p$  or  $q$  is true. Thus  $p \vee (\neg p)$  is a tautology.

We already said that two statements are *equivalent* if their truth tables are the same. Here is an example of three statements which are equivalent:

$$\neg h \wedge \neg b \equiv \neg b \wedge \neg h \equiv \neg(b \vee h).$$

We notice that the first equivalence follows from the commutativity of  $\wedge$ . For the second equivalence, this is one of De Morgan law! (which can be seen through a truth table).

## Equivalent Expressions

Consider the following three statements

- Alice is not married but Bob is not single  
 $\neg h \wedge \neg b$
  - Bob is not single and Alice is not married  
 $\neg b \wedge \neg h$
  - Neither Bob is single nor Alice is married  
 $\neg(b \vee h)$
  - These three statements are equivalent  
 $\neg h \wedge \neg b \equiv \neg b \wedge \neg h \equiv \neg(b \vee h)$
- 

## Equivalent Expressions

The three statements

$$\neg h \wedge \neg b \equiv \neg b \wedge \neg h \equiv \neg(b \vee h)$$

are equivalent.

$$b \ h \ \neg b \ \neg h \ b \vee h \ (\neg h \wedge \neg b) \ (\neg b \wedge \neg h) \ \neg(b \vee h)$$

TT	F	F	T	F	F	F
TF	F	T	T	F	F	F
FT	T	F	T	F	F	F
FF	T	T	F	T	T	T

---

The term **logical equivalence (law)** is new to us, but in fact, we already saw several examples of such equivalences. We speak of logical equivalences to describe laws that express transformations from one logical expression to another equivalent one. Let us summarize those we know already:

- De Morgan laws:  $\neg(p \wedge q) \equiv \neg p \vee \neg q$ , and  $\neg(p \vee q) \equiv \neg p \wedge \neg q$ ,
- Commutativity of  $\vee$  and  $\wedge$ :  $p \wedge q \equiv q \wedge p$ ,  $p \vee q \equiv q \vee p$ ,
- The negation of true is false, the negation of false is true:  $\neg T = F$ ,  $\neg F = T$ .

Here are two other ones that we saw implicitly. Suppose that  $\mathcal{T}$  is a tautology (statement always true) and  $C$  is a contradiction (statement always false). Then

- The negation of a tautology is a contradiction:  $\neg \mathcal{T} \equiv C$ .
- The negation of a contradiction is a tautology:  $\neg C \equiv \mathcal{T}$ .

Here are three new logical equivalences:

- Double negation:  $\neg(\neg p) \equiv p$ .
- Idempotent:  $p \wedge p \equiv p$ ,  $p \vee p \equiv p$ .
- Absorption: The first one is  $p \vee (p \wedge q) \equiv p$ . Let us see why, using a truth table:

$p$	$q$	$p \wedge q$	$p \vee (p \wedge q)$
$T$	$T$	$T$	$T$
$T$	$F$	$F$	$T$
$F$	$T$	$F$	$F$
$F$	$F$	$F$	$F$

The second absorption law is  $p \wedge (p \vee q) \equiv p$  (see Exercise 10).

## Logical Equivalences

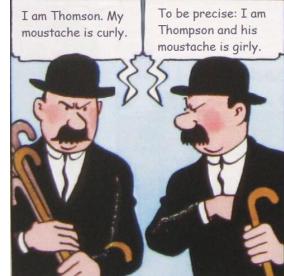
- Useful laws to **transform** one logical expression to an equivalent one.
- **Axioms** ( $T \equiv$  tautology,  $C \equiv$  contradiction):

$$\neg T \equiv F \quad \neg F \equiv T \quad \neg T \equiv C \equiv F \quad \neg C \equiv T \equiv T$$

- De Morgan:
- $$\neg(p \wedge q) \equiv \neg p \vee \neg q$$
- $$\neg(p \vee q) \equiv \neg p \wedge \neg q$$

- Commutativity:

$$p \wedge q \equiv q \wedge p \quad p \vee q \equiv q \vee p$$



© Herge

## Logical Equivalence Laws

- **double negation:**  $\neg(\neg p) \equiv p$
- **idempotent:**  $p \wedge p \equiv p$  and  $p \vee p \equiv p$
- **absorption:**  $p \vee (p \wedge q) \equiv p$  and  $p \wedge (p \vee q) \equiv p$

We will next add three more logical operators. We already know 3 of them:

- $\wedge$  conjunction (and)
- $\vee$  disjunction (or)
- $\neg$  negation (not)

Our fourth logic operator is a [conditional statement](#):

 $\rightarrow$  if then

that is  $p \rightarrow q$  means “if  $p$  then  $q$ ”. You have to be a bit careful with this operator! It sounds a lot like a statement you may have encountered while programming, but actually it is a bit different: if  $p$  is true, then  $p \rightarrow q$  takes the truth value of  $q$ , this part is not too surprising. However, if  $p$  is false, then it is assumed that  $p \rightarrow q$  is [true by default](#), also called [vacuously true!!](#)

Now that we have a new logic operator, let us see how it relates to those we know.

**Theorem 2** (The Conversion Theorem.). *We have*

$$p \rightarrow q \equiv \neg p \vee q.$$

*Proof.* This can be seen using the truth tables:

$p$	$q$	$p \rightarrow q$	$p$	$q$	$\neg p$	$\neg p \vee q$
T	T	T	T	T	F	T
T	F	F	T	F	F	F
F	T	T	F	T	T	T
F	F	T	F	F	T	T

Alternatively

$p$	$q$	$\neg q$	$p \wedge \neg q$	$\neg(p \wedge \neg q)$
T	T	F	F	T
T	F	T	T	F
F	T	F	F	T
F	F	T	F	T

and then use De Morgan law.  $\square$

## Conditional Statement →

Known operators:  $\wedge$  conjunction (and),  $\vee$  disjunction (or),  $\neg$  negation.

- if  $p$  then  $q$ :  $p \rightarrow q$

By definition, when  $p$  is false,  
 $p \rightarrow q$  is true. This is called  
**vacuously true** or **true by default**.

$p$	$q$	$p \rightarrow q$
T	T	T
T	F	F
F	T	T
F	F	T

```
gap>
gap> a:=10;; if (a>5) then Print("yes"); fi;
yes
gap> a:=1;; if (a>5) then Print("yes"); fi;
gap>
gap>
```



## Conversion Theorem

**Theorem:**  $p \rightarrow q \equiv \neg p \vee q$

### Proof

- $p \rightarrow q$  means  $\neg(p \wedge \neg q)$  (it cannot be that  $p$  is true but  $q$  is false).
- Apply DeMorgan's law

$p$	$q$	$p \rightarrow q$	$\neg p$	$\neg p \vee q$
T	T	T	F	T
T	F	F	F	F
F	T	T	T	T
F	F	T	T	T

Given the conditional statement if  $p \rightarrow q$ , it can be modified in three different ways to give raise to some other statements:

- The **converse** of  $p \rightarrow q$  is  $q \rightarrow p$ .
  - The **inverse** of  $p \rightarrow q$  is  $\neg p \rightarrow \neg q$ .
  - The **contrapositive** of  $p \rightarrow q$  is  $\neg q \rightarrow \neg p$ .

The notion of converse and contrapositive often appear in the context of proof techniques. The reason why the contrapositive appears in the context of proof is because of this theorem:

**Theorem 3.** *We have*

$$\neg q \rightarrow \neg p \equiv p \rightarrow q,$$

that is the contrapositive of  $p \rightarrow q$  is equivalent to  $p \rightarrow q$ !

*Proof.* Recall the Conversion Theorem:  $p \rightarrow q \equiv \neg p \vee q$ . Let us start with  $\neg q \rightarrow \neg p$ :

$$\begin{aligned}
 & \neg q \rightarrow \neg p \\
 & \equiv \neg(\neg q) \vee \neg p \quad \text{Conversion Theorem} \\
 & \equiv q \vee \neg p \quad \text{double negation} \\
 & \equiv \neg p \vee q \quad \text{commutativity} \\
 & \equiv p \rightarrow q \quad \text{Conversion Theorem.}
 \end{aligned}$$

Another proof using truth tables is found in Exercise 13.

The contrapositive of  $p \rightarrow q$  is also called **only if**, which is our 5th logic operator ( $\triangleq$  means “equal by definition”)

only if  $\triangleq \neg q \rightarrow \neg p$ ,

even though it is equivalent to  $p \rightarrow q$ .

**Example 14.** The formulas are saying

$$(p \text{ only if } q) = (\neg q \rightarrow \neg p) \equiv (p \rightarrow q).$$

Let us see how this translates into sentences:

“Bob pays taxes only if his income is more than 1000 SD”.

We notice where is the “only if”. Then  $\neg q \rightarrow \neg p$  is

"if Bob's income is less than 1000 SD, then he does not pay taxes."

## Conditional Statements

The **converse** of  $p \rightarrow q$  is  $q \rightarrow p$ , the **inverse** of  $p \rightarrow q$  is  $\neg p \rightarrow \neg q$ , the **contrapositive** of  $p \rightarrow q$  is  $\neg q \rightarrow \neg p$ .

**Theorem:**  $\neg q \rightarrow \neg p \equiv p \rightarrow q$

Proof

$$\begin{aligned} & \neg q \rightarrow \neg p \\ & \equiv \neg(\neg q) \vee \neg p \\ & \equiv q \vee \neg p \\ & \equiv \neg p \vee q \\ & \equiv p \rightarrow q \end{aligned}$$



© Mimi and Eunice (<http://mimiandeunice.com>)

## Only If

- $p$  **only if**  $q \triangleq \neg q \rightarrow \neg p$
- $\neg q \rightarrow \neg p$  is the **contrapositive** of  $p \rightarrow q$ .
- $(\text{if not } q \text{ then not } p) \equiv (p \rightarrow q)$ . Why?

- Example: ‘Bob pays taxes only if his income  $\geq \$1000$ ’  
 $\triangleq$  ‘if Bob’s income  $< \$1000$  then he does not pay taxes’  
 $\equiv$  ‘if Bob pays tax then his income  $\geq \$1000$ ’

Finally  $p \rightarrow q$  is

“if Bob pays taxes, then his income is more than 1000 SD.

$p$     $q$

The terms “sufficient” and “necessary” might sound more or less like having the same meaning, but they have different specific meanings in the world of logic:

**Definition 15.** When  $p \rightarrow q$ , we call  $p$  a **sufficient condition** for  $q$ , while  $q$  is called a **necessary condition** for  $p$ .

Our list of logical operators will now be complete, with the addition of the **biconditional** operator:

$p \leftrightarrow q$  biconditional of  $p$  and  $q$ .

By definition  $p \leftrightarrow q$  means  $(p \rightarrow q) \wedge (q \rightarrow p)$ . This statement, also referred to as **if and only if**, appears very often during proofs.

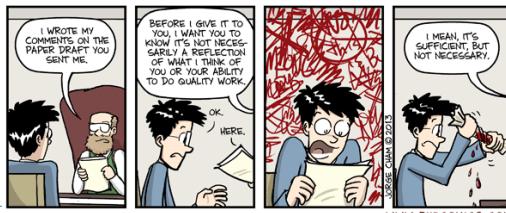
Finally, when combining logic operators, there may be doubts on which operator to apply first. The use of parentheses helps in deciding which operators should be computed first. If several  $\neg$  are used, start with the rightmost one (the one that applies directly on the proposition). If several  $\rightarrow$  are used instead (or other logical operators connecting two propositions), start from the leftmost one.

A good exercise to see whether the different logical equivalence laws seen so far are handled is to prove that  $p \vee q \rightarrow r \equiv (p \rightarrow r) \wedge (q \rightarrow r)$ . The details are given in Exercise 14.

## Sufficient and Necessary Conditions

When  $p \rightarrow q$ , p is called a **sufficient condition** for q , q is a **necessary condition** for p.

- Being Japanese is a sufficient condition for being Asian.  
≡ if someone is Japanese then s/he will be an Asian
- Being Asian is a necessary condition for being Japanese.  
≡ 'if someone is not Asian, he can not be Japanese'



© Jorge Cham

## Example

- Let f: 'you fix my ceiling', p: 'I pay my rent'.
  - 'you fix my ceiling or I won't pay my rent'
    - $f \vee \neg p \equiv p \rightarrow f$
    - 'If you do not fix my ceiling then I won't pay my rent'
      - $\neg f \rightarrow \neg p \equiv p \rightarrow f$
      - 'I will pay my rent only if you fix my ceiling'
        - $\neg f \rightarrow \neg p \equiv p \rightarrow f$

## Biconditional $\leftrightarrow$

- The **biconditional** of  $p$  and  $q$ :  $p \leftrightarrow q \triangleq (p \rightarrow q) \wedge (q \rightarrow p)$ 
  - True only when  $p$  and  $q$  have identical truth values
- If and only if (iff)*

Known operators:

- $\wedge$  conjunction (**and**),
- $\vee$  disjunction (**or**),
- $\neg$  negation,
- $\rightarrow$  conditional (**if then**)

$p$	$q$	$p \rightarrow q$	$q \rightarrow p$	$p \leftrightarrow q$
T	T	T	T	T
T	F	F	T	F
F	T	T	F	F
F	F	T	T	T

## Operator Precedence

- From high to low:  $( )$ ,  $\neg$ ,  $\wedge\vee$ ,  $\rightarrow$ ,  $\leftrightarrow$
- When equal priority instances of *binary connectives* are not separated by  $( )$ , the *leftmost one has precedence*. E.g.  $p \rightarrow q \rightarrow r \equiv (p \rightarrow q) \rightarrow r$
- When instances of  $\neg$  are not separated by  $( )$ , the *rightmost one has precedence*:  
E.g.  $\neg\neg\neg p \equiv \neg(\neg(\neg p))$



All animals are equal, but some are more equal than others – George Orwell, Animal Farm

© to the artist, <http://pda-animalfarm.wikispaces.com/Animalism>

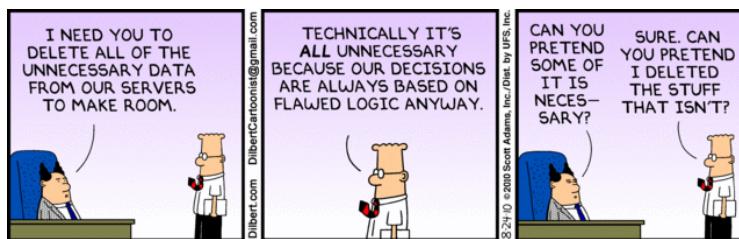
## Example

- Show that  $p \vee q \rightarrow r \equiv (p \rightarrow r) \wedge (q \rightarrow r)$

$$\begin{aligned}
 & p \vee q \rightarrow r \\
 & \equiv (p \vee q) \rightarrow r && \text{(operator precedence)} \\
 & \equiv \neg(p \vee q) \vee r && \text{(why?)} \\
 & \equiv (\neg p \wedge \neg q) \vee r && \text{(DeMorgan's)} \\
 & \equiv (\neg p \vee r) \wedge (\neg q \vee r) && \text{(why?)} \\
 & \equiv (p \rightarrow r) \wedge (q \rightarrow r) && \text{(why?)}
 \end{aligned}$$

## Summary

- Useful logical equivalence laws
  - Proving equivalence using these laws
- Conditional & Biconditional statements
  - Sufficient and necessary conditions
- Operator precedence



© Scott Adams

Before continuing, it is probably a good time to summarize shortly what we did. We defined 6 logical operators:

$\wedge$	conjunction (and)	$p \wedge q$
$\vee$	disjunction (or)	$p \vee q$
$\neg$	negation (not)	$\neg p$
$\rightarrow$	conditional (if then) only if	$p \rightarrow q$ $p \text{ only if } q$
$\leftrightarrow$	biconditional (if and only if)	$p \leftrightarrow q$

We also saw several important logical equivalences, De Morgan Laws and the Conversion Theorem:

$$\neg(p \vee q) \equiv \neg p \wedge \neg q, \quad \neg(p \wedge q) \equiv \neg p \vee \neg q, \quad p \rightarrow q \equiv \neg p \vee q.$$

Now that we have these basic tools, we will use them to construct more sophisticated logical constructs, such as *arguments*.

**Definition 16.** An **argument** is a sequence of statements. The last statement is called the **conclusion**, all the previous statements are **premises**, or **assumptions/hypotheses**.

**Definition 17.** A **valid argument** is an argument where the conclusion is true if the premises are true.

We may rephrase this definition of valid argument in the language of logic: a series of statements forms a valid argument if and only if the conjunction of premises (that is, premises connected by a  $\wedge$ ) implying the conclusion is a tautology. Recall that a tautology is a proposition which is always true. Here is an example of conjunction of premises implying a conclusion:

$$((\text{premise}) \wedge (\text{premise})) \rightarrow \text{conclusion}$$

Thus if the premises are true, then the truth value of

$$((\text{premise}) \wedge (\text{premise}))$$

is true, and thus the truth value of

$$((\text{premise}) \wedge (\text{premise})) \rightarrow \text{conclusion}$$

is true as well.

## Valid Argument

An **argument** is a sequence of statements. The last statement is called the **conclusion**, all the previous statements are **premises** (or **assumptions/ hypotheses**).

**A valid argument** is an argument where the *conclusion is true if the premises are true.*

Example:

'if you pay up in full  
then I will deliver it'; } premises  
'you pay up in full'; } conclusion  
'I will deliver it'; }



© belongs to the cartoonist

## Valid Argument

**A valid argument** is an argument where the *conclusion is true if the premises are true.*

- A series of statements form a valid argument if and only if 'the conjunction of premises implying the conclusion' is a tautology
- $((\text{premise}) \wedge (\text{premise})) \rightarrow \text{conclusion}$

If instead any of the premises is false, then

$$((\text{premise}) \wedge (\text{premise}))$$

takes the value false, and by definition of the conditional operator  $\rightarrow$ , the truth value of

$$((\text{premise}) \wedge (\text{premise})) \rightarrow \text{conclusion}$$

is true as well. Thus

$$((\text{premise}) \wedge (\text{premise})) \rightarrow \text{conclusion}$$

always takes true as truth value, and is indeed a tautology.

Here is a typical template of what an argument looks like:

$p \rightarrow q;$
$p;$
$\therefore q;$

The premises are  $p \rightarrow q; p;$ , and the conclusion is  $q$ .

How do we know that this argument is valid? We check that it is indeed a tautology...which is something that we do as usual, using a truth table, where the first columns contain the premises, and the last column contains  $((p \rightarrow q) \wedge p) \rightarrow q$ , and we need that the last column contains the truth value true.

In fact, it is even easier than a normal truth table, because here, the only case we really care about is the case where all premises are true (the corresponding rows are called [critical rows](#))! Again, this is because if any premise is wrong, then the truth value of  $((p \rightarrow q) \wedge p) \rightarrow q$  is always true by default.

Let us compute the first row of the truth table:

$p$	$q$	$(p \rightarrow q)$	$(p \rightarrow q) \wedge q$	$((p \rightarrow q) \wedge p) \rightarrow q$
T	T	T	T	T

This first row is a critical row, because all premises namely  $p \rightarrow q$  and  $p$ , both are true. There is only one critical row, because the only other entry with  $p$  true is when  $q$  is false, in which case the other premise  $p \rightarrow q$  is false. On this first row, the truth value of the conclusion is true, this is thus a tautology, and we do have a valid argument.

## Valid Argument Template

$p \rightarrow q;$	premises
$p;$	
$\therefore q$	conclusion

- By definition, a valid argument satisfies: “If the premises are true, then the conclusion is true”
- Also:  $((p \rightarrow q) \wedge p) \rightarrow q$  is a tautology.

truth values of premises  
and of the conclusion

*critical rows* are rows  
with *all premises true*

if in all critical rows the  
conclusion is true, then  
the *argument is valid*  
(otherwise it is *invalid*).  
*No need to calculate*

p	q	$p \rightarrow q$	$(p \rightarrow q) \wedge p$	$((p \rightarrow q) \wedge p) \rightarrow q$
T	T	T	T	T
T	F	F	F	T
F	T	T	F	T
F	F	T	F	T

## Counter Example

If in all critical rows the conclusion is true, then the  
*argument is valid* (otherwise it is *invalid*).

A *critical row with a false conclusion* is a *counter example* that invalidates the argument (=makes the argument not valid).

- A counter example indicates a situation where the conclusion does not follow from the premises.

Note that we could use another simplification, which is to replace the last column by the conclusion, namely to compute

$p$	$q$	$(p \rightarrow q)$	$(p \rightarrow q) \wedge q$	$q$
$T$	$T$	$T$	$T$	$T$

Why? because once all the premises are true, then the truth value of  $((p \rightarrow q) \wedge p) \rightarrow q$  is that of  $q$ ...this is true in general, once

$$((\text{premise}) \wedge (\text{premise}))$$

is true, then the truth value of

$$((\text{premise}) \wedge (\text{premise})) \rightarrow \text{conclusion}$$

is exactly the truth value of the conclusion!

It is enough to have one critical row with a truth value false, to invalidate the argument: this row provides a [counter-example](#), that is a situation where all the premises are true, yet the conclusion does not follow.

**Example 15.** Consider the argument:

$$\begin{aligned} S &= (f \wedge a \rightarrow r); \\ f; \\ \neg a; \\ \therefore \neg r; \end{aligned}$$

To check whether this argument is valid, we need to check the truth tables when the premises are all true. The premises are  $S$ ,  $f$  and  $\neg a$ . When is  $S$  true? it is true when  $f \wedge a$  is false, and when  $f \wedge a$  is true and  $r$  is true.

$a$	$r$	$f$	$\neg a$	$f \wedge a$	$S$	$\neg r$
				$F$	$T$	
$T$				$T$	$T$	

Note that we put the value of  $\neg r$ , namely the conclusion, in the last column.

Then we need  $f$  and  $\neg a$  to be true, that is  $f$  is true and  $a$  is false. This means that  $f \wedge a$  is false, thus the second row in the above table cannot be a critical row. Now the first row is one, irrespectively of the truth value of  $r$ .

## Invalid Argument Example

'if it is falling *and* directly above me then I'll run'

'It is falling'

'it is *not* directly above me'

'I will *not* run'

$S = (f \wedge a \rightarrow r);$

f;

$\neg a;$

$\therefore \neg r$  conclusion

a	r	f	$\neg a$	$f \wedge a$	$S$	$\neg r$
T	T	T	F	T	T	F
T	T	F	F	F	T	F
T	F	T	F	T	F	T
T	F	F	F	F	T	T
F	T	I	T	F	T	F
F	T	F	T	F	T	F
F	F	T	F	T	T	T
F	F	F	F	F	F	F

Critical rows

Invalid argument : conclusion on 5<sup>th</sup> row is false!

## Fallacy

A **fallacy** is an error in reasoning that results in an invalid argument.

Fallacy 1: **converse error**.

Example

- If it is Christmas, then it is a holiday.
- It is a holiday. Therefore, it is Christmas!

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

Fallacy 2: **inverse error**.

Example

- If it is raining, then I will stay at home.
- It is not raining. Therefore I would not stay at home!

$$\begin{array}{l} p \rightarrow q; \\ \neg p; \\ \therefore \neg q \end{array}$$

Thus the two critical rows that we need to consider are:

$a$	$r$	$f$	$\neg a$	$f \wedge a$	$S$	$\neg r$
$F$	$T$	$T$	$T$	$F$	$T$	$F$
$F$	$F$	$T$	$T$	$F$	$T$	$T$

and the first row gives a counter example!

**Definition 18.** A **fallacy** is an error in reasoning that results in an invalid argument.

Here are examples of fallacies:

A **converse error** consists of

$$p \rightarrow q; q; \therefore p.$$

This is not a valid argument because the row of

$p$	$q$	$p \rightarrow q$
$F$	$T$	$T$

gives a counter-example.

An **inverse error** consists of

$$p \rightarrow q; \neg p; \therefore \neg q.$$

This is not a valid argument because the row of

$p$	$q$	$p \rightarrow q$	$\neg p$	$\neg q$
$F$	$T$	$T$	$T$	$F$

gives a counter-example.

It is also possible that an argument is invalid, but still it may lead to a correct conclusion, e.g. by coincidence...

**Definition 19.** A **rule of inference** is a logical construct which takes premises, analyzes their synthax, and returns a conclusion.

We already saw

$$\boxed{p \rightarrow q; p; \therefore q; \text{ Modus Ponens}}$$

Similarly, we have

$$\boxed{p \rightarrow q; \neg q; \therefore \neg p; \text{ Modus Tollens.}}$$

## Invalid Argument, Correct Conclusion

- An argument may be invalid, but it may still draw a correct conclusion (e.g. by coincidence)
- Example
  - If New York is a big city, then New York has tall buildings
  - New York has tall buildings
    - So New York is a big city
- So what happened?
  - We have just made an invalid argument
    - Converse error!
  - But conclusion is true (a fact true by itself)

## Inference Rules

A **rule of inference** is a logical construct which takes premises, analyzes their syntax and returns a conclusion.

We already saw

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

**Modus ponens**  
(method of affirming)

$$\begin{array}{l} p \rightarrow q; \\ \neg q; \\ \therefore \neg p \end{array}$$

**Modus tollens**  
(method of denying)



Indeed, the premises are  $p \rightarrow q$  and  $\neg q$ , the conclusion is  $\neg p$  and

$p$	$q$	$(p \rightarrow q)$	$\neg q$	$\neg p$
$T$	$T$	$T$	$F$	
$T$	$F$	$F$	$T$	
$F$	$T$	$T$	$F$	
$F$	$F$	$T$	$T$	$T$

and the only critical row is the last one, for which the conclusion is true. We did not fill up the first three rows on purpose, since they are not critical.

Here are some more inference rules:

$$\boxed{p \wedge q; \therefore q.} \quad (2.1)$$

In its truth table, we only care about the entry when  $p \wedge q$  is true, for which it must be that both  $p$  and  $q$  are true, thus  $q$  is true as well and the argument is valid.

$$\boxed{p; q; \therefore p \wedge q.}$$

In its truth table, we only care about the row where both  $p$  and  $q$  are true. In this case,  $p \wedge q$  is true.

$$\boxed{p; \therefore p \vee q.}$$

If  $p$  is true, then  $p \vee q$  is true (irrespectively of the truth value of  $q$ ).

$$\boxed{p \vee q; \neg p; \therefore q.}$$

Again, in its truth table, we care about the rows where  $p \vee q$  is true. There are three such rows (exclude the one where both  $p$  and  $q$  are false). Now the row for which  $\neg p$  is true is the one where  $p$  is false, this means we have only one critical row, and if  $p$  is false,  $q$  has to be true (for  $p \vee q$  to be true). Two more such rules are found in Exercises 16 and 17.

The dilemma inference rule is:

$$\boxed{p \vee q; p \rightarrow r; q \rightarrow r; \therefore r.}$$

For  $p \vee q$  to be true, we exclude both  $p$  and  $q$  false. When  $p$  is true, then  $r$  has to be true. When  $p$  is false,  $r$  may take any value since  $p \rightarrow r$  is automatically true.

## More Inference Rules

Conjunctive  
Simplification  
(particularizing)

$$\begin{array}{l} p \wedge q; \\ \therefore p \end{array}$$

Conjunctive  
Addition  
(specializing)

$$\begin{array}{l} p; \\ q; \\ \therefore p \wedge q \end{array}$$

Disjunctive  
addition  
(generalization)

$$\begin{array}{l} p; \\ \therefore p \vee q \end{array}$$

Disjunctive  
Syllogism  
(case  
elimination)

$$\begin{array}{l} p \vee q; \\ \neg p; \\ \therefore q \end{array}$$

Rule of  
contradiction

$$\begin{array}{l} \neg p \rightarrow C; \\ \therefore p \end{array}$$

Alternative  
rule of  
contradiction

$$\begin{array}{l} \neg p \rightarrow F; \\ \therefore p \end{array}$$

## Inference Rule: Dilemma

Dilemma (case by  
case discussions)

$$\begin{array}{l} p \vee q; \\ p \rightarrow r; \\ q \rightarrow r; \\ \therefore r \end{array}$$

$p$	$q$	$r$	$(p \rightarrow r)$
$T$	$T$	$T$	$T$
$T$	$F$	$T$	$T$
$F$	$T$	$T$	$T$
$F$	$T$	$F$	$T$

Among these rows, the 4th one is not critical, since when  $q$  is true and  $r$  is false, then  $q \rightarrow r$  is false. The 3 first rows are all critical, and the conclusion  $r$  takes the value true, thus the argument is valid.

The hypothetical syllogism rule is:

$$\boxed{p \rightarrow q; q \rightarrow r; \therefore p \rightarrow r.}$$

For a change, we will prove this rule using equivalences instead of a truth table.

$$\begin{aligned}
& (p \rightarrow q) \wedge (q \rightarrow r) \\
\equiv & (p \rightarrow q) \wedge (\neg q \vee r) \text{ conversion Theorem} \\
\equiv & [(p \rightarrow q) \wedge \neg q] \vee [(p \rightarrow q) \wedge r] \text{ distributivity} \\
\equiv & [(p \rightarrow q) \wedge \neg q] \vee (p \rightarrow q) \wedge [(p \rightarrow q) \wedge \neg q] \vee r) \text{ distributivity}
\end{aligned}$$

Let us simplify the first term  $[(p \rightarrow q) \wedge \neg q] \vee (p \rightarrow q)$ : we see that  $(p \rightarrow q)$  appears twice, let us call give it a name, say  $a = (p \rightarrow q)$  to see better what this statement looks like:

$$(a \wedge \neg q) \vee a.$$

We see that whenever  $a$  is true, this expression is true (because of  $\vee a$ ). Now whenever  $a$  is false, then  $a \wedge \neg q$  is false, no matter what is  $\neg q$ . Thus

$$(a \wedge \neg q) \vee a \equiv a$$

and

$$([(p \rightarrow q) \wedge \neg q] \vee (p \rightarrow q)) \equiv p \rightarrow q.$$

Thus we can go back to our computations, and find that

$$(p \rightarrow q) \wedge (q \rightarrow r) \equiv (p \rightarrow q) \wedge ([(p \rightarrow q) \wedge \neg q] \vee r).$$

We now look at the other term, namely

$$[(p \rightarrow q) \wedge \neg q] \vee r.$$

## Inference Rule: Hypothetical Syllogism

Hypothetical syllogism:

- Example

$$\boxed{\begin{array}{l} p \rightarrow q; \\ q \rightarrow r; \\ \therefore p \rightarrow r \end{array}}$$

If I do not wake up, then I cannot go to work.  
 If I cannot go to work, then I will not get paid.  
 Therefore, if I do not wake up, then I will not get paid.

## Proof of Hypothetical Syllogism

$$p \rightarrow q; q \rightarrow r; \therefore p \rightarrow r$$

$$\begin{aligned} & (p \rightarrow q) \wedge (q \rightarrow r) && \text{(hypotheses; assumed true)} \\ & \equiv (p \rightarrow q) \wedge (\neg q \vee r) && \text{(Conversion Theorem)} \\ & \equiv [(p \rightarrow q) \wedge \neg q] \vee [(p \rightarrow q) \wedge r] && \text{(Distributive)} \\ & \equiv [(p \rightarrow q) \wedge \neg q] \vee [(p \rightarrow q) \wedge ((p \rightarrow q) \wedge \neg q) \vee r] && \text{(Distributive)} \\ & \equiv (p \rightarrow q) \wedge [(p \rightarrow q) \wedge \neg q] \vee r && \text{(Recall absorption law: } a \vee (a \wedge b) \equiv a, \text{ hence } [(p \rightarrow q) \wedge \neg q] \vee (p \rightarrow q) \equiv p \rightarrow q) \\ & \equiv (p \rightarrow q) \wedge [((\neg p \vee q) \wedge \neg q) \vee r] && \text{(Conversion)} \\ & \equiv (p \rightarrow q) \wedge [((\neg p \wedge \neg q) \vee (q \wedge \neg q)) \vee r] && \text{(Distributive)} \\ & \equiv (p \rightarrow q) \wedge [((\neg p \wedge \neg q) \vee F) \vee r] && \text{(Negation)} \\ & \equiv (p \rightarrow q) \wedge [(\neg p \wedge \neg q) \vee r] && \text{(Unity)} \\ & \equiv (p \rightarrow q) \wedge [(\neg p \vee r) \wedge (\neg q \vee r)] && \text{(Distributive)} \\ & \equiv [(p \rightarrow q) \wedge (\neg q \vee r)] \wedge (\neg p \vee r) && \text{(Commutative; Associative)} \\ & \therefore (\neg p \vee r) \equiv p \rightarrow r && \text{(Conjunctive simplification; conversion)} \end{aligned}$$

Using the conversion theorem, we have  $(p \rightarrow q) \equiv \neg p \vee q$ , thus

$$\begin{aligned} & [(p \rightarrow q) \wedge \neg q] \vee r \\ \equiv & [(\neg p \vee q) \wedge \neg q] \vee r \\ \equiv & [(\neg p \wedge \neg q) \vee (q \wedge \neg q)] \vee r \text{ distributivity} \end{aligned}$$

We can now simplify  $(q \wedge \neg q)$ , since this expression is always false (a statement and its contrary cannot be true at the same time). Then we have

$$(\neg p \wedge \neg q) \vee F,$$

which is true exactly when  $(\neg p \wedge \neg q)$  is true, and false exactly when  $(\neg p \wedge \neg q)$  is false, thus we get

$$\begin{aligned} & [(p \rightarrow q) \wedge \neg q] \vee r \\ \equiv & (\neg p \wedge \neg q) \vee r \\ \equiv & (\neg p \vee r) \wedge (\neg q \vee r) \text{ distributivity}. \end{aligned}$$

Let us now combine everything together:

$$\begin{aligned} & (p \rightarrow q) \wedge (q \rightarrow r) \\ \equiv & (p \rightarrow q) \wedge [(\neg p \vee r) \wedge (\neg q \vee r)] \\ \equiv & (p \rightarrow q) \wedge (\neg p \vee r) \wedge (\neg q \vee r) \\ \equiv & [(\neg p \vee r) \wedge (\neg q \vee r)] \wedge (p \rightarrow q) \\ \therefore & (\neg p \vee r) \\ \equiv & p \rightarrow r \text{ conversion theorem}. \end{aligned}$$

The second last “therefore” statement follows from the inference rule (2.1):

$$\boxed{p \wedge q; \therefore q.}$$

## Exercises for Chapter 2

**Exercise 8.** Decide whether the following statements are propositions. Justify your answer.

1.  $2 + 2 = 5$ .
2.  $2 + 2 = 4$ .
3.  $x = 3$ .
4. Every week has a Sunday.
5. Have you read “Catch 22”?

**Exercise 9.** Show that

$$\neg(p \vee q) \equiv \neg p \wedge \neg q.$$

This is the second law of De Morgan.

**Exercise 10.** Show that the second absorption law  $p \wedge (p \vee q) \equiv p$  holds.

**Exercise 11.** These two laws are called distributivity laws. Show that they hold:

1. Show that  $(p \wedge q) \vee r \equiv (p \vee r) \wedge (q \vee r)$ .
2. Show that  $(p \vee q) \wedge r \equiv (p \wedge r) \vee (q \wedge r)$ .

**Exercise 12.** Verify  $\neg(p \vee \neg q) \vee (\neg p \wedge \neg q) \equiv \neg p$  by

- constructing a truth table,
- developing a series of logical equivalences.

**Exercise 13.** Using a truth table, show that:

$$\neg q \rightarrow \neg p \equiv p \rightarrow q.$$

**Exercise 14.** Show that  $p \vee q \rightarrow r \equiv (p \rightarrow r) \wedge (q \rightarrow r)$ .

**Exercise 15.** Are  $(p \rightarrow q) \vee (q \rightarrow r)$  and  $p \rightarrow r$  equivalent statements?

**Exercise 16.** Show that this argument is valid:

$$\boxed{\neg p \rightarrow F; \therefore p.}$$

**Exercise 17.** Show that this argument is valid, where  $C$  denotes a contradiction.

$$\boxed{\neg p \rightarrow C; \therefore p.}$$

**Exercise 18.** Determine whether the following argument is valid:

$$\begin{aligned} &\neg p \rightarrow r \wedge \neg s \\ &t \rightarrow s \\ &u \rightarrow \neg p \\ &\neg w \\ &u \vee w \\ &\therefore t \rightarrow w. \end{aligned}$$

**Exercise 19.** Determine whether the following argument is valid:

$$\begin{aligned} &p \\ &p \vee q \\ &q \rightarrow (r \rightarrow s) \\ &t \rightarrow r \\ &\therefore \neg s \rightarrow \neg t. \end{aligned}$$

## Examples for Chapter 2

To practice some of the logical concepts we have studied so far, we will consider the well known puzzles of the island of knights and knaves.

The assumptions are that on this island:

- Knights always tell the truth.
- Knaves always lie.

As a visitor, you will be told some statements, and you have to decide whether the islanders you are speaking with are knights or knaves.

The general method will be to rewrite the claims of the islanders using propositional logic, and then to use truth tables to figure out the truth!

**Example 16.** You meet two islanders  $A$  and  $B$ .  $A$  says: “I am a knave but he is not”. You need to decide what are  $A$  and  $B$ .

To do so, we will use logic. Let us call  $p$  the statement “ $A$  is a knight”, and  $q$  the statement “ $B$  is a knight”. With that, we rephrase the statement of  $A$ .

$A$  says: “ $\neg p \wedge q$ ”. The truth table of  $\neg p \wedge q$  is

$p$	$\neg p$	$q$	$\neg p \wedge q$
$T$	$F$	$T$	$F$
$T$	$F$	$F$	$F$
$F$	$T$	$T$	$T$
$F$	$T$	$F$	$F$

Now either  $A$  is a knight, or  $A$  is a knave. If  $A$  is a knight, then the statement  $p$  is true, and  $A$  must always tell the truth, thus  $\neg p \wedge q$  must be true as well. This cannot be from the truth table.

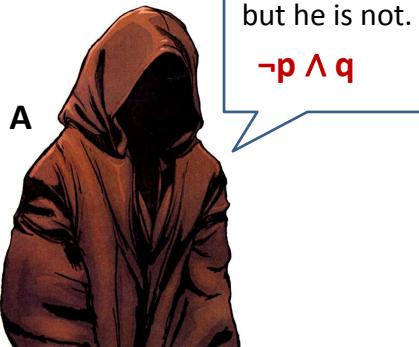
If  $A$  is a knave, then the statement  $p$  is not true, and  $A$  must always lie, thus  $\neg p \wedge q$  must be false as well. This happens in the last row of the table, where  $p$  is false,  $q$  is false, and  $\neg p \wedge q$  is false, thus we conclude that  **$A$  is a knave, and  $B$  is a knave**.

## Knights & Knaves I



Art belongs to the artist

## Knights & Knaves I



- $p = \text{"A is a knight"}$
- $q = \text{"B is a knight"}$

p	q	$\neg p \wedge q$
T	T	F
T	F	F
F	T	T
F	F	F

- If A is a knight, then  $p = \text{true}$ , and  $\neg p \wedge q$  must be true.
- If A is a knave, then  $p = \text{false}$ , and  $\neg p \wedge q$  must false.

**Example 17.** You meet two islanders  $A$  and  $B$ .  $A$  says: “If I am a knight then so is he”. You need to decide what are  $A$  and  $B$ .

Again, let us call  $p$  the statement “ $A$  is a knight”, and  $q$  the statement “ $B$  is a knight”. Then  $A$  says: “ $p \rightarrow q$ ”, with truth table

$p$	$q$	$p \rightarrow q$
$T$	$T$	$T$
$T$	$F$	$F$
$F$	$T$	$T$
$F$	$F$	$T$

If  $A$  is a knight, then the statement  $p$  is true, and  $A$  must always tell the truth, thus  $p \rightarrow q$  must be true as well. This is possible with the first row

If  $A$  is a knave, then the statement  $p$  is not true, and  $A$  must always lie, thus  $p \rightarrow q$  must be false as well. This cannot happen. We conclude that  **$A$  is a knight, and  $B$  is a knight**.

**Example 18.** You meet two islanders  $A$  and  $B$ .  $A$  says: “I am a knave or  $B$  is a knight”. You need to decide what are  $A$  and  $B$ .

As twice above, let us call  $p$  the statement “ $A$  is a knight”, and  $q$  the statement “ $B$  is a knight”. Then  $A$  says: “ $\neg p \vee q$ ”. Its truth table is

$p$	$\neg p$	$q$	$\neg p \vee q$
$T$	$F$	$T$	$T$
$T$	$F$	$F$	$F$
$F$	$T$	$T$	$T$
$F$	$T$	$F$	$T$

If  $A$  is a knight, then the statement  $p$  is true, and  $A$  must always tell the truth, thus  $\neg p \vee q$  must be true as well. This happens in the first row of the table.

If  $A$  is a knave, then the statement  $p$  is not true, and  $A$  must always lie, thus  $\neg p \vee q$  must be false as well. This cannot happen, thus we conclude that  **$A$  is a knight, and  $B$  is a knight**.

## Knights & Knaves II



If I am a knight  
then so is he.  
 $p \rightarrow q$

- $p = \text{"A is a knight"}$
- $q = \text{"B is a knight"}$

$p$	$q$	$p \rightarrow q$
T	T	T
T	F	F
F	T	T
F	F	T

- If A is a knight, then  $p = \text{true}$ , and  $p \rightarrow q$  must be true.
- If A is a knave, then  $p = \text{false}$ , and  $p \rightarrow q$  must false.

## Knights & Knaves III



I am a Knave  
or B is a Knight  
 $\neg p \vee q$

- $p = \text{"A is a knight"}$
- $q = \text{"B is a knight"}$

$p$	$q$	$\neg p \vee q$
T	T	T
T	F	F
F	T	T
F	F	T

- If A is a knight, then  $p = \text{true}$ , and  $\neg p \vee q$  must be true.
- If A is a knave, then  $p = \text{false}$ , and  $\neg p \vee q$  must false.

**Example 19.** You meet one islander  $A$ .  $A$  says: “I am a knave”. You need to decide what is  $A$ . This case is actually similar to the “liar’s paradox”. You can write the truth table as above, and see that nothing can be decided from it, which is the definition of a paradox! If  $A$  is knight, he tells the truth and says that he is a knave, which is not possible. So  $A$  must be a knave. But if  $A$  is a knave, he lies, and so he must be a knight, which is not possible either.

We next give another example of how to combine different statements with inference rules to extract information:

**Example 20.** During a murder investigation, you have gathered the following clues:

1. if the knife is in the store room, then we saw it when we cleared the store room;
2. the murder was committed at the basement or inside the apartment;
3. if the murder was committed at the basement, then the knife is in the yellow dust bin;
4. we did not see a knife when we cleared the store room;
5. if the murder was committed outside the building, then we are unable to find the knife;
6. if the murder was committed inside the apartment, then the knife is in the store room.

The question is: “where is the knife?”

First, we assigned symbols to the above clues:

- $s$  : the knife is in the store room;
- $c$  : we saw the knife when we clear the store room;
- $b$  : the murder was committed at the basement;
- $a$  : murder was committed inside the apartment;
- $y$  : the knife is in the yellow dust bin;
- $o$  : the murder was committed outside the building;
- $u$  : we are unable to find the knife;

## Knights & Knaves IV

- $p = \text{"A is a knight"}$



$p$	$\neg p$
T	F
F	T

It's a paradox!

## The Murder Clues

1. if the knife is in the **s**tore room, then we saw it when we **c**leared the store room;
2. the murder was committed at the **b**asement or inside the **a**partment;
3. if the murder was committed at the **b**asement, then the knife is in the **y**ellow dust bin;
4. we did not see a knife when we **c**leared the store room;
5. if the murder was committed **o**utside the building, then we are **u**nable to find the knife;
6. if the murder was committed inside the **a**partment, then the knife is in the **s**tore room.



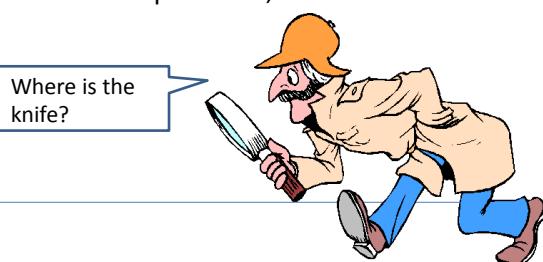
Next we rewrite the clues inside these symbols, and logical operators:

1.  $s \rightarrow c$
2.  $b \vee a$
3.  $b \rightarrow y$
4.  $\neg c$
5.  $o \rightarrow u$
6.  $a \rightarrow s$

From 1. and 4., we deduce that  $\neg s$ . From 6. and  $\neg s$ , we deduce  $\neg a$ . Once we have  $\neg a$ , with 2., we deduce  $b$ . Once we have  $b$ , with 3., we deduce  $y$ , that is, the knife is in the yellow bin!

## Statements

1. if the knife is in the **s**tore room, then we saw it when we **c**leared the **s**tore room;
2. the **m**urder was committed at the **b**asement or inside the **a**partment;
3. if the murder was committed at the **b**asement, then the knife is in the **y**ellow dust bin;
4. we did not see a knife when we **c**leared the store room;
5. if the murder was committed **o**utside the building, then we are **u**nable to find the knife;
6. if the murder was committed inside the **a**partment, then the knife is in the **s**tore room.



## Applying Inference Rules

- |                                  |                      |                                    |
|----------------------------------|----------------------|------------------------------------|
| 1. if <b>s</b> , then <b>c</b> ; | 1. $s \rightarrow c$ | The knife is in<br>the yellow bin! |
| 2. <b>b</b> or <b>a</b> ;        | 2. $b \vee a$        |                                    |
| 3. if <b>b</b> , then <b>y</b> ; | 3. $b \rightarrow y$ |                                    |
| 4. not <b>c</b> ;                | 4. $\neg c$          |                                    |
| 5. if <b>o</b> , then <b>u</b> ; | 5. $o \rightarrow u$ |                                    |
| 6. if <b>a</b> , then <b>s</b>   | 6. $a \rightarrow s$ |                                    |
|                                  | 7. $\neg s$          | 1, 4; modus tollens                |
|                                  | 8. $\neg a$          | 6, 7; modus tollens                |
|                                  | 9. $b$               | 2, 8; case elimination             |
|                                  | $\therefore y$       | 3, 9; modus ponens                 |





# Chapter 3

## Predicate Logic

*“Logic will get you from A to B. Imagination will take you everywhere.” A. Einstein*

In the previous chapter, we studied propositional logic. This chapter is dedicated to another type of logic, called *predicate logic*.

Let us start with a motivating example.

**Example 21.** Consider the following two statements:

- Every SCE student must study discrete mathematics.
- Jackson is an SCE student.

It looks “logical” to deduce that therefore, Jackson must study discrete mathematics. However, this cannot be expressed by propositional logic...you may try it, but you can already notice that none of the logical operators we have learnt are applicable here.

We need new tools!

**Definition 20.** A **predicate** is a statement that contains variables (predicate variables), and they may be true or false depending on the values of these variables.

**Example 22.**  $P(x) = “x^2 \text{ is greater than } x”$  is a predicate. It contains one predicate variable  $x$ . If we choose  $x = 1$ ,  $P(1)$  is “1 is greater than 1”, which is a proposition (always false).

## Limitation Of Propositional Logic

- Every SCE student must study discrete mathematics
  - Jackson is a SCE student
    - So Jackson must study discrete mathematics
      - This idea *can't be expressed* with propositional logic
  - What propositional logic allows to express:
    - If Jackson is a SCE student he must study discrete mathematics
    - Jackson is a SCE student
      - So Jackson must study discrete mathematics
- 

## Predicates

- Is the statement “ $x^2$  is greater than  $x$ ” a proposition?
- Define  $P(x) = 'x^2 \text{ is greater than } x'$ .
- Is  $P(1)$  a proposition?  $P(1) = "1^2 \text{ is greater than } 1"$  (F)

A **predicate** is a statement that contains variables (**predicate variables**) and that may be true or false depending on the values of these variables.

- $P(x)$  is a predicate.

```
#include <stdio.h>
void main()
{
    int a,b;
    a=10;
    b=3;
    printf("Is %d equal to %d ? %d\n",a,b,a==b);
    printf("Is %d different from %d? %d\n",a,b,a!=b);
}
```

Since a predicate takes value true or false once instantiated (that is, once its variables are taking values), we may alternatively say that a predicate instantiated becomes a proposition.

It is needed to explicit which are the values that a predicate variable can possibly take.

**Definition 21.** The domain of a predicate variable is the collection of all possible values that the variable may take.

**Example 23.** Consider the predicate  $P(x) = "x^2 \text{ is greater than } x"$ . Then the domain of  $x$  could be for example the set  $\mathbb{Z}$  of all integers. It could alternatively be the set  $\mathbb{R}$  of real numbers. Whether instantiations of a predicate are true or false may depend on the domain considered.

When several predicate variables are involved, they may or not have different domains.

**Example 24.** Consider the predicate  $P(x, y) = "x > y"$ , in two predicate variables. We have  $\mathbb{Z}$  (the set of integers) as domain for both of them.

- Take  $x = 4, y = 3$ , then  $P(4, 3) = "4 > 3"$ , which is a proposition taking the value true.
- Take  $x = 1, y = 2$ , then  $P(1, 2) = "1 > 2"$ , which is a proposition taking the value false.
- Note that in general  $P(x, y) \neq P(y, x)$ !

We now introduce two quantifiers (describing “parts or quantities” from a domain), the universal quantification and the existential quantification.

**Definition 22.** A universal quantification is a quantifier meaning “given any” or “for all”. We use the following symbol:

$$\boxed{\forall \text{ (universal quantification)}}$$

**Example 25.** Here is a formal way to say that for all values that a predicate variable  $x$  can take in a domain  $D$ , the predicate is true:

$$\underbrace{\forall x}_{\text{for all } x \text{ belonging to } D} \quad \underbrace{\in D}_{\text{, } P(x) \text{ (is true)}}$$

For example

$$\underbrace{\forall x}_{\text{for all } x \text{ belonging to the real numbers}} \quad \underbrace{\in \mathbb{R}}_{\text{, } x^2 \geq 0.}$$

## Predicate Instantiated/Domain

A **predicate** is a statement that contains variables (**predicate variables**) and that may be true or false depending on the values of these variables.

- A predicate instantiated (where variables are evaluated in specific values) is a proposition.

The **domain** of a predicate variable is the collection of all possible values that the variable may take.

- e.g. the domain of  $x$  in  $P(x)$ : integer
- Different variables may have different domains.

- Predicate logic extends (is more powerful than) propositional logic.

4/12

## Example

- Let  $P(x, y) = "x > y"$ .  
Domain: integers, i.e. both  $x$  and  $y$  are integers.
- $P(4, 3)$  means " $4 > 3$ ", so  $P(4, 3)$  is TRUE;
- $P(1, 2)$  means " $1 > 2$ ", so  $P(1, 2)$  is FALSE;
- $P(3, 4)$  is false (in general,  $P(x,y)$  and  $P(y,x)$  not equal).

## Quantification

- Statements like
  - Some birds are angry.
  - On the internet, no one knows who you are.
  - The square of any real number is nonnegative.



© Rovio

## Universal Quantification

A **universal quantification** is a quantifier (something that tells the amount or quantity) meaning "given any" or "for all".

**Symbol:**  $\forall$

- E.g. " $\forall x \in D P(x)$  is true" iff " $P(x)$  is true for every  $x$  in  $D$ ".
  - $\forall$  universal quantifiers, "for all", "for every"
  - $\in$  - "is a member (or) element of", "belonging to"
  - $D$  – domain of predicate variable
- The square of any real number is nonnegative.

$$\forall x \in \mathbb{R}, x^2 \geq 0.$$

## Existential Quantification

An **existential quantification** is a quantifier (something that tells the amount or quantity) meaning “there exists”, “there is at least one” or “for some”.

**Symbol:**  $\exists$

- E.g. “ $\exists x \in D, P(x)$  is true” iff “ $P(x)$  is true for *at least one*  $x$  in  $D$ ”.
  - $\exists$  existential quantifier, “there exists”
- **Some** birds are angry.
  - $D = \{\text{birds}\}$ ,  $P(x) = "x \text{ is angry}"$ .

---

8/12

## Nested Quantification (I)

- A proposition may contain multiple quantifiers
    - **All** rabbits are faster than **all** tortoises.”
    - Domains:  $R = \{\text{rabbits}\}$ ,  $T = \{\text{tortoises}\}$
    - Predicate  $C(x, y)$ : Rabbit  $x$  is faster than tortoise  $y$
  - In symbols
    - $\forall x \in R (\forall y \in T, C(x, y))$  or  $\forall x \in R, \forall y \in T, C(x, y)$
  - In words
    - For any rabbit  $x$ , and for any tortoise  $y$ ,  $x$  is faster than  $y$ .
-

**Definition 23.** An **existential quantification** is a quantifier meaning “there exists”, “there is at least one” or “for some”. We use the following symbol:

$$\exists \text{ (existential quantification)}$$

**Example 26.** Here is a formal way to say that for some values that a predicate variable  $x$  can take in a domain  $D$ , the predicate is true:

$$\underbrace{\exists x}_{\text{for some } x} \quad \underbrace{\in D}_{\text{belonging to } D}, \quad P(x) \text{ (is true)}$$

For example, for  $D = \{ \text{birds} \}$ ,  $P(x) = "x \text{ is angry}"$ ,

$$\underbrace{\exists x}_{\text{Some birds}} \quad \underbrace{\in D}_{\text{are angry}}, \quad \underbrace{P(x) \text{ (is true)}}_{\text{.}}$$

The term nested quantification refers to statements involving several quantifiers. Here is a series of examples.

**Example 27.** All statements involve two predicate variables  $x$  and  $y$ , where  $x$  has for domain  $R = \{ \text{rabbits} \}$ , while  $y$  has for domain  $T = \{ \text{tortoises} \}$ . The predicate used is  $C(x, y) = \text{"Rabbit } x \text{ is faster than tortoise } y"$ .

- In logic symbolism, we write “All rabbits are faster than all tortoises”:

$$\underbrace{\forall x}_{\text{For any rabbit } x} \quad \underbrace{\in R}_{\text{, for any } x}, \quad \underbrace{\forall y}_{\in T}, \quad \underbrace{C(x, y) \text{ (is true)}}_{x \text{ is faster than } y}.$$

- In logic symbolism, we write “Every rabbit is faster than some tortoise”:

$$\underbrace{\forall x}_{\text{For any rabbit } x} \quad \underbrace{\in R}_{\text{, there is a tortoise } y}, \quad \underbrace{\exists y}_{\in T}, \quad \underbrace{C(x, y) \text{ (is true)}}_{x \text{ is faster than } y}.$$

- In logic formalism, we write “There is a rabbit which is faster than all tortoises”:

$$\underbrace{\exists x}_{\text{There exists a rabbit } x} \quad \underbrace{\in R}_{\text{such that}}, \quad \underbrace{\forall y}_{\in T}, \quad \underbrace{C(x, y) \text{ (is true)}}_{x \text{ is faster than } y}.$$

## Nested Quantification (II)

- Another example
    - “**Every rabbit** is faster than **some tortoise**.”
    - Domains: R={rabbits}, T={tortoises}.
    - Predicate C(x, y): Rabbit x is faster than tortoise y
  - In symbols
    - $\forall x \in R (\exists y \in T, C(x, y))$  or  $\forall x \in R, \exists y \in T, C(x, y)$
  - In words:
    - **For any** rabbit x, **there exists a (some)** tortoise y, such that x is faster than y.
- 

## Nested Quantification (III)

- Another example
    - “**There is a rabbit** which is faster than **all tortoises**.”
    - Domains: R={rabbits}, T={ tortoises}.
    - Predicate C(x, y): Rabbit x is faster than tortoise y.
  - In symbols (note the ordering in nesting)
    - $\exists x \in R (\forall y \in T, C(x, y))$
  - In words:
    - **There exists** a rabbit x, such that for any tortoise y, this rabbit x is faster than y.
-

The same way we assigned truth values to propositions, we may assign truth values to quantified statements.

- $(\forall x \in D, P(x))$  is true exactly when  $P(x)$  is true for every  $x \in D$ . Thus it is false whenever there is at least one  $x$  for which  $P(x)$  is false. Formally, for  $D = \{x_1, \dots, x_n\}$ , we have the following equivalence:

$$(\forall x \in D, P(x)) \equiv (P(x_1) \wedge P(x_2) \wedge \dots \wedge P(x_n)).$$

- $(\exists x \in D, P(x))$  is true exactly when  $P(x)$  is true for at least one  $x \in D$ . Thus it is false when  $P(x)$  is false for all  $x \in D$ . Formally, for  $D = \{x_1, \dots, x_n\}$ , we have the following equivalence:

$$(\exists x \in D, P(x)) \equiv (P(x_1) \vee P(x_2) \vee \dots \vee P(x_n)).$$

Since  $(\exists x \in D, P(x))$  takes truth values, it can also be negated, that is

$$\neg(\exists x \in D, P(x)) \equiv \forall x \in D, \neg P(x),$$

where the equivalence follows from the fact that  $(\exists x \in D, P(x))$  is false whenever for all  $x \in D$   $P(x)$  is false. Similarly

$$\neg(\forall x \in D, P(x)) \equiv \exists x \in D, \neg P(x) \quad (3.1)$$

since  $(\forall x \in D, P(x))$  is false whenever there is one  $x \in D$  for which  $P(x)$  is false. See Exercise 23 for negating a statement involving several quantifiers.

We can similarly assign truth values to combinations of predicates, or negation of combinations of predicates. The equivalence

$$\neg(\forall x \in D, P(x) \wedge Q(x)) \equiv \exists x \in D, \neg(P(x) \wedge Q(x))$$

holds, setting  $P'(x) = P(x) \wedge Q(x)$  and using (3.1) on  $P'(x)$ . Now  $(P(x) \wedge Q(x))$  is a proposition for any instantiation of  $x$ , thus we can apply De Morgan laws:

$$\neg(\forall x \in D, P(x) \wedge Q(x)) \equiv \exists x \in D, \neg P(x) \vee \neg Q(x).$$

Suppose now you are given a statement involving quantifiers, whose truth table has to be determined. There are several ways to do so: (1) Method of Exhaustion, (2) Method of Case, and (3) Method of Logic Derivation.

**Method of exhaustion:** if the domain contains a small number of elements, try them all! For example, if  $D = \{5, 6, 7, 8, 9\}$ , and  $P(x) = "x \in D, x^2 = x"$ , then just compute  $x^2$  for all the values of  $x \in D$  to conclude that this false.

## Truth Value of Quantified Statements

Statement	When true	When false
$\forall x \in D, P(x)$	$P(x)$ is true for every $x$ .	There is one $x$ for which $P(x)$ is false.
$\exists x \in D, P(x)$	There is one $x$ for which $P(x)$ is true.	$P(x)$ is false for every $x$ .

Assume that  $D$  consists of  $x_1, x_2, \dots, x_n$

- $\forall x \in D, P(x) \equiv P(x_1) \wedge P(x_2) \wedge \dots \wedge P(x_n)$
- $\exists x \in D, P(x) \equiv P(x_1) \vee P(x_2) \vee \dots \vee P(x_n)$

## Negation of Quantification

- ‘Not all SCE students study hard’ = ‘There is at least one SCE student who does not study hard’  

$$\neg (\forall x \in D, P(x)) \equiv \exists x \in D, \neg P(x)$$
- Negation of a universal quantification becomes an existential quantification.
- ‘It is not the case that some students in this class are from NUS.’ = ‘All students in this class are not from NUS’  

$$\neg (\exists x \in D, P(x)) \equiv \forall x \in D, \neg P(x)$$
- Negation of an existential quantification becomes an universal quantification.

## Negation of Quantification

$$\begin{aligned}
 & \neg (\forall x \in D, P(x) \wedge Q(x)) \\
 & \equiv \exists x \in D \neg (P(x) \wedge Q(x)) \quad (\text{negation of quantification}) \\
 & \equiv \exists x \in D (\neg P(x) \vee \neg Q(x)) \quad (\text{DeMorgan})
 \end{aligned}$$

- Example: Not all students in this class are using Facebook and (also) Google+
    - There is some (at least one) student in this class who is not using Facebook or not using Google+ (or may be using neither)
- 

## How To Determine Truth Value

Statement	When true	When false
$\forall x \in D, P(x)$	P(x) is true for every x.	There is one x for which P(x) is false.
$\exists x \in D, P(x)$	There is one x for which P(x) is true.	P(x) is false for every x.

- Systematic approaches:
    - Method of **exhaustion**
    - Method of **case**
    - Method of **logic derivation**
-

## Method of Exhaustion

- Let  $D=\{5,6,7,8,9\}$ . Is  $\exists x \in D, x^2=x$  true or false?
    - $5^2=25 \neq 5$ ,  $6^2=36 \neq 6$ ,  $7^2=49 \neq 7$ ,  $8^2=64 \neq 8$ ,  $9^2=81 \neq 9$
    - So, **false!**
  - Limitation?
    - Domain may be too large to try out all options
      - E.g., all integers
- 

## Method of Case

- Positive examples to **prove existential quantification**
- Let  $Z$  denote all integers. Is  $\exists x \in Z, x^2=x$  true or false?
  - Take  $x = 0$  or  $1$  and we have it. True.
- Counterexample to **disprove universal quantification**
- Let  $R$  denote all reals. Is  $\forall x \in R, x^2 > x$  true or false?
  - Take  $x=0.3$  as a counterexample. False.

Positive example is **not** a proof of universal quantification

Negative example **is not** disproof of existential quantification

May be **hard** to find suitable “cases” even if such cases do exist!

---



**Method of Case.** Suppose you want to show that the truth value of  $(\exists x, P(x))$  is true. For this, you just need to find one case, one instantiation of  $x$ , for which  $P(x)$  is true.

**Example 28.**  $P(x) = \exists x \in \mathbb{Z}, x^2 = x$  is true, take  $x = 1$  for example. Thus such an  $x$  exists.

Similarly, if you want to show that the truth value of  $(\forall x, P(x))$  is false, it is enough to find one counterexample.

**Example 29.**  $P(x) = \forall x \in \mathbb{R}, x^2 > x$  is false, take  $x = 0.3$  for example. Thus  $P(x)$  cannot be true for all  $x$ .

However, you **cannot** show that  $(\exists x, P(x))$  is false using some examples, you need to prove that you cannot find a single  $x$  in your domain for which  $P(x)$  is true! Vice-versa, you cannot show that  $(\forall x, P(x))$  is true by giving some examples, you need to show that this is always true, for every  $x$  in the domain considered.

To do this, again, if the domain is small, one may use the exhaustion method of trying all options, but if the domain is big (or infinite), like  $\mathbb{Z}$ , we need another method.

**Method of Logic Derivation.** This method consists of using logical steps to transform one logical expression into another.

**Example 30.** Suppose you want to know the truth value of  $\exists x, (P(x) \vee Q(x))$ ,  $x$  has for domain  $D = \{x_1, \dots, x_n\}$ . Then  $\exists x, (P(x) \vee Q(x))$  is true if there is an  $x_i \in D$  for which  $P(x_i) \vee Q(x_i)$  is true that is

$$(\exists x, (P(x) \vee Q(x))) \equiv (P(x_1) \vee Q(x_1)) \vee \dots \vee (P(x_n) \vee Q(x_n))$$

but now this new expression becomes true exactly when at least one  $P(x_i)$  or  $Q(x_j)$  is true, that is

$$(\exists x, (P(x) \vee Q(x))) \equiv (\exists x, P(x)) \vee (\exists x, Q(x)).$$

When trying to derive logic steps involving quantifiers, one should be **very careful** with the ordering of the quantifiers...A typical example is that

$$\forall x, \exists y, P(x, y) \equiv \exists y, \forall x, P(x, y)$$

does **not hold** in general!

**Example 31.** Consider the predicate  $P(x, y) = "x \text{ admires } y"$ . Then  $\forall x, \exists y, P(x, y)$  means that everyone admires someone, while  $\exists y, \forall x, P(x, y)$  means that there exists one person who is admired by everyone!

## Method Of Logic Derivation

Consider an (arbitrary) domain with  $n$  members. Is  $\exists x(P(x) \vee Q(x))$  logically equivalent to  $\exists xP(x) \vee \exists xQ(x)$ ?

$$\begin{aligned}
 & \exists x(P(x) \vee Q(x)) \\
 & \equiv [P(x_1) \vee Q(x_1)] \vee \dots \vee [P(x_n) \vee Q(x_n)] \\
 & \equiv [P(x_1) \vee \dots \vee P(x_n)] \vee [Q(x_1) \vee \dots \vee Q(x_n)] \\
 & \equiv \exists xP(x) \vee \exists xQ(x)
 \end{aligned}$$


---

## Order Of Nesting Matters

- Is  $\forall x \exists y P(x,y) \equiv \exists y \forall x P(x,y)$  in general?
    - LHS:  $y$  can *vary* with respect to  $x$ ,  $y$  is *fixed* with respect to  $x$
    - Let  $P(x, y) = "x \text{ admires } y"$ . LHS = “Every person admires some people”, RHS = “Some people are admired by everyone”
  - Consider  $x, y \in \mathbb{R}^+$ , and let  $P(x,y)$  be  $xy=1$ .
    - Is  $\forall x \exists y P(x,y) \equiv \exists y \forall x P(x,y)$ ?
  - Consider  $X=\{9, 10, 15\}$ ,  $Y=\{2, 3\}$ . Let  $Q(x,y)$ :  $y$  divides  $x$ 
    - Then, is  $\forall x \in X \exists y \in Y Q(x,y) \equiv \exists y \in Y \forall x \in X Q(x,y)$ ?
-

We have seen so far how quantified statements have truth values, how they can be combined with an AND operator, or negated with a negation operator. We next discuss how they can be combined with the “if then” conditional operator:

$$\boxed{\forall x \in D, (P(x) \rightarrow Q(x))}$$

which means, for all  $x$  in its domain  $D$ , if  $P(x)$  is true, then  $Q(x)$  is true.

**Example 32.** Take  $P(x) = "x > 1"$ ,  $Q(x) = "x^2 > 1"$ , and  $x$  has for domain the real numbers  $\mathbb{R}$ . Then

$$\forall x \in \mathbb{R}, (P(x) \rightarrow Q(x))$$

becomes for all  $x \in \mathbb{R}$ , if  $x > 1$  then  $x^2 > 1$ .

Attached to the conditional operator were several of its variations, its contrapositive, its inverse, and its negation. We can similarly define these for quantified statements.

$$\begin{aligned} \forall x \in D, (\neg Q(x) \rightarrow \neg P(x)) &\quad \text{contrapositive} \\ \forall x \in D, (Q(x) \rightarrow P(x)) &\quad \text{converse} \\ \forall x \in D, (\neg P(x) \rightarrow \neg Q(x)) &\quad \text{inverse} \end{aligned}$$

See Exercise 25 for a proof that a conditional proposition is equivalent to its contrapositive. See also Exercise 26 to compute the negation of a conditional quantification, namely

$$\neg(\forall x, P(x) \rightarrow Q(x)).$$

## Conditional Quantification (I)

- For all real number  $x$ , if  $x > 1$  then  $x^2 > 1$ 
    - i.e., any real number greater than 1 has a square larger than 1
  - In symbolic form
    - Let  $P(x)$  denote “ $x > 1$ ”
    - Let  $Q(x)$  denote “ $x^2 > 1$ ”
    - Let  $R$  denote the domain, the collection of all real numbers
      - $\forall x (P(x) \rightarrow Q(x))$
- 

## Conditional Quantification (III)

- Given a conditional quantification
    - Such as  $\forall x \in A (P(x) \rightarrow Q(x))$
  - Then we define
 

<i>– contrapositive</i>	$\forall x \in A, \neg Q(x) \rightarrow \neg P(x)$
<i>– converse</i>	$\forall x \in A, Q(x) \rightarrow P(x)$
<i>– inverse</i>	$\forall x \in A, \neg P(x) \rightarrow \neg Q(x)$
  - Note: A conditional proposition is logically equivalent to its contrapositive
-

Our motivating Example 21 to start predicate logic was:

**Example 33.** Consider the following two statements: (1) Every SCE student must study discrete mathematics. (2) Jackson is an SCE student. It looks “logical” to deduce that therefore, Jackson must study discrete mathematics.

We will now develop inference rules, that will allow us to express this example (see Exercise 28).

Consider a predicate variable  $x$  taking value in the domain  $D$ . Then

$$\boxed{\forall x \in D, P(x); \therefore P(c) \text{ for any } c \in D.}$$

As we did for propositional logic, we look at when the premises are true. When  $\forall x, P(x)$  is true,  $P(x)$  is true for any choice of  $x$  in the domain  $D$ , in particular it is true for any choice of  $c$ , therefore  $P(c)$  is true for any  $c \in D$ . *This rule says that if  $P(x)$  is true for any  $x$  in a domain, one is allowed to instantiate  $P(x)$  in  $x = c$  for any choice of  $c \in D$ .*

**Example 34.** Suppose we have the following premises: (1) No cat can catch Jerry; (2) Tom is a cat. We want to deduce that therefore Tom cannot catch Jerry. We define two predicates:  $\text{Cat}(x) = "x \text{ is a cat}"$ ,  $\text{Catch}(x) = "x \text{ can catch Jerry}"$ . The second premise, Tom is a cat, is the easiest to write, it becomes:  $\text{Cat}(\text{Tom}) = "\text{Tom is a cat}"$ . Now for the first premise, suppose we want to say “cats are catching Jerry”, this would be if  $x$  is a cat, then  $x$  catches Jerry, that is

$$\text{Cat}(x) \rightarrow \text{Catch}(x),$$

keeping in mind that we have not yet assigned a quantifier to this statement. To say that “cats are not catching Jerry”, then this would be if  $x$  is cat, then  $x$  cannot catch Jerry, that is

$$\text{Cat}(x) \rightarrow \neg \text{Catch}(x),$$

and finally, to say that “no cat can catch Jerry”, we add a universal quantifier:

$$\forall x (\text{Cat}(x) \rightarrow \neg \text{Catch}(x)).$$

We then have the following premises in predicate logic:

1.  $\forall x (\text{Cat}(x) \rightarrow \neg \text{Catch}(x));$
2.  $\text{Cat}(\text{Tom});$

We can then instantiate the first premise with  $x = \text{Tom}$ , to get

$$(\text{Cat}(\text{Tom}) \rightarrow \neg \text{Catch}(\text{Tom})).$$

## Universal Instantiation

$$\forall x P(x)$$

$$\therefore P(c)$$

where  $c$  is **any** element of the domain.

**Example:**

$\text{Cat}(x)$ :  $x$  is Cat,  $\text{Catch}(x)$ :  $x$  can catch Jerry

- No cat can catch Jerry.
- Tom is a cat. Therefore  
Tom cannot catch



1.  $\forall x [\text{Cat}(x) \rightarrow \neg \text{Catch}(x)]$  Hypothesis

2.  $\text{Cat}(\text{Tom})$  Hypothesis

3.  $\text{Cat}(\text{Tom}) \rightarrow \neg \text{Catch}(\text{Tom})$

Universal Instantiation on 1

4.  $\neg \text{Catch}(\text{Tom})$

Modus ponens on 2 and 3

© William Hanna and Joseph Barbera

## Universal Generalization

$P(c)$  for **any arbitrary**  $c$  from the domain.

$$\therefore \forall x P(x)$$

**Example**

- $P(x) = ``x^2 \text{ is non-negative}''$ .
- $P(c)$  for an arbitrary real  $c$ .
- Therefore  $P(x)$  for all  $x$ .

But since the second premise says that  $\text{Cat}(\text{Tom})$ ; modus ponens ( $p \rightarrow q; p; \therefore q$ ) tells us that therefore  $\neg\text{Catch}(\text{Tom})$ .

Consider a predicate variable  $x$  taking value in the domain  $D$ . Then

$$\boxed{P(c) \text{ for an arbitrary } c \in D; \therefore \forall x P(x) \in D}.$$

This just means that if a predicate is true for an arbitrary element  $c \in D$ , then it is true for all  $x \in D$ . Indeed, if whichever premise you look at is true, then the conclusion is true. *This rule allows us to infer  $P(x)$  for all  $x \in D$  based on  $P(c)$  being true for an arbitrary instance  $c \in D$ .*

**Example 35.** Consider the premise for an arbitrary real number  $x$ ,  $x^2 \geq 0$ . Therefore the square of any real number is non-negative. Set  $P(x) = "x^2 \geq 0"$ . In predicate logic, we have for any arbitrary  $c \in \mathbb{R}$ ,  $P(c)$ . Therefore  $\forall x P(x)$ .

In fact, we have already used this rule implicitly...In Exercise 2, we showed that if  $n^2$  is even, then  $n$  is even, for  $n$  an integer. The way we did it, is that we showed the result for one arbitrary  $n$ , and concluded this is true for all of them! See Exercise 29 for a more complicated example.

Consider a predicate variable  $x$  taking value in the domain  $D$ . Then

$$\boxed{\exists x \in D, P(x); \therefore P(c) \text{ for some } c \in D}.$$

We look at when the premises are true. When  $\exists x, P(x)$  is true,  $P(c)$  is true for at least one choice of  $c$  in the domain  $D$ . *This rule allows to instantiate  $P(x)$  in some values of  $c$  for which  $P(c)$  is true.*

**Example 36.** The premises are: (1) if any student gets  $> 80$  in the exam, then (s)he gets an  $A$ , (2) there are students who get  $> 80$  in the exam, (3) Sam is such a student. We want to conclude that therefore Sam gets an  $A$ . Set the predicates  $A(x) = "x \text{ gets an } A"$ ,  $M(x) = "x \text{ gets } > 80 \text{ in the exam}"$ , the domain  $D$  is  $D = \{ \text{students} \}$ . In predicate logic, (1) and (2) are respectively given by

1.  $\forall x, M(x) \rightarrow A(x)$ .
2.  $\exists x, M(x)$ .

Now that Sam is such a student can be expressed using the above rule, to obtain  $M(\text{Sam})$ . Then we can instantiate 1. with Sam, to get  $M(\text{Sam}) \rightarrow A(\text{Sam})$ , which combined with  $M(\text{Sam})$  leads to therefore  $A(\text{Sam})$ .

## Existential Instantiation

$\exists x P(x)$   
 $\therefore P(c)$  for some  $c$  in the domain.

### Example

- If any student gets >80 in the final exam, then (s)he gets an A.
  - There are students who get >80 in the final exam, Sam is such a student.
  - Therefore, Sam gets an A.
- |   |                                |
|---|--------------------------------|
| <p>1. <math>\forall x [M(x) \rightarrow A(x)]</math> Hypothesis</p> <p>2. <math>\exists x M(x)</math> Hypothesis</p> <p>3. <math>M(Sam)</math> Hypothesis + Existential instantiation</p> <p>4. <math>M(Sam) \rightarrow A(Sam)</math></p> <p>5. <math>A(Sam)</math> Universal instantiation on 1</p> | <p>Modus ponens on 4 and 3</p> |
|---|--------------------------------|

## Existential Generalization

$P(c)$   
 $\therefore \exists x P(x)$   
for  $c$  some specific element of domain.

Domain={all people},  
Sell( $x$ ) = “ $x$  is selling stocks”.

- $\forall x \text{Sell}(x) \rightarrow \exists x \text{Sell}(x)$
1.  $\forall x \text{Sell}(x)$  Hypothesis
  2.  $\text{Sell}(c)$  Universal instantiation
  3.  $\exists x \text{Sell}(x)$  Existential generalization



### Example

- If everyone is selling stocks, then someone is selling stocks.

© belongs to the cartoonist

Consider a predicate variable  $x$  taking value in the domain  $D$ . Then

$$\boxed{P(c) \text{ for some specific } c \in D; \therefore \exists x, P(x)}.$$

We look at when the premises are true. When  $P(c)$  is true for some specific  $c \in D$ ,  $\exists x$  for which  $P(x)$  is true. *This rule allows to go from one instantiation  $P(c)$  to deduce that there is at least one  $x$  for which  $P(x)$  is true.*

**Example 37.** Suppose we want to show formally that if everyone is selling stocks, there must be someone selling stocks. Consider the predicate  $\text{Sell}(x) = "x \text{ is selling stocks}"$ . Then we want to show that if  $\forall x \text{ Sell}(x)$  is true, then  $\exists x \text{ Sell}(x)$  is true as well. From  $\forall x \text{ Sell}(x)$  is true, we instantiate it in one  $c$  in the domain (here the domain is { people }). This gives  $\text{Sell}(c)$ . Now using the above rule, we know there must exist at least an  $x$  for which  $\text{Sell}(x)$  is true, as desired.

These 4 rules may look either “obvious” or “contrived”, but they are needed to write down things formally, in particular whenever formal methods are involved, e.g., if you need to program formal verifications!

We now come to the last part of predicate logic, namely, we will see how to apply the logic rules we have seen to justify different proof techniques. We will discuss three proof technique: direct proof, induction, proof by contradiction.

**Direct Proof.** As the name suggests, these are proofs that are performed without particular techniques. Some claim has to be shown, and there is a specific way to do so in this particular context.

**Example 38.** Suppose we want to show that  $\sum_{i=0}^n i = \frac{n(n+1)}{2}$ ,  $\forall n \in \mathbb{N}$ . Write down the following array:

$$\begin{array}{ccccccc} 1 & 2 & 3 & \dots & n-1 & n \\ n & n-1 & n-2 & \dots & 2 & 1 \end{array}$$

If you sum up the entries of the first row, you get  $\sum_{i=0}^n i$ , and if you sum up the entries of the second row, you also get  $\sum_{i=0}^n i$ . Thus if you sum up all entries in this array, you get  $2 \sum_{i=0}^n i$ . Now if we sum up the first column, we get  $n+1$ , if we sum up the second column we get  $n+1$ , ..., and for the  $n$ th column we also get  $n+1$ , so the total is  $n+1$  times  $n$  columns.

# Proof Techniques

A **valid proof** is a valid argument, i.e. the conclusion follows from the given assumptions.

## Three techniques:

- Direct proof
- Proof by induction
- Proof by contradiction.

Proof by example:  
The author gives only the case  $n = 2$  and suggests that it contains most of the ideas of the general proof.

Proof by intimidation: 'Trivial.'

Proof by cumbersome notation: Best done with access to at least four alphabets and special symbols.

Proof by exhaustion: An issue or two of a journal devoted to your proof is useful.

Proof by omission: 'The reader may easily supply the details.' 'The other 253 cases are analogous.' '...'

Proof by obfuscation: A long plotless sequence of true and/or meaningless syntactically related statements.

Proof by wishful citation: The author cites the negation, converse, or generalization of a theorem from literature to support his claims.

© PROOF TECHNIQUES by A. H. Zemanian, The Physics Teacher, May 1994.

## Proof Technique: Direct Proof

- Prove that  $\forall n \in \mathbb{Z}, \sum_{i=0}^n i = \frac{n(n+1)}{2}$
- Define  $S = \sum_{i=0}^n i = \underbrace{0+1+2+\dots+n-1+n}_{n+1 \text{ terms}}$ 
  - Note:  $S = \sum_{i=0}^n i = \underbrace{n+n-1+\dots+2+1+0}_{n+1 \text{ terms}}$
- Sum up:  $2S = \underbrace{n+n+\dots+n+n+n}_{\Rightarrow 2S = (n+1)n}$
- Thus:  $S = \frac{n(n+1)}{2}$



Leonhard Euler  
(1707-1783)

If we sum up the elements horizontally, we got  $2 \sum_{i=0}^n i$ , while if we sum up the elements vertically, we got  $n(n + 1)$ . But the sum does not change when we count differently, thus:

$$\sum_{i=0}^n i = \frac{n(n + 1)}{2}.$$

The legend attributes this proof technique to young Euler, who apparently was punished for not behaving in the classroom. His teacher would have asked him to compute the sum of integers from 1 to 100, and Euler would have, or so the legend says, came up with this technique so as to have to compute all the additions!

**Mathematical Induction.** This is a proof technique to show statements of the form  $\forall n, P(n)$ . We first explain the technique, then give a proof of why this proof technique is valid, and finally provide an example. A proof by mathematical induction follows two steps:

1. Basis step: You need to show that  $P(1)$  is true.
2. Inductive step: You assume that  $P(k)$  is true, and have to prove that  $P(k + 1)$  is then true.

When both steps are complete, we have proved that  $\forall n, P(n)$  is true. Why is that the case? From the inductive step, we have that

$$P(k) \rightarrow P(k + 1)$$

for any  $k$ , therefore this is true for when we instantiate in  $k = 1$ , that is

$$P(1) \rightarrow P(2).$$

But from the basis step, we know that  $P(1)$  is true, thus combining  $(P(1) \rightarrow P(2)); P(1)$ ; we get that therefore  $P(2)$  holds. We can repeat this process with  $k = 2$  to deduce that  $P(3)$  holds, and so on and so forth.

## Mathematical Induction

- Prove propositions of the form:  $\forall n P(n)$
- The proof consists of two steps:
  - **Basis Step:** The proposition  $P(1)$  is shown to be true
  - **Inductive Step:**
    - Assume  $P(k)$  is true (when  $n=k$ ), then, prove  $P(k+1)$  is true (when  $n=k+1$ ).
- When both steps are complete, we have proved that " $\forall n P(n)$ " is true

## Why Does it Work?

- From step 2:  $P(1) \rightarrow P(2)$  by Universal Instantiation.
- From step 1:  $P(1)$
- Applying *modus ponen*:  $P(2)$ .
- Repeat the process to get  $P(3), P(4), P(5)$ , etc. So, all  $P(k)$  are true! i.e.,  $\forall k P(k)$

Analogy with climbing Ladders.



**Inductive step**

$$\boxed{[P(1) \wedge \forall k (P(k) \rightarrow P(k+1))] \rightarrow \forall n P(n)}$$

**Basis**

**Hypothesis**

(valid argument)

**Example 39.** We want to prove

$$\sum_{i=0}^n i = \frac{n(n+1)}{2}, \forall n \in \mathbb{N}$$

using mathematical induction. Then set

$$P(n) = \text{"} \sum_{i=0}^n i = \frac{n(n+1)}{2} \text{"}, \forall n \in \mathbb{N}.$$

- Basis step:  $P(1) = 1 = \frac{1(1+1)}{2}$ .
- Inductive step: suppose that  $P(k)$  is true, that is  $\sum_{i=0}^k i = \frac{k(k+1)}{2}$  is true for all  $k$ . Now we need to show that  $P(k+1)$  holds.

$$\begin{aligned} \sum_{i=0}^{k+1} i &= \sum_{i=0}^k i + (k+1) \\ &= \frac{k(k+1)}{2} + (k+1) \\ &= \frac{k(k+1) + 2(k+1)}{2} = \frac{(k+1)(k+2)}{2}. \end{aligned}$$

There  $P(n)$  is true for all  $n$ .

**Proof by Contradiction.** We want to prove that  $P(n) \rightarrow Q(n)$  is true. In a proof by contradiction, we assume by contradiction that  $P(n) \rightarrow Q(n)$  is false, that is, that  $\neg(P(n) \rightarrow Q(n))$  is true. The only way this might happen, is if  $P(n)$  is true and  $Q(n)$  is false. Thus we start with  $P(n)$  true and  $Q(n)$  false. If from there we deduce a contradiction, that is a statement of the form  $C \wedge \neg C$ , which is always false, what we have proven is

$$\neg(P(n) \rightarrow Q(n)) \rightarrow C \wedge \neg C$$

is true. This is equivalent to  $P(n) \rightarrow Q(n)$ . To see that, set  $S(n) = \text{"} P(n) \rightarrow Q(n) \text{"}$ , and look at the truth table:

$S$	$C$	$\neg S$	$C \wedge \neg S$	$(\neg S) \rightarrow (C \wedge \neg S)$
$T$	$T$	$F$	$F$	$T$
$T$	$F$	$F$	$F$	$T$
$F$	$T$	$T$	$F$	$F$
$F$	$F$	$T$	$F$	$F$

Therefore, to prove  $P(n) \rightarrow Q(n)$  (or any other statement), it suffices to instead prove the conditional statement  $\neg(P(n) \rightarrow Q(n)) \rightarrow C \wedge \neg C$ , which is done by direct proof, by assuming  $\neg(P(n) \rightarrow Q(n))$  and deduce  $C \wedge \neg C$ . One difficulty is to figure out what is  $C$  given the proof to effectuate.

**Example 40.** Suppose we want to prove that: if  $n^2$  is even, then  $n$  is even, for  $n$  integer. Set  $P(n)$ =“ $n^2$  is even”, and  $Q(n)$ =“ $n$  is even”. We want to prove that  $P(n) \rightarrow Q(n)$ , which is equivalent to  $\neg(P(n) \rightarrow Q(n)) \rightarrow C \wedge \neg C$ . Suppose  $\neg(P(n) \rightarrow Q(n))$ , that means  $P(n)$  is true and  $Q(n)$  is false:  $n^2$  is even, and  $n$  is not even (equivalently  $n$  is odd). Now if  $n$  is odd, then  $n = 2k + 1$ , and  $n^2 = (2k + 1)^2 = 4k^2 + 4k + 1 = 2(2k^2 + 2k) + 1$ , that is  $n^2$  is odd. Thus  $C = “n^2 \text{ is even}”$ , and we have just shown that  $n^2$  is odd, that is  $C \wedge \neg C$ , a contradiction!

We may alternatively use that  $P(n) \rightarrow Q(n)$  is equivalent to  $\neg Q(n) \rightarrow \neg P(n)$ . This would be a proof using contrapositive.

**Example 41.** Suppose we want to prove that: if  $n^2$  is even, then  $n$  is even, for  $n$  integer. Set  $P(n)$ =“ $n^2$  is even”, and  $Q(n)$ =“ $n$  is even”. We want to prove that  $P(n) \rightarrow Q(n)$ , which is equivalent to  $\neg Q(n) \rightarrow \neg P(n)$ . Suppose that  $\neg Q(n)$ , that is:  $n$  is not even, or  $n$  is odd. Now if  $n$  is odd, then  $n = 2k + 1$ , and  $n^2 = (2k + 1)^2 = 4k^2 + 4k + 1 = 2(2k^2 + 2k) + 1$ , that is  $n^2$  is odd, which is equivalent to  $\neg P(n)$ , which concludes the proof.

## Example: Mathematical Induction

- Prove that  $\forall n \in N, \sum_{i=0}^n i = \frac{n(n+1)}{2}$
- Let  $P(n)$  denote  $\left[ \sum_{i=0}^n i = \frac{n(n+1)}{2} \right]$
- **Inductive step.** Assume  $P(k)$  true,  $k > 0$ :  $\sum_{i=0}^k i = \frac{k(k+1)}{2}$

**Basis step:**  $P(1)$  is true

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

Prove  $P(k+1)$  true :

$$\begin{aligned} \sum_{i=0}^{k+1} i &= \sum_{i=0}^k i + (k+1) = \frac{k(k+1)}{2} + (k+1) \\ &= \frac{(k+1)(k+2)}{2} = \frac{(k+1)[(k+1)+1]}{2} \end{aligned}$$

So,  $P(n)$  is true for  $n=k+1$  and thus true for all  $n$ :  
 $\forall n P(n)$  is true

## Proof Technique: Contradiction

- Prove that: If  $n^2$  is even, then  $n$  is even, for  $n$  integer.
- Lets assume:  $n^2$  is even but  $n$  is not even (assume the negation of the statement is true).
- $n$  is not even  $\Leftrightarrow n$  is odd, i.e.,  $n = 2k+1$ ,  $k$  integer.
  - Then  $n^2 = (2k+1)^2$   
 $= 4k^2 + 4k + 1$   
 $= 2(2k^2 + 2k) + 1$  (odd)
- This is in contradiction with the assumption.
- Hence, the assumption is false, thus the negation of the assumption is true.

## Exercises for Chapter 3

**Exercise 20.** Consider the predicates  $M(x, y) = “x \text{ has sent an email to } y”$ , and  $T(x, y) = “x \text{ has called } y”$ . The predicate variables  $x, y$  take values in the domain  $D = \{\text{students in the class}\}$ . Express these statements using symbolic logic.

1. There are at least two students in the class such that one student has sent the other an email, and the second student has called the first student.
2. There are some students in the class who have emailed everyone.

**Exercise 21.** Consider the predicate  $C(x, y) = “x \text{ is enrolled in the class } y”$ , where  $x$  takes values in the domain  $S = \{\text{students}\}$ , and  $y$  takes values in the domain  $D = \{\text{courses}\}$ . Express each statement by an English sentence.

1.  $\exists x \in S, C(x, \text{MH1812})$ .
2.  $\exists y \in D, C(\text{Carol}, y)$ .
3.  $\exists x \in S, (C(x, \text{MH1812}) \wedge C(x, \text{CZ2002}))$ .
4.  $\exists x \in S, \exists x' \in S, \forall y \in D, ((x \neq x') \wedge (C(x, y) \leftrightarrow C(x', y)))$ .

**Exercise 22.** Consider the predicate  $P(x, y, z) = “xyz = 1”$ , for  $x, y, z \in \mathbb{R}$ ,  $x, y, z > 0$ . What are the truth values of these statements? Justify your answer.

1.  $\forall x, \forall y, \forall z, P(x, y, z)$ .
2.  $\exists x, \exists y, \exists z, P(x, y, z)$ .
3.  $\forall x, \forall y, \exists z, P(x, y, z)$ .
4.  $\exists x, \forall y, \forall z, P(x, y, z)$ .

**Exercise 23.** 1. Express

$$\neg(\forall x, \forall y, P(x, y))$$

in terms of existential quantification.

2. Express

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

in terms of universal quantification.

**Exercise 24.** Consider the predicate  $C(x, y) = "x \text{ is enrolled in the class } y"$ , where  $x$  takes values in the domain  $S = \{\text{students}\}$ , and  $y$  takes values in the domain  $C = \{\text{courses}\}$ . Form the negation of these statements:

1.  $\exists x, (C(x, \text{MH1812}) \wedge C(x, \text{CZ2002}))$ .
2.  $\exists x \exists y, \forall z, ((x \neq y) \wedge (C(x, z) \leftrightarrow C(y, z)))$ .

**Exercise 25.** Show that  $\forall x \in D, P(x) \rightarrow Q(x)$  is equivalent to its contrapositive.

**Exercise 26.** Show that

$$\neg(\forall x, P(x) \rightarrow Q(x)) \equiv \exists x, P(x) \wedge \neg Q(x).$$

**Exercise 27.** Let  $y, z$  be positive integers. What is the truth value of " $\exists y, \exists z, (y = 2z \wedge (\text{y is prime}))$ ".

**Exercise 28.** Write in symbolic logic "Every SCE student studies discrete mathematics. Jackson is an SCE student. Therefore Jackson studies discrete mathematics".

**Exercise 29.** Here is an optional exercise about universal generalization. Consider the following two premises: (1) for any number  $x$ , if  $x > 1$  then  $x - 1 > 0$ , (2) every number in  $D$  is greater than 1. Show that therefore, for every number  $x$  in  $D$ ,  $x - 1 > 0$ .

**Exercise 30.** Let  $q$  be a positive real number. Prove or disprove the following statement: if  $q$  is irrational, then  $\sqrt{q}$  is irrational.

**Exercise 31.** Prove using mathematical induction that the sum of the first  $n$  odd positive integers is  $n^2$ .

**Exercise 32.** Prove using mathematical induction that  $n^3 - n$  is divisible by 3 whenever  $n$  is a positive integer.



# Chapter 4

## Set Theory

“A set is a Many that allows itself to be thought of as a One.”  
(Georg Cantor)

In the previous chapters, we have often encountered “sets”, for example, prime numbers form a set, domains in predicate logic form sets as well. Defining a set formally is a pretty delicate matter, for now, we will be happy to consider an intuitive definition, namely:

**Definition 24.** A **set** is a collection of abstract objects.

A set is typically determined by its distinct elements, or members, by which we mean that the order does not matter, and if an element is repeated several times, we only care about one instance of the element. We typically use the bracket notation  $\{\}$  to refer to a set.

**Example 42.** The sets  $\{1, 2, 3\}$  and  $\{3, 1, 2\}$  are the same, because the ordering does not matter. The set  $\{1, 1, 1, 2, 3, 3, 3\}$  is also the same set as  $\{1, 2, 3\}$ , because we are not interested in repetition: either an element is in the set, or it is not, but we do not count how many times it appears.

One may specify a set *explicitly*, that is by listing all the elements the set contains, or *implicitly*, using a predicate description as seen in predicate logic, of the form  $\{x, P(x)\}$ . Implicit descriptions tend to be preferred for infinite sets.

**Example 43.** The set  $A$  given by  $A = \{1, 2\}$  is an explicit description. The set  $\{x, x \text{ is a prime number}\}$  is implicit.

## Set

A **set** is a collection of abstract objects

- Examples: prime numbers, domain in predicate logic
- Determined by (distinct) elements/members.
  - E.g.  $\{1, 2, 3\} = \{3, 1, 2\} = \{1, 3, 2\} = \{1, 1, 1, 2, 3, 3, 3\}$
- Two common ways to specify a set
  - **Explicit:** Enumerate the members  
e.g.  $A = \{2, 3\}$
  - **Implicit:** Description using predicates  $\{x \mid P(x)\}$   
e.g.  $A = \{x \mid x \text{ is a prime number}\}$

## Membership & Subset

We write  $x \in S$  iff  $x$  is an element (member) of  $S$ .

- e.g.  $A = \{x \mid x \text{ is a prime number}\}$  then  $2 \in A, 3 \in A, 5 \in A, \dots, 1 \notin A, 4 \notin A, 6 \notin A, \dots$

A set  $A$  is a **subset** of the set  $B$ , denoted by  $A \subseteq B$  iff every element of  $A$  is also an element of  $B$ . i.e.,

- $A \subseteq B \triangleq \forall x(x \in A \rightarrow x \in B)$
- $A \not\subseteq B \triangleq \neg(A \subseteq B)$ 
  - $\equiv \neg \forall x(x \in A \rightarrow x \in B)$
  - $\equiv \exists x(x \in A \wedge x \notin B)$

Subset versus Membership:  $S = \{\text{rock, paper, scissors}\}$

$R = \{\text{rock}\}, R \subseteq S, \text{rock} \in S$

Given a set  $S$ , one may be interested in elements belonging to  $S$ , or in subset of  $S$ . The two concepts are related, but different.

**Definition 25.** A set  $A$  is a **subset** of a set  $B$ , denoted by  $A \subseteq B$ , if and only if every element of  $A$  is also an element of  $B$ . Formally

$$A \subseteq B \iff \forall x(x \in A \rightarrow x \in B).$$

Note the two notations  $A \subset B$  and  $A \subseteq B$ : the first one says that  $A$  is a subset of  $B$ , while the second emphasizes that  $A$  is a subset of  $B$ , possibly equal to  $B$ . The second notation is typically preferred if one wants to emphasize that one set is possibly equal to the other.

To say that  $A$  is not a subset of  $S$ , we use the negation of  $\forall x(x \in A \rightarrow x \in B)$ , which is (using the rules we have studied in predicate logic! namely negation of universal quantifier, conversion theorem, and De Morgan's law)  $\exists x(x \in A \wedge x \notin B)$ . The notation is  $A \not\subseteq B$ .

For an element  $x$  to be an element of a set  $S$ , we write  $x \in S$ . This is a notation that we used already in predicate logic. Note the difference between  $x \in S$  and  $\{x\} \subseteq S$ : in the first expression,  $x$  is in element of  $S$ , while in the second, we consider the subset  $\{x\}$ , which is emphasized by the bracket notation.

**Example 44.** Consider the set  $S = \{\text{rock, paper, scissors}\}$ , then  $R = \{\text{rock}\}$  is a subset of  $S$ , while  $\text{rock} \in S$ , it is an element of  $S$ .

**Definition 26.** The **empty set** is a set that contains no element. We denote it  $\emptyset$  or  $\{\}$ .

There is a difference between  $\emptyset$  and  $\{\emptyset\}$ : the first one is an empty set, the second one is a set, which is not empty since it contains one element: the empty set!

**Definition 27.** The **empty set** is a set that contains no element. We denote it  $\emptyset$  or  $\{\}$ .

**Example 45.** We say that two sets  $A$  and  $B$  are equal, denoted by  $A = B$ , if and only if  $\forall x, (x \in A \leftrightarrow x \in B)$ .

To say that two sets  $A$  and  $B$  are not equal, we use the negation from predicate logic, which is:

$$\neg(\forall x, (x \in A \leftrightarrow x \in B)) \equiv \exists x((x \in A \wedge x \notin B) \vee (x \in B \wedge x \notin A)).$$

## Empty Set

The set that contains no element is called the ***empty set*** or ***null set***.

- The empty set is denoted by  $\emptyset$  or by { }.
- Note:  $\emptyset \neq \{\emptyset\}$

## Set Equality

$$A=B \triangleq \forall x(x \in A \leftrightarrow x \in B)$$

- Two sets A, B are equal *iff* they have the same elements.

$$\begin{aligned} A \neq B &\triangleq \neg \forall x(x \in A \leftrightarrow x \in B) \\ &\equiv \exists x [(x \in A \wedge x \notin B) \vee (x \in B \wedge x \notin A)] \end{aligned}$$

- Two sets are not equal if they do not have identical members, i.e., there is some element in one of the sets which is absent in the other.
- Example:

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

This makes our earlier example  $\{1, 2, 3\} = \{1, 1, 1, 2, 3, 3, 3\}$  easier to justify than what we had intuitively before: both sets are equal because whenever a number belongs to one, it belongs to the other.

**Definition 28.** The **cardinality** of a set  $S$  is the number of distinct elements of  $S$ . If  $|S|$  is finite, the set is said to be finite. It is said to be infinite otherwise.

We could say the number of elements of  $S$ , but then this may be confusing when elements are repeated as in  $\{1, 2, 3\} = \{1, 1, 1, 2, 3, 3, 3\}$ , while there is no ambiguity for distinct elements. There  $|S| = |\{1, 2, 3\}| = 3$ . The set of prime numbers is infinite, while the set of even prime numbers is finite, because it contains only 2.

**Definition 29.** The **power set**  $P(S)$  of a set  $S$  is the set of all subsets of  $S$ :

$$P(S) = \{A, A \subseteq S\}.$$

If  $S = \{1, 2, 3\}$ , then  $P(S)$  contains  $S$  and the empty set  $\emptyset$ , and all subsets of size 1, namely  $\{1\}$ ,  $\{2\}$ , and  $\{3\}$ , and all subsets of size 2, namely  $\{1, 2\}$ ,  $\{1, 3\}$ ,  $\{2, 3\}$ .

The cardinality of  $P(S)$  is  $2^n$  when  $|S| = n$ . This is not such an obvious result, it may be derived in several ways, one of them being the so-called binomial theorem, which says that

$$(x + y)^n = \sum_{j=0}^n \binom{n}{j} x^j y^{n-j},$$

where  $\binom{n}{j}$  counts the number of ways to choose  $j$  elements out of  $n$ . The notation  $\sum_{j=0}^n$  means that we sum for the values of  $j$  going from 0 to  $n$ . See Exercise 33 for a proof of the binomial theorem. When  $n = 3$ , evaluating in  $x = y = 1$ , we have

$$2^3 = \binom{3}{0} + \binom{3}{1} + \binom{3}{2} + \binom{3}{3}$$

and we see that  $\binom{3}{0}$  says we pick no element from 3, there is one way, and it corresponds to the empty set, then  $\binom{3}{1}$  is telling us that we have 3 ways to choose a single subset, this is for  $\{1\}$ ,  $\{2\}$ , and  $\{3\}$ ,  $\binom{3}{2}$  counts  $\{1, 2\}$ ,  $\{1, 3\}$ ,  $\{2, 3\}$  and  $\binom{3}{3}$  counts the whole set  $\{1, 2, 3\}$ .

When dealing with sets, it is often useful to draw Venn diagrams to show how sets are interacting. They are useful to visualize “unions” and “intersections”.

## Cardinality

The **cardinality**  $|S|$  of  $S$  is the number of elements in  $S$ .

- e.g. for  $S=\{1, 3\}$ ,  $|S|=2$

If  $|S|$  is finite,  $S$  is a finite set; otherwise,  $S$  is infinite.

- The set of positive integers is an infinite set.
- The set of prime numbers is an infinite set.
- The set of even prime numbers is a finite set.
- Note:  $|\emptyset| = 0$

## Power Set

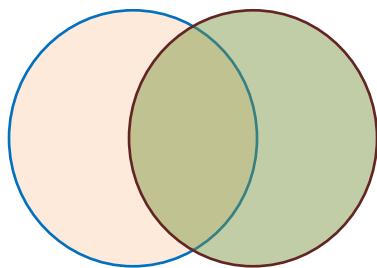
The **power set**  $P(S)$  of a given set  $S$  is the set of all subsets of  $S$ :  $P(S) = \{ A \mid A \subseteq S\}$ .

- Example
  - For  $S=\{1,2,3\}$   
 $P(S)=\{\emptyset,\{1\},\{2\},\{3\},\{1,2\},\{1,3\},\{2,3\},\{1,2,3\}\}$
- If a set  $A$  has  $n$  elements, then  $P(s)$  has  $2^n$  elements.
  - Hint: Try to leverage *the Binomial theorem*

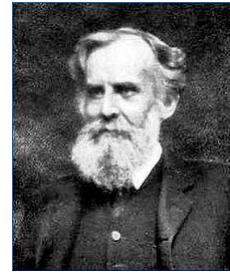
$$(x+y)^n = \binom{n}{0} x^n y^0 + \binom{n}{1} x^{n-1} y^1 + \binom{n}{2} x^{n-2} y^2 + \cdots + \binom{n}{n-1} x^1 y^{n-1} + \binom{n}{n} x^0 y^n,$$

## Venn Diagram

A Venn diagram is used to show/visualize the possible relations among a collection of sets.



Pictures from wikipedia



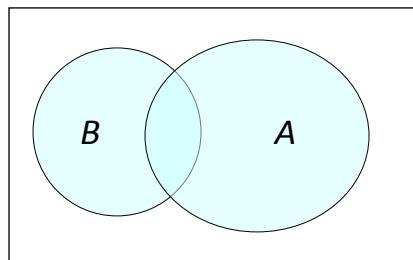
John Venn  
(1834-1923)



## Union and Intersection

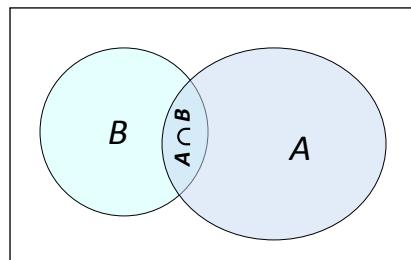
The *union of sets A and B* is the set of those elements that are either in A or in B, or in both.

$$A \cup B \triangleq \{x \mid x \in A \vee x \in B\}$$



The *intersection of the sets A and B* is the set of all elements that are in both A and B.

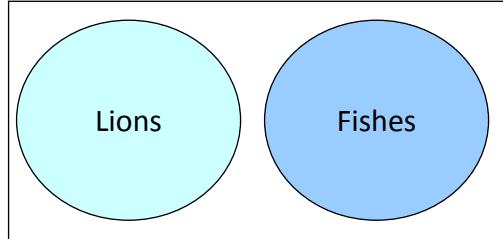
$$A \cap B \triangleq \{x \mid x \in A \wedge x \in B\}$$



## Disjoint Sets

Sets  $A$  and  $B$  are *disjoint* iff  $A \cap B = \emptyset$

- $|A \cap B| = 0$



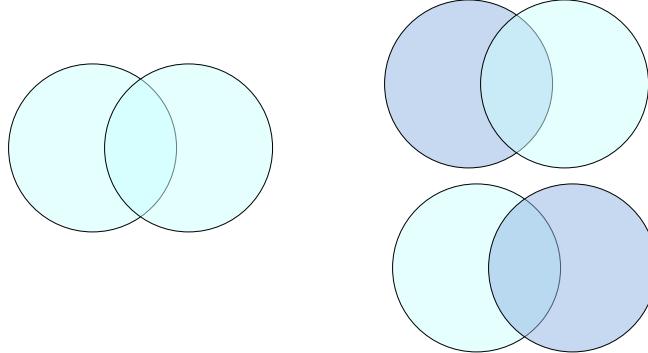
$$\text{Lions} \cap \text{Fishes} = \emptyset$$

© photographer



## Cardinality Of Union

$$|A \cup B| = |A| + |B| - |A \cap B|$$



**Definition 30.** The **union** of the sets  $A$  and  $B$  is by definition

$$A \cup B = \{x, x \in A \vee x \in B\}.$$

The **intersection** of the sets  $A$  and  $B$  is by definition

$$A \cap B = \{x, x \in A \wedge x \in B\}.$$

When the intersection of  $A$  and  $B$  is empty, we say that  $A$  and  $B$  are **disjoint**.

The cardinality of the union and intersection of the sets  $A$  and  $B$  are related by:

$$|A \cup B| = |A| + |B| - |A \cap B|.$$

This is true, because to count the number of elements in  $A \cup B$ , we start by counting those in  $A$ , and then add those in  $B$ . If  $A$  and  $B$  were disjoint, then we are done, otherwise, we have double counted those in both sets, so we must subtract those in  $A \cap B$ .

**Definition 31.** The **difference of  $A$  and  $B$** , also called **complement of  $B$  with respect to  $A$**  is the set containing elements that are in  $B$  but not in  $B$ :

$$A - B = \{x, x \in A \wedge x \notin B\}.$$

The **complement of  $A$**  is the complement of  $A$  with respect to the universe  $U$ :

$$\bar{A} = U - A = \{x, x \notin A\}.$$

The universe  $U$  is the set that serves as a framework for all our set computations, the biggest set in which all the other sets we are interested in lie. Note that  $\bar{\bar{A}} = A$ .

**Definition 32.** The **Cartesian product**  $A \times B$  of the sets  $A$  and  $B$  is the set of all ordered pairs  $(a, b)$ , where  $a \in A, b \in B$ :

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

**Example 46.** Take  $A = \{1, 2\}$ ,  $B = \{x, y, z\}$ . Then

$$A \times B = \{(a, b), a \in \{1, 2\} \wedge b \in \{x, y, z\}\}$$

thus  $a$  can be either 1 or 2, and for each of these 2 values,  $b$  can be either  $x$ ,  $y$  or  $z$ :

$$A \times B = \{(1, x), (1, y), (1, z), (2, x), (2, y), (2, z)\}.$$

Note that  $A \times B \neq B \times A$ , and that a Cartesian product can be formed from  $n$  sets  $A_1, \dots, A_n$ , which is denoted by  $A_1 \times A_2 \times \dots \times A_n$ .

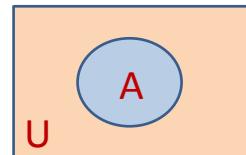
## Set Difference & Complement

The difference of A and B (or complement of B with respect to A) is the set containing those elements that are *in A but not in B*.

$$A - B \triangleq \{x \mid x \in A \wedge x \notin B\}$$

The *complement of A* is the complement of A with respect to U.

$$\bar{A} = U - A \triangleq \{x \mid x \notin A\}$$



## Cartesian Product

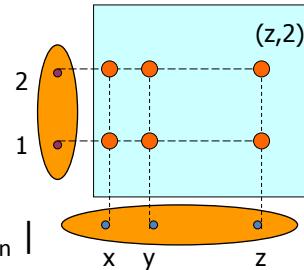
The *Cartesian product*  $A \times B$  of the sets A and B is the set of all *ordered pairs*  $(a, b)$  where  $a \in A$  and  $b \in B$ .

$$A \times B \triangleq \{(a, b) \mid a \in A \wedge b \in B\}$$



René Descartes  
(1596-1650)

- Example:  $A = \{1, 2\}$ ,  $B = \{x, y, z\}$   
 $A \times B = \{(1, x), (1, y), (1, z), (2, x), (2, y), (2, z)\}$   
 $B \times A = \{(x, 1), (x, 2), (y, 1), (y, 2), (z, 1), (z, 2)\}$
- In general:  $A_1 \times A_2 \times \dots \times A_n \triangleq \{(a_1, a_2, \dots, a_n) \mid a_i \in A_i \text{ for } i=1, 2, \dots, n\}$
- $|A_1 \times A_2 \times \dots \times A_n| = |A_1| |A_2| \dots |A_n|$



Picture from wikipedia

**Definition 33.** A collection of nonempty sets  $\{A_1, \dots, A_n\}$  is a **partition** of a set  $A$  if and only if

1.  $A = A_1 \cup A_2 \cup \dots \cup A_n$
2. and  $A_1, \dots, A_n$  are mutually disjoint:  $A_i \cap A_j = \emptyset$ ,  $i \neq j$ ,  $i, j = 1, 2, \dots, n$ .

**Example 47.** Consider  $A = \mathbb{Z}$ ,  $A_1 = \{\text{even numbers}\}$ ,  $A_2 = \{\text{odd numbers}\}$ . Then  $A_1, A_2$  form a partition of  $A$ .

We next derive a series of set identities:

$$A \cap \bar{B} = A - B.$$

By Definition 31,  $A - B = \{x, x \in A \wedge x \notin B\}$ . Then  $A \cap \bar{B} = \{x, x \in A \wedge x \in \bar{B}\}$ , but by the definition of  $\bar{B}$ ,  $A \cap \bar{B} = \{x, x \in A \wedge x \notin B\}$ , which completes the proof.

We have the set theoretic version of De Morgan's law:

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

We have  $\overline{A \cap B} = \{x, x \notin A \cap B\} = \{x, \neg(x \in A \wedge x \in B)\}$ , and using the usual De Morgan's law, we get  $\overline{A \cap B} = \{x, x \notin A \vee x \notin B\}$  as desired.

Applying de Morgan's law on  $\overline{A \cap \bar{B}}$ , and  $\bar{B} = B$  we get:

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

Recall that  $U$  denotes the universe set, the one to which belongs all the sets that we are manipulating. In particular,  $A \subset U$ . We have

$$A \cup \emptyset = A, A \cap U = A, A \cup U = U, A \cap \emptyset = \emptyset, A \cup A = A, A \cap A = A.$$

Furthermore, the order in which  $\cup$  or  $\cap$  is done does not matter:

$$A \cup B = B \cup A, A \cap B = B \cap A, A \cup (B \cup C) = (A \cup B) \cup C, A \cap (B \cap C) = (A \cap B) \cap C.$$

Distributive laws hold as well:

$$A \cap (B \cup C) = (A \cap B) \cup (A \cap C), A \cup (B \cap C) = (A \cup B) \cap (A \cup C).$$

For example,  $A \cap (B \cup C) = \{x, x \in A \wedge (x \in B \vee x \in C)\}$  and we can apply the distribute law from propositional logic to get the desired result. And finally

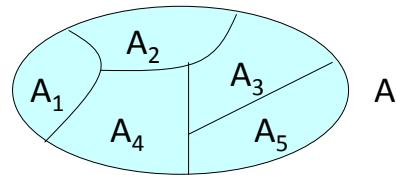
$$A \cup (A \cap B) = A, A \cap (A \cup B) = A.$$

This follows from the fact that  $A \cap B$  is a subset of  $A$ , while  $A$  is a subset of  $A \cup B$ .

## Partition

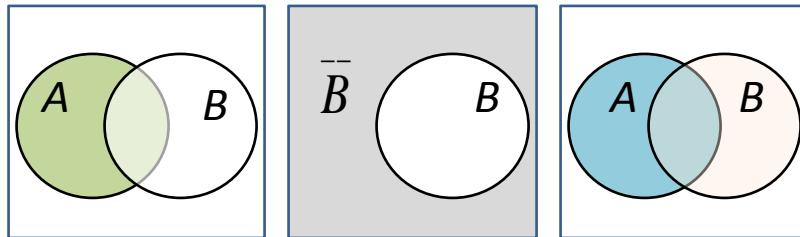
A collection of nonempty sets  $\{A_1, A_2, \dots, A_n\}$  is a **partition** of a set  $A$ , iff  $A = A_1 \cup A_2 \cup \dots \cup A_n$  and  $A_1, A_2, \dots, A_n$  are *mutually disjoint*, i.e.

$$A_i \cap A_j = \emptyset \text{ for all } i, j = 1, 2, \dots, n, \text{ and } i \neq j.$$



## Set Identities

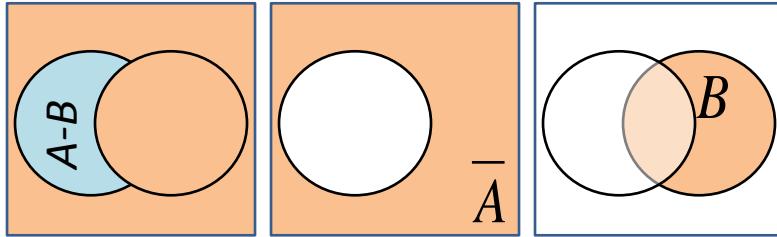
$$A \cap \bar{B} = A - B$$



Compare  $A \cap \bar{B}$  with  $A - B = \{x \mid x \in A \wedge x \notin B\}$   
(not a formal proof)

## Set Identities

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



- Consider  $\overline{A - B} = A \cap \overline{B}$
- Apply DeMorgan's Law  $\overline{X \cap Y} = \overline{X} \cup \overline{Y}$  with  $X=A$  and  $Y=\overline{B}$

## Set Identities

### Identity

$$A \cup \emptyset = A$$

$$A \cap U = A$$

$$A \cup U = U$$

$$A \cap \emptyset = \emptyset$$

$$A \cup A = A$$

$$A \cap A = A$$

$$\overline{\overline{A}} = A$$

### Name

Identity laws

Domination laws

Idempotent laws

Double Complement laws

## Set Identities

Identity	Name
$A \cup B = B \cup A$	
$A \cap B = B \cap A$	Commutative laws
$A \cup (B \cup C) = (A \cup B) \cup C$	
$A \cap (B \cap C) = (A \cap B) \cap C$	Associative laws
$A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$	
$A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$	Distributive laws
$\overline{A \cup B} = \overline{A} \cap \overline{B}$	
$\overline{A \cap B} = \overline{A} \cup \overline{B}$	De Morgan's laws

## Set Identities

Identity	Name
$A \cup (A \cap B) = A$	
$A \cap (A \cup B) = A$	Absorption laws
$A - B = A \cap \overline{B}$	Alternate Representation for set difference

Suppose that you want to prove that two sets  $A$  and  $B$  are equal. We will discuss 3 possible methods to do so:

1. Double inclusion:  $A \subseteq B$  and  $B \subseteq A$ .
2. Set identities.
3. Membership tables.

**Example 48.** To show that  $(B - A) \cup (C - A) = (B \cup C) - A$ , we show the double inclusion.

- Take an element  $x \in (B - A) \cup (C - A)$ , then either  $x \in (B - A)$ , or  $x \in (C - A)$ . Then  $x \in B \wedge x \notin A$ , or  $x \in C \wedge x \notin B$ . Then either way,  $x \in B \cup C \wedge x \notin A$ , that is  $x \in (B \cup C) - A$ , and  $(B - A) \cup (C - A) \subseteq (B \cup C) - A$  is shown.
- Now take an element  $x \in (B \cup C) - A$ , that is  $x \in B \cup C$  but  $x \notin A$ . Then  $x \in B$  and not in  $A$ , or  $x \in C$  and not in  $A$ . Then  $x \in B - A$  or  $x \in C - A$ . Thus either way,  $x \in (B - A) \cup (C - A)$ , which shows that  $(B - A) \cup (C - A) \supseteq (B \cup C) - A$

**Example 49.** We show that  $(A - B) - (B - C) = A - B$  using set identities.

$$\begin{aligned} (A - B) - (B - C) &= (A - B) \cap \overline{(B - C)} \\ &= (A \cap \bar{B}) \cap (\bar{B} \cap \bar{C}) \\ &= (A \cap \bar{B}) \cap (\bar{B} \cup C) \\ &= [(A \cap \bar{B}) \cap \bar{B}] \cup [(A \cap \bar{B}) \cap C] \end{aligned}$$

where the third equality is De Morgan's law, and the 4th one is distributivity. We also notice that the first term can be simplified to get  $(A \cap \bar{B})$ . We then apply distributivity again:

$$(A \cap \bar{B}) \cup [(A \cap \bar{B}) \cap C] = [A \cup [(A \cap \bar{B}) \cap C]] \cap [\bar{B} \cup [(A \cap \bar{B}) \cap C]].$$

Since  $(A \cap \bar{B}) \cap C$  is a subset of  $A$ , then the first term is  $A$ . Similarly, since  $(A \cap \bar{B}) \cap C$  is a subset of  $\bar{B}$ , the second term is  $\bar{B}$ . Therefore

$$(A - B) - (B - C) = A \cap \bar{B} = A - B.$$

The third method is a membership table, where columns of the table represent different set expressions, and rows take combinations of memberships in constituent sets: 1 means membership, and 0 non-membership. For two sets to be equal, they need to have identical columns.

## Proving Set Equality

- Recall. Two sets are **equal if and only if** they contain exactly the same elements, i.e.,  **$A \subseteq B$  and  $B \subseteq A$**
- Three methods to prove set equality:
  - Show that each set is a subset of the other
  - Apply set identities theorems
  - Use membership table

## Each Others' Subset

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

For any  $x \in LHS$ ,  $x \in (B-A)$  or  $x \in (C-A)$  [or both].

$$\begin{aligned}
 \text{when } x \in B - A &\implies (x \in B) \wedge (x \notin A) \\
 &\implies (x \in B \cup C) \wedge (x \notin A) \\
 &\implies x \in (B \cup C) - A \\
 \text{when } x \in C - A &\implies (x \in C) \wedge (x \notin A) \\
 &\implies (x \in B \cup C) \wedge (x \notin A) \\
 &\implies x \in (B \cup C) - A
 \end{aligned}$$

Therefore,  $LHS \subseteq RHS$

## Each Others' Subset

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

For any  $x \in \text{RHS}$ ,  $x \in (B \cup C)$  and  $x \notin A$ .

when  $x \in B$  and  $x \notin A$

$$\begin{aligned} (x \in B) \wedge (x \notin A) &\implies x \in B - A \\ &\implies x \in (B - A) \cup (C - A) \end{aligned}$$

when  $x \in C$  and  $x \notin A$ ,

$$\begin{aligned} (x \in C) \wedge (x \notin A) &\implies x \in C - A \\ &\implies x \in (B - A) \cup (C - A) \end{aligned}$$

Therefore,  $\text{RHS} \subseteq \text{LHS}$

With  $\text{LHS} \subseteq \text{RHS}$  and  $\text{RHS} \subseteq \text{LHS}$ , we can conclude that  $\text{LHS} = \text{RHS}$

## Using Set Identities

Show that  $(A-B)-(B-C)=A-B$

$$\begin{aligned} (A-B)-(B-C) &= (A \cap \overline{B}) \cap (\overline{B} \cap \overline{C}) && \text{(By alternate representation for set difference)} \\ &= (A \cap \overline{B}) \cap (\overline{B} \cup C) && \text{(By De Morgan's laws)} \\ &= [(A \cap \overline{B}) \cap \overline{B}] \cup [(A \cap \overline{B}) \cap C] && \text{(By Distributive laws)} \\ &= [A \cap (\overline{B} \cap \overline{B})] \cup [A \cap (\overline{B} \cap C)] && \text{(By Associative laws)} \\ &= (A \cap \overline{B}) \cup [A \cap (\overline{B} \cap C)] && \text{(By Idempotent laws)} \\ &= A \cap [\overline{B} \cup (\overline{B} \cap C)] && \text{(By Distributive laws)} \\ &= A \cap \overline{B} && \text{(By Absorption laws)} \\ &= A - B && \text{(By the alternate representation for set difference)} \end{aligned}$$

## Using Membership Tables

Similar to truth table (in propositional logic)

- Columns for different set expressions
- Rows for all combinations of memberships in constituent sets
- “1” = *membership*, “0” = *non-membership*
- Two sets are equal, iff they have **identical columns**

Prove that  $(A \cup B) - B = A - B$

A	B	$A \cup B$	$(A \cup B) - B$	$A - B$
0	0	0	0	0
0	1	1	0	0
1	0	1	1	1
1	1	1	0	0

**Example 50.** To prove  $(A \cup B) - B = A - B$ , we create a table

A	B	$A \cup B$	$(A \cup B) - B$	$A - B$
0	0			
0	1			
1	0			
1	1			

The first row, if  $x$  is not in  $A$  and not in  $B$ , it will not be in any of the sets, therefore the first row contains only zeroes. If  $x$  is only in  $B$ , then it belongs to  $A \cup B$ , but not in the others, since  $B$  is removed. So the second row has only a 1 in  $A \cup B$ . Then if  $x$  is only in  $A$ , it belongs to all the three sets. Finally, if  $x$  is in both  $A$  and  $B$ , it is in their intersection, therefore it belongs to  $A \cup B$ , but not in the 2 others, since  $B$  is removed.

## Exercises for Chapter 4

**Exercise 33.** 1. Show that

$$\binom{n}{k} + \binom{n}{k-1} = \binom{n+1}{k}$$

for  $1 \leq k \leq l$ , where by definition

$$\binom{n}{k} = \frac{n!}{k!(n-k)!}, \quad n! = n \cdot (n-1) \cdot (n-2) \cdots 2 \cdot 1.$$

2. Prove by mathematical induction that

$$(x+y)^n = \sum_{k=0}^n \binom{n}{k} x^{n-k} y^k.$$

You will need 1. for this!

3. Deduce that the cardinality of the power set  $P(S)$  of a finite set  $S$  with  $n$  elements is  $2^n$ .

**Exercise 34.** Let  $P(C)$  denote the power set of  $C$ . Given  $A = \{1, 2\}$  and  $B = \{2, 3\}$ , determine:

$$P(A \cap B), \quad P(A), \quad P(A \cup B), \quad P(A \times B).$$

**Exercise 35.** Prove by contradiction that for two sets  $A$  and  $B$

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

**Exercise 36.** Let  $P(C)$  denote the power set of  $C$ . Prove that for two sets  $A$  and  $B$

$$P(A) = P(B) \iff A = B.$$

**Exercise 37.** Let  $P(C)$  denote the power set of  $C$ . Prove that for two sets  $A$  and  $B$

$$P(A) \subseteq P(B) \iff A \subseteq B.$$

**Exercise 38.** Show that the empty set is a subset of all non-null sets.

**Exercise 39.** Show that for two sets  $A$  and  $B$

$$A \neq B \equiv \exists x[(x \in A \wedge x \notin B) \vee (x \in B \wedge x \notin A)].$$

**Exercise 40.** Prove that for the sets  $A, B, C, D$

$$(A \times B) \cup (C \times D) \subseteq (A \cup C) \times (B \cup D).$$

Does equality hold?

**Exercise 41.** Does the equality

$$(A_1 \cup A_2) \times (B_1 \cup B_2) = (A_1 \times B_1) \cup (A_2 \times B_2)$$

hold?

**Exercise 42.** For all sets  $A, B, C$ , prove that

$$\overline{(A - B) - (B - C)} = \bar{A} \cup B.$$

using set identities.

**Exercise 43.** This exercise is more difficult. For all sets  $A$  and  $B$ , prove  $(A \cup B) \cap \overline{A \cap B} = (A - B) \cup (B - A)$  by showing that each side of the equation is a subset of the other.

**Exercise 44.** The symmetric difference of  $A$  and  $B$ , denoted by  $A \oplus B$ , is the set containing those elements in either  $A$  or  $B$ , but not in both  $A$  and  $B$ .

1. Prove that  $(A \oplus B) \oplus B = A$  by showing that each side of the equation is a subset of the other.
2. Prove that  $(A \oplus B) \oplus B = A$  using a membership table.

**Exercise 45.** In a fruit feast among 200 students, 88 chose to eat durians, 73 ate mangoes, and 46 ate litchis. 34 of them had eaten both durians and mangoes, 16 had eaten durians and litchis, and 12 had eaten mangoes and litchis, while 5 had eaten all 3 fruits. Determine, how many of the 200 students ate none of the 3 fruits, and how many ate only mangoes?

# Chapter 5

## Combinatorics

*“I think you’re begging the question,” said Haydock, “and I can see looming ahead one of those terrible exercises in probability where six men have white hats and six men have black hats and you have to work it out by mathematics how likely it is that the hats will get mixed up and in what proportion. If you start thinking about things like that, you would go round the bend. Let me assure you of that!” (Agatha Christie, *The Mirror Crackd*)*

This chapter is dedicated to combinatorics, which refers broadly to different ways of counting objects.

Suppose for example that you have two slots to be filled, and for the first slot, there are  $n_1$  choices, while there are  $n_2$  choices for the second slot. How many ways are there to fill up both slots? Well, for the first slot, we have  $n_1$  choices, now for each of these, we still have  $n_2$  choices, for a total of  $n_1 n_2$  choices.

**Example 51.** Suppose you have 3 choices for the main course, and 2 choices for the dessert. How many choices of menus do you have? Well, you can pick any of the 3 main courses, so 3 choices here. Next, for main course 1, you can choose 2 desserts, then for main course 2, you can choose 2 desserts, and finally for main course 3, you can still choose 2 desserts, which makes it a total of 6 menus.

More generally, if there are  $k$  slots, and  $n_1$  choices for the 1st slot,  $n_2$  choices for the second slot, until  $n_k$  choices for the  $k$ th slot, we get a total of  $n_1 \cdot n_2 \cdots n_k$  choices.

## Principle Of Counting

- There are two slots to be filled, there are  $n_1$  choices for slot 1 and  $n_2$  choices for slot 2.
  - E.g., you have 3 choices for the main course and 2 choices for dessert.
- The total number of unique choices to fill the slots is  $n_1 n_2$
- In general:  $n_1, n_2, \dots, n_k$  choices for  $k$ -slots
- $n_1 * n_2 * \dots * n_k$  ways
  - (cardinality of the cartesian product of sets)



© image from <http://key-boxes.com/page/2>

## Cardinality Of Power Set

- Consider a set A with n elements.
- Each of these n elements are either in a subset of A or not: 2 choices
  - Such a choice needs to be made for each of the n elements
- Thus  $2 * 2 * \dots * 2 = 2^n$  choices.
  - We saw another derivation using the Binomial theorem.

Now  $n_1 \cdot n_2 \cdots n_k$  is also the cardinality of the cartesian product of  $k$  sets, where the set  $i$  has  $n_i$  elements.

**Example 52.** Suppose you have 3 choices for the main course, and 2 choices for the dessert. How many choices of menus do you have? An alternative way to view this question is to explicitly list all the choices:

main course 1, dessert 1	main course 1, dessert 2
main course 2, dessert 1	main course 2, dessert 2
main course 3, dessert 1	main course 3, dessert 2

This makes a total of 6 menus. You notice that when we list all the options, we get a cartesian product of two sets, the set  $\{ \text{main course 1, main course 2, main course 3} \}$ , and the set  $\{ \text{dessert 1, dessert 2} \}$ .

We recall that given a set  $A$ , its power set  $P(A)$  is the set of all subsets of  $A$ . We already saw in the previous chapter that the cardinality of  $P(A)$  is  $2^n$ . Here is another way of proving it. Write  $A = \{a_1, \dots, a_n\}$ . Now list all subsets of  $A$ , and to each subset, associate a binary vector of length  $n$ , where every coefficient is either 0 or 1: the first coefficient is 1 if  $a_1$  is in the subset, and 0 otherwise, similarly, the second coefficient is 1 if  $a_2$  is in the subset, and 0 otherwise, and so on and so forth. Since every element is in a given subset or not, we do obtain all possible binary vectors of length  $n$ , and there are  $2^n$  of them.

**Example 53.** Consider the set  $A = \{1, 2\}$ .

$\emptyset$	00
{1}	10
{2}	01
$A$	11

Now suppose that there are  $n$  elements, to be put in  $r$  slots. If elements can be repeated, we are in the scenario we have just seen, and there are  $n^r$  choices. Now if elements cannot be repeated, then, we have  $n$  choices for the first slot,  $n - 1$  choices for the second slot, and so on and so forth, until  $n - (r - 1)$  choices for the last slot. We thus get

$$n(n - 1)(n - 2) \cdots (n - r + 1). \tag{5.1}$$

This is for example what happens when picking cards from a deck of cards, once the cards are not put back in the deck.

## Filling $r$ Slots With $n$ Choices

- There are  $n$  elements, with which to fill  $r$  slots.
- When elements **can be** repeated:
  - Using the principle of counting:  $n * n * \dots * n = n^r$  choices
- When elements **cannot be** repeated:
  - $n$  choices for first slot,
  - $n-1$  choices for second slot,...
  - $n-(r-1)$  choices for last slot
  - In total:  $n(n-1)(n-2)\dots(n-r+1)$  choices



- E.g., sequence of choice of cards from a deck of cards

© to the artist

## Permutation: $P(n,r)$

A **permutation** is an arrangement of all or part of a set of objects, *with* regard to the order of the arrangement.

Number of permutations of  $n$  objects taken  $r$  at a time:

$$P(n,r) = n(n-1)(n-2)\dots(n-r+1) = n!/(n-r)!$$

where  $n! = n * (n-1) * (n-2) * \dots * 2 * 1$  (called  **$n$  factorial**).

If  $r = n$ , we notice that all the  $n$  elements are attributed to the  $n$  slots, which gives a *permutation* of the  $n$  elements. This also shows that the number of permutations of  $n$  elements is

$$n(n-1)(n-2) \cdots 2 \cdot 1 = n!.$$

The above scenario, when there are  $n$  elements but only  $r$  slots, and elements cannot be repeated, is called *permutations of  $n$  objects taken  $r$  at a time*, that is an arrangement where ordering matters, and the number  $P(n, r)$  of such permutations is

$$P(n, r) = n(n-1)(n-2) \cdots (n-r+1) = \frac{n!}{(n-r)!}.$$

Permutations can be defined also if some of the objects are repeated, that is, we still have  $n$  elements, but  $n = k_1 + k_2 + \cdots + k_r$ , that is, there are  $k_1$  elements of type 1,  $k_2$  elements of type 2, until  $k_r$  elements of type  $r$ . How many permutations do we have in this case? To count this, we can proceed as follows: place  $k_1$  elements out of  $n$  places, then place  $k_2$  elements in  $n - k_1$  places, etc until you place  $k_r$  elements in the remaining places. This means that we have

$$\binom{n}{k_1} \binom{n-k_1}{k_2} \cdots \binom{k_r}{k_r} \quad (5.2)$$

where we recall that  $\binom{n}{k}$  counts the number of ways of choosing  $k$  elements out of  $n$ .

We also call  $\binom{n}{r}$  a *combination*, that is a say of selecting objectings without considering the order of the selection. We have that

$$\binom{n}{r} = C(n, r) = \frac{n!}{r!(n-r)!}.$$

Indeed, recall that when we had  $r$  slots and  $n$  objects, we have  $P(n, r)$  ways of placing the objects, where

$$P(n, r) = n(n-1)(n-2) \cdots (n-r+1) = \frac{n!}{(n-r)!}.$$

Now we still have  $r$  slots,  $n$  objects to choose from, but this time, we do not care about the ordering, and there are  $r!$  possible ordering for each combination:

$$r!C(n, r) = P(n, r).$$

## Permutation

In general: The number of distinguishable permutations from a collection of objects, where first object appears (repeats)  $k_1$  times, second object  $k_2$  times, ... for  $r$  distinct objects:

$$n!/(k_1! k_2! \dots k_r!)$$

## Combination: $C(n,r)$ or $\binom{n}{r}$

A **combination** is a selection of all or part of a set of objects, *without* regard to the order in which objects are selected.

E.g. Team of 4 people from a group of 10

Number of combinations of  $n$  objects taken  $r$  at a time

$$\binom{n}{r} = C(n,r) = n!/r!(n-r)!$$

- There are  $r!$  possible orderings within each combination
- So  $r! C(n,r) = P(n,r)$  by definition of permutation

**Example 54.** From a committee of 8 people, in how many ways can you choose

- a chair and a vice-chair (one person cannot hold more than one position): it is  $P(8, 2) = \frac{8!}{6!} = 8 \cdot 7$ . Indeed, once the chair is chosen (8 choices), we have 7 choices for the vice chair.
- a subcommittee of 2 people: it is  $C(8, 2) = \frac{8!}{2!6!} = 28$ . Indeed, in this case, any 2 persons among the 8 people will do, irrespectively of the ordering. This means that we choose person 1 with person 2, person 1 with person 3, etc until person 1 with person 8 (7 possibilities), or person 2 with person 3, etc until person 2 with person 8 (6 possibilities), or person 3 with person 4, ..., person 3 with person 8 (5 possibilities), and by continuing the list, we get  $7 + 6 + 5 + 4 + 3 + 2 + 1 = 28$ .

This example also illustrates that  $r!C(n, r) = P(n, r)$ . Indeed, we know that  $P(8, 2)$  takes into account the ordering, therefore, if say person 1 is chair and person 2 is vice chair, it counts for 1 choice, while it counts for 2 choices for the subcommittee, since person 1 and person 2, and person 2 and person 1, represent the same subcommittee.

We now finish the computations of (5.2):

$$\binom{n}{k_1} \binom{n-k_1}{k_2} \cdots \binom{k_r}{k_r} = \frac{n!}{k_1!(n-k_1)!} \frac{(n-k_1)!}{k_2!(n-k_1-k_2)!} \cdots \frac{(n-k_1-\cdots-k_{r-1})!}{k_r!}$$

where we notice that we can cancel out numerator and denominator to finally obtain

$$\frac{n!}{k_1!k_2!\cdots k_r!}.$$

Some experiments may not have a deterministic outcome, e.g., tossing a coin, or throwing a dice, in which case, different techniques are needed.

**Definition 34.** The set of possible outcomes of random trial is called a [sample space](#).

For example, if you toss a coin, the set of possible outcomes is { head, tail }. If you toss a coin twice, you may be interested in the number of heads, for which the set of possible outcomes is {0,1,2 }, or in the actually sequence of heads/tails, for which the set of possible outcomes is { head head, head tail, tail head, tail tail }, or in whether the two twosses matched, for an outcome which belongs to { yes, no }.

## Example

From a committee of 8 people, in how many ways can you choose:

- a chair and vice-chair (one person cannot hold more than one position?)
  - $P(8,2)$
  - a subcommittee of 2 people?
  - $C(8,2)$
- 

## Sample Space

- Some experiments may not have a deterministic outcome, e.g., tossing of a coin, throwing a dice.

The set of possible outcomes of a random trial is called the *sample space*.

- E.g, for coin toss, the sample space is {Head, Tail}
  - Two coins tossed
    - Record the number of heads {0,1,2}
    - Record sequence of heads/tails {HH, HT, TH, TT}
    - Record if the two tosses matched {Yes, No}
-

**Definition 35.** An **event** is a set of outcomes of a random trial, or in other words, a subset of the sample space.

There are different types of events.

- An *impossible event* refers to an outcome which is actually not possible, that is, which does not belong to the sample space. For example, if you roll a dice, the number that you will get belongs to the sample space  $\{1, 2, 3, 4, 5, 6\}$ , therefore, roll a dice and obtain 7 is an impossible event.
- A *certain event* is on the contrary an event which always happens, which corresponds to the whole sample space, such as: roll a dice, and get a number which is less than 10. Any number in the sample space  $\{1, 2, 3, 4, 5, 6\}$  is less than 10.
- Two events are said to be *mutually exclusive* when they cannot happen at the same time. For example, you roll a dice, the events “get an even number” and “get a number divisible by 5” are mutually exclusive, they cannot happen both at the same time.

We now want to define the notion of *probability of an event*. Informally, this represents the likelihood that an event will occur, or the ratio of the number of wanted outcomes, by the number of possible outcomes. For example, if you toss a fair coin, the probability of a head (that is, the likelihood that a head happens), is  $1/2$ . This comes from the fact that you have two possible outcomes, head and tail, and both of them are equally likely (so 2 at the denominator). On the numerator, you want a head (so 1 at the numerator).

You may also want to repeat a given random experiment, say  $n$  times. When you look at a particular event  $E$ , over the  $n$  times, it will appear  $n_E$  times. Therefore the *frequency* of occurrence of  $E$  over these  $n$  trials is

$$f_E = \frac{n_E}{n}.$$

The notion of frequency is different from that of probability of an event, but it is also related. For example, if you toss a coin 10 times, and you are interested in counting the number of occurrences of the event  $E = \text{head}$ , maybe you get  $n_E = 6$ , therefore  $f_E = 6/10$ .

## Events

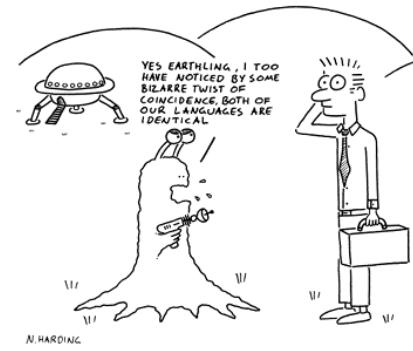
An **event** is a set of outcomes of a random trial (a subset of the sample space).

- **Impossible event** : event not in the sample space.  
– E.g. Roll a dice, get more than 6 on top.
- **Certain event**: event which is the whole sample space.  
– E.g. Roll a dice, get a number less than 10.
- **Mutually exclusive events**: events which cannot happen at the same time.  
– E.g. Roll a dice, get an even number which is divisible by 5.



## Probability Of Events

- The **likelihood** that an event will occur  
– When tossing a fair coin, the probability of a head is 0.5.
- **Empirical interpretation**
  - Repeat an experiment  $n$  times
  - An event  $E$  occurs  $n_E$  times
  - Then  $f_E = n_E/n$  is the **frequency** of occurrence of  $E$  in  $n$  trials
    - such a frequency measure is not necessarily a proof of the probability of event  $E$ , but can be an “indicator”



© to the artist

The frequency is then slightly different than what you may have expected, since with a probability of  $1/2$  and 10 trials, you may have expected the number of heads to be 5. Therefore the frequency is an “indicator”.

We next introduce a formal definition of probability.

**Definition 36.** A **probability space** is a sample space  $A$ , with some events  $E_i \in A$ , and probability measure  $P$ , satisfying 3 axioms:

1.  $0 \leq P(E_i) \leq 1$  for every event  $E_i \in A$ .
2.  $P(A) = 1$  and  $P(\emptyset) = 0$ .
3.  $P(E_1 \cup E_2 \cup \dots \cup E_k) = P(E_1) + P(E_2) + \dots + P(E_k)$  when  $E_i \cap E_j = \emptyset$ ,  $i \neq j$ .

**Example 55.** Suppose that the random trial is tossing a fair coin: The sample space is { head, tail }. Because the coin is fair (unbiased), head and tail are equally likely events ( $P(\text{head})=P(\text{tail})$ ), and they are mutually exclusive ( $P(\text{head})+P(\text{tail})=1$ ). Therefore, we obtain formally what our intuition told us, namely, that  $P(\text{head})=P(\text{tail})=1/2$ .

More generally, if there are  $n$  equally likely mutually exclusive (and spanning the sample space) events  $E_1, \dots, E_n$ , then we get:  $P(E_1) = \dots = P(E_n)$  and  $P(E_1) + \dots + P(E_n) = 1$  therefore  $P(E_i) = 1/n$  for all every event  $E_i$ .

**Example 56.** Suppose that you choose 4 cards from a deck of 52 cards. What is the probability of getting 4 kings? There are  $C(52, 4)$  ways of choosing 4 cards. Now all are equally likely, but only one of these choices has all four kings. Therefore the probability of getting all 4 kings is

$$\frac{1}{C(52, 4)}.$$

## Axioms For A Probability Space

For a sample space  $A$ , and some event  $E_i \in A$ :

- Axiom 1:  $0 \leq p(E_i) \leq 1$  for every event  $E_i \in A$
  - Axiom 2:  $P(A) = 1$  and  $P(\emptyset) = 0$   
(Event A must happen everytime the experiment is done since every event belongs to A)
  - Axiom 3:  $P(E_1 \cup E_2 \cup \dots \cup E_k) = P(E_1) + P(E_2) + \dots + P(E_k)$   
when  $E_i \cap E_j = \emptyset$  for  $i \neq j$   
(i.e.,  $E_i$  and  $E_j$  are *mutually exclusive*)
- 

## Equally Likely Outcomes: Symmetry

- Example: Tossing an unbiased coin
    - Sample space is  $\{H,T\}$  ( $H=$ head,  $T=$ tail)
    - No reason for H T to occur more often than T:  $P(H)=P(T)$
    - H and T are mutually exclusive events:  $P(H)+P(T)=1$
    - Thus  $P(H)=P(T)=0.5$
  - In general: if there are  $n$  equally likely mutually exclusive (and spanning the sample space) events, then the elementary probability of each such event is  $1/n$
-

## Exercises for Chapter 5

**Exercise 46.** A set menu proposes 2 choices of starters, 3 choices of main dishes, and 2 choices of desserts. How many possible set menus are available?

**Exercise 47.** Consider the set  $A = \{1, 2, 3\}$ ,  $P(A)$  = power set of  $A$ .

- Compute the cardinality of  $P(A)$  using the binomial theorem approach.
- Compute the cardinality of  $P(A)$  using the counting approach.

**Exercise 48.** • If you toss two coins, what is the probability of getting 2 heads?

- If you toss three coins, what is the probability of getting exactly 2 heads?

**Exercise 49.** Ten fair coins are tossed together. What is the probability that there were at least seven heads?

**Exercise 50.** Snow white is going to a party with the seven dwarves. Each of the eight of them owns a red dress and a blue dress. If each of them is likely to choose either colored dress randomly and independently of the other's choices, what is the chance that all of them go to the party wearing the same colored dress?

## Examples for Chapter 5

Let us recall one storage application, given in Example 8.

**Example 57.** Suppose you want to store 200GB of data, and the shop is selling disks of 100 GB each. Then you can buy 4 disks, store half of your data (let us call it  $D_1$ ) on disk 1, the other half (say  $D_2$ ) on disk 2, then copy the content of disk 1 to disk 3, and the content of disk 2 to disk 4. We get thus the following data allocation:

$$\text{disk 1 : } D_1, \text{ disk 2 : } D_2, \text{ disk 3 : } D_1, \text{ disk 4 : } D_2.$$

This strategy does tolerate any one disk failure. It does not tolerate any two disks failures, as we already know. However, what is the probability of actually losing the data in case of two disks failures? The number of patterns with 2 failures is  $C(4, 2) = 6$ , while the number of patterns creating a data loss is 2, therefore the probability is

$$\frac{2}{6} = \frac{1}{3}.$$

Our next example is a famous puzzle, **the Hat Problem**. It is famous because it made the news!

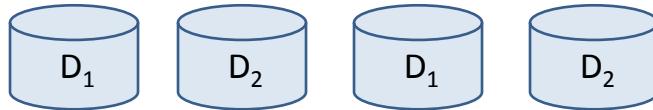
<http://www.nytimes.com/2001/04/10/science/why-mathematicians-now-care-about-their-hat-color.html>

Here is how the puzzle goes:  $N$  players enter a room. A red or blue hat is placed on each person's head. The color of each hat is determined by a fair coin toss, that is  $P(\text{blue})=P(\text{red})$ , with the outcome of one coin toss having no effect on the others. Each person can see the other players' hats but not his own.

- The players must simultaneously guess the color of their own hats or pass.
- The group shares a prize if at least one player guesses correctly, and no player guesses incorrectly.
- No communication of any sort is allowed, except for any initial strategy session before the game begins. For a strategy to be acceptable, it must always result in at least one prisoner making a guess.

## Example (II)

$$D = (D_1, D_2)$$



- If one hard disk fails, your data is safe.
- What is the probability of losing your data in case two hard disks fail?

$$\frac{2}{C(4,2)} = \frac{2}{6} = \frac{1}{3}$$


---

## The Hat Problem



N players enter a room

- A red or blue hat is placed on each person's head.  
–  $P(\text{red})=P(\text{blue})=1/2$ , independently.
  - Each player sees the other hats but not his own.  
– The players must simultaneously guess the color of their own hats or pass.
  - Win if at least one player guesses correctly and no players guess incorrectly.  
– No communication is allowed, except for any initial strategy session before the game begins.
-

**Example 58.** If  $N = 1$ , there is only one player, thus all he can do is make a guess, the probability of winning is simply  $1/2$ .

**Example 59.** If  $N = 2$ , there are two players. If both players guess randomly, their chance of winning is only  $1/4$ , because they try to guess one combination of colors among 4 possible combinations. A single player only making a guess is better, the probability of winning is then  $1/2$ .

Suppose now that  $N = 3$ . Let us assign the number 0 to a hat of blue color, and 1 to a hat of red color. There are 8 possible hat assignments:

$$000, 100, 010, 110, 001, 101, 011, 111.$$

Therefore, if a player sees two hats of the same color, he guesses the opposite color (this is an acceptable strategy, there must always be two hats of the same color). Otherwise he passes. This corresponds to

$$100, 010, 110, 001, 101, 011,$$

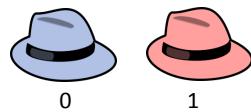
where only one player will see two hats of the same color, will guess his color to be the other one, and will be right with probability  $6/8 = 3/4$ . The game will be won with probability  $3/4 > 1/2$ .

In fact, we can prove that this strategy is optimal!

(1) First, we need to note that the number of correct guesses and the number of incorrect guesses are the same. The reason is because the probability of making a correct/incorrect guess is  $1/2$ . For example, when  $N = 2$ , when both players are making a guess, there is one chance of winning (2 correct guesses), and three chances of losing (2 wrong guesses, 1 wrong guess and 1 correct guess, twice). Thus the total number of correct guesses is 4, and the number of incorrect guesses is 4 as well! For another example, when  $N = 3$ , the proposed strategy comprises 6 correct guesses (one for each of the 6 wins), and 6 wrong guesses (3 incorrect guesses per each loss, and there are two losses). More precisely, for the correct guesses 100, 010, 001, correspond the 3 incorrect guesses 000, and for the correct guesses 011, 101, 110, correspond the 3 incorrect guesses 111.

(2) Suppose we could get a better strategy, where we get 7 wins instead of 6, and thus 1 loss. For a winning strategy, no incorrect guess can be made, therefore we need at least 7 right guesses, which means in turn, using the above argument, 7 wrong guesses. But this is not possible to have 7 wrong guesses in one loss, and only 3 players.

## The Hat Problem



000 100 010 110 001 101 011 111

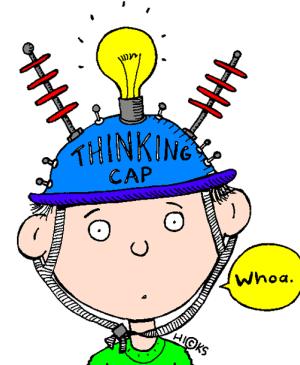
Strategy: If the other two guys have the same hat color, “guess the opposite”, if they have different colors, stay silent!

– *Chance of winning with this strategy:*  $3/8+3/8=0.75$

---

## The Hat Problem

- Optimal strategy?
  - Number of correct guesses = number of incorrect guesses
  - Better strategy: 7 wins & 1 loss
  - At least 7 correct guesses, impossible to have 7 incorrect guesses in one loss and 3 players
- 





# Chapter 6

## Linear Recurrences

*“Everything goes, everything comes back; eternally rolls the wheel of being.” (Friedrich Nietzsche)*

This chapter is dedicated to linear recurrences, a special type of equations that defines a sequence, that is a series of terms of the form

$$a_0, a_1, a_2, a_3, \dots, a_n, a_{n+1}, \dots$$

recursively, that is such that each term  $a_n$  is defined as a function of the preceding terms. A recursive linear recurrence must be accompanied by initial conditions, that is information about some of the first terms such as  $a_0$  or  $a_0, a_1$ .

**Example 60.** The Fibonacci sequence is defined by

$$f_n = f_{n-1} + f_{n-2}, \quad f_0 = 0, \quad f_1 = 1.$$

To compute  $f_2$ , we have

$$f_2 = f_1 + f_0 = 1.$$

To compute  $f_3$ , we have

$$f_3 = f_2 + f_1 = 1 + 1 = 2.$$

We will see in this chapter two methods to solve linear recurrences involving one or two preceding terms.

## Recurrence Relation

A *recurrence relation* is an equation that *recursively defines a sequence*, i.e., each term of the sequence is defined as a function of the preceding terms

A recursive formula must be accompanied by *initial conditions* (information about the beginning of the sequence).

---

## Fibonacci Sequence

$$f_n = f_{n-1} + f_{n-2} \text{ with } f_0 = 0, f_1 = 1$$

- 0, 1, 1, 2, 3, 5, 8, 13, 21, 34, 55, ...



Leonardo  
Pisano Bigollo  
(c. 1170-c.  
1250)

---

Image from wikipedia

The first method is called *backtracking*, and consists of taking a linear recurrence defining  $a_n$ , and replace the terms  $a_{n-1}, a_{n-2}, \dots$  with the relation that defines  $a_n$ , but where  $n$  is replaced by  $n - 1, n - 2$ , etc. This is best illustrated on an example.

**Example 61.** Consider the linear recurrence

$$a_n = a_{n-1} + 3, \quad a_1 = 2.$$

Then

$$\begin{aligned} a_{n-1} &= a_{n-2} + 3 \\ a_{n-2} &= a_{n-3} + 3 \\ a_{n-3} &= a_{n-4} + 3 \end{aligned}$$

and so on and so forth. Therefore

$$\begin{aligned} a_n &= a_{n-1} + 3 \\ &= (a_{n-2} + 3) + 3 = a_{n-2} + 2 \cdot 3 \\ &= (a_{n-3} + 3) + 6 = a_{n-3} + 3 \cdot 3 \\ &= \dots \\ &= a_1 + 3(n-1). \end{aligned}$$

The last equality follows because a generic term is of the form  $a_{n-i} + 3i$ , therefore when  $n - i = 1$ ,  $i = n - 1$ . By plugging the initial condition, we conclude

$$a_n = 2 + 3(n-1).$$

Once the solution has been found, you may wonder how to check whether this is the right answer. One way to do it is by proving it by induction!

**Example 62.** Let us provide a proof by induction for the above example. Define  $P(n) = "a_n = 2 + 3(n-1)"$ . Then  $P(1) = "a_1 = 2"$ , which is the initial condition, is true. Suppose  $P(k) = "a_k = 2 + 3(k-1)"$  is true. We want to prove  $P(k+1)$ .

$$\begin{aligned} a_{k+1} &= a_k + 3 \\ &= 2 + 3(k-1) + 3 \\ &= 2 + 3k \end{aligned}$$

as desired.

## Solving Recurrence Relation

- **Backtracking** is a technique for finding explicit formula for recurrence relation
  - E.g., say  $a_n = a_{n-1} + 3$  and  $a_1 = 2$
  - $$\begin{aligned} a_n &= a_{n-1} + 3 = (a_{n-2} + 3) + 3 = a_{n-2} + 2 \cdot 3 \\ &= (a_{n-3} + 3) + 2 \cdot 3 = a_{n-3} + 3 \cdot 3 \\ &= (a_{n-2} + 3) + 3 = a_{n-2} + 2 \cdot 3 \\ &\dots \\ &= a_1 + (n-1) \cdot 3 \\ a_n &= 2 + (n-1) \cdot 3 \end{aligned}$$
- 

## Homogeneous Relation Of Degree $d$

A **linear homogeneous relation** of degree  $d$  is of the form  

$$a_n = c_1 a_{n-1} + c_2 a_{n-2} + \dots + c_d a_{n-d}$$

### Examples

- The Fibonacci sequence
- The relation:  $a_n = 2a_{n-1}$  (degree 1)
- But **not** the relation:  $a_n = 2a_{n-1} + 1$

The **characteristic equation** of the above relation is

$$x^d = c_1 x^{d-1} + c_2 x^{d-2} + \dots + c_d$$


---

**Definition 37.** A *linear homogeneous relation* of degree  $d$  is of the form

$$a_n = c_1 a_{n-1} + c_2 a_{n-2} + \dots + c_d a_{n-d}.$$

Its *characteristic equation* is

$$x^d = c_1 x^{d-1} + c_2 x^{d-2} + \dots + c_d.$$

The characteristic equation is obtained from  $a_n = c_1 a_{n-1} + c_2 a_{n-2} + \dots + c_d a_{n-d}$  by replacing  $a_i$  by  $x^i$ :

$$x^n - c_1 x^{n-1} - c_2 x^{n-2} - \dots - c_d x^{n-d} = 0,$$

then factor out  $x^{n-d}$  to get

$$x^{n-d}(x^d - c_1 x^{d-1} - c_2 x^{d-2} - \dots - c_d) = 0.$$

**Example 63.** The Fibonacci sequence  $f_n = f_{n-1} + f_{n-2}$  is a homogeneous relation. Let us compute its characteristic equation:

$$x^n - x^{n-1} - x^{n-2} = 0 \iff x^{n-2}(x^2 - x - 1) = 0$$

therefore  $x^2 - x - 1 = 0$  is the characteristic equation.

Let us focus on quadratic characteristic equations, that is of the form

$$x^2 - c_1 x - c_2 = 0$$

which corresponds to linear recurrences of the form

$$a_n = c_1 a_{n-1} + c_2 a_{n-2}.$$

Suppose that  $x^2 - c_1 x - c_2 = 0$  has two distinct real roots  $s_1, s_2$ , then

$$s_1^2 - c_1 s_1 - c_2 = 0, \quad s_2^2 - c_1 s_2 - c_2 = 0.$$

Therefore

$$s_1^n - c_1 s_1^{n-1} - c_2 s_1^{n-2} = 0, \quad s_2^n - c_1 s_2^{n-1} - c_2 s_2^{n-2} = 0$$

and we have that if  $s$  is a solution of  $x^2 - c_1 x - c_2 = 0$  then  $s^n$  is a solution of  $a_n = c_1 a_{n-1} + c_2 a_{n-2}$ . This tells us that solutions of  $a_n$  are composed of  $s_1^n, s_2^n$ . Note the term "composed" is used, because if a sequence  $a'_n$  also satisfies the recurrence of  $a_n$ , then  $a_n + a'_n$  satisfies the recurrence of  $a_n$  as well, as does multiples of  $a_n$  (see Exercise 53).

## Theorem

If the *characteristic equation*  $x^2 - c_1x - c_2 = 0$  (of the recurrence relation  $a_n = c_1a_{n-1} + c_2a_{n-2}$ ) has

- two distinct roots  $s_1, s_2$ , then the explicit formula for the sequence  $a_n$  is

$$u.s_1^n + v.s_2^n$$

- a single root  $s$ , then the explicit formula for  $a_n$  is

$$u.s^n + v.n.s^n$$

where  $u$  &  $v$  are determined by initial conditions.

## Example

Determine the number of bit strings (i.e., comprising 0/1s) of length  $n$  that contains *no adjacent 0s*.

- $C_n$  = this number of bit strings
- A binary string with no adjacent 0s is constructed by
  - Adding “1” to any string  $w$  of length  $n-1$  satisfying the condition, or
  - Adding “10” to any string  $v$  of length  $n-2$  satisfying the condition
- Thus  $C_n = C_{n-1} + C_{n-2}$  where  $C_1=2$  (0,1),  $C_2=3$  (01, 10, 11)

This means that the final solution is really a composition of  $s_1^n, s_2^n$ , namely a solution for  $a_n = c_1 a_{n-1} + c_2 a_{n-2}$  is given by

$$a_n = us_1^n + vs_2^n,$$

where  $u, v$  depend on the initial conditions (that is on  $a_0, a_1$ ).

Suppose now that  $x^2 - c_1x - c_2 = 0$  has one double real root  $s$ , that is  $x^2 - c_1x - c_2 = (x - s)^2$ . Then

$$s^2 - c_1s - c_2 = 0,$$

and  $s^n$  is a solution of  $a_n = c_1 a_{n-1} + c_2 a_{n-2}$  as for the case of two distinct roots. We obtained the characteristic equation from

$$x^n - c_1 x^{n-1} - c_2 x^{n-2} = x^{n-2}(x^2 - c_1x - c_2) = 0.$$

If  $s$  is a root of this equation, then  $s$  is a root of its derivative:

$$nx^{n-1} - c_1(n-1)x^{n-2} - c_2(n-2)x^{n-3} = 0.$$

Therefore  $s$  satisfies both

$$s^n = c_1 s^{n-1} + c_2 s^{n-2}, \quad ns^n = c_1(n-1)s^{n-1} + c_2(n-2)s^{n-2}. \quad (6.1)$$

If we combine  $s^n$  and  $ns^n$ , as we did for  $s_1^n$  and  $s_2^n$ , we get

$$a_n = us^n + vns^n,$$

and  $a_{n+1} = us^{n+1} + v(n+1)s^{n+1}$ . We are left to check

$$\begin{aligned} c_1 a_{n+1} + c_2 a_{n+2} &= c_1(us^{n+1} + v(n+1)s^{n+1}) + c_2(us^{n+2} + v(n+2)s^{n+2}) \\ &= u[c_1 s^{n+1} + c_2 s^{n+2}] + v[c_1(n+1)s^{n+1} + c_2(n+2)s^{n+2}] \\ &= us^{n+2} + v(n+2)s^{n+2} \\ &= a_{n+2} \end{aligned}$$

using (6.1), which is consistent with our recurrence relation.

**Example 64.** Suppose we want to determine the number of bit strings of length  $n$  that contains no adjacent zeroes. We denote this number by  $C_n$ . We first observe that there are two ways of obtaining such sequences from a smaller sequence. One may take any string:

- of length  $n - 1$  satisfying the condition and add a 1 (one cannot add a 0, since the sequence may finish by 0),
- of length  $n - 2$  satisfying the condition and add 10 (one cannot add 00, or 01 since the string could finish by 0, and 11 is included above).

## Example

- Now solve  $C_n = C_{n-1} + C_{n-2}$  where  $C_1=2, C_2=3$
- Characteristic equation:  $x^2 - x - 1 = 0$
- Its roots are  $(1 + \sqrt{5})/2$   
 $(1 - \sqrt{5})/2$
- Thus

Recall roots of quadratic eqn.  
 $a.x^2 + b.x + c = 0$   
 $x = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}$ ,

$$C_n = u \cdot \left(\frac{1+\sqrt{5}}{2}\right)^n + v \cdot \left(\frac{1-\sqrt{5}}{2}\right)^n$$

## Example

Initial conditions give us:

$$C_1 = u \cdot \left(\frac{1+\sqrt{5}}{2}\right) + v \cdot \left(\frac{1-\sqrt{5}}{2}\right) = 2$$

$$\text{i.e., } \frac{u+v}{2} + \frac{(u-v)\sqrt{5}}{2} = 2$$

$$C_2 = u \cdot \left(\frac{1+\sqrt{5}}{2}\right)^2 + v \cdot \left(\frac{1-\sqrt{5}}{2}\right)^2 = 3$$

$$\text{i.e., } \frac{3(u+v)}{2} + \frac{(u-v)\sqrt{5}}{2} = 3$$

Solving, we get

$$u = \frac{\sqrt{5}+3}{2\sqrt{5}}$$

$$v = \frac{\sqrt{5}-3}{2\sqrt{5}}$$

Therefore the linear recurrence involved is

$$C_n = C_{n-1} + C_{n-2}.$$

The characteristic equation is obtained from

$$x^n - x^{n-1} - x^{n-2} = x^{n-2}(x^2 - x - 1) = 0,$$

it is  $x^2 - x - 1 = 0$ . To find its roots, we compute

$$\frac{1 \pm \sqrt{1 - 4(-1)}}{2} = \frac{1 \pm \sqrt{5}}{2}.$$

The solution is then

$$C_n = u\left(\frac{1+\sqrt{5}}{2}\right)^n + v\left(\frac{1-\sqrt{5}}{2}\right)^n.$$

We are left to find  $u, v$  based on the initial conditions. They are  $C_1 = 2$  (the strings are 0 and 1), while  $C_2 = 3$  (the strings are 11,01,10). This gives us two equations for two unknowns:

$$\begin{aligned} u\left(\frac{1+\sqrt{5}}{2}\right) + v\left(\frac{1-\sqrt{5}}{2}\right) &= 2 \\ u\left(\frac{1+\sqrt{5}}{2}\right)^2 + v\left(\frac{1-\sqrt{5}}{2}\right)^2 &= 3. \end{aligned}$$

which can be simplified to

$$\begin{aligned} \frac{u+v}{2} + \sqrt{5}\frac{u-v}{2} &= 2 \\ 3\frac{u+v}{2} + \sqrt{5}\frac{u-v}{2} &= 3. \end{aligned}$$

Set  $a = (u+v)/2$ ,  $b = \sqrt{5}(u-v)/2$ . We need to solve

$$a+b=2, 3a+b=3 \Rightarrow b=\frac{3}{2}, a=\frac{1}{2}.$$

Thus

$$v = \frac{\sqrt{5}-3}{2\sqrt{5}}, u = 1 - \frac{\sqrt{5}-3}{2\sqrt{5}} = \frac{\sqrt{5}+3}{2\sqrt{5}}.$$

## Exercises for Chapter 6

**Exercise 51.** Consider the linear recurrence  $a_n = 2a_{n-1} - a_{n-2}$  with initial conditions  $a_1 = 3$ ,  $a_0 = 0$ .

- Solve it using the backtracking method.
- Solve it using the characteristic equation.

**Exercise 52.** What is the solution of the recurrence relation

$$a_n = a_{n-1} + 2a_{n-2}$$

with  $a_0 = 2$  and  $a_1 = 7$ ?

**Exercise 53.** Let  $a_n = c_1a_{n-1} + c_2a_{n-2} + \dots + c_k a_{n-k}$  be a linear homogeneous recurrence. Assume both sequences  $a_n, a'_n$  satisfy this linear homogeneous recurrence. Show that  $a_n + a'_n$  and  $\alpha a_n$  also satisfy it, for  $\alpha$  some constant.

## Examples for Chapter 6

Linear Recurrence relations are often useful to analyze algorithms, such as divide-and-conquer algorithms. We will illustrate this using the game of Hanoi tower.

In this game, the goal is to move  $n$  disks ranked from the largest at the bottom to the smallest on top from one post to another. The only permitted action is to remove the top disk from a post and drop it onto another post. The rule is that a larger disk can never lie above a smaller disk on any post. When  $n = 3$  disks, the Hanoi tower game can be solved in 7 steps. But say one would like to know how many steps it would take to solve it for  $n = 50$  disks, how could this be figured out?

The method is to derive a linear recurrence relation, and then to solve it. To find a linear recurrence relation, notice that to solve the Hanoi tower game for  $n = 3$  disks, the following steps are done:

1. Solve a Hanoi tower game for  $n = 2$  disks,
2. Move the largest disk,
3. Solve another Hanoi tower game for  $n = 2$  disks.

In fact, this is true in general, which yields the linear recurrence

$$T_n = 2T_{n-1} + 1,$$

where  $T_n$  denotes the number of steps for  $n$  disks.

We solve this linear recurrence using backtracking. Note that

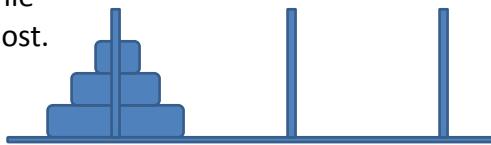
$$T_{n-1} = 2T_{n-2} + 1, \quad T_{n-2} = 2T_{n-3} + 1, \quad T_{n-3} = 2T_{n-4} + 1, \dots$$

Then

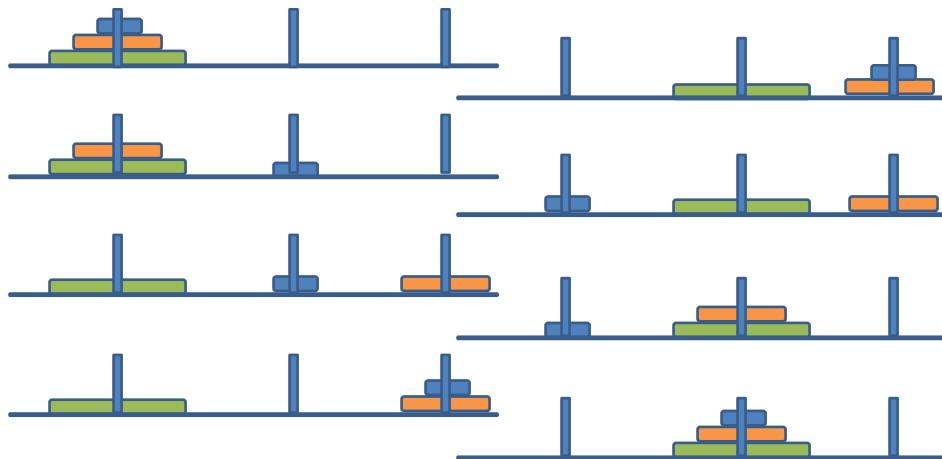
$$\begin{aligned} T_n &= 2T_{n-1} + 1 \\ &= 2(2T_{n-2} + 1) = 4T_{n-2} + 3 \\ &= 4(2T_{n-3} + 1) + 3 = 8T_{n-3} + 7 \\ &= 8(2T_{n-4} + 1) + 7 = 16T_{n-4} + 15 \\ &= \dots \end{aligned}$$

## Hanoi Tower

- Goal: move all  $n$  disks in the same order, but on a different post.
- Only permitted action: remove the top disk from a post and drop it onto another post.
- Rule: a larger disk can never lie above a smaller disk on any post.

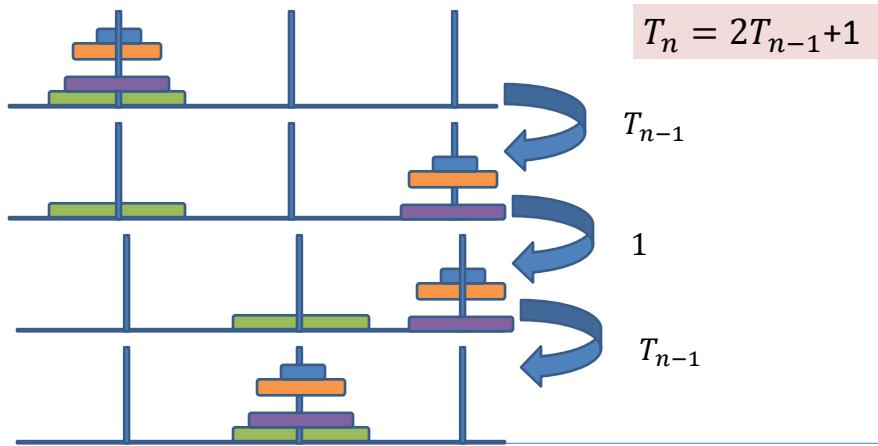


## Hanoi Tower ( $n=3$ )



## Find a Recurrence

- $T_n$  = minimum number of steps needed to move an  $n$ -disk tower from one post to another



## Backtracking

$$T_1 = 1 \quad T_n = 2T_{n-1} + 1$$

## Backtracking

$$T_1 = 1 \quad T_n = 2T_{n-1} + 1$$

$$\begin{aligned} T_n &= 2T_{n-1} + 1 = 2(2T_{n-2} + 1) + 1 && \leftarrow 3 \\ &= 4T_{n-2} + 3 = 4(2T_{n-3} + 1) + 3 && \leftarrow 7 \\ &= 8T_{n-3} + 7 && \leftarrow 15 \end{aligned}$$

$$T_n = 2^n - 1$$

## Induction

- $P(n) = ``T_n = 2^n - 1''$
- Basis step:  $P(1) = ``T_1 = 1''$
- Inductive step: suppose  $P(n)$  is true.
- To show,  $P(n+1)$ .
- $T_{n+1} = 2T_n + 1 = 2(2^n - 1) + 1 = 2^{n+1} - 1$

We notice that a general term is of the form

$$2^i T_{n-i} + (2^i - 1),$$

therefore when  $n - i = 1$ ,  $i = n - 1$ , and we get

$$T_n = 2^{n-1} T_1 + (2^{n-1} - 1)$$

with  $T_1 = 1$ . Thus finally

$$T_n = 2^n - 1.$$

Once a solution has been found by backtracking, it is advised to confirm that the solution is sound, by performing a proof by induction. Here  $P(n) = "T_n = 2^n - 1"$ . The basis step is  $P(1) = "T_1 = 2^1 - 1 = 1"$  which is true. The induction step is to assume that  $P(k) = "T_k = 2^k - 1"$  is true. We need to prove  $P(k + 1)$ . But

$$T_{k+1} = 2T_k + 1 = 2 \cdot 2^k + 1 = 2^{k+1}.$$

as desired!



# Chapter 7

## Complex Numbers

*“I tell you, with complex numbers you can do anything.” (J. Derbyshire, Prime Obsession: Bernhard Riemann and the Greatest Unsolved Problem in Mathematics)*

So far, the largest set of numbers we have seen is that of real numbers. This will change in this chapter, with the introduction of complex numbers. They were introduced around 1545 by the mathematician Gerolamo Cardano, in order to obtain closed form expressions for roots of cubic polynomial equations, which need square roots of negative numbers, which do not exist (meaning do not exist if only real numbers are considered).

Recall that there is no real number  $z$  such that  $z^2 = -1$  (the square of a real number is always positive).

**Definition 38.** Define an **imaginary unit  $i$**  such that

$$i^2 = -1.$$

More generally, we define an **imaginary number  $z$**  to be of the form

$$z = iy, \quad y \in \mathbb{R}$$

therefore

$$z^2 = (iy)^2 = i^2y^2 = -y^2.$$

We may often use the notation  $i = \sqrt{-1}$  to emphasize that  $i$  is not an index and is instead the imaginary unit, however, it is really best to avoid writing negative roots to avoid confusion in the computations...

**Example 65.** We have  $(4i)^2 = 16i^2 = -16$ .

## Definition of $i$

- There is no real number  $z$  such that  $z^2 = -1$ .

Define an **imaginary unit  $i$**  (denoted also  $j$ ) such that

$$i^2 = -1 \text{ (that is) } i = \sqrt{-1}.$$

Define an **imaginary number  $z$**  to be of the form

$$z = iy = y\sqrt{-1}$$

for  $y$  any real number.

- Then the imaginary number  $z$  is such that

$$z^2 = -y^2.$$


---

## Computations with $i$

To avoid confusion, write  $i$  instead of a negative root!!

**Powers of  $i$ :**  $i^0 = 1, i^1 = i, i^2 = -1, i^3 = -i, i^4 = 1.$

**Inverse of  $i$ :**  $i z = 1$  thus  $z = i^{-1} = -i.$

Powers of  $z=iy$ ,  $y$  real:  $z^2 = i^2y^2 = -y^2, z^3 = -iy^3$

Example:  $(4i)^2 = -16.$

---

Let us compute some powers of  $i$ :

$$i^0 = 1, \quad i^1 = i, \quad i^2 = -1, \quad i^3 = i(i^2) = -i, \quad i^4 = (i^2)^2 = (-1)^2 = 1.$$

To compute the inverse of  $i$ , we need to find an imaginary number  $z$  such that

$$iz = 1 \rightarrow z = i^{-1} = -i.$$

Correspondingly we get powers of imaginary numbers of the form  $z = iy$ ,  $y \in \mathbb{R}$ :

$$z^0 = 1, \quad z^1 = z, \quad z^2 = (iy)^2 = -y^2, \quad z^3 = iy(-y^2) = -iy^3, \quad z^4 = (z^2)^2 = (-y^2)^2 = y^4.$$

**Definition 39.** We define a **complex number  $z$**  to be of the form

$$z = a + ib, \quad a, b \in \mathbb{R}.$$

We call  $a$ , and write  $\Re(z)$  the **real part** of  $z$ . We call  $b$ , and write  $\Im(z)$  the **imaginary part** of  $z$ .

**Example 66.** Take  $z = 3+5i$ , then its real part is  $\Re(z) = 3$  and its imaginary part is  $\Im(z) = 5$ .

**Definition 40.** We define the **conjugate** of a complex number  $z = a + ib$ ,  $a, b \in \mathbb{R}$ , to be

$$\bar{z} = a - ib, \quad a, b \in \mathbb{R}.$$

Note that for  $z = a + ib$ ,  $a, b \in \mathbb{R}$ :

$$\overline{\bar{z}} = \overline{a - ib} = a + ib = z.$$

Also

$$z\bar{z} = (a + ib)(a - ib) = (a - ib)(a + ib) = a^2 + iab - iab - (i^2)b^2 = a^2 + b^2.$$

**Example 67.** Take  $z = 3+5i$ , then its conjugate is  $\bar{z} = 3-5i$  and  $z\bar{z} = 9+25$ .

We discuss next how to visualize a complex number geometrically. To do so, we associate to a complex number  $z = a + ib$  a pair comprising its real part and its imaginary part, namely  $(a, b)$ . Now we see  $(a, b)$  in the two-dimensional real plane, where the real part corresponds to the  $x$ -axis, while the imaginary part corresponds to the  $y$ -axis.

## Complex Numbers

A **complex number**  $z$  is of the form

$$z = a + bi, \text{ where } a, b \text{ are real numbers.}$$

We call  $a = \operatorname{Re}(z)$  the **real part of  $z$** , and  $b = \operatorname{Im}(z)$  the **imaginary part of  $z$** .

Example:  $3+5i$ ,  $\operatorname{Re}(3+5i)=3$ ,  $\operatorname{Im}(3+5i)=5$ .

---

## Conjugate

For  $z=a+ib$  a complex number, its **conjugate**  $\bar{z}$  is

$$\bar{z} = a - ib.$$

- We have  $\bar{\bar{z}}=z$ .
- We have  $z\bar{z} = \bar{z}z = a^2 + b^2$ .

Example:  $\overline{3+5i} = 3-5i$ ,  $(\overline{3+5i})(3+5i) = 9+25$ .

---

We next define operations on complex numbers.

**Addition.** Take two complex numbers  $a + ib$ ,  $c + id$ , their sum is given by

$$(a + ib) + (c + id) = (a + bc) + i(b + d).$$

**Multiplication.** Take two complex numbers  $a + ib$ ,  $c + id$ , their product is given by

$$(a + ib)(c + id) = ac + aid + ibc - bd = (ac - bd) + i(ad + bc).$$

**Division.** Take two complex numbers  $a + ib$ ,  $c + id$ , their ratio is given by

$$\frac{a + ib}{c + id}.$$

To be able to handle this case, the technique is to multiply both the numerator and denominator by the conjugate of the denominator, namely  $c - id$ :

$$\frac{a + ib}{c + id} = \frac{(a + ib)(c - id)}{(c + id)(c - id)} = \frac{(a + ib)(c - id)}{c^2 + d^2}.$$

Now we are in familiar territories since the denominator is a real number:

$$\frac{a + ib}{c + id} = \frac{(ac - iad + ibc + db)}{c^2 + d^2} = \frac{ac + db}{c^2 + d^2} + i \frac{bc - ad}{c^2 + d^2}.$$

We saw above that a complex number  $z = a + ib$  can be represented as the point  $(a, b)$  in the 2-dimensional real plane. Recall that a point on a circle of radius one centered around the origin can be written as  $(\cos \theta, \sin \theta)$ , where  $\theta$  is the angle from the  $x$ -axis counter clockwise. Now to be able to write similarly an arbitrary point in the 2-dimensional real plane, note that this point will be on a circle of radius  $r$ , where  $r$  is the length of the vector  $(a, b)$ . Therefore, this point can be written as

$$(a, b) = (r \cos \theta, r \sin \theta),$$

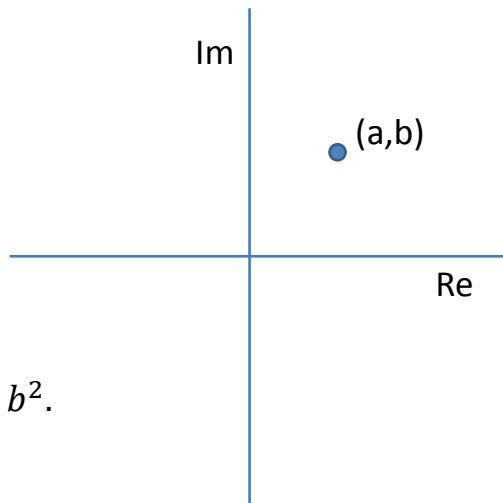
which are called [polar coordinates](#). Alternatively, we may write

$$z = a + ib = r \cos \theta + i \sin \theta.$$

We call  $r$  the [modulus](#) of  $z$ , and  $\theta$  the [argument](#) of  $z$ .

## Complex Plane

Geometrically:  
 $a+ib$ , or  $(a,b)$



Also  
 $\overline{(a + ib)}(a + ib) = a^2 + b^2.$

## Complex Numbers Operations

**Addition:**  $(a + ib) + (c + id) = (a + c) + i(b + d)$

**Multiplication:**  $(a + ib)(c + id) = (ac - bd) + i(ad + bc)$

**Division:**  $\frac{a+ib}{c+id} = \frac{(a+ib)(c-id)}{(c+id)(c+id)}$

Once we have a complex number written in polar coordinates, we may use Euler formula to write the complex number in exponential form.

[Euler formula](#) says that

$$e^{i\theta} = \cos \theta + i \sin \theta$$

for  $\theta$  any real number (in radians).

*Proof.* The proof that is provided now is the most classical one, and it relies on Taylor series for the different quantities involved:

$$\begin{aligned} e^x &= \sum_{n=0}^{\infty} \frac{x^n}{n!} = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \dots \\ \sin x &= \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n+1)!} x^{2n+1} = x - \frac{x^3}{3!} + \frac{x^5}{5!} + \dots \\ \cos x &= \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n)!} x^{2n} = 1 - \frac{x^2}{2!} + \frac{x^4}{4!} + \dots \end{aligned}$$

Recall that a Taylor series for a function  $f(x)$  is

$$f(x) = f(a) + \frac{f'(a)}{1}(x-a) + \frac{f''(a)}{2!}(x-a)^2 + \dots$$

where  $a$  is taken to be zero in our case.

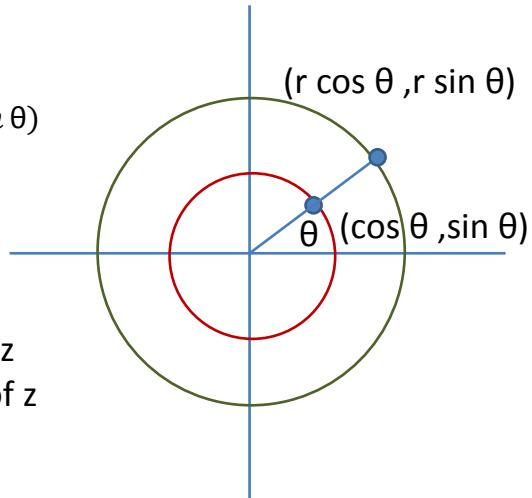
Thus

$$\begin{aligned} e^{ix} &= \sum_{n=0}^{\infty} \frac{(ix)^n}{n!} \\ &= \sum_{n=0}^{\infty} \frac{(ix)^{2n}}{(2n)!} + \sum_{n=0}^{\infty} \frac{(ix)^{2n+1}}{(2n+1)!} \\ &= \sum_{n=0}^{\infty} \frac{(i^2)^n}{(2n)!} x^{2n} + \sum_{n=0}^{\infty} \frac{i(i^2)^n}{(2n+1)!} x^{2n+1} \\ &= \cos x + i \sin x \end{aligned}$$

as needed. □

## Polar Coordinates

$$z = a + ib = r(\cos \theta + i \sin \theta)$$



$r$  is called the **modulus** of  $z$   
 $\theta$  is called the **argument** of  $z$

## Euler Formula

For  $\theta$  any real number (in radians)

$$e^{i\theta} = \cos \theta + i \sin \theta.$$

(Recall that  $2\pi$  radians = 360 degrees)

Euler identity:  $e^{i\pi} + 1 = 0.$



Leonhard Euler  
(1707-1783)

Image from wikipedia

As a corollary, we get [Euler identity](#):

$$e^{i\pi} + 1 = 0.$$

Indeed choose  $\theta$  to be  $\pi$  (=180 degrees), then

$$e^{i\pi} = \cos \pi + i \sin \pi = -1,$$

since  $\cos \pi = -1$  and  $\sin \pi = 0$ .

Thanks to Euler Formula, we can rewrite a complex number  $z$  as:

$$z = a + ib = r(\cos \theta + i \sin \theta) = re^{i\theta}.$$

Notice now that if we compute  $z\bar{z}$ , we get

$$z\bar{z} = (a + ib)(a - ib) = re^{i\theta}re^{-i\theta}$$

that is

$$z\bar{z} = a^2 + b^2 = r^2$$

therefore the modulus of  $z$ , denoted by  $|z|$ , satisfies

$$|z| = \sqrt{z\bar{z}} = \sqrt{a^2 + b^2} = r.$$

The argument (or phase) of  $z$  is  $\theta$ , let us try to express it as a function of  $a, b$ . For that, remember that given a right triangle (drawn on the  $x$ -axis) with angle  $\theta$  that goes from the  $x$ -axis counter clockwise, we have that  $\tan \theta = \frac{b}{a}$  (the ratio of the opposite side and the adjacent one). Therefore  $\theta = \tan^{-1} \frac{b}{a}$  where the range of  $\tan^{-1}$  is  $(-\pi, \pi]$ , and one has to be careful that there are special cases depending on the sign of  $a, b$ :

$$\arg(z) = \begin{cases} \tan^{-1} \frac{b}{a} & a > 0 \\ \tan^{-1} \frac{b}{a} + \pi & a < 0, b \geq 0 \\ \tan^{-1} \frac{b}{a} - \pi & a < 0, b < 0 \\ \frac{\pi}{2} & a = 0, b > 0 \\ \frac{-\pi}{2} & a = 0, b < 0 \\ \text{indeterminate} & a = 0, b = 0. \end{cases}$$

**Example 68.** Take  $z = 3 + 3\sqrt{3}i$ , then its modulus  $|z|$  is

$$|z| = \sqrt{3^2 + (3\sqrt{3})^2} = \sqrt{9 + 27} = 6.$$

Then for the phase

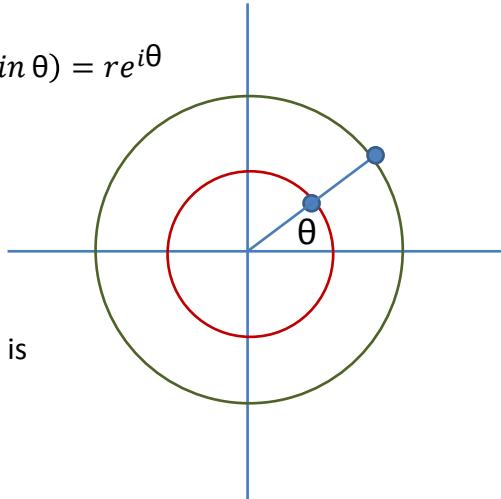
$$\arg z = \tan^{-1} \sqrt{3} = \frac{\pi}{3}.$$

## Converting among Forms

$$z = a + ib = r(\cos \theta + i \sin \theta) = re^{i\theta}$$

The **modulus**  $|z|$  of  $z$  is

$$|z| = \sqrt{z\bar{z}} = \sqrt{a^2 + b^2} = r.$$



The **argument** (or **phase**) of  $z$  is

$$\arg(z) = \theta = \tan^{-1} \frac{b}{a}.$$

## Example

$$z = 3 + 3\sqrt{3}i$$

- $|z| = \sqrt{9 + 27} = 6.$
- $\arg(z) = \tan^{-1} \frac{3\sqrt{3}}{3} = \tan^{-1} \sqrt{3} = \frac{\pi}{3}$

Thus the exponential form of  $z = 3 + 3\sqrt{3}i$  is  $6e^{i\frac{\pi}{3}}$

while its polar form is:  $6 \left( \cos \frac{\pi}{3} + i \sin \frac{\pi}{3} \right).$

Thus in exponential form, we get:

$$z = 3 + 3\sqrt{3}i = 6e^{i\frac{\pi}{3}}$$

while the polar form is:

$$z = 3 + 3\sqrt{3}i = 6e^{i\frac{\pi}{3}} = 6 \left( \cos \frac{\pi}{3} + i \sin \frac{\pi}{3} \right)$$

**Definition 41.** A [root of unity](#) is a complex number  $z$  such that  $z^n = 1$ .

Suppose we want to compute the 2nd roots of unity, we are looking for complex numbers  $z$  satisfying

$$z^2 = 1 \iff z^2 - 1 = 0 \iff (z - 1)(z + 1) = 0$$

and therefore there are two 2nd roots of unity: 1 and  $-1$ . Next we want to compute the 3rd roots of unity, we look at the equation

$$z^3 = 1 \iff z^3 - 1 = 0 \iff (z - 1)(z^2 + z + 1) = 0.$$

Thus 1 is a root, and we need to identify the other two (it is a polynomial of degree 3, therefore we expect three roots). Consider  $z = e^{2\pi i/3}$ , then

$$z^3 = (e^{2\pi i/3})^3 = e^{2\pi i} = 1$$

and  $z$  is indeed a 3rd root of unity. So is  $(e^{2\pi i/3})^2$  since

$$[(e^{2\pi i/3})^2]^3 = (e^{2\pi i})^2 = 1.$$

Note that apart 1, the other 2 roots are complex, which means that they do not exist in  $\mathbb{R}$ ! (so one says that there are no roots in the case of real roots). In general, the  $n$  roots of  $z^n = 1$  are

$$(e^{2\pi i/n})^k, \quad k = 1, \dots, n.$$

This is because

$$[(e^{2\pi i/n})^k]^n = (e^{2\pi i})^k = 1$$

thus we have  $n$  distinct solutions, and so we got all of them!

We may ask the same question with another real number than 1, namely, what are the roots of  $z^n = a$ , for  $a$  a real number. The roots are

$$\sqrt[n]{a}(e^{2\pi i/n})^k, \quad k = 1, \dots, n.$$

To check it, it is enough to compute the  $n$ th power, and see that we get indeed  $a$ :

$$(\sqrt[n]{a}(e^{2\pi i/n})^k)^n = a(e^{2\pi i})^k = a,$$

and we have found the  $n$  distinct roots.

## N<sup>th</sup> Roots of Unity

- What are the roots of  $z^2 = 1$ ?

1 and  $-1$ .

- What are the roots of  $z^3 = 1$ ?

$$e^{\frac{2\pi i}{3}}, \left(e^{\frac{2\pi i}{3}}\right)^2, 1$$

- What are the roots of  $z^n = 1$ ?

$$\left(e^{\frac{2\pi i}{n}}\right)^k, k = 1, \dots, n.$$


---

## N<sup>th</sup> Roots

- What are the roots of  $z^2 = 2$ ?

$\sqrt{2}$  and  $-\sqrt{2}$ .

- What are the roots of  $z^3 = 2$ ?

$$\sqrt[3]{2}e^{\frac{2\pi i}{3}}, \sqrt[3]{2}\left(e^{\frac{2\pi i}{3}}\right)^2, \sqrt[3]{2}$$

- What are the roots of  $z^n = a$ , for a real?

$$\sqrt[n]{a}\left(e^{\frac{2\pi i}{n}}\right)^k, k = 1, \dots, n.$$


---

## Exercises for Chapter 7

**Exercise 54.** Set  $i = \sqrt{-1}$ . Compute

$$i^5, \frac{1}{i^2}, \frac{1}{i^3}.$$

**Exercise 55.** Set  $i = \sqrt{-1}$ . Compute the real part and the imaginary part of

$$\frac{(1+2i)-(2+i)}{(2-i)(3+i)}.$$

**Exercise 56.** Set  $i = \sqrt{-1}$ . Compute  $d, e \in \mathbb{R}$  such that

$$4 - 6i + d = \frac{7}{i} + ei.$$

**Exercise 57.** For  $z_1, z_2 \in \mathbb{C}$ , prove that

- $\overline{z_1 + z_2} = \bar{z}_1 + \bar{z}_2$ .
- $\overline{z_1 \cdot z_2} = \bar{z}_1 \cdot \bar{z}_2$ .

**Exercise 58.** Consider the complex number  $z$  in polar form:  $z = re^{i\theta}$ . Express  $re^{-i\theta}$  as a function of  $z$ .

**Exercise 59.** Prove that

$$(\cos x + i \sin x)^n = \cos nx + i \sin nx,$$

for  $n$  an integer.

**Exercise 60.** Compute  $|e^{i\theta}|$ ,  $\theta \in \mathbb{R}$ .

**Exercise 61.** Prove the so-called triangle inequality:

$$|a + b| \leq |a| + |b|, \quad a, b \in \mathbb{C}.$$

**Exercise 62.** Compute the two roots of  $4i$ , that is

$$\sqrt{4i}.$$



# Chapter 8

## Linear Algebra

*“Algebra is generous; she often gives more than is asked of her.”*  
*(Jean D’Alembert)*

This chapter is called linear algebra, but what we will really see is the definition of a matrix, a few basic properties of matrices, and how to compute (reduced) row echelon form, with its applications. There is much more to linear algebra!

Let us start by defining a matrix.

**Definition 42.** A [matrix](#) is a rectangular array containing numbers, also called coefficients. We say that the matrix is an  $m \times n$  matrix to specify that the array comprises  $m$  rows and  $n$  columns. If the matrix is called  $A$ , we usually write its coefficients as  $a_{ij}$ , where  $i$  tells us that this coefficient is on the  $i$ th row, and  $j$  that it is on the  $j$ th column:

$$A = \begin{pmatrix} a_{11} & \dots & a_{1n} \\ a_{21} & \dots & a_{2n} \\ \vdots & & \\ a_{m1} & \dots & a_{mn} \end{pmatrix}$$

Once the shape of the matrix is given, we need to specify where the coefficients belong to, namely whether they are real numbers, complex numbers, integer numbers. They could also be integers modulo  $n$ .

**Definition 43.** A  $1 \times n$  matrix is called a [row vector](#). An  $m \times 1$  matrix is called a [column vector](#).

## Matrices

An  **$m \times n$  (real) matrix** is a rectangular array whose coefficients are (real) numbers.

$$A = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix}.$$

If  $m=1$  or if  $n=1$ , we call a  $1 \times n$  matrix a **row vector**, and an  $m \times 1$  matrix a **column vector**.

Example:

$A = \begin{pmatrix} 1 & 7 \\ 5 & 2 \end{pmatrix}$  is a  $2 \times 2$  real matrix,  $(2 \ 3)$  is a  $1 \times 2$  row vector.

## Matrix Addition

**Addition** of  $m \times n$  matrices is done componentwise:

$$\begin{aligned} & \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix} + \begin{pmatrix} b_{11} & \cdots & b_{1n} \\ \vdots & \ddots & \vdots \\ b_{m1} & \cdots & b_{mn} \end{pmatrix} \\ &= \begin{pmatrix} a_{11} + b_{11} & \cdots & a_{1n} + b_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} + b_{m1} & \cdots & a_{mn} + b_{mn} \end{pmatrix} \end{aligned}$$

Example:

$$A = \begin{pmatrix} 1 & 7 \\ 5 & 2 \end{pmatrix}, B = \begin{pmatrix} 1 & -2 \\ 0 & 3 \end{pmatrix} \text{ then } A + B = \begin{pmatrix} 2 & 5 \\ 5 & 5 \end{pmatrix}$$

**Example 69.** The matrix

$$A = \begin{pmatrix} 1 & 7 \\ 5 & 2 \end{pmatrix}$$

is a  $2 \times 2$  matrix with real coefficients, while  $(2, 3)$  is a  $1 \times 2$  row vector.

We are of course interested in performing operations on matrices. The easiest one is probably **matrix addition**. For this, we need two matrices  $A$  and  $B$  both with the same number of rows and columns, say, both of them are  $n \times m$  matrices:

$$A = \begin{pmatrix} a_{11} & \dots & a_{1n} \\ a_{21} & \dots & a_{2n} \\ \vdots & & \\ a_{m1} & \dots & a_{mn} \end{pmatrix}, \quad B = \begin{pmatrix} b_{11} & \dots & b_{1n} \\ b_{21} & \dots & b_{2n} \\ \vdots & & \\ b_{m1} & \dots & b_{mn} \end{pmatrix}.$$

Then  $A + B$  is computed componentwise, namely

$$A + B = \begin{pmatrix} a_{11} & \dots & a_{1n} \\ a_{21} & \dots & a_{2n} \\ \vdots & & \\ a_{m1} & \dots & a_{mn} \end{pmatrix} + \begin{pmatrix} b_{11} & \dots & b_{1n} \\ b_{21} & \dots & b_{2n} \\ \vdots & & \\ b_{m1} & \dots & b_{mn} \end{pmatrix} = \begin{pmatrix} a_{11} + b_{11} & \dots & a_{1n} + b_{1n} \\ a_{21} + b_{21} & \dots & a_{2n} + b_{2n} \\ \vdots & & \\ a_{m1} + b_{m1} & \dots & a_{mn} + b_{mn} \end{pmatrix}$$

**Example 70.** Consider the following two matrices:

$$A = \begin{pmatrix} 1 & 7 \\ 5 & 2 \end{pmatrix}, \quad B = \begin{pmatrix} 1 & -2 \\ 0 & 3 \end{pmatrix}$$

Then

$$A + B = \begin{pmatrix} 1 & 7 \\ 5 & 2 \end{pmatrix} + \begin{pmatrix} 1 & -2 \\ 0 & 3 \end{pmatrix} = \begin{pmatrix} 1+1 & 7-2 \\ 5 & 2+3 \end{pmatrix} = \begin{pmatrix} 2 & 5 \\ 5 & 5 \end{pmatrix}.$$

**Definition 44.** Given an  $m \times n$  matrix  $A$  its **transpose matrix**  $A^T$  is an  $n \times m$  matrix obtaining by interchanging the rows and columns of  $A$ :

$$A = \begin{pmatrix} a_{11} & \dots & a_{1n} \\ a_{21} & \dots & a_{2n} \\ \vdots & & \\ a_{m1} & \dots & a_{mn} \end{pmatrix}, \quad A^T = \begin{pmatrix} a_{11} & \dots & a_{m1} \\ a_{12} & \dots & a_{m2} \\ \vdots & & \\ a_{1n} & \dots & a_{mn} \end{pmatrix}$$

**Example 71.**

$$A = \begin{pmatrix} 1 & 7 \\ 5 & 2 \end{pmatrix}, \quad A^T = \begin{pmatrix} 1 & 5 \\ 7 & 2 \end{pmatrix}.$$

## Transpose

The transpose  $A^T$  of an  $m \times n$  matrix A is the  $n \times m$  matrix obtained by interchanging the rows and columns of A:

$$A = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix}, \text{ then } A^T = \begin{pmatrix} a_{11} & \cdots & a_{m1} \\ \vdots & \ddots & \vdots \\ a_{1n} & \cdots & a_{mn} \end{pmatrix}$$

Example:

$$A = \begin{pmatrix} 1 & 7 \\ 5 & 2 \end{pmatrix}, \text{ then } A^T = \begin{pmatrix} 1 & 5 \\ 7 & 2 \end{pmatrix}$$


---

## Scalar Multiplication

$$A = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix} \text{ then } s \cdot A = A \cdot s = \begin{pmatrix} sa_{11} & \cdots & sa_{1n} \\ \vdots & \ddots & \vdots \\ sa_{m1} & \cdots & sa_{mn} \end{pmatrix}$$

for s a (real) scalar.

Example:

$$A = \begin{pmatrix} 1 & 7 \\ 5 & 2 \end{pmatrix}, \text{ then } 2A = \begin{pmatrix} 2 & 14 \\ 10 & 4 \end{pmatrix}$$


---

The next matrix operation is **scalar multiplication**. The term scalar refers to a  $1 \times 1$  matrix. We have that an  $n \times m$  matrix  $A$  multiplied by a scalar  $s$  is defined componentwise, namely:

$$A = \begin{pmatrix} a_{11} & \dots & a_{1n} \\ a_{21} & \dots & a_{2n} \\ \vdots & & \vdots \\ a_{m1} & \dots & a_{mn} \end{pmatrix} \Rightarrow sA = \begin{pmatrix} sa_{11} & \dots & sa_{1n} \\ sa_{21} & \dots & sa_{2n} \\ \vdots & & \vdots \\ sa_{m1} & \dots & sa_{mn} \end{pmatrix}.$$

**Example 72.**

$$A = \begin{pmatrix} 1 & 7 \\ 5 & 2 \end{pmatrix}, \quad 2A = \begin{pmatrix} 2 & 10 \\ 14 & 4 \end{pmatrix}.$$

We next recall the definition of scalar product, which will be needed to define matrix multiplication. The term "scalar product" reflects the fact that we perform a "product", and that the result is a "scalar", so it is an operation that takes two vectors, and results in a  $1 \times 1$  matrix.

**Definition 45.** The **scalar product** of a  $1 \times n$  vector  $v$  with an  $n \times 1$  column vector  $w$  is

$$v \cdot w = (v_1, \dots, v_n) \begin{pmatrix} w_1 \\ w_2 \\ \vdots \\ w_n \end{pmatrix} = \sum_{i=1}^n v_i w_i.$$

Sometimes, we may say the scalar product of  $v$  and  $w$  without specifying whether  $v$  and  $w$  are row or columns vectors, but for the scalar product to be valid, one need to be row and the other column. Note that  $v \cdot w = w \cdot v$ .

**Example 73.** The scalar product of  $(2, 3)$  and  $(2, -1)$  is

$$(2, 3) \begin{pmatrix} 2 \\ -1 \end{pmatrix} = 2 \cdot 2 - 3 = 1.$$

We are now ready to define the **multiplication of two matrices**  $A$  and  $B$ , where  $A$  is an  $m \times n$  matrix, and  $B$  is an  $n \times r$  matrix. Then for

$$A = \begin{pmatrix} a_{11} & \dots & a_{1n} \\ a_{21} & \dots & a_{2n} \\ \vdots & & \vdots \\ a_{m1} & \dots & a_{mn} \end{pmatrix}, \quad B = \begin{pmatrix} b_{11} & \dots & b_{1r} \\ b_{21} & \dots & b_{2r} \\ \vdots & & \vdots \\ b_{n1} & \dots & b_{nr} \end{pmatrix}$$

## Scalar Product

The **scalar product** of a  $1 \times n$  row vector  $v$  with a  $n \times 1$  column vector  $w$  is defined to be

$$\begin{aligned} v \cdot w &= (v_1, \dots, v_n) \begin{pmatrix} w_1 \\ \vdots \\ w_n \end{pmatrix} \\ &= \sum_{i=1}^n v_i w_i \end{aligned}$$

Example: the scalar product of  $(2 \ 3)$  and  $(2 \ -1)$  is

$$(2 \ 3) \begin{pmatrix} 2 \\ -1 \end{pmatrix} = 4 - 3 = 1.$$


---

## Matrix Multiplication

The **product** of an  $m \times n$  matrix A with a  $n \times r$  matrix B is

$$AB = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix} \begin{pmatrix} b_{11} & \cdots & b_{1r} \\ \vdots & \ddots & \vdots \\ b_{n1} & \cdots & b_{nr} \end{pmatrix} = C$$

where  $c_{ij}$  is the scalar product of the row  $i$  of A and the column  $j$  of B.

Example:

$$A = \begin{pmatrix} 1 & 7 \\ 5 & 2 \end{pmatrix}, B = \begin{pmatrix} 2 & 0 \\ -1 & -1 \end{pmatrix}, AB = \begin{pmatrix} -5 & -7 \\ 8 & -2 \end{pmatrix}$$


---

we have  $AB = C$ , where  $c_{ij}$  is the scalar product of the row  $i$  of  $A$  and the column  $j$  of  $B$ . Note that typically  $AB \neq BA$ , therefore the ordering of the multiplication is very important! Furthermore, dimensions must be compatible for multiplication, namely to compute  $AB$ , we need the number of columns of  $A$  to be equal to the number of rows of  $B$ .

**Example 74.**

$$A = \begin{pmatrix} 1 & 7 \\ 5 & 2 \end{pmatrix}, \quad B = \begin{pmatrix} 2 & 0 \\ -1 & -1 \end{pmatrix}.$$

Then

$$AB = C$$

where

$$c_{11} = (1, 7) \begin{pmatrix} 2 \\ -1 \end{pmatrix} = -5, \quad c_{12} = (1, 7) \begin{pmatrix} 0 \\ -1 \end{pmatrix} = -7,$$

and

$$c_{21} = (5, 2) \begin{pmatrix} 2 \\ -1 \end{pmatrix} = 8, \quad c_{22} = (5, 2) \begin{pmatrix} 0 \\ -1 \end{pmatrix} = -2.$$

Some matrices have a special shape, this is the case of diagonal matrices.

**Definition 46.** An  $n \times n$  matrix  $A$  is said to be **diagonal** if its coefficients  $a_{ij}$  are 0 whenever  $i \neq j$ .

**Example 75.** The matrix

$$\begin{pmatrix} 2 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -3 \end{pmatrix}$$

is diagonal.

The matrix identity is a special type of diagonal matrix.

**Definition 47.** The  $n$ -dimensional **identity matrix**  $I_n$  is a diagonal matrix whose diagonal coefficients are all 1.

**Example 76.** The identity matrix  $I_3$  is

$$I_3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

## Diagonal Matrices and Identity

An  $n \times n$  (square) matrix  $A$  is called **diagonal** if all its coefficients  $a_{ij}$  are 0 whenever  $i \neq j$ :

$$\begin{pmatrix} a_{11} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & a_{nn} \end{pmatrix}$$

The  $n \times n$  (square) **identity matrix  $I$**  is a diagonal matrix with  $a_{ii} = 1$  for all  $i$ :

$$I_n = \begin{pmatrix} 1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & 1 \end{pmatrix}$$

## Square Matrices and Inverse

An  $n \times n$  (square) matrix  $A$  is **invertible** (has an inverse) if there exists an  $n \times n$  matrix  $A^{-1}$  such that

$$AA^{-1} = A^{-1}A = I_n.$$

Example 1:

$$A = \begin{pmatrix} 3 & 1 \\ 5 & 2 \end{pmatrix}, A^{-1} = \begin{pmatrix} 2 & -1 \\ -5 & 3 \end{pmatrix}, AA^{-1} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

Example 2:

$$A = \begin{pmatrix} 3 & 1 \\ 0 & 0 \end{pmatrix}, B = \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix}, AB = \begin{pmatrix} * & * \\ 0 & 0 \end{pmatrix}$$

Thanks to the definition of identity matrix, we may introduce that of inverse of a matrix. For a real scalar, we say that the inverse of  $x$  when  $x \neq 0$  is  $x^{-1}$ , and  $x^{-1}$  is such that  $xx^{-1} = 1$ . The identity matrix plays the role of 1 for matrices.

**Definition 48.** An  $n \times n$  matrix  $A$  is [invertible](#) if there exists a matrix  $A^{-1}$  such that  $AA^{-1} = A^{-1}A = I_n$ . The matrix  $A^{-1}$  is called the inverse of  $A$ .

The definition mentions that multiplication of  $A$  by its inverse both on the right and on the left must give  $I_n$ . In practice, we will typically show only one of the two, the reason being that it can be proven that for a square matrix (square means that the number of rows is the same at the number of columns), the existence of an inverse on one side actually implies that of an inverse on the other side, and both of them are the same.

For real numbers  $x$ , as recalled above, only when  $x = 0$  it is not possible to compute  $x^{-1}$ . For matrices, many of them are not invertible!

**Example 77.** The matrix

$$A = \begin{pmatrix} 3 & 1 \\ 5 & 2 \end{pmatrix}$$

is invertible, because its inverse  $A^{-1}$  is

$$A^{-1} = \begin{pmatrix} 2 & -1 \\ -5 & 3 \end{pmatrix}.$$

Indeed, one may check that  $AA^{-1} = I_2$ , or  $A^{-1}A = I_2$ . On the other hand

$$A = \begin{pmatrix} 3 & 1 \\ 0 & 0 \end{pmatrix}$$

has no inverse. No matter which matrix  $B$  one takes,  $AB$  will always have a row of zeroes, and cannot possibly be equal to the identity matrix.

In the above example,  $A^{-1}$  is given and one can just check that  $AA^{-1}$  is indeed  $I_2$ . We next see a general technique, which among other things allows to compute the inverse of a matrix.

**Definition 49.** An  $m \times n$  matrix  $A$  is in [row echelon form](#) if

1. The nonzero rows (if any) in  $A$  lie above all zero rows.
2. The first nonzero entry (in a nonzero row) lies to the right of the first nonzero entry in the row immediately above it.

## Row Echelon Form

An  $m \times n$  matrix A is in **row echelon form** if

1. The nonzero rows (if any) in A lie above all zero rows.
2. The first nonzero entry (in a nonzero row) lies to the right of the first nonzero entry in the row immediately above it.

Example:

$$A = \begin{pmatrix} 3 & 1 \\ 0 & 2 \end{pmatrix}, B = \begin{pmatrix} 2 & 9 \\ 0 & 3 \\ 0 & 0 \end{pmatrix}$$

---

## Elementary Row Operations

1. **Row switching:** Switch row i with row j.
2. **Row multiplication:** Multiply each element in row i by a nonzero k.
3. **Row addition:** Replace row i by the sum of row i and a nonzero multiple k of row j.

Any  $m \times n$  matrix can be transformed into a row echelon form (not uniquely) using elementary row operations.

---

**Example 78.** The matrices

$$A = \begin{pmatrix} 3 & 1 \\ 0 & 2 \end{pmatrix}, \quad B = \begin{pmatrix} 2 & 9 \\ 0 & 3 \\ 0 & 0 \end{pmatrix}$$

are in row echelon form.

Now a matrix is not necessarily in row echelon form. However, there is a series of operations which we are allowed to do on a matrix, to bring it into such a form. These operations are called [elementary row operations](#), and comprise:

1. **Row switching:** switch row  $i$  with row  $j$ .
2. **Row multiplication:** multiply each element in row  $i$  by a nonzero  $k$ . Note that  $k$  can be of the form  $1/k'$ ,  $k' \neq 0$ , therefore division is allowed as well.
3. **Row addition:** replace row  $i$  by the sum of row  $i$  and a nonzero multiple  $k$  of row  $j$ .

Note that it is always possible to bring a matrix into a row echelon form. Roughly, this is because either a column contains only zeroes, or only zeroes and one non-zero entry, in which case it is fine. If it contains two non-zero entries, then one can be used to cancel out the other one using elementary row operations.

**Example 79.** Consider the matrix

$$\begin{pmatrix} 1 & 2 & 3 \\ 1 & 1 & -4 \\ 7 & 0 & 2 \end{pmatrix}.$$

To obtain its row echelon form, we first take care of the first column. For this we replace (row 2) by (row 2)- (row 1), and then we replace (row 3) by (row 3) -7 (row 1), to get

$$\begin{pmatrix} 1 & 2 & 3 \\ 0 & -1 & -7 \\ 0 & -14 & -19 \end{pmatrix}.$$

## Example

$$\begin{pmatrix} 1 & 2 & 3 \\ 1 & 1 & -4 \\ 7 & 0 & 2 \end{pmatrix} \rightarrow \begin{pmatrix} 1 & 2 & 3 \\ 0 & -1 & -7 \\ 0 & -14 & -19 \end{pmatrix} \rightarrow \begin{pmatrix} 1 & 2 & 3 \\ 0 & -1 & -7 \\ 0 & 0 & 79 \end{pmatrix}$$


---

## Reduced Row Echelon Form

An  $m \times n$  matrix A is in **reduced row echelon form** if

1. A is in echelon form.
2. The first nonzero entry (in a nonzero row) is 1, and all other entries in the column are zero.

“Gaussian Elimination”

Example:

$$A = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, B = \begin{pmatrix} 1 & 0 & 3 \\ 0 & 1 & 2 \end{pmatrix}$$



Picture from wikipedia

The only step missing now is to change the second column, which is done by replacing (row 3) by (row 3)-14(row 2):

$$\begin{pmatrix} 1 & 2 & 3 \\ 0 & -1 & -7 \\ 0 & 0 & 79 \end{pmatrix}.$$

The row echelon form of a matrix is not unique, this can be seen by the fact that one is allowed to multiply a row by a non-zero constant. For example, the above matrix has last row  $(0, 0, 79)$ , any multiple of this row would also give a row echelon form.

One may however further reduce a matrix into what is called a reduced echelon form, in which case it becomes unique. This procedure is also called Gaussian elimination.

**Definition 50.** An  $m \times n$  matrix  $A$  is in [reduced row echelon form](#) if

1.  $A$  is in echelon form.
2. The first nonzero entry (in a nonzero row) is 1, and all other entries in the column are zero.

**Example 80.** We continue Example 79, with the matrix

$$\begin{pmatrix} 1 & 2 & 3 \\ 1 & 1 & -4 \\ 7 & 0 & 2 \end{pmatrix}$$

which has row echelon form

$$\begin{pmatrix} 1 & 2 & 3 \\ 0 & -1 & -7 \\ 0 & 0 & 79 \end{pmatrix}.$$

We first divide the last row by 79, and add 2(row 2) to the first row:

$$\begin{pmatrix} 1 & 0 & -11 \\ 0 & -1 & -7 \\ 0 & 0 & 1 \end{pmatrix}.$$

We are left to add 11(row 3) to the first row, 7(row 3) to the second row, and multiply the second row by -1, to get  $I_3$ .

It may look surprising to find the identity matrix at the end, but this happens in fact whenever the matrix  $A$  we started with is invertible.

## Example

$$\begin{pmatrix} 1 & 2 & 3 \\ 1 & 1 & -4 \\ 7 & 0 & 2 \end{pmatrix} \rightarrow \begin{pmatrix} 1 & 2 & 3 \\ 0 & -1 & -7 \\ 0 & -14 & -19 \end{pmatrix} \rightarrow \begin{pmatrix} 1 & 2 & 3 \\ 0 & -1 & -7 \\ 0 & 0 & 79 \end{pmatrix}$$

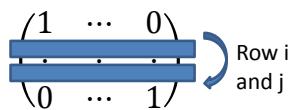
$$\begin{pmatrix} 1 & 2 & 3 \\ 0 & -1 & -7 \\ 0 & 0 & 1 \end{pmatrix} \rightarrow \begin{pmatrix} 1 & 0 & -11 \\ 0 & -1 & -7 \\ 0 & 0 & 1 \end{pmatrix} \rightarrow \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

Any  $m \times n$  matrix can be transformed into a unique reduced row echelon form using elementary row operations.

## Elementary Matrices

Elementary row operation = Multiplication by **elementary matrix**

1. Switch row  $i$  with row  $j$
2. Multiply each element in row  $i$  by a nonzero  $k$ :
3. Replace row  $i$  by the sum of row  $i$  and a nonzero multiple  $k$  of row  $j$ .



$$\begin{pmatrix} 1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & 1 \end{pmatrix}$$

Row  $i$   
and  $j$

Let us try to understand why, which will also give us an algorithm to find the inverse of a matrix. First of all, we notice that all the elementary operations used for obtaining a row echelon form can be expressed by matrices, which are called elementary matrices.

1. **Row switching:** row  $i$  with row  $j$  are switched if the matrix is multiplied by a matrix obtained from the identity matrix by switching its row  $i$  with its row  $j$ .
2. **Row multiplication:** multiply each element in row  $i$  by a nonzero  $k$ . This is done by multiplication by a diagonal matrix, where all diagonal coefficients are 1 but for the  $i$ th row which contains a  $k$ .
3. **Row addition:** replace row  $i$  by the sum of row  $i$  and a nonzero multiple  $k$  of row  $j$ . This is done by using a matrix which has 1 on the diagonal, except for the row  $i$ , which further contains a  $k$  in the  $j$ th column.

Now take a matrix  $A$ , and suppose that it is invertible, in particular it is square, that is  $A$  is an  $n \times n$  matrix. Once the matrix is in row echelon form, only two things can happen: either all the diagonal coefficients are non-zero, in which case the reduced form will give the identity matrix, or at least one row is a whole zero row. This really results from the fact the the number of rows and columns are the same. Next form an augmented matrix, which contains  $A$  and  $I_n$ , namely  $(A|I_n)$ . Multiply the matrix on the left with elementary matrices, say  $M_1, \dots, M_l$ , to get a reduced row echelon form of  $A$ :

$$M_l M_{l-1} \cdots M_2 M_1 (A|I_n).$$

Now if  $A$  is invertible  $M_l M_{l-1} \cdots M_2 M_1 \cdot A = I_n$ , therefore  $M_l M_{l-1} \cdots M_2 M_1 = A^{-1}$  and

$$M_l M_{l-1} \cdots M_2 M_1 (A|I_n) = (I_n|A^{-1}).$$

If multiplication by elementary matrices result in a matrix with at least one row which is zero, then it is impossible to obtain the identity matrix on the left hand side of  $(A|I_n)$  and the matrix cannot be invertible.

#### **Recipe to compute the inverse of a matrix $A$ .**

1. Write the augmented matrix  $(A|I_n)$ .
2. Compute its reduced row echelon form, to obtain  $(I_n|A^{-1})$

## Gauss-Jordan Elimination

(Elementary Matrices)  $A = A$  in Reduced Echelon Form.

If  $A$  invertible, then  $A$  in Reduced Echelon Form = identity

(Elementary Matrices)  $(A \ I_n) = (I_n \quad A^{-1})$

Recipe to compute  $A^{-1}$

1. Write the matrix  $(A \ I_n)$
  2. Compute its reduced echelon form.
- 

## Example

Example 1:

$$A = \begin{pmatrix} 3 & 1 \\ 5 & 2 \end{pmatrix}, \begin{pmatrix} 3 & 1 & 1 & 0 \\ 5 & 2 & 0 & 1 \end{pmatrix}, \begin{pmatrix} 3 & 1 & 1 & 0 \\ 0 & 1 & -5 & 3 \end{pmatrix}, \\ \begin{pmatrix} 3 & 0 & 6 & -3 \\ 0 & 1 & -5 & 3 \end{pmatrix}, \begin{pmatrix} 1 & 0 & 2 & -1 \\ 0 & 1 & -5 & 3 \end{pmatrix}, \\ A^{-1} = \begin{pmatrix} 2 & -1 \\ -5 & 3 \end{pmatrix}.$$

Example 2:

$$A = \begin{pmatrix} 3 & 1 \\ 0 & 0 \end{pmatrix}$$


---

**Example 81.** Consider the matrix

$$A = \begin{pmatrix} 3 & 1 \\ 5 & 2 \end{pmatrix}$$

To compute its inverse, create the augmented matrix:

$$\begin{pmatrix} 3 & 1 & 1 & 0 \\ 5 & 2 & 0 & 1 \end{pmatrix}$$

To obtain a row echelon form, replace (row 2) by -5(row 1)+3(row 2), which in matrix form is given by

$$\begin{pmatrix} 1 & 0 \\ -5 & 3 \end{pmatrix} \begin{pmatrix} 3 & 1 & 1 & 0 \\ 5 & 2 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 3 & 1 & 1 & 0 \\ 0 & 1 & -5 & 3 \end{pmatrix}$$

and we can tell that the matrix is invertible. Then replace (row 1) by (row 1)-(row 2):

$$\begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ -5 & 3 \end{pmatrix} \begin{pmatrix} 3 & 1 & 1 & 0 \\ 5 & 2 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 3 & 1 & 1 & 0 \\ 0 & 1 & -5 & 3 \end{pmatrix} = \begin{pmatrix} 3 & 0 & -6 & -3 \\ 0 & 1 & -5 & 3 \end{pmatrix}$$

and we are left by dividing the first row by 3:

$$\begin{pmatrix} 1/3 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ -5 & 3 \end{pmatrix} \begin{pmatrix} 3 & 1 & 1 & 0 \\ 5 & 2 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & -2 & -1 \\ 0 & 1 & -5 & 3 \end{pmatrix}.$$

Another application of the (reduced) row echelon form is solving systems of linear equations.

**Definition 51.** A system of  $m$  linear equations in  $n$  unknowns is of the form:

$$\left\{ \begin{array}{lcl} a_{11}x_1 + \dots + a_{1n}x_n & = & b_1 \\ \vdots & & \vdots \\ a_{i1}x_1 + \dots + a_{in}x_n & = & b_i \\ \vdots & & \vdots \\ a_{m1}x_1 + \dots + a_{mn}x_n & = & b_m \end{array} \right.$$

It is said to be **homogeneous** when  $b_1 = \dots = b_m = 0$ .

## Systems of Linear Equations

A **system of m linear equations in n unknowns**:

$$\begin{cases} a_{11}x_1 + \cdots + a_{1n}x_n = b_1 \\ a_{i1}x_1 + \cdots + a_{in}x_n = b_i \\ a_{m1}x_1 + \cdots + a_{mn}x_n = b_m \end{cases}$$

Alternatively:

$$\begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix}$$

If  $b_1 = b_2 = \cdots = b_m = 0$  then the system is **homogeneous**.

## Solutions to Linear Equations

$$A x = b \quad A = mxn \text{ matrix. Find } x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$$

(Elementary Matrices)  $A = A$  in Reduced Echelon Form.

(Elementary Matrices)  $(A \ b) \leftrightarrow$  (Elementary Matrices)

$(Ax) = (\text{Elementary Matrices})b$

Recipe to solve  $Ax = b$

1. Write the matrix  $(A \ b)$
2. Compute its reduced echelon form.

A system of linear equations is **consistent** if it has at least one solution.

In matrix form, such a system of linear equations can be rewritten as:

$$\underbrace{\begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \\ a_{i1} & \dots & a_{in} \\ \vdots & & \\ a_{m1} & \dots & a_{mn} \end{pmatrix}}_A \underbrace{\begin{pmatrix} x_1 \\ \vdots \\ x_i \\ \vdots \\ x_n \end{pmatrix}}_x = \underbrace{\begin{pmatrix} b_1 \\ \vdots \\ b_i \\ \vdots \\ b_m \end{pmatrix}}_b$$

We now form an augmented matrix, similarly to what was done for the inverse computation, which contains  $A$  and  $b$ , namely  $(A|b)$ , and we multiply the matrix  $(A|b)$  on the left with elementary matrices, say  $M_1, \dots, M_l$ , to get a reduced row echelon form of  $A$ :

$$M_l M_{l-1} \cdots M_2 M_1 (A|b).$$

If  $A$  is invertible  $M_l M_{l-1} \cdots M_2 M_1 \cdot A = I_n$ , and as before

$$M_l M_{l-1} \cdots M_2 M_1 (A|I_n) = (I_n|A^{-1}b).$$

This means that

$$Ax = b \iff M_l M_{l-1} \cdots M_2 M_1 Ax = M_l M_{l-1} \cdots M_2 M_1 b$$

which becomes

$$x = A^{-1}b$$

when  $A$  is invertible. If  $A$  is not invertible, the reduced echelon form still allows us to write the system  $Ax = b$  in such a form that it is easy to read the solution  $x$  from it.

#### **Recipe to solve a system $Ax = b$ of linear equations.**

1. Write the augmented matrix  $(A|b)$ .
2. Compute its reduced row echelon form.

Let us first discuss the solutions when the system is homogeneous, that is  $Ax = 0$ . First note that  $x = 0$  is always a solution. If  $m < n$ , namely there are less equations than unknowns, there will be infinitely many solutions. Some unknowns will be unconstrained, they can take any value.

## Homogenous Systems of Equations I

$A x = 0$     $A = mxn$  matrix. Find  $x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$

$x = 0$  is always a solution!

If  $m$  equations  $< n$  unknowns: infinity of solutions

$$\begin{pmatrix} a_{11} & \cdots & a_{1n} \\ a_{m1} & \cdots & a_{mn} \\ 0 & \cdots & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix}$$

Example: (1 2)  $\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = 0$ .

---

## Homogenous Systems of Equations II

$A x = 0$     $A = mxn$  matrix. Find  $x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$

If  $m$  equations  $= n$  unknowns: if  $A$  not invertible

$$\begin{pmatrix} a_{11} & \cdots & a_{1n} \\ a_{m1} & \cdots & a_{mn} \\ 0 & \cdots & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix}$$

If  $A$  invertible, then  $A$  in Reduced Echelon Form = identity.

Thus  $x = 0$  is the only solution!

if  $m$  equations  $> n$  unknowns :  $x = 0$  is always a solution!

---

**Example 82.** The system

$$(1, 2) \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = 0$$

has infinitely many solutions, no matter what is the value of  $x_1$ , there is a corresponding value for  $x_2$  which satisfies this equation.

If  $m = n$ , we have the same number of equations as unknowns, two things can happen: if  $A$  is invertible, then  $x = 0$  is the only solution, if  $A$  is not invertible, this means that when computing its row echelon form, rows of zeros will appear, and in fact, some of the equations were redundant. Therefore we will have as before infinitely many solutions.

The situation is similar when  $m > n$ . We still have  $x = 0$  has a solution. Then when we compute the row echelon form of  $A$ , it could be that we still end up with many redundant equations, in which case infinitely many solutions still could happen, but if the number of non-zero rows is still bigger than  $n$ , then  $x = 0$  is the only solution.

Let us now see what happens when the system is not homogeneous. The key thing, as noted for the homogeneous case, is to see the number of equations left once the matrix  $A$  is in reduced form.

**Definition 52.** The [rank](#) of a matrix  $A$  is the number of non-zero rows in an echelon form of  $A$ .

For a system  $Ax = b$  of linear questions, we thus get:

- If  $\text{rank}(A) < \text{rank}(A|b)$ , then the system has no solution. This is because when we compute the row echelon form of  $(A|b)$ , the part in  $A$  will have rows of zeroes, which are non-zero for  $(A|b)$ , corresponding to an equation of the form 0 is equal to something non-zero, which is not possible.
- If  $\text{rank}(A) = \text{rank}(A|b) < n$ , then the system has infinitely many solutions. The fact that both ranks are the same means that there are solutions. Now less than  $n$  means less equations than unknowns.
- If  $\text{rank}(A) = \text{rank}(A|b) = n$ , there is a unique solution. This corresponds to the case where  $A$  is invertible.

## Non-Homogenous Systems

The **rank** of a matrix A is the number of nonzero rows in an echelon form of A.

$$A x = b$$

- If  $\text{rank}(A) < \text{rank}(A|b)$  then the system has no solution.
- If  $\text{rank}(A) = \text{rank}(A|b) < n$  then the system is consistent and has infinitely many solutions.
- If  $\text{rank}(A) = \text{rank}(A|b) = n$  then the system is consistent and has a unique solution. This is when  $A^{-1}$ .

## Examples

$$A = \begin{pmatrix} 1 & 1 \\ 2 & 2 \end{pmatrix}, b = \begin{pmatrix} 1 \\ 3 \end{pmatrix}$$

$$A = \begin{pmatrix} 1 & 1 \\ 2 & 2 \end{pmatrix}, b = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$$

$$A = \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix}, b = \begin{pmatrix} 1 \\ 3 \end{pmatrix}$$

Note that  $\text{rank}(A)$  is always less or equal  $\text{rank}(A|b)$ , this is because removing columns from  $(A|b)$  cannot increase the number of non-zero rows, thus all cases have been considered.

**Example 83.** Consider the system  $Ax = b$  with

$$A = \begin{pmatrix} 1 & 1 \\ 2 & 2 \end{pmatrix}, \quad b = \begin{pmatrix} 1 \\ 3 \end{pmatrix}.$$

The rank of  $A$  is 1, that of  $(A|b)$  is 2, thus no solution. Now with

$$A = \begin{pmatrix} 1 & 1 \\ 2 & 2 \end{pmatrix}, \quad b = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$$

the rank of  $A$  stays 1, but that of  $(A|b)$  is 1 as well, so we have infinitely many solutions. Finally, with

$$A = \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix}, \quad b = \begin{pmatrix} 1 \\ 3 \end{pmatrix}$$

the matrix  $A$  is invertible, so there is a unique solution, given by  $A^{-1}b$ .

## Exercises for Chapter 8

**Exercise 63.** Compute the sum  $A + B$  of the matrices  $A$  and  $B$ , where  $A$  and  $B$  are as follows:

1.

$$A = \begin{pmatrix} 2 & \sqrt{2} \\ -1 & 3 \end{pmatrix}, \quad B = \begin{pmatrix} 0 & \sqrt{2} \\ 4 & 2 \end{pmatrix}$$

where  $A, B$  are matrices with coefficients in  $\mathbb{R}$ .

2.

$$A = \begin{pmatrix} 2+i & -1 \\ -1+i & 3 \end{pmatrix}, \quad B = \begin{pmatrix} -i & 1 \\ -1 & 2 \end{pmatrix}$$

where  $A, B$  are matrices with coefficients in  $\mathbb{C}$ , and  $i = \sqrt{-1}$ .

3.

$$A = \begin{pmatrix} 0 & 1 & 2 \\ 2 & 0 & 2 \end{pmatrix}, \quad B = \begin{pmatrix} 1 & 0 & 2 \\ 0 & 1 & 1 \end{pmatrix}$$

where  $A, B$  are matrices with coefficients that are integers  $\pmod{3}$ .

What are the dimensions of the matrices involved?

**Exercise 64.** 1. Compute the transpose  $A^T$  of  $A$  for

$$A = \begin{pmatrix} 1 & 0 & 2 \\ 0 & 1 & 1 \end{pmatrix}.$$

2. Show that  $(A + B)^T = A^T + B^T$ .

**Exercise 65.** Compute

$$2A + BC + B^2 + AD$$

where

$$A = \begin{pmatrix} 2 & 0 \\ -1 & 1 \end{pmatrix}, \quad B = \begin{pmatrix} 1 & -1 \\ 0 & 2 \end{pmatrix}, \quad C = \begin{pmatrix} -1 & 0 \\ 2 & 1 \end{pmatrix}$$

are real matrices and  $D = I_2$  is the 2-dimensional identity matrix.

**Exercise 66.** Consider the complex matrix

$$A = \begin{pmatrix} 0 & i \\ -1 & 0 \end{pmatrix},$$

where  $i = \sqrt{-1}$ . What is  $A^l$ , for  $l \geq 1$ .

**Exercise 67.** 1. Let  $S$  be the set of  $3 \times 3$  diagonal real matrices. Is  $S$  closed under matrix addition?

2. Consider the real matrix

$$A = \begin{pmatrix} 2 & 3 \\ -1 & 1 \end{pmatrix}.$$

Compute a matrix  $B$  such that  $A + B$  is diagonal, and a matrix  $C$  such that  $AC$  is diagonal.

**Exercise 68.** Let  $A$  and  $B$  be  $n \times n$  matrices which satisfy

$$A^2 + AB + A - I_n = 0,$$

where  $I_n$  means the  $n \times n$  identity matrix, and  $0$  the  $n \times n$  zero matrix. Show that  $A$  is invertible.

**Exercise 69.** Compute, if it exists, the inverse  $A^{-1}$  of the matrix  $A$ , where  $A$  is given by

•

$$A = \begin{pmatrix} 2 & 3 & -2 \\ -1 & 1 & 2 \\ 3 & 7 & 2 \end{pmatrix}$$

for  $A$  a real matrix.

•

$$A = \begin{pmatrix} 1 & 1+i \\ 1-i & 1 \end{pmatrix}$$

for  $A$  a complex matrix and  $i = \sqrt{-1}$ .

•

$$A = \begin{pmatrix} 2 & 3 \\ 1 & 1 \end{pmatrix}$$

for  $A$  a matrix with coefficients modulo 5.

**Exercise 70.** Write the following system of linear equations in a matrix form and solve it.

$$\begin{cases} x_1 + x_2 - 2x_3 = 1 \\ 2x_1 - 3x_2 + x_3 = -8 \\ 3x_1 + x_2 + 4x_3 = 7 \end{cases}$$

**Exercise 71.** Write the following system of linear equations in a matrix form and solve it.

$$\begin{cases} x_1 - x_2 + x_3 - x_4 = 2 \\ x_1 - x_2 + x_3 + x_4 = 0 \\ 4x_1 - 4x_2 + 4x_3 = 4 \\ -2x_1 + 2x_2 - 2x_3 + x_4 = -3 \end{cases}$$

## Exercises for Chapter 8

We will provide some applications of matrices to cryptography. We start by explaining the notion of cipher and give some examples.

It is said that the roman general Caesar used to communicate secretly with his army commanders using the following cipher:

$$e_K : x \rightarrow e_K(x) = x + K \pmod{26}, \quad K = 3.$$

What it means is the following thing: one can map letters from  $A$  to  $Z$  to the integers 0 to 25. Then to say  $A$ , which is 0, encrypt it  $e_K(0) = K = 3$ , which is  $D$ . Therefore to say  $A$ , Caesar would write in his message  $D$ , and similarly all his messages would be encrypted. We call  $e_K$  an encryption function, and  $K$  a secret key. If the key is known, it is easy to recover the original message.

A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q
0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
R	S	T	U	V	W	X	Y	Z								
17	18	19	20	21	22	23	24	25								

To recover a message, one uses a decryption function  $d_K$ . In this case

$$d_K(y) = y - K \pmod{26}, \quad K = 3.$$

Indeed

$$d_K(e_K(x)) = e_K(x) - K = x + K - K = x \pmod{26}.$$

Suppose Caesar wrote  $YHQL$ ,  $YLGL$ ,  $YLFL$ . In numbers, it becomes  $24, 7, 16, 11, 24, 11, 6, 11, 24, 11, 5, 11$  now we apply the decryption function on this, to find  $21, 4, 13, 8, 21, 8, 3, 8, 21, 8, 2, 8$  that is  $VENI, VIDI, VICI$ . (This is a famous quote by Caesar, in latin, it means “I came, I saw, I conquered”).

Caesar’s cipher described above may look too simple to break. After all, one could just try all possible  $K$ , there are only 26 of them, and figure out which one works out. A variation of this cipher is called affine cipher. Now the encryption looks like this

$$e_K : x \rightarrow e_K(x) = k_1 x + k_2 \pmod{26}, \quad K = (k_1, k_2).$$

The question is then, how to choose  $K = (k_1, k_2)$ ? Well, the first important thing is that decryption must be possible, which is not possible for any choice of keys! For example, pick the key  $K = (13, 7)$ . Then

$$e_K : x \rightarrow e_K(x) = 13x + 7 \pmod{26}, \quad K = (13, 7).$$

## Caesar's Cipher

To send secret messages to his generals, Caesar is said to have used the following cipher.

$$e_K: x \rightarrow e_K(x) = x + K \bmod 26, K=3$$

Map A to 0,...,Z to 25 and decipher this message from Caesar: YHQL YLGL YLFL



Caesar belongs to Goscinny and Uderzo.

## Affine Cipher

Caesar's cipher is a well-defined cipher because there is a function  $d_K$  such that  $d_K(e_K(x)) = x$  for every  $x$  integer mod 26.

$$e_K: x \rightarrow e_K(x) = k_1 x + k_2 \bmod 26, \\ K = (k_1, k_2)$$



Choose the best key (if any):  
 $K = (7, 13)$  or  $K = (13, 7)$



Alice belongs to Disney, Sponge Bob to Hillenburg

To decrypt, let us try  $d_K(y) = ay + b$ , then

$$d_K(e_K(x)) = a(e_K(x)) + b = a(13x + 7) + b.$$

To be able to find  $x$ , we need to be able to solve  $a(13x + 7) + b = x$ , that is  $13ax + 7a + b = x$ ,  $7a + b = 0$ , and  $13a = 1$ . But there is no such  $a \pmod{26}$ !

Let us try instead the key  $K = (7, 13)$ . Then

$$e_K : x \rightarrow e_K(x) = 7x + 13 \pmod{26}, K = (7, 13).$$

To decrypt, let us try  $d_K(y) = ay + b$ , then

$$d_K(e_K(x)) = a(e_K(x)) + b = a(7x + 13) + b.$$

To be able to find  $x$ , we need to be able to solve  $a(7x + 13) + b = x$ , that is  $7ax + 13a + b = x$ ,  $13a + b = 0$ , and  $7a = 1$ . But now this is possible: take  $a = 15$ , then  $105 = 104 + 1 \equiv 1 \pmod{26}$ . Then  $13 \cdot 15 = -b \equiv 13 \pmod{26}$  and

$$d_K(y) = 15y + 13.$$

Let us now move to an encryption scheme which uses matrices:

$$e_K : x \rightarrow e_K(x) = K_1x + K_2 \pmod{26}, K = (K_1, K_2),$$

with

$$x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, K_1 = \begin{pmatrix} 5 & 4 \\ 4 & 2 \end{pmatrix}, K_2 = \begin{pmatrix} 4 \\ 2 \end{pmatrix}.$$

We encounter the same problem as before, namely, we need to make sure that  $d_K$  exists. Above, the trouble happened when the encryption was  $13x + k_2$ , because 13 is not invertible  $\pmod{26}$ . Here similarly we need to make sure that  $K_1$  is invertible. Let us try to compute the inverse of  $K_1$ :

$$\begin{pmatrix} 5 & 4 & 1 & 0 \\ 4 & 2 & 0 & 1 \end{pmatrix}$$

Replace (row 2) by  $-5(\text{row 2}) + 4(\text{row 1})$ :

$$\begin{pmatrix} 5 & 4 & 1 & 0 \\ 0 & 6 & 0 & -5 \end{pmatrix}$$

and the matrix is not invertible, because 6 is not invertible  $\pmod{26}$ , namely, it is not possible to find an element  $x \pmod{26}$  such that  $6x \equiv 1 \pmod{26}$ .

## Matrix Encryption

$$e_K: x \rightarrow e_K(x) = K_1 x + K_2 \pmod{26}, K = (K_1, K_2)$$

$$x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, K_1 = \begin{bmatrix} 5 & 4 \\ 4 & 2 \end{bmatrix}, K_2 = \begin{pmatrix} 4 \\ 2 \end{pmatrix}$$

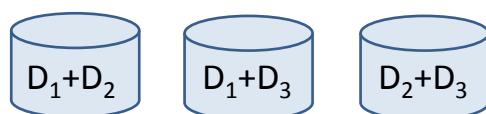
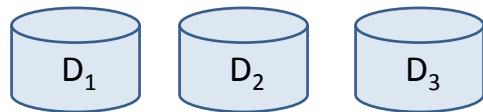
$$x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, K_1 = \begin{bmatrix} 5 & 4 \\ 4 & 1 \end{bmatrix}, K_2 = \begin{pmatrix} 4 \\ 1 \end{pmatrix}$$

Map A to 0,...,Z to 25 and decipher this message using the right key: OJJMGI

---

## Data Storage (II)

$$D = (D_1, D_2, D_3)$$



Write the data stored in matrix form as a function of the data  $(D_1, D_2, D_3)$

To tolerate two failures, we need each  $D_i$  to be present at least 3 times.

---

Let us try another encryption scheme which uses matrices:

$$e_K : x \rightarrow e_K(x) = K_1 x + K_2 \pmod{26}, \quad K = (K_1, K_2),$$

with

$$x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, \quad K_1 = \begin{pmatrix} 5 & 4 \\ 4 & 1 \end{pmatrix}, \quad K_2 = \begin{pmatrix} 4 \\ 1 \end{pmatrix}.$$

Let us try to compute the inverse of  $K_1$ :

$$\begin{pmatrix} 5 & 4 & 1 & 0 \\ 4 & 1 & 0 & 1 \end{pmatrix}$$

Replace (row 2) by -5(row 2)+4(row 1):

$$\begin{pmatrix} 5 & 4 & 1 & 0 \\ 0 & 11 & 0 & -5 \end{pmatrix}$$

and this time, 11 is invertible  $\pmod{26}$ , namely  $11 \cdot (-7) \equiv 1 \pmod{26}$ . We then multiply the second row by -7:

$$\begin{pmatrix} 5 & 4 & 1 & 0 \\ 0 & 1 & -2 & 9 \end{pmatrix}$$

We then replace (row 1) by -4(row 2):

$$\begin{pmatrix} 5 & 0 & 9 & -10 \\ 0 & 1 & -2 & 9 \end{pmatrix}$$

Finally 5 is invertible,  $5 \cdot 21 = 5 \cdot (-5) \equiv 1 \pmod{26}$ . This gives

$$\begin{pmatrix} 1 & 0 & 7 & -2 \\ 0 & 1 & -2 & 9 \end{pmatrix}$$

and

$$K_1^{-1} = \begin{pmatrix} 7 & -2 \\ -2 & 9 \end{pmatrix}$$

To decipher for example  $OJ$ , we map it to integers, namely 14, 9, then

$$\begin{pmatrix} 14 \\ 9 \end{pmatrix} - \begin{pmatrix} 4 \\ 1 \end{pmatrix} = \begin{pmatrix} 10 \\ 8 \end{pmatrix}$$

and we apply  $K_1^{-1}$  to get

$$\begin{pmatrix} 7 & -2 \\ -2 & 9 \end{pmatrix} \begin{pmatrix} 10 \\ 8 \end{pmatrix} = \begin{pmatrix} 2 \\ 0 \end{pmatrix}$$

that is the message  $CA$ . In a similar way, we find that  $OJJMGI$  gives  $CANLAH$ .

### Data Storage (III)

$$(D_1 \quad D_2 \quad D_3) \begin{pmatrix} 1 & 0 & 0 & 1 & 1 & 0 \\ 0 & 1 & 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & 1 & 0 & 1 \end{pmatrix}$$

Suppose now you can choose any strategy, where the first 3 disks do not have to be the data itself:

$$(D_1 \quad D_2 \quad D_3) \begin{pmatrix} a_{11} & & & & & a_{16} \\ & a_{31} & & & & \\ & & a_{36} & & & \end{pmatrix}$$


---

### Data Storage (IV)

Show that you cannot do better using a matrix of a more general form.

- Compute the reduced row echelon form of the generic matrix
    - Either the first 3x3 block is the identity matrix: this is the case where the data is stored in the first disks
    - Or there are zero columns: this is worse, this means that some disks store “useless” data.
-

We finally come back once more to our example of data storage. Suppose that you have some data  $D$ , split into 3 parts  $D = (D_1, D_2, D_3)$ . We saw that to tolerate 2 failures, we need at least 6 disks, assuming that the first 3 disks are storing  $D_1, D_2, D_3$  respectively. For example, one way of storing the data could be

$$\begin{array}{ll} \text{disk 1: } D_1 & \text{disk 4: } D_1 + D_2 \\ \text{disk 2: } D_2 & \text{disk 5: } D_1 + D_3 \\ \text{disk 3: } D_3 & \text{disk 6: } D_2 + D_3 \end{array}$$

This way of storing data can be represented using matrices:

$$(D_1, D_2, D_3) \begin{pmatrix} 1 & 0 & 0 & 1 & 1 & 0 \\ 0 & 1 & 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & 1 & 0 & 1 \end{pmatrix}$$

Now we want to check that any arbitrary strategy where we could have chosen any way of combining the data, not necessarily storing  $D_1, D_2, D_3$  in the first 3 disks, cannot improve the one we already got. Write

$$(D_1, D_2, D_3) \underbrace{\begin{pmatrix} a_{11} & a_{12} & a_{13} & a_{14} & a_{15} & a_{16} \\ a_{21} & a_{22} & a_{23} & a_{24} & a_{25} & a_{26} \\ a_{31} & a_{32} & a_{33} & a_{34} & a_{35} & a_{36} \end{pmatrix}}_A$$

To see this, visualize that whatever is stored in our 6 disks is obtained by combining the three rows of this matrix  $A$ . We can then try to combine them in any way we want using elementary operations (swap rows, add one row to another). Since every operation can be reversed, we can always go back and forth from one form to another. Then we compute the reduced row echelon form of this matrix, and two things can happen: either the first 3 columns are  $I_3$ , in which case we are back to the strategy we already know, or one column becomes zero, which is worse, since this corresponds to a disk storing no data.



# Chapter 9

## Relations

“” ()

The topic of our next chapter is relations, it is about having 2 sets, and connecting related elements from one set to another.

**Definition 53.** Let  $A$  and  $B$  be two sets. A [binary relation](#)  $R$  from  $A$  to  $B$  is a subset of the cartesian product  $A \times B$ . Given  $x, y \in A \times B$ , we say that  $x$  is related to  $y$  by  $R$ , also written  $(xRy) \leftrightarrow (x, y) \in R$ .

**Example 84.** Suppose that you have two sets  $A = \{1, 2\}$  and  $B = \{1, 2, 3\}$ , and the relation is given by  $(x, y) \in R \leftrightarrow x - y$  is even. Since the relation is a subset of  $A \times B$ , we start by computing the cartesian product  $A \times B$ :

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

Then in this list of pairs, we select those which satisfy the relation  $R$ . For example, for  $(1, 2)$ , we have  $x = 1$  and  $y = 2$ , we compute  $x - y = 1 - 2 = -1$ , which is odd, thus it does not belong to  $R$ . We try out similarly all the pairs in  $A \times B$  to get

$$R = \{(1, 1), (1, 3), (2, 2)\}.$$

This may be visualized using a diagram: draw a circle to represent the set  $A$ , and this circle contains two points, one for 1 and one for 2. Similarly, draw a circle to represent  $B$ , and points of 1, 2, 3. Then an arrow from  $A$  to  $B$  connects  $x$  in  $A$  with  $y$  in  $B$  if  $x - y$  is even.

# Binary Relations between Two Sets

Let  $A$  and  $B$  be sets. A **binary relation**  $R$  from  $A$  to  $B$  is a subset of  $A \times B$ . Given  $(x, y)$  in  $A \times B$ ,  $x$  is related to  $y$  by  $R$  ( $x R y$ )  $\leftrightarrow (x, y) \in R$ .

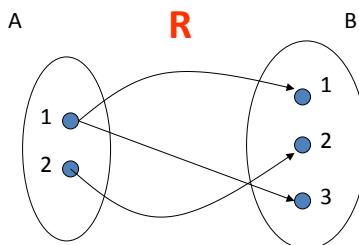
**Example.**  $A=\{1,2\}$ ,  $B=\{1,2,3\}$ ,  $(x,y) \in R \leftrightarrow (x-y)$  is even.

- $A \times B = \{(1,1), (1,2), (1,3), (2,1), (2,2), (2,3)\}$
  - $(1,1) \in R, (1,3) \in R, (2,2) \in R.$

**Examples.**  $x > y$ ,  $x$  owes  $y$ ,  $x$  divides  $y$

## Graphically

- **Example.**  $A=\{1,2\}$ ,  $B=\{1,2,3\}$ ,  $(x,y) \in R \leftrightarrow (x-y)$  is even.
  - $A \times B = \{(1,1), (1,2), (1,3), (2,1), (2,2), (2,3)\}$
  - $(1,1) \in R$ ,  $(1,3) \in R$ ,  $(2,2) \in R$ .



**Definition 54.** Let  $R$  be a relation from the set  $A$  to the set  $B$ . The [inverse relation](#)  $R^{-1}$  from  $B$  to  $A$  is defined as

$$R^{-1} = \{(y, x) \in B \times A, (x, y) \in R\}.$$

What it says is that for every pair  $(x, y)$  in  $R$ , you take it, flip the role of  $x$  and  $y$  to get  $(y, x)$ , which then belongs to  $R^{-1}$ .

**Example 85.** Consider the sets  $A = \{2, 3, 4\}$ ,  $B = \{2, 6, 8\}$ , with the relation  $(x, y) \in R \leftrightarrow x \text{ divides } y$ . Let us look at it step by step. First we compute the cartesian product  $A \times B$ :

$$A \times B = \{(2, 2), (2, 6), (2, 8), (3, 2), (3, 6), (3, 8), (4, 2), (4, 6), (4, 8)\}.$$

Then we check for which pair  $(x, y)$  it is true that  $x \mid y$ . For example, if  $(x, y) = (2, 2)$ , then  $2 \mid 2$  and  $(2, 2) \in R$ , but for  $(x, y) = (3, 2)$ , 3 does not divide 2, and  $(3, 2)$  is not in  $R$ . Trying out all the pairs, we get

$$R = \{(2, 2), (2, 6), (2, 8), (3, 6), (4, 8)\}.$$

Now for every pair  $(x, y) \in R$ , we flip the role of  $x$  and  $y$  to get

$$R^{-1} = \{(2, 2), (6, 2), (8, 2), (6, 3), (8, 4)\}.$$

In this case, there is a nice interpretation of what  $R^{-1}$  means:  $(x, y) \in R \leftrightarrow x \mid y$ , but  $x \mid y \iff y$  is a multiple of  $x$  and  $R^{-1}$  describes the relation  $(y, x) \in R^{-1} \leftrightarrow y$  is a multiple of  $x$ . If one draws a diagram, then to go from  $R$  to  $R^{-1}$ , all is needed is to change the direction of the arrows!

Apart diagrams, another convenient way to represent a relation is to use a matrix representation. Take a binary relation  $R$  from the set  $A = \{a_1, \dots, a_m\}$  to the set  $B = \{b_1, b_2, \dots, b_n\}$ . Create a matrix whose rows are indexed by the elements of  $A$  (thus  $m$  rows) and whose columns are indexed by the elements of  $B$  (thus  $n$  columns). Now the entry  $(i, j)$  of the matrix, corresponding to the  $i$ th row and  $j$ th column, contains  $a_i R b_j$ , that is, a truth value (True or False), depending on whether it is true or not that  $a_i R b_j$  (that is,  $a_i$  is related to  $b_j$ ).

**Example 86.** Take  $A = \{2, 3, 4\}$ ,  $B = \{2, 6, 8\}$  and the relation  $R$  defined by  $(x, y) \in R \leftrightarrow x \text{ divides } y$ . Then the rows of the matrix are indexed by 2, 3, 4, and the columns by 2, 6, 8. We thus get

$$\begin{pmatrix} 2R2 & 2R6 & 2R8 \\ 3R2 & 3R6 & 3R8 \\ 4R2 & 4R6 & 4R8 \end{pmatrix} = \begin{pmatrix} T & T & T \\ F & T & F \\ F & F & T \end{pmatrix}.$$

## Inverse of a Binary Relation

Let  $R$  be a relation from  $A$  to  $B$ . The inverse relation  $R^{-1}$  from  $B$  to  $A$  is defined as:  $R^{-1} = \{(y,x) \in B \times A \mid (x,y) \in R\}$ .

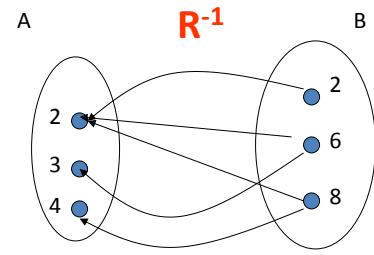
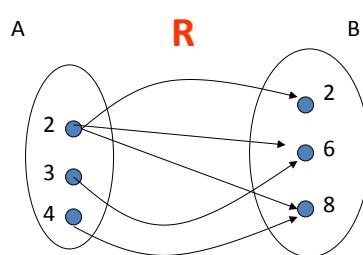
**Example.**  $A=\{2,3,4\}$ ,  $B=\{2,6,8\}$ ,  $(x, y) \in R \leftrightarrow x \text{ divides } y$ .

- $A \times B = \{(2,2), (2,6), (2,8), (3,2), (3,6), (3,8), (4,2), (4,6), (4,8)\}$
- $(2,2) \in R, (2,6) \in R, (2,8) \in R, (3,6) \in R, (4,8) \in R$
- $(2,2) \in R^{-1}, (6,2) \in R^{-1}, (8,2) \in R^{-1}, (6,3) \in R^{-1}, (8,4) \in R^{-1}$
- $(y, x) \in R^{-1} \leftrightarrow y \text{ is a multiple of } x$ .

## Graphically

**Example.**  $A=\{2,3,4\}$ ,  $B=\{2,6,8\}$ ,  $(x, y) \in R \leftrightarrow x \text{ divides } y$ .

- $(2,2) \in R, (2,6) \in R, (2,8) \in R, (3,6) \in R, (4,8) \in R$
- $(2,2) \in R^{-1}, (6,2) \in R^{-1}, (8,2) \in R^{-1}, (6,3) \in R^{-1}, (8,4) \in R^{-1}$



## Matrix Representation (I)

$$A = (a_1, a_2, a_3), B = (b_1, b_2, b_3, b_4), \\ R = \{(a_1, b_2), (a_2, b_1), (a_3, b_1), (a_3, b_4)\}$$

$a_i R b_j$  is represented by true, false else:

$F$	$T$	$F$	$F$
$T$	$F$	$F$	$F$
$T$	$F$	$F$	$T$

**Example.**  $A=\{2,3,4\}, B=\{2,6,8\}$ ,  
 $(x, y) \in R \leftrightarrow x \text{ divides } y$ .

$T$	$T$	$T$
$F$	$T$	$F$
$F$	$F$	$T$

6/15

## Matrix Representation (II)

$R$  relation from  $A$  to  $B$ :  $R^{-1} = \{(y, x) \in B \times A \mid (x, y) \in R\}$ .

$$A = (a_1, a_2, a_3), B = (b_1, b_2, b_3, b_4), \\ R = \{(a_1, b_2), (a_2, b_1), (a_3, b_1), (a_3, b_4)\} \\ R^{-1} = \{(b_2, a_1), (b_1, a_2), (b_1, a_3), (b_4, a_3)\}$$

$$a_i R b_j = \text{true} \quad \begin{bmatrix} F & T & F & F \\ T & F & F & F \\ T & F & F & T \end{bmatrix} \quad b_i R^{-1} a_j = \text{true} \quad \begin{bmatrix} F & T & T \\ T & F & F \\ F & F & F \\ F & F & T \end{bmatrix}$$

The matrix of  $R^{-1}$  is the transpose of the matrix of  $R$ .

## Composition of Relations

Given  $R$  in  $A \times B$ , and  $S$  in  $B \times C$ , the **composition** of  $R$  and  $S$  is a relation on  $A \times C$  defined by

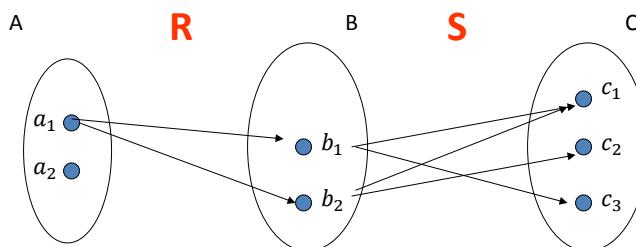
$$R \circ S = \{(a, c) \in A \times C \mid \exists b \in B, aRb \text{ and } bSc\}.$$

**Example.**  $A = \{a_1, a_2\}, B = \{b_1, b_2\}, C = \{c_1, c_2, c_3\}$

- $R = \{(a_1, b_1), (a_1, b_2)\}$
- $S = \{(b_1, c_1), (b_2, c_1), (b_1, c_3), (b_2, c_2)\}$
- What is  $R \circ S$  ?
  
- $R \circ S = \{(a_1, c_1), (a_1, c_3), (a_1, c_2)\}$

## Graphically

- **Example.**  $A = \{a_1, a_2\}, B = \{b_1, b_2\}, C = \{c_1, c_2, c_3\}$
- $R = \{(a_1, b_1), (a_1, b_2)\}$
- $S = \{(b_1, c_1), (b_2, c_1), (b_1, c_3), (b_2, c_2)\}$
- $R \circ S = \{(a_1, c_1), (a_1, c_3), (a_1, c_2)\}$



We may ask next how to interpret the inverse relation  $R^{-1}$  on its matrix. First of all, if  $R$  goes from  $A = \{a_1, \dots, a_m\}$  to  $B = \{b_1, b_2, \dots, b_n\}$ , then  $R^{-1}$  goes from  $B$  to  $A$ . This means that the rows of the matrix of  $R^{-1}$  will be indexed by the set  $B = \{b_1, b_2, \dots, b_n\}$ , while its columns by the set  $A = \{a_1, \dots, a_m\}$ . Then, by definition of  $R^{-1}$ , whenever there was a T (true) in row  $i$  and column  $j$ , this meant that  $(a_i, b_j) \in R$ , thus  $(b_j, a_i) \in R^{-1}$ , and this becomes a T (true) in row  $j$  and column  $i$ . If you take the first row of the matrix of  $R$ , whenever  $(a_1, b_j) \in R$ , for the column  $j$ ,  $(b_j, a_1) \in R^{-1}$ , and a true in the first row of  $R$  becomes a true in the first column of  $R^{-1}$ , and the other entries which are false in the first row of  $R$  similarly become false in the first column of  $R^{-1}$ . This shows that the matrix of  $R^{-1}$  is the transpose of  $R$ ! (recall that the transpose of a matrix is obtained by switching rows and columns).

**Example 87.** We continue the above example with  $A = \{2, 3, 4\}$ ,  $B = \{2, 6, 8\}$  and the relation  $R$  defined by  $(x, y) \in R \leftrightarrow x \text{ divides } y$ . We have that  $R = \{(2, 2), (2, 6), (2, 8), (3, 6), (4, 8)\}$  thus  $R^{-1} = \{(2, 2), (6, 2), (8, 2), (6, 3), (8, 4)\}$ . Then the matrix of  $R$  and  $R^{-1}$  are respectively given by

$$\begin{pmatrix} T & T & T \\ F & T & F \\ F & F & T \end{pmatrix}, \quad \begin{pmatrix} T & F & F \\ T & T & F \\ T & F & T \end{pmatrix}.$$

We continue to explore properties of relations.

**Definition 55.** Given two relations  $R \in A \times B$  and  $S \in B \times C$ , the [composition](#) of  $R$  and  $S$  is a relation on  $A \times C$  defined by

$$R \circ S = \{(a, c) \in A \times C, \exists b \in B, aRb, bSc\}.$$

What it says is that for  $(a, c)$  to be part of your relation  $R \circ S$ , we need to find an element  $b \in B$ , with the property that  $a$  is in relation with  $b$ , and  $b$  is in relation with  $c$ . It is probably best visualize on a diagram: draw 3 circles for  $A, B, C$ , and arrows from  $A$  to  $B$  using the relation  $R$ , and arrows from  $B$  to  $C$  using the relation  $S$ . If you can find a path following those arrows from  $a$  to  $c$ , then  $(a, c)$  is in  $R \circ S$ .

**Example 88.** Consider the sets  $A = \{a_1, a_2\}$ ,  $B = \{b_1, b_2\}$ ,  $C = \{c_1, c_2, c_3\}$ , with relations defined by

$$R = \{(a_1, b_1), (a_1, b_2)\}, \quad S = \{(b_1, c_1), (b_2, c_1), (b_1, c_3), (b_2, c_2)\}.$$

## Reflexivity

A relation  $R$  on a set  $A$  is **reflexive** if every element of  $A$  is related to itself:  $\forall x \in A, xRx$

Examples.

1.  $A=\mathbb{Z}$ ,  $xRy \Leftrightarrow x=y$  : reflexive
2.  $A=\mathbb{Z}$ ,  $xRy \Leftrightarrow x>y$  : not reflexive
3. Reflexivity on the matrix representing  $R$ ?

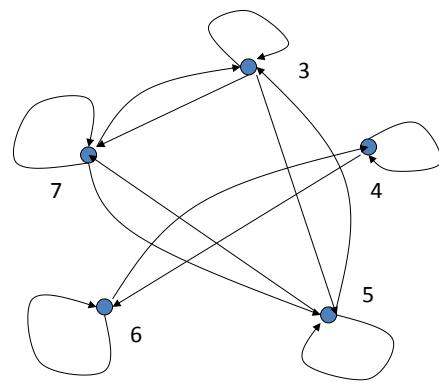


Escher, M C (1948); 'Drawing Hands'

## Graphically

$A = \{3, 4, 5, 6, 7\}$ ,  $xRy \Leftrightarrow (x-y)$  is even

- $R$  reflexive



To compute  $R \circ S$ , start with  $(a_1, b_1)$  and look for pairs starting with  $b_1$  in  $S$ :  $(b_1, c_1)$  and  $(b_1, c_3)$ . Therefore,  $(a_1, b_1)$  combined with  $(b_1, c_1)$  gives the pair  $(a_1, c_1)$  and the pair  $(a_1, b_1)$  combined with  $(b_1, c_3)$  gives  $(a_1, c_3)$ . We do the same with  $(a_1, b_2)$  and pairs starting with  $b_2$  in  $S$  to find  $(a_1, c_2)$ , and

$$R \circ S = \{(a_1, c_1), (a_1, c_2), (a_1, c_3)\}.$$

So far, we were looking at binary relations from  $A$  to  $B$ . Next we focus on relations where  $A = B$ , that is we have relations from a set into itself.

**Definition 56.** A relation  $R$  on a set  $A$  is **reflexive** if every element of  $A$  is related to itself:  $\forall x \in A, xRx$ .

**Example 89.** If  $A$  is the set  $\mathbb{Z}$  of integers, and the relation  $R$  is defined by  $xRy \leftrightarrow x = y$ , then this relation is reflexive, because it is true that  $x$  is always in relation with itself ( $xRx \leftrightarrow x = x$  is always true).

But  $xRy \leftrightarrow x > y$  is not reflexive, because it is never true that  $xRx$  (we never have  $x > x$ ).

On the matrix representation of  $R$ , reflexivity is shown by having T (true) on the diagonal of the matrix. If one represents a relation on itself with a diagram, reflexivity will be seen by having arrows looping on every element of the diagram!

**Definition 57.** A relation  $R$  on a set  $A$  is **symmetric** if  $(x, y) \in R$  implies  $(y, x) \in R$ :  $\forall x \forall y \in A, xRy \rightarrow yRx$ .

On a diagram, this is visualized with having a second arrow between 2 elements of  $A$  in the other direction whenever you have one arrow in one direction.

**Example 90.** If  $A$  is the set  $\mathbb{Z}$  of integers, and the relation  $R$  is defined by  $xRy \leftrightarrow x = y$ , then this relation is symmetric, because it is true that if  $x$  is in relation with  $y$  then  $y$  is in relation with  $x$  ( $xRy \leftrightarrow x = y$  implies  $y = x \leftrightarrow yRx$ ).

But  $xRy \leftrightarrow x > y$  is not symmetric, because it is never true that  $xRy$  implies  $yRx$  (we never have  $x > y$  that implies  $y > x$ ).

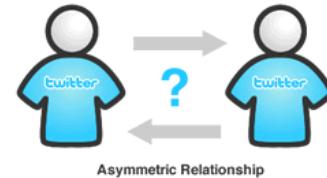
**Definition 58.** A relation  $R$  on a set  $A$  is **transitive** if  $(x, y) \in R$  and  $(y, z) \in R$  implies  $(x, z) \in R$ :  $\forall x \forall y \forall z \in A, xRy \wedge yRz \rightarrow xRz$ .

## Symmetry

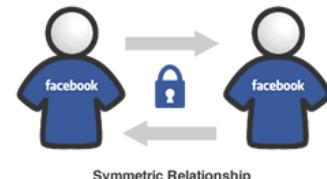
A relation  $R$  on a set  $A$  is **symmetric** if  $(x, y) \in R$  implies  $(y, x) \in R$ :  $\forall x \forall y xRy \rightarrow yRx$

### Examples.

1.  $A=\mathbb{Z}$ ,  $xRy \leftrightarrow x=y$  : symmetric



2.  $A=\mathbb{Z}$ ,  $xRy \leftrightarrow x>y$  : not symmetric

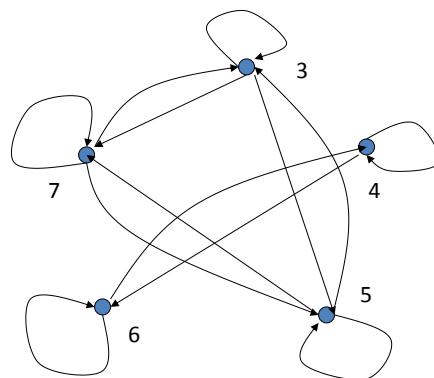


<http://bokardo.com/archives/relationship-symmetry-in-social-networks-why-facebook-will-go-fully-asymmetric/>

## Graphically

$A = \{3, 4, 5, 6, 7\}$ ,  $xRy \leftrightarrow |x-y|$  is even

- $R$  reflexive
- $R$  symmetric



**Example 91.** If  $A$  is the set  $\mathbb{Z}$  of integers, and the relation  $R$  is defined by  $xRy \leftrightarrow x = y$ , this relation is transitive, because it is true that if  $x$  is in relation with  $y$  and  $y$  is in relation with  $z$  then  $x$  is in relation with  $z$  ( $xRy \leftrightarrow x = y$  and  $y = z \leftrightarrow yRz$  implies that  $x = y = z$  that is  $x = z \leftrightarrow xRz$ ).

Also  $xRy \leftrightarrow x > y$  is transitive, because if  $xRy \leftrightarrow x > y$  and  $yRz \leftrightarrow y > z$ , then we have  $x > y > z$  that is  $x > z \leftrightarrow xRz$ .

If a relation  $R$  on a set  $A$  turns out to satisfy the 3 properties we have just seen: reflexivity, symmetry, and transitivity, then this relation is special, and thus gets a special name:

**Definition 59.** A relation  $R$  on a set  $A$  is an **equivalence relation** if  $R$  is reflexive, symmetric and transitive. The **equivalence class** of  $a$  in  $A$  is

$$[a] = \{x \in A, aRx\}.$$

There is a reason for this name: an equivalence relation is so strong, it so strongly ties together elements that are in relation with each other, that instead of looking at elements one by one, we can just consider all those elements in relation with each other as one entity, called equivalence class.

**Example 92.** Consider the set  $A = \{3, 4, 5, 6, 7\}$  with the relation  $xRy \leftrightarrow (x - y)$  is even. Then  $R$  is reflexive: indeed,  $xRx$  is always true, since  $(x - x) = 0$  which is even. Also  $R$  is symmetric: indeed,  $xRy \leftrightarrow (x - y)$  is even implies that  $-(x - y) = y - x$  is also even, and then  $(y - x)$  is even  $\leftrightarrow yRx$ . Finally it is transitive: if  $xRy \leftrightarrow (x - y)$  is even, and  $yRz \leftrightarrow (y - z)$  is even, then  $(x - z) = (x - y) + (y - z)$  which is even (sum of two even numbers is even), thus  $(x - z)$  is even  $\leftrightarrow xRz$ . The equivalence class of  $[3]$  is the set of elements in relation with 3, that is  $[3] = \{3, 5, 7\}$ , similarly  $[4] = \{4, 6\}$ .

It turns out that equivalence classes partition  $A$  (for  $A$  a set with  $R$  a relation which is an equivalence relation). See Exercise 83.

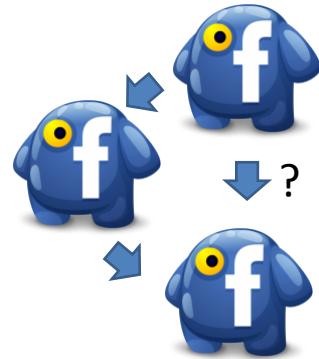
The above example does form an equivalence relation, but it probably does not explain well the concept of equivalence relation, so let us try to get a better feeling using something that we already know (even though we do not know yet that these are equivalence classes!) namely, integers modulo  $n$ .

## Transitivity

A relation R on a set A is **transitive** if  $(x, y) \in R$  and  $(y, z) \in R$  implies  $(x, z) \in R$ :  $\forall x \forall y \forall z xRy \wedge yRz \rightarrow xRz$

### Examples.

1.  $A = \mathbb{Z}$ ,  $xRy \leftrightarrow x=y$  : transitive
2.  $A = \mathbb{Z}$ ,  $xRy \leftrightarrow x > y$  : transitive



<http://www.apkdad.com/tag/atrium-for-facebook-apk/>

## Equivalence Relation

A relation R on a set A is **an equivalence relation** if

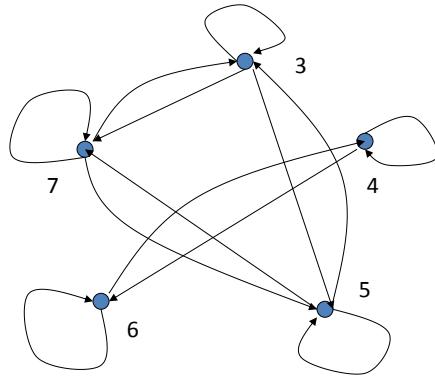
1. R is reflexive:  $\forall x \in A, xRx$
2. R is symmetric:  $\forall x \forall y xRy \rightarrow yRx$
3. R is transitive:  $\forall x \forall y \forall z xRy \wedge yRz \rightarrow xRz$

**Equivalence class of a in A:**  $[a] = \{x \in A \mid aRx\}$   
for R an equivalence relation.

## Example

$A = \{3, 4, 5, 6, 7\}$ ,  $xRy \Leftrightarrow (x-y)$  is even

- $R$  reflexive
- $R$  symmetric
- $R$  transitive
- $[3] = \{3, 5, 7\}, [4] = \{4, 6\}$

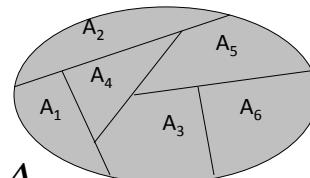


## Equivalence Classes

Partition of a set  $A$ :

$$A_i \cap A_j = \emptyset \quad \text{whenever} \quad i \neq j$$

$$A_1 \cup A_2 \cup A_3 \cup A_4 \cup A_5 \cup A_6 = A$$



Equivalence classes of  $A$  form a partition of  $A$ .

## Integers mod n (I)

$$a \equiv b \pmod{n} \Leftrightarrow a = qn + b$$

$\equiv \pmod{n}$  is **an equivalence relation**:

1.  $\equiv \pmod{n}$  is reflexive:  $\forall x \in A, x \equiv x \pmod{n}$
  2.  $\equiv \pmod{n}$  is symmetric:  $\forall x \forall y x \equiv y \pmod{n} \rightarrow y \equiv x \pmod{n}$
  3.  $\equiv \pmod{n}$  is transitive:  $\forall x \forall y \forall z x \equiv y \pmod{n} \wedge y \equiv z \pmod{n} \rightarrow x \equiv z \pmod{n}$ .
- 

## Integers mod n (II)

Equivalence class of  $[0] = \{0, n, 2n, 3n, \dots, -n, -2n, -3n, \dots\}$

Equivalence class of  $[1] = \{1, n+1, 2n+1, 3n+1, \dots, -n+1, -2n+1, \dots\}$

**Example.** Integers mod 4



- Integers mod  $n$  can be represented as elements between 0 and  $n-1$ :  $\{0, 1, 2, \dots, n-1\}$
-

**Example 93.** The relation  $\equiv \pmod{n}$  is an equivalence relation on  $\mathbb{Z}$ .

- It is reflexive:  $x \equiv x \pmod{n}$  is always true.
- It is symmetric:  $x \equiv y \pmod{n}$  means that  $x = qn + y$  for some integer  $q$ , thus  $y = -qn + x$  and  $y \equiv x \pmod{n}$ .
- It is transitive: if  $x \equiv y \pmod{n}$  and  $y \equiv z \pmod{n}$  then we have  $x = qn + y$  and  $y = rn + z$  thus  $x = qn + y = qn + rn + z = n(q+r) + z$  and  $x \equiv z \pmod{n}$ .

Now what is the equivalence class of 0? it is formed by all multiples of  $n$ :

$$[0] = \{\dots, -2n, -n, 0, n, 2n, \dots\},$$

and similarly the equivalence class of 1 is all multiples of  $n$ , plus 1, and we see that there are exactly  $n$  equivalence classes, which partition  $\mathbb{Z}$ :

$$[0], [1], [2], \dots, [n-1].$$

This is why when we do operations modulo  $n$ , we are allowed to pick one element per equivalence class, namely  $0, 1, \dots, n-1$  and work with them!!

We add one more property to those we know: reflexivity, symmetry, and transitivity.

**Definition 60.** A relation  $R$  on a set  $A$  is **antisymmetric** if  $(x, y) \in R$  and  $(y, x) \in R$  implies  $x = y$ :  $\forall x \forall y, xRy \wedge yRx \rightarrow x = y$ .

Note that symmetry and antisymmetric are not related, despite their name, see Exercise 80.

**Example 94.** If  $A$  is the set  $\mathbb{Z}$  of integers, and the relation  $R$  is defined by  $xRy \leftrightarrow x = y$ , this relation is antisymmetric, because it is true that if  $x$  is in relation with  $y$  and  $y$  is in relation with  $x$  then  $x = y$  ( $xRy \leftrightarrow x = y$  and  $y = x \leftrightarrow yRx$  implies that  $x = y$ ).

Also  $xRy \leftrightarrow x > y$  is antisymmetric, because we have a statement which is vacuously true!! if  $xRy \leftrightarrow x > y$  and  $yRx \leftrightarrow y > x$ , well, this statement is always false...when we have a  $p \rightarrow q$  where  $p$  is false then  $p \rightarrow q$  is true (apply here with  $p = "xRy \wedge yRx"$  and  $q = "x = y"$ ).

Consider two sets  $B$  and  $C$  and the relation  $B$  is in relation with  $C \leftarrow B \subseteq C$ . Then  $B \subseteq C$  and  $C \subseteq B$  implies that  $B = C$ ! this is what we used to show set equality (double inclusion), and this shows that this relation is antisymmetric!

## Antisymmetry

A relation  $R$  on a set  $A$  is **antisymmetric** if  $(x, y) \in R$  and  $(y, x) \in R$  implies  $x=y$ :  $\forall x \forall y xRy \wedge yRx \rightarrow x = y$

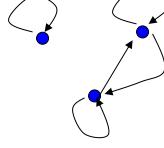
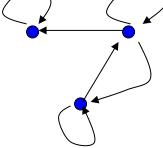
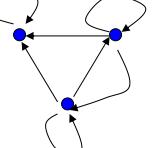
### Examples.

1.  $A=\mathbb{Z}$ ,  $xRy \leftrightarrow x=y$  : antisymmetric

2.  $A=\mathbb{Z}$ ,  $xRy \leftrightarrow x>y$  : vacuously true

3.  $B \subseteq C \leftrightarrow B \subseteq C$  : antisymmetric

## Examples

			
Reflexive?	Y	Y	Y
Symmetric?	Y	N	N
Antisymmetric?	N	N	N
Transitive?	Y	N	Y

**Definition 61.** A relation  $R$  on a set  $A$  is a **partial order** if  $R$  is reflexive, antisymmetric and transitive.

The word partial order can be explained by the antisymmetry property. It is not possible to have "a loop" between two elements, namely a relation from one element to another, and back.

**Example 95.** If  $A$  is the set  $\mathbb{Z}$  of integers, and the relation  $R$  is defined by  $xRy \leftrightarrow x \leq y$ , this relation is a partial order:

- It is reflexive:  $x \leq x$  always.
- It is antisymmetric:  $x \leq y$  and  $y \leq x$  implies that  $x = y$ .
- It is transitive: if  $x \leq y$  and  $y \leq z$  then  $x \leq y \leq z$  and thus  $x \leq z$  as needed.

A set with a relation  $R$  may not satisfy the transitivity property, but then, one may wonder whether it is possible to "complete" the set with more elements to obtain the transitivity property. This gives rise to the notion of transitive closure:

**Definition 62.** Consider a relation  $R$  on a set  $A$ . The **transitive closure** of  $R$  is the binary relation  $R^t$ , that satisfies the properties:

- $R^t$  is transitive,
- $R \subseteq R^t$ ,
- If  $S$  is any other transitive relation that contains  $R$ , then  $R^t \subset S$ .

The first property says the property of transitivity is satisfied, the second one that  $R$  is contained in  $R^t$  and the third one says  $R^t$  is minimal with this property!

## Partial Order

A relation  $R$  on a set  $A$  is a **partial order** if  $R$  is reflexive, antisymmetric and transitive.

**Example.**  $A=\mathbb{Z}$ ,  $xRy \Leftrightarrow x \leq y$

Notion of partial order useful for scheduling problems across possibly different domains.

---

## Transitive Closure

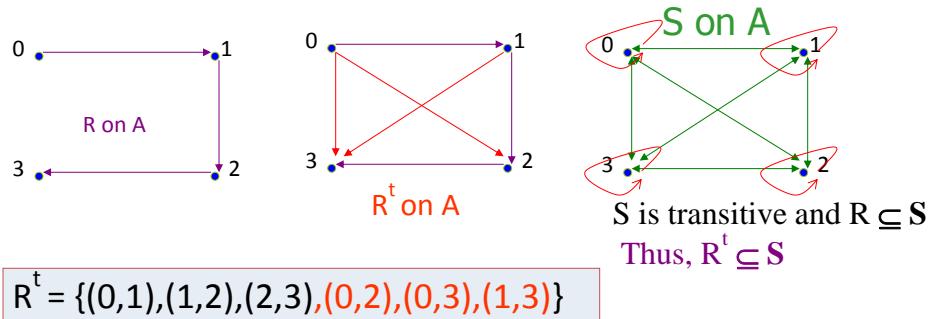
Let  $A$  be a set and  $R$  a binary relation on  $A$ .  
The **transitive closure** of  $R$  is the binary relation  $R^t$  on  $A$  that satisfies the following three properties:

1.  $R^t$  is Transitive
  2.  $R \subseteq R^t$
  3. If  $S$  is any other transitive relation that contains  $R$ , then  $R^t \subseteq S$
-

## Example

Let  $A = \{0,1,2,3\}$

Consider a relation  $R = \{(0,1), (1,2), (2,3)\}$  on  $A$ .



## Non-binary Relations

Let  $A_1, \dots, A_n$  be sets. A **n-ary relation  $R$**  is a subset of  $A_1 \times \dots \times A_n$ .  $a_1, \dots, a_n$  are related if  $(a_1, \dots, a_n) \in R$ .

## Exercises for Chapter 9

**Exercise 72.** Consider the sets  $A = \{1, 2\}$ ,  $B = \{1, 2, 3\}$  and the relation  $(x, y) \in R \iff (x - y)$  is even. Compute the inverse relation  $R^{-1}$ . Compute its matrix representation.

**Exercise 73.** Consider the sets  $A = \{2, 3, 4\}$ ,  $B = \{2, 6, 8\}$  and the relation  $(x, y) \in R \iff x \mid y$ . Compute the matrix of the inverse relation  $R^{-1}$ .

**Exercise 74.** Let  $R$  be a relation from  $\mathbb{Z}$  to  $\mathbb{Z}$  defined by  $xRy \leftrightarrow 2|(x - y)$ . Show that if  $n$  is odd, then  $n$  is related to 1.

**Exercise 75.** This exercise is about composing relations.

1. Consider the sets  $A = \{a_1, a_2\}$ ,  $B = \{b_1, b_2\}$ ,  $C = \{c_1, c_2, c_3\}$  with the following relations  $R$  from  $A$  to  $B$ , and  $S$  from  $B$  to  $C$ :

$$R = \{(a_1, b_1), (a_1, b_2)\}, \quad S = \{(b_1, c_1), (b_2, c_1), (b_1, c_3), (b_2, c_2)\}.$$

What is the matrix of  $R \circ S$ ?

2. In general, what is the matrix of  $R \circ S$ ?

**Exercise 76.** Consider the relation  $R$  on  $\mathbb{Z}$ , given by  $aRb \iff a - b$  divisible by  $n$ . Is it symmetric?

**Exercise 77.** Consider a relation  $R$  on any set  $A$ . Show that  $R$  symmetric if and only if  $R = R^{-1}$ .

**Exercise 78.** Consider the set  $A = \{a, b, c, d\}$  and the relation

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

Is this relation reflexive? symmetric? transitive?

**Exercise 79.** Consider the set  $A = \{0, 1, 2\}$  and the relation  $R = \{(0, 2), (1, 2), (2, 0)\}$ . Is  $R$  antisymmetric?

**Exercise 80.** Are symmetry and antisymmetry mutually exclusive?

**Exercise 81.** Consider the relation  $R$  given by divisibility on positive integers, that is  $xRy \leftrightarrow x|y$ . Is this relation reflexive? symmetric? antisymmetric? transitive? What if the relation  $R$  is now defined over non-zero integers instead?

**Exercise 82.** Consider the set  $A = \{0, 1, 2, 3, 4, 5, 6, 7, 8\}$ . Show that the relation  $xRy \leftrightarrow 2|(x - y)$  is an equivalence relation.

**Exercise 83.** Show that given a set  $A$  and an equivalence relation  $R$  on  $A$ , then the equivalence classes of  $R$  partition  $A$ .

**Exercise 84.** Consider the set  $A = \{2, 3, 4, 5, 6, 7, 8, 9, 10\}$  and the relation

$$xRy \leftrightarrow \exists c \in \mathbb{Z}, y = cx.$$

Is  $R$  an equivalence relation? Is  $R$  a partial order?



# Chapter 10

## Functions

*“One of the most important concepts in all of mathematics is that of function.” (T.P. Dick and C.M. Patton)*

Functions...finally a topic that most of you must be familiar with. However here, we will not study derivatives or integrals, but rather the notions of one-to-one and onto (or injective and surjective), how to compose functions, and when they are invertible.

Let us start with a formal definition.

**Definition 63.** Let  $X$  and  $Y$  be sets. A **function**  $f$  from  $X$  to  $Y$  is a rule that assigns every element  $x$  of  $X$  to a unique  $y$  in  $Y$ . We write  $f : X \rightarrow Y$  and  $f(x) = y$ . Formally, using predicate logic:

$$(\forall x \in X, \exists y \in Y, y = f(x)) \wedge (\forall x_1, x_2 \in X, f(x_1) \neq f(x_2) \rightarrow x_1 \neq x_2).$$

Then  $X$  is called the **domain** of  $f$ , and  $Y$  is called the **codomain** of  $f$ . The element  $y$  is the **image** of  $x$  under  $f$ , while  $x$  is the **preimage** of  $y$  under  $f$ . Finally, we call **range** the subset of  $Y$  with preimages.

**Example 96.** Consider the assignment rule  $f : X = \{a, b, c\} \rightarrow Y = \{1, 2, 3, 4\}$  which is defined by:  $f = \{(a, 2), (b, 4), (c, 2)\}$ . We first check that this is a function. For every element in  $X$ , we do have an assignment:  $f(a) = 2$ ,  $f(b) = 4$ ,  $f(c) = 2$ . Then the condition that whenever  $f(x_1) \neq f(x_2)$  it must be that  $x_1 \neq x_2$  is also satisfied. The domain of  $f$  is  $X$ , the codomain of  $f$  is  $Y$ . The preimage of 2 is  $\{a, c\}$  because  $f(a) = f(c) = 2$ . For the range, we look at  $Y$ , and among 1, 2, 3, 4, only 2 and 4 have a preimage, therefore the range is  $\{2, 4\}$ .

## Function

Let  $X$  and  $Y$  be sets. A **function**  $f$  from  $X$  to  $Y$  is a rule that assigns every element  $x$  of  $X$  to a unique  $y$  in  $Y$ . We write  $f: X \rightarrow Y$  and  $f(x) = y$

$$(\forall x \in X \exists y \in Y, y = f(x)) \wedge (\forall x_1, x_2 \in X, f(x_1) \neq f(x_2) \rightarrow x_1 \neq x_2)$$

$X$  = domain,  $Y$  = codomain

$y$  = image of  $x$  under  $f$ ,

$x$  = preimage of  $y$  under  $f$

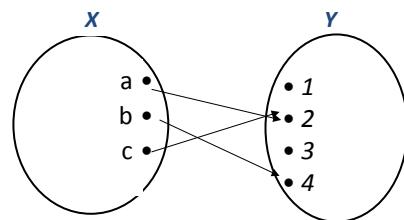
range = subset of  $Y$  with preimages

## Example 1

$$(\forall x \in X \exists y \in Y, y = f(x)) \wedge (\forall x_1, x_2 \in X, f(x_1) \neq f(x_2) \rightarrow x_1 \neq x_2)$$

Arrow Diagram of  $f$ :

Domain  $X=\{a,b,c\}$ ,  
 Co-domain  $Y=\{1,2,3,4\}$   
 $f=\{(a,2),(b,4),(c,2)\}$ ,  
 preimage of 2 is {a,c}  
 Range={2,4}



**Example 97.** The rule  $f$  that assigns the square of an integer to this integer is a function. Indeed, every integer has an image: its square. Also whenever two squares are different, it must be that their square roots were different. We write

$$f : \mathbb{Z} \rightarrow \mathbb{Z}, \quad f(x) = x^2.$$

Its domain is  $\mathbb{Z}$ , its codomain is  $\mathbb{Z}$  as well, but its range is  $\{0, 1, 4, 9, 16, \dots\}$ , that is the set of squares in  $\mathbb{Z}$ .

**Definition 64.** Let  $f$  be a function from  $X$  to  $Y$ ,  $X, Y$  two sets, and consider the subset  $S \subset X$ . The **image of the subset**  $S$  is the subset of  $Y$  that consists of the images of the elements of  $S$ :  $f(S) = \{f(s), s \in S\}$

We next move to our first important definition, that of one-to-one.

**Definition 65.** A function  $f$  is **one-to-one** or **injective** if and only if  $f(x) = f(y)$  implies  $x = y$  for all  $x, y$  in the domain  $X$  of  $f$ . Formally:

$$\forall x, y \in X (f(x) = f(y) \rightarrow x = y).$$

In words, this says that all elements in the domain of  $f$  have different images.

**Example 98.** Consider the function  $f : \mathbb{R} \rightarrow \mathbb{R}$ ,  $f(x) = 4x - 1$ . We want to know whether each element of  $\mathbb{R}$  has a different image. Yes, this is the case, why? well, visually, this function is a line, so one may "see" that two distinct elements have distinct images, but let us try a proof of this. We have to show that  $f(x) = f(y)$  implies  $x = y$ . Ok, let us take  $f(x) = f(y)$ , that is two images that are the same. Then  $f(x) = 4x - 1$ ,  $f(y) = 4y - 1$ , and thus we must have  $4x - 1 = 4y - 1$ . But then  $4x = 4y$  and it must be that  $x = y$ , as we wanted. Therefore  $f$  is injective.

**Example 99.** Consider the function  $g : \mathbb{R} \rightarrow \mathbb{R}$ ,  $g(x) = x^2$ . Do we also have that two distinct reals have distinct images? Well no... because 1 and  $-1$  are both sent to 1...so this function is not injective! If  $g(x) = g(y) = 1$ , we cannot conclude that  $x = y$ , in fact this is wrong, it could be that  $x = -y$ .

The other definition that always comes in pair with that of one-to-one/injective is that of onto.

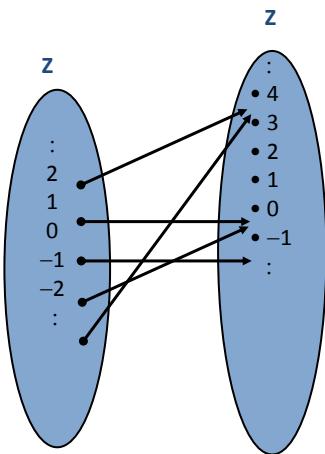
## Example 2

Let  $f$  be the function from  $\mathbf{Z}$  to  $\mathbf{Z}$  that assigns the square of an integer to this integer.

Then,  $f: \mathbf{Z} \rightarrow \mathbf{Z}, f(x) = x^2$

Domain and co-domain of  $f: \mathbf{Z}$

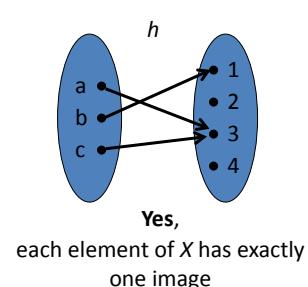
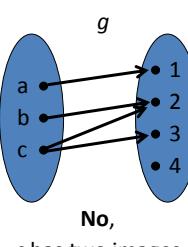
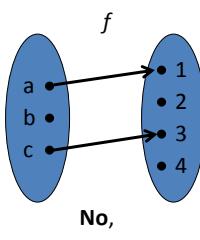
$\text{Range}(f) = \{0, 1, 4, 9, 16, 25, \dots\}$



## Functions Vs Non-functions

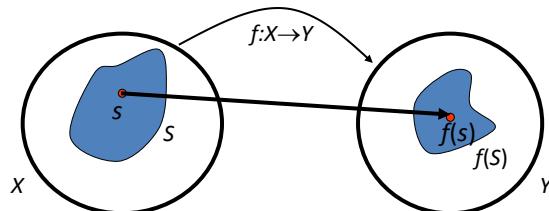
$$(\forall x \in X \exists y \in Y, y = f(x)) \wedge (\forall x_1, x_2 \in X, f(x_1) \neq f(x_2) \rightarrow x_1 \neq x_2)$$

$X = \{a,b,c\}$  to  $Y = \{1,2,3,4\}$



## Image of a Set

Let  $f$  be a function from  $X$  to  $Y$  and  $S \subseteq X$ . The **image of  $S$**  is the subset of  $Y$  that consists of the images of the elements of  $S$ :  $f(S) = \{f(s) \mid s \in S\}$



## One-To-One Function

A function  $f$  is **one-to-one** (or **injective**), if and only if  $f(x) = f(y)$  implies  $x = y$  for all  $x$  and  $y$  in the domain of  $f$ .

**In words:**

*"All elements in the domain of  $f$  have different images"*

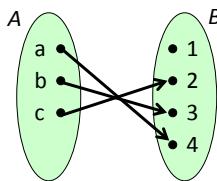
**Mathematical Description:**

$f:A \rightarrow B$  is **one-to-one**  $\Leftrightarrow \forall x_1, x_2 \in A (f(x_1) = f(x_2) \rightarrow x_1 = x_2)$

or

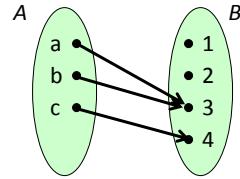
$f:A \rightarrow B$  is **one-to-one**  $\Leftrightarrow \forall x_1, x_2 \in A (x_1 \neq x_2 \rightarrow f(x_1) \neq f(x_2))$

## Example: One-to-One (Injective)



one-to-one

(all elements in A have a different image)



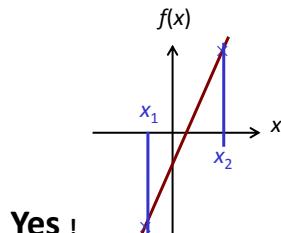
not one-to-one

(a and b have the same image)

## Example: One-To-One (Injective)

$$f: \mathbb{R} \rightarrow \mathbb{R}, f(x) = 4x - 1$$

("Does each element in  $\mathbb{R}$  have a different image ?")



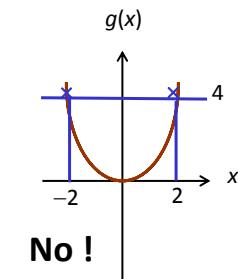
Yes !

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

Take some  $x_1, x_2 \in \mathbb{R}$  with  $f(x_1) = f(x_2)$ .

Then  $4x_1 - 1 = 4x_2 - 1 \Rightarrow 4x_1 = 4x_2 \Rightarrow x_1 = x_2$

$$g: \mathbb{R} \rightarrow \mathbb{R}, g(x) = x^2$$



No !

Take  $x_1 = 2$  and  $x_2 = -2$ .

Then  $g(x_1) = 2^2 = 4 = g(x_2)$   
and  $x_1 \neq x_2$

**Definition 66.** A function  $f$  is **onto** or **surjective** if and only if for every element  $y \in Y$ , there is an element  $x \in X$  with  $f(x) = y$ :

$$\forall y \in Y, \exists x \in X, f(x) = y.$$

In words, each element in the co-domain of  $f$  has a pre-image.

**Example 100.** Consider again the function  $f : \mathbb{R} \rightarrow \mathbb{R}$ ,  $f(x) = 4x - 1$ . We want to know whether each element of  $\mathbb{R}$  has a preimage. Yes, it has, let us see why: we want to show that there exists  $x$  such that  $f(x) = 4x - 1 = y$ . Given  $y$ , we have the relation  $x = (y + 1)/4$  thus this  $x$  is indeed sent to  $y$  by  $f$ .

**Example 101.** Consider again the function  $g : \mathbb{R} \rightarrow \mathbb{R}$ ,  $g(x) = x^2$ . Does each element in  $\mathbb{R}$  have a preimage? well, again no... Because  $\mathbb{R}$  contains all the negative real numbers, and it is not possible to square a real number and get something negative... Formally, if  $y = -1$ , there is no  $x \in \mathbb{R}$  such that  $g(x) = x^2 = -1$ .

We next combine the definitions of one-to-one and onto, to get:

**Definition 67.** A function  $f$  is a **one-to-one correspondence** or **bijection** if and only if it is both one-to-one and onto (or both injective and surjective).

An important example of bijection is the identity function.

**Definition 68.** The **identity function**  $i_A$  on the set  $A$  is defined by:

$$i_A : A \rightarrow A, i_A(x) = x.$$

**Example 102.** Consider the function  $f : \mathbb{R} \rightarrow \mathbb{R}$ ,  $f(x) = 4x - 1$ , which we have just studied in two examples. We know it is both injective (see Example 98) and surjective (see Example 100), therefore it is a bijection.

Bijections have a special feature: they are invertible, formally:

**Definition 69.** Let  $f : A \rightarrow B$  be a bijection. Then the **inverse function** of  $f$ ,  $f^{-1} : B \rightarrow B$  is defined elementwise by:  $f^{-1}(b)$  is the unique element  $a \in A$  such that  $f(a) = b$ . We say that  $f$  is invertible.

Note the importance of the hypothesis:  $f$  must be a bijection, otherwise the inverse function is not well defined. For example, if  $f$  is not one-to-one, then  $f^{-1}(b)$  will have more than one value, and thus is not properly defined.

Note that given a bijection  $f : A \rightarrow B$  and its inverse  $f^{-1} : B \rightarrow A$ , we can write formally the above definition as:

$$\forall b \in B, \forall a \in A (f^{-1}(b) = a \iff b = f(a)).$$

## Onto Functions

A function  $f$  from  $X$  to  $Y$  is **onto** (or **surjective**), if and only if for every element  $y \in Y$  there is an element  $x \in X$  with  $f(x) = y$ .

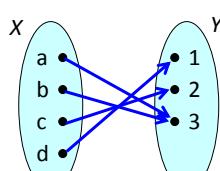
**In words:**

“Each element in the co-domain of  $f$  has a pre-image”

**Mathematical Description:**

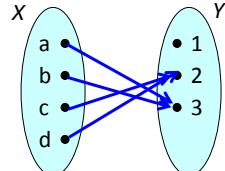
$$f: X \rightarrow Y \text{ is onto} \Leftrightarrow \forall y \exists x, f(x) = y$$

## Example: Onto (Surjective)



onto

(all elements in  $Y$  have a  
pre-image)



not onto

(1 has no pre-image)

## Example: Onto (Surjective)

$$g: \mathbb{R} \rightarrow \mathbb{R}, g(x) = x^2$$

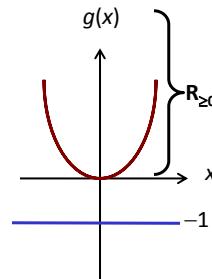
("Does each element in  $\mathbb{R}$  have a pre-image ?")

**No !**

**To show:**  $\exists y \in \mathbb{R}$  such that  $\forall x \in \mathbb{R} g(x) \neq y$

Take  $y = -1$

Then any  $x \in \mathbb{R}$  holds  $g(x) = x^2 \neq -1 = y$



But  $g: \mathbb{R} \rightarrow \mathbb{R}_{\geq 0}$ ,  $g(x) = x^2$ , (where  $\mathbb{R}_{\geq 0}$  denotes the set of non-negative real numbers) is onto !

## One-to-one Correspondence

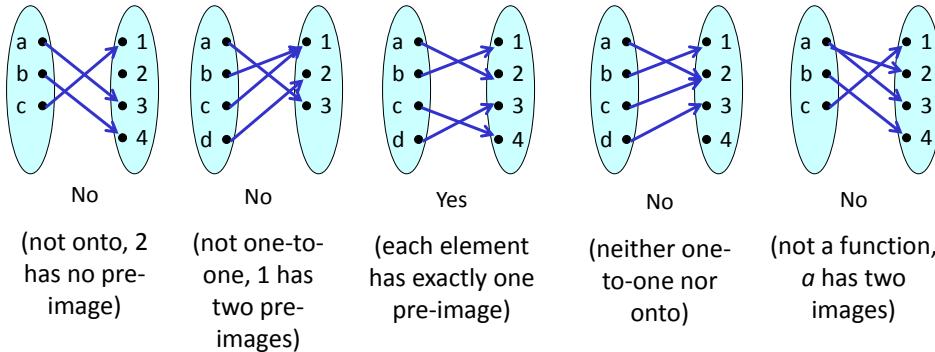
A function  $f$  is a **one-to-one correspondence** (or **bijection**), if and only if it is both one-to-one and onto

**In words:**

"No element in the co-domain of  $f$  has two (or more) pre-images" (*one-to-one*) **and**

"Each element in the co-domain of  $f$  has a pre-image" (*onto*)

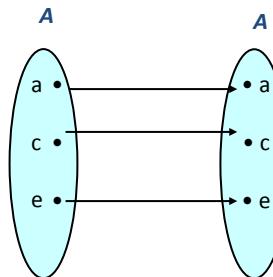
## Example: Bijection



## Identity Function

The **identity function** on a set  $A$  is defined as:  
 $i_A : A \rightarrow A, i_A(x) = x$ .

**Example.** Any identity function is a bijection.  
e.g. for  $A = \{a, c, e\}$ :



**Example 103.** Let us look again at our two previous examples, namely,  $f(x) = 4x - 1$  and  $g(x) = x^2$ . Then  $g(x)$ , for  $g : \mathbb{R} \rightarrow \mathbb{R}$  is not a bijection, so it cannot have an inverse. Now  $f(x)$  is a bijection, so we can compute its inverse. Suppose that  $y = f(x)$ , then

$$y = 4x - 1 \iff y + 1 = 4x \iff x = \frac{y+1}{4},$$

and  $f^{-1}(y) = \frac{y+1}{4}$ .

We saw that for the notion of inverse  $f^{-1}$  to be defined, we need  $f$  to be a bijection. The next result shows that  $f^{-1}$  is a bijection as well.

**Proposition 1.** *If  $f : X \rightarrow Y$  is a one-to-one correspondence, then  $f^{-1} : Y \rightarrow X$  is a one-to-one correspondence.*

*Proof.* To prove this, we just apply the definition of bijection, namely, we need to show that  $f^{-1}$  is an injection, and a surjection. Let us start with injection.

- $f^{-1}$  is an injection: we have to prove that if  $f^{-1}(y_1) = f^{-1}(y_2)$ , then  $y_1 = y_2$ . All right, then  $f^{-1}(y_1) = f^{-1}(y_2) = x$  for some  $x$  in  $X$ . But  $f^{-1}(y_1) = x$  means that  $y_1 = f(x)$ , and  $f^{-1}(y_2) = x$  means that  $y_2 = f(x)$ , by definition of the inverse of function. But this shows that  $y_1 = y_2$ , as needed.
- $f^{-1}$  is a surjection: by definition, we need to prove that any  $x \in X$  has a preimage, that is, there exists  $y$  such that  $f^{-1}(y) = x$ . Because  $f$  is a bijection, there is some  $y$  such that  $y = f(x)$ , therefore  $x = f^{-1}(y)$  as needed.

□

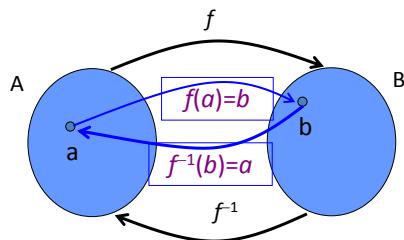
Suppose that you have two functions  $f$  and  $g$ . It may be possible to compose them to obtain a third function, here is how:

**Definition 70.** Let  $f : A \rightarrow B$  be a function, and  $g : B \rightarrow C$  be a function. Then the **composition** of  $f$  and  $g$  is a new function denoted by  $g \circ f$ , and defined by:  $g \circ f : A \rightarrow C$ ,  $(g \circ f)(a) = g(f(a))$ .

Note that the codomain of  $f$  is  $B$ , which is the domain of  $g$ . Under this condition, the composition  $g \circ f$  consists of applying first  $f$ , and then apply  $g$  on the result. Therefore,  $g \circ f \neq f \circ g$  in general!

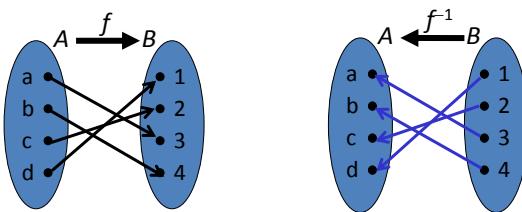
## Inverse Function

Let  $f:A \rightarrow B$  be a one-to-one correspondence (bijection). Then the **inverse function of  $f$** ,  $f^{-1}:B \rightarrow A$ , is defined by:  $f^{-1}(b) = \text{that unique element } a \in A \text{ such that } f(a)=b$ . We say that  $f$  is **invertible**.



## Example 1

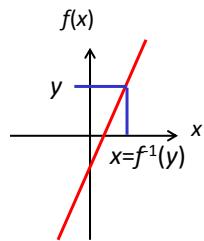
Find the inverse function of the following function:



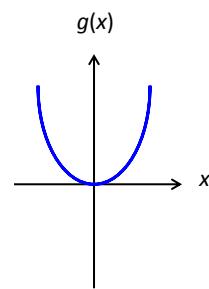
Let  $f:A \rightarrow B$  be a one-to-one correspondence and  $f^{-1}:B \rightarrow A$  its inverse. Then  $\forall b \in B \ \forall a \in A (f^{-1}(b)=a \Leftrightarrow b=f(a))$

## Example 2

What is the inverse of  $f: \mathbb{R} \rightarrow \mathbb{R}$ ,  
 $f(x) = 4x - 1$ ?



What is the inverse of  $g: \mathbb{R} \rightarrow \mathbb{R}$ ,  
 $g(x) = x^2$ ?



Let  $y \in \mathbb{R}$ . Calculate  $x$  with  $f(x) = y$ :

$$y = 4x - 1 \Leftrightarrow (y+1)/4 = x$$

Hence,  $f^{-1}(y) = (y+1)/4$

## One-to-one Correspondence

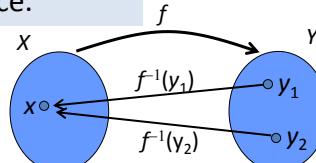
**Theorem 1:** If  $f: X \rightarrow Y$  is a one-to-one correspondence,  
then  $f^{-1}: Y \rightarrow X$  is a one-to-one correspondence.

**Proof:**

(a)  $f^{-1}$  is one-to-one:

Take  $y_1, y_2 \in Y$  such that  $f^{-1}(y_1) = f^{-1}(y_2) = x$ .

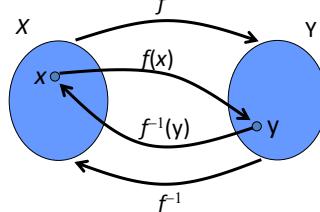
Then  $f(x) = y_1$  and  $f(x) = y_2$ , thus  $y_1 = y_2$ .



(b) To show  $f^{-1}$  is onto:

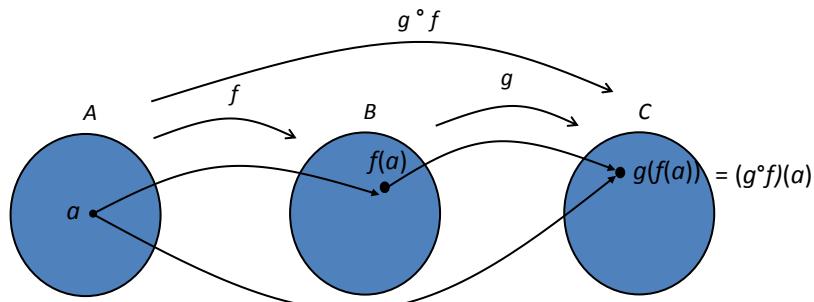
Take some  $x \in X$ , and let  $y = f(x)$ .

Then  $f^{-1}(y) = x$ .



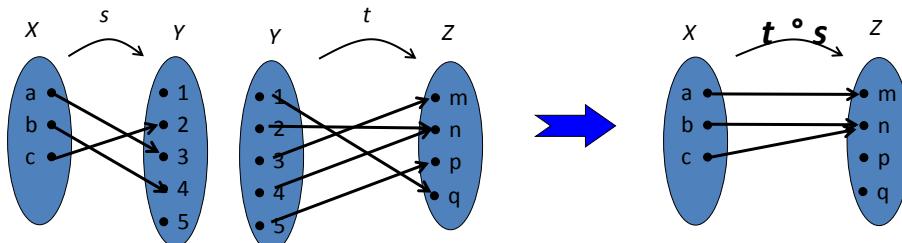
## Composition of Functions

Let  $f:A \rightarrow B$  and  $g:B \rightarrow C$  be functions. The **composition** of the functions  $f$  and  $g$ , denoted as  $g \circ f$ , is defined by:  
 $g \circ f: A \rightarrow C, (g \circ f)(a) = g(f(a))$



## Examples

**Example:** Given functions  $s:X \rightarrow Y$  and  $t:Y \rightarrow Z$ . Find  $t \circ s$  and  $s \circ t$ .



**Example**  $f:Z \rightarrow Z$ ,  $f(n)=2n+3$ ,  $g:Z \rightarrow Z$ ,  $g(n)=3n+2$ . What is  $g \circ f$  and  $f \circ g$ ?

$$(f \circ g)(n) = f(g(n)) = f(3n + 2) = 2(3n + 2) + 3 = 6n + 7$$

$$(g \circ f)(n) = g(f(n)) = g(2n + 3) = 3(2n + 3) + 2 = 6n + 11$$

$f \circ g \neq g \circ f$  (no **commutativity** for the composition of functions !)

**Example 104.** Consider  $f : \mathbb{Z} \rightarrow \mathbb{Z}$  and  $g : \mathbb{Z} \rightarrow \mathbb{Z}$  defined by  $f(n) = 2n + 3$ ,  $g(n) = 3n + 2$ . We have

$$(f \circ g)(n) = f(g(n)) = f(3n + 2) = 2(3n + 2) + 3 = 6n + 7,$$

while

$$(g \circ f)(n) = g(f(n)) = g(2n + 3) = 3(2n + 3) + 2 = 6n + 11.$$

Suppose now that you compose two functions  $f, g$ , and both of them turn out to be injective. The next result tells us that the combination will be as well!

**Proposition 2.** *Let  $f : X \rightarrow Y$  and  $g : Y \rightarrow Z$  be two injective functions. Then  $g \circ f$  is also injective.*

*Proof.* What we need to do is check the injectivity of a function, so we do this as usual: we check that  $g \circ f(x_1) = g \circ f(x_2)$  implies  $x_1 = x_2$ . Typically, to be able to prove this, you will have to keep in mind assumptions, namely that both  $f$  and  $g$  are injective. So let us start. We have  $g \circ f(x_1) = g \circ f(x_2)$  or equivalently  $g(f(x_1)) = g(f(x_2))$ . But we know that  $g$  is injective, so this implies  $f(x_1) = f(x_2)$ . Next we use that  $f$  is injective, thus  $x_1 = x_2$ , as needed!  $\square$

Let us ask the same question with surjectivity, namely whether the composition of two surjective functions gives a function which is surjective too. Here is the answer:

**Proposition 3.** *Let  $f : X \rightarrow Y$  and  $g : Y \rightarrow Z$  be two surjective functions. Then  $g \circ f$  is also surjective.*

*Proof.* The codomain of  $g \circ f$  is  $Z$ , therefore we need to show that every  $z \in Z$  has a preimage  $x$ , namely that there always exists an  $x$  such that  $g \circ f(x) = z$ . Again, we keep in mind that  $f$  and  $g$  are both surjective. Since  $g$  is surjective, we know there exists  $y \in Y$  such that  $g(y) = z$ . Now again, since  $f$  is surjective, we know there exists  $x \in X$  such that  $f(x) = y$ . Therefore there exist  $x, y$  such that  $z = g(y) = g(f(x))$  as needed.  $\square$

## One-to-one Propagation

**Theorem 2:** Let  $f:X \rightarrow Y$  and  $g:Y \rightarrow Z$  be both one-to-one functions.  
Then  $g \circ f$  is also one-to-one.

**Proof: to show:**  $\forall x_1, x_2 \in X ((g \circ f)(x_1) = (g \circ f)(x_2) \Rightarrow x_1 = x_2)$

Suppose  $x_1, x_2 \in X$  with  $(g \circ f)(x_1) = (g \circ f)(x_2)$ .

Then  $g(f(x_1)) = g(f(x_2))$ .

Since  $g$  is one-to-one, it follows  $f(x_1) = f(x_2)$ .

Since  $f$  is one-to-one, it follows  $x_1 = x_2$ .

## Onto Propagation

**Theorem 3:** Let  $f:X \rightarrow Y$  and  $g:Y \rightarrow Z$  be both onto functions.  
Then  $g \circ f$  is also onto.

**Proof: to show:**  $\forall z \in Z \exists x \in X$  such that  $(g \circ f)(x) = z$

Let  $z \in Z$ .

Since  $g$  is onto  $\exists y \in Y$  with  $g(y) = z$ .

Since  $f$  is onto  $\exists x \in X$  with  $f(x) = y$ .

Hence, with  $(g \circ f)(x) = g(f(x)) = g(y) = z$ .

We conclude this chapter on functions, by discussing the pigeonhole principle.

**Definition 71.** The [pigeonhole principle](#) states the following: if you have  $k$  pigeonholes, and  $n$  pigeons, but the number  $n$  of pigeons is more than the number  $k$  of pigeonholes, then at least one pigeonhole contains at least two pigeons.

Here is a simple illustration: if you have 4 pigeons and 3 pigeonholes:

1. Put the first pigeon in the first pigeonhole, if the second pigeon is also here, then we are done, we have at least one pigeonhole with at least 2 pigeons.
2. If the second pigeon went into the second pigeonhole, repeat the argument: if the third pigeon is also here, then we are done, we have at least one pigeonhole with at least 2 pigeons.
3. If the third pigeon went into the third pigeonhole, then at this time, you have 3 pigeonholes, each containing one pigeon, therefore no matter where the fourth pigeon will go, we have at least one pigeonhole with at least 2 pigeons!

This principle is attributed to the mathematician Dirichlet, and is actually very powerful. It is a consequence of the fact that a function from a finite set (the number of pigeons) to a smaller finite set (the number of pigeonholes) cannot be one-to-one, meaning that there must be at least two elements (two pigeons) in the domain, that have the same image (the same pigeonhole) in the co-domain!

## Pigeonhole Principle



$k$  pigeonholes,  $n$  pigeons,  $n > k$   
at least one pigeonhole  
contains at least two pigeons



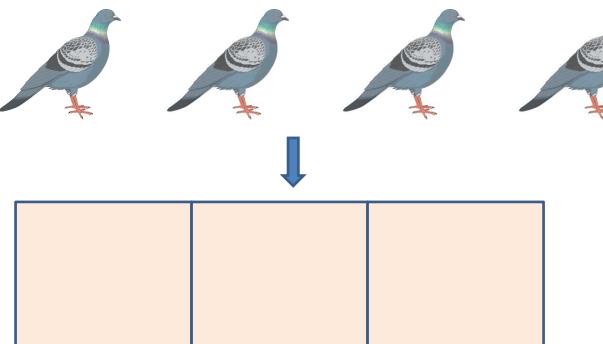
Peter Gustav  
Lejeune Dirichlet  
(1805-1859)



image belongs to the artist, Dirichlet portrait comes from wiki

## Pigeonhole Principle

A function from one finite set to a smaller finite set cannot be one-to-one: there must be at least two elements in the domain that have the same image in the co-domain.



## Examples

Consider Thorin and his 12 dwarf companions.

- At least two of the dwarves were born on the same day of the week.
- They go to sleep at the Prancing Pony Inn. Thorin gets a room of his own, but the others got to share 4 rooms. Then there are at least 3 dwarves sleeping in at least one of them.



This image belongs to the Hobbit movie

**Example 105.** Consider Thorin and his 12 dwarf companions.

- At least two of the dwarves were born on the same day of the week: this is consequence of the pigeonhole principle. You have 7 days of the week, and more than 7 dwarves, therefore 7 of them at most could be born each on one day of the week, but the 8th one will necessarily have to share the same day of the week as birthday.
- They sleep at the Prancing Pony Inn, Thorin gets a room of his own (of course, he is the chief!) but the 12 others got to share 4 rooms. Then at least 3 dwarves sleep in at least one room. This is again a consequence of the pigeonhole principle. Imagine room 1, room 2, room 3 and room 4, and 12 dwarves have to fit. The first 4 dwarves could choose room 1, 2, 3, and 4, and be alone in each room. But then the next 4 dwarves will add up, and we will have 2 dwarves in each room. Then no matter how, at least 3 dwarves will end up in one room!

## Exercises for Chapter 10

**Exercise 85.** Consider the set  $A = \{a, b, c\}$  with power set  $P(A)$  and  $\cap : P(A) \times P(A) \rightarrow P(A)$ . What is its domain? its co-domain? its range? What is the cardinality of the pre-image of  $\{a\}$ ?

**Exercise 86.** Show that  $\sin : \mathbb{R} \rightarrow \mathbb{R}$  is not one-to-one.

**Exercise 87.** Show that  $\sin : \mathbb{R} \rightarrow \mathbb{R}$  is not onto, but  $\sin : \mathbb{R} \rightarrow [-1, 1]$  is.

**Exercise 88.** Is  $h : \mathbb{Z} \rightarrow \mathbb{Z}$ ,  $h(n) = 4n - 1$ , onto (surjective)?

**Exercise 89.** Is  $f : \mathbb{R} \rightarrow \mathbb{R}$ ,  $f(x) = x^3$ , a bijection (one-to-one correspondence)?

**Exercise 90.** Consider  $f : \mathbb{R} \rightarrow \mathbb{R}$ ,  $f(x) = x^2$  and  $g : \mathbb{R} \rightarrow \mathbb{R}$ ,  $g(x) = x + 5$ . What is  $g \circ f$ ? What is  $f \circ g$ ?

**Exercise 91.** Consider  $f : \mathbb{Z} \rightarrow \mathbb{Z}$ ,  $f(n) = n + 1$  and  $g : \mathbb{Z} \rightarrow \mathbb{Z}$ ,  $g(n) = n^2$ . What is  $g \circ f$ ? What is  $f \circ g$ ?

**Exercise 92.** Given two functions  $f : X \rightarrow Y$ ,  $g : Y \rightarrow Z$ . If  $g \circ f : X \rightarrow Z$  is one-to-one, must both  $f$  and  $g$  be one-to-one? Prove or give a counter-example.

**Exercise 93.** Show that if  $f : X \rightarrow Y$  is invertible with inverse function  $f^{-1} : Y \rightarrow X$ , then  $f^{-1} \circ f = i_X$  and  $f \circ f^{-1} = i_Y$ .

**Exercise 94.** If you pick five cards from a deck of 52 cards, prove that at least two will be of the same suit.

**Exercise 95.** If you have 10 black socks and 10 white socks, and you are picking socks randomly, you will only need to pick three to find a matching pair.

# Chapter 11

## Graph Theory

*“The origins of graph theory are humble, even frivolous.” (N. Biggs, E. K. Lloyd, and R. J. Wilson)*

Let us start with a formal definition of what is a graph.

**Definition 72.** A [graph](#)  $G = (V, E)$  is a structure consisting of a set  $V$  of vertices (also called nodes), and a set  $E$  of edges, which are lines joining vertices.

One way to denote an edge  $e$  is to explicit the 2 vertices that are connected by this edge, say if the edge  $e$  links the vertex  $u$  to the vertex  $v$ , we write  $e = \{u, v\}$ .

**Definition 73.** Two vertices  $u, v$  in a graph  $G$  are [adjacent](#) in  $G$  if  $\{u, v\}$  is an edge of  $G$ . If  $e = \{u, v\}$  is an edge of  $G$ , then  $e$  is called [incident](#) with the vertices  $u$  and  $v$ .

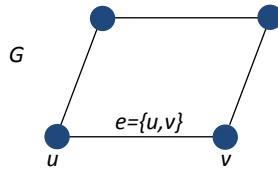
Graphs can be of several types.

**Definition 74.** A [simple](#) graph  $G$  is a graph that has no [loop](#), that is no edge  $\{u, v\}$  with  $u = v$  and no parallel edges between any pair of vertices.

**Example 106.** Consider the table of fictitious flights among different cities. Suppose all you want to know is whether there is a direct flight between any two cities, and you are not interested in the direction of the flight. Then you can draw a graph whose vertices are the cities, and there is an edge between city  $A$  and city  $B$  exactly when there is at least one flight going either from city  $A$  to the city  $B$ , or from city  $B$  to city  $A$ .

## Definitions

A **graph**  $G = (V, E)$  is a structure consisting of a set  $V$  of vertices (nodes) and a set  $E$  of edges (lines joining vertices).

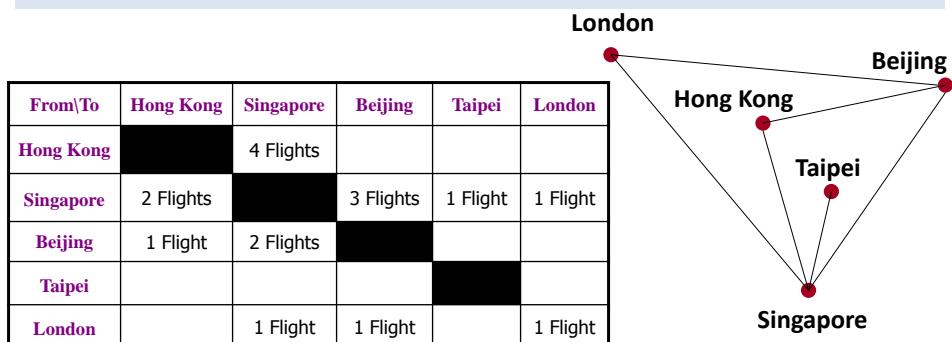


- Two vertices  $u$  and  $v$  are **adjacent** in  $G$  if  $\{u, v\}$  is an edge of  $G$ .
- If  $e = \{u, v\}$ , the edge  $e$  is called **incident** with the vertices  $u$  and  $v$ .

Graphs are useful to represent data.

## Simple Graphs

A **simple** graph is a graph that has no **loop** (=edge  $\{u, v\}$  with  $u=v$ ) and no parallel edges between any pair of vertices.



Draw a graph to see whether there are direct flights between any two cities (in either direction)

Take for example Hong Kong: there will be one edge between Hong Kong and Singapore (this is read in the first row), and one edge from Hong Kong to Beijing (this is read in the first column). This graph is simple, there is no loop, and no parallel edge.

**Definition 75.** A [multigraph](#)  $G$  is a graph that has no loop and at least two parallel edges between some pair of vertices.

**Example 107.** We continue with our fictitious example of flights. Suppose now we want to know how many flights are there that operate between two cities (we are still not interested in the direction). In this case, we will need parallel edges to represent multiple flights. For example, let us consider Hong Kong and Singapore: there are 4 flights from Hong Kong to Singapore, and 2 flights from Singapore to Hong Kong, thus a total of 6 edges between the two vertices representing these cities.

Very often, a relation from a vertex  $A$  to  $B$  does not yield one from  $B$  to  $A$ , in this case, edges become arrows.

**Definition 76.** A [directed graph](#)  $G$ , also called digraph for short, is a graph where edges  $\{u, v\}$  are ordered ( $\{u, v\}$  and  $\{v, u\}$  are not the same), that is edges have a direction. The graph is called [undirected](#) otherwise. Parallel edges are allowed in [directed multigraph](#). Loops are allowed for both directed and directed multigraphs.

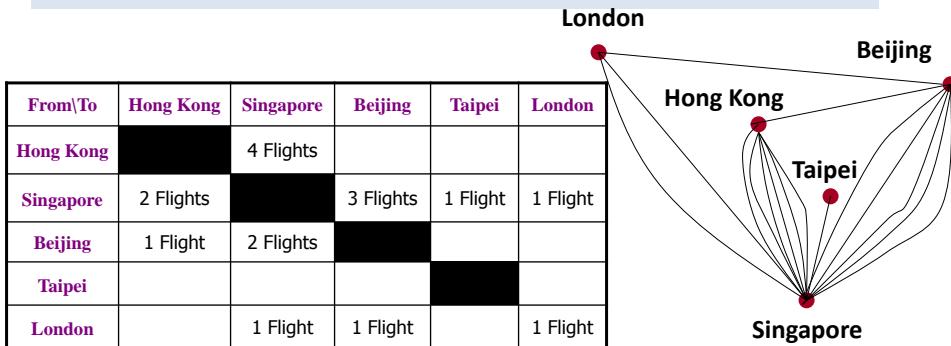
**Example 108.** On our example of fictitious flights, a directed graph corresponds to put arrows for flights going in a given direction, for example, there is one arrow from London to Beijing, but none from Beijing to London.

People attribute the origin of graph theory to the Königsberg bridge question, solved by the mathematician Euler in 1736. The question was, given the map of Königsberg, which contains 7 bridges, is it possible to find a walk that goes through the 7 bridges without crossing a bridge twice? You may look at the map and give it a try yourself, to convince yourself of the answer. The answer turns out to be no, it is not possible. We will see why next. But first, we will give a name to such walks, in honour of Euler:

**Definition 77.** A [Euler path/trail](#) is a walk on the edges of a graph which uses each edge in the graph exactly once. A [Euler circuit/cycle](#) is a walk on the edges of a graph which starts and ends at the same vertex, and uses each edge in the graph exactly once.

## Multigraphs

A **multigraph** is a graph that has no loop and at least 2 parallel edges between some pair of vertices.



Draw a graph with an edge for each flight that operates between two cities (in either direction).

4/15

## Directed (Multi)graphs

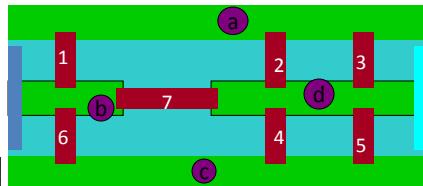
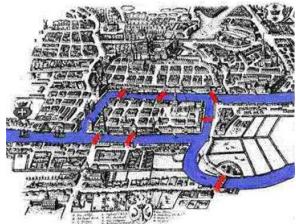
A **directed** graph is a graph where edges  $\{u,v\}$  are ordered, that is, edges have a direction. Parallel edges are allowed in **directed multigraphs**. Loops are allowed for both.

From\To	Hong Kong	Singapore	Beijing	Taipei	London
Hong Kong		4 Flights			
Singapore	2 Flights		3 Flights	1 Flight	1 Flight
Beijing	1 Flight	2 Flights			
Taipei					
London		1 Flight	1 Flight		1 Flight

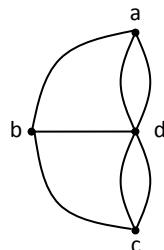
Draw a graph to see whether there are direct flights between any two cities (direction matters)

## Origin: The Bridges of Königsberg

- Königsberg (now Kaliningrad, Russia) has 7 bridges.



- People tried (without success) to find a way to walk all 7 bridges without crossing a bridge twice.



*Leonhard Euler* introduced Graphs in 1736 to solve the Königsberg Bridge problem (no solution and why).

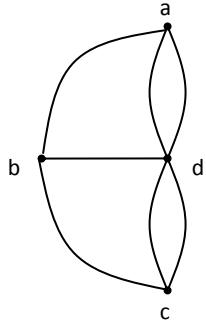
## Euler Path and Circuit

A **Euler path** (Eulerian trail) is a walk on the edges of a graph which uses each edge in the original graph exactly once.

The beginning and end of the walk may or not be the same vertex.

A **Euler circuit** (Eulerian cycle) is a walk on the edges of a graph which starts and ends at the same vertex, and uses each edge in the original graph exactly once.

## Euler Circuit



- Suppose the beginning and end are the same node  $u$ .
- The graph must be **connected**.
- At every vertex  $v \neq u$ , we reach  $v$  along one edge and go out along another, thus the number of edges incident at  $v$  (called the **degree** of  $v$ ) is even.
- The node  $u$  is visited once the first time we leave, and once the last time we arrive, and possibly in between (back and forth), thus the degree of  $u$  is even.
- Since the Königsberg Bridges graph has odd degrees, no solution!

8/15

## Euler Theorem

The **degree** of a vertex is the number of edges incident with it.

**Theorem.** Consider a connected graph  $G$ .

1. If  $G$  contains an Euler path that starts and ends at the same node, then all nodes of  $G$  have an even degree.
  2. If  $G$  contains an Euler path, then exactly two nodes of  $G$  have an odd degree.
- 
- Suppose  $G$  as an Euler path, which starts at  $v$  and finishes at  $w$ .
  - Add the edge  $\{v,w\}$ .
  - Then by the first part of the theorem, all nodes have even degree, but for  $v$  and  $w$  which have odd degrees.

The Euler circuit/cycle is simply an Euler path/trail whose start and end are the same vertex.

**Definition 78.** The [degree of a node](#) is the number of edges incident with it.

With this definition of degree, we next answer the question of the bridges of Knigsberg. We start with the more constrained case of starting and finishing at the same vertex. We assume the graph  $G$  is [connected](#), which means that there is always a way to walk from any vertex to any other (possibly using several times the same vertex or the same edge).

**Theorem 4.** *Consider a connected graph  $G$ . Then  $G$  contains an Euler circuit/cycle, if and only if all nodes of  $G$  have an even degree.*

*Proof.* We prove only that if  $G$  contains an Euler circuit, then all nodes have an even degree. We start the walk at vertex  $u$ . Now for any vertex  $v$  which is not  $u$ , we need to walk in  $v$  using some edge, and walk out of  $u$  using another edge. We may come back to  $v$ , but for every come back, we still need one edge to come in, and one to walk out. Therefore the degree of  $v$  must be even! As for the starting point  $u$ , it is visited once the first time we leave, and the last time we arrive (2 edges), and any possible back and forth counts for 2 edges as well, which shows that indeed, it must be that all nodes have an even degree!  $\square$

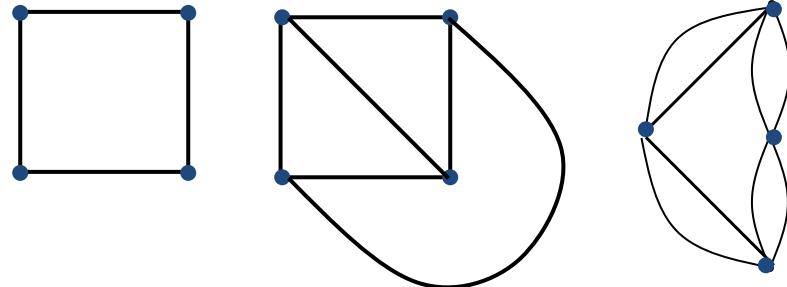
Since the Knigsberg bridge graph has odd degrees, it has no Euler cycle. We next extend the argument to an Euler path.

**Theorem 5.** *Consider a connected graph  $G$ . Then  $G$  contains an Euler path if and only if exactly two nodes of  $G$  have an odd degree.*

*Proof.* We only prove that if  $G$  has an Euler path, then exactly two nodes of  $G$  have an odd degree. Suppose thus that  $G$  has an Euler path, which starts at  $v$  and finishes at  $w$ . Create a new graph  $G'$ , which is formed from  $G$  by adding one edge between  $v$  and  $w$ . Now  $G'$  has an Euler cycle, and so we know by the previous theorem that  $G'$  has the property that all its vertices have an even degree. Therefore the degrees of  $v$  and  $w$  in  $G$  are odd, while all the others are even and we are done. The other direction is left as an exercise (see Exercise 96).  $\square$

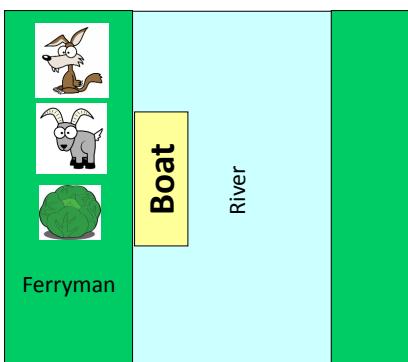
## Examples

Note: Euler Theorem actually states an if and only if.



## Example: Wolf, Goat, Cabbage

*A classical puzzle that involves graphs:*



From the left bank of the river, the ferryman is to transport the wolf, the goat and the cabbage to the right bank.

The boat is only big enough to transport one object/animal at a time, other than himself.

The wolf cannot be left alone with the goat, and the goat cannot be left alone with the cabbage.

**How should the ferryman proceed?**

Here is a classical puzzle.

**Example 109.** A ferryman needs to transport a wolf, a goat and a cabbage from one side of a river to the other, the boat is big enough for himself and one object/animal at a time. How should the ferryman proceed, knowing that the wolf cannot be left alone with the goat, and the goat cannot be left alone with the cabbage? One way to solve this is to use a graph that represents the different possible states of this system. The initial start up point is a state where  $wgcf$  ( $w$ =wolf,  $g$ =goat,  $c$ =cabbage and  $f$ =ferryman) are on the left side of the river, with nothing on the right side. The first step, the ferryman has no choice, he takes the goat on the other side, which leads to a second step, with  $wc$  on the left bank, and  $gf$  on the right bank. The ferryman returns, he then can choose: he either takes the cabbage or the wolf. This creates two branches in the graph. Each branch leads to a couple of states, after which (see the graph itself) we reach a state where  $g$  is on the left, and  $wfc$  is on the right, leading to the end of the puzzle. A solution is a path in this graph that represents the different states of the system. In this example, it is a fairly easy graph, and there are 2 paths, each of the same length!

Here are two more types of graphs which are important, in that you are very likely to encounter them.

**Definition 79.** A [complete graph](#) with  $n$  vertices is a simple graph that has every vertex connected to every other distinct vertex.

**Definition 80.** A [bipartite graph](#) is a graph whose vertices can be partitioned into 2 (disjoint) subset  $V$  and  $W$  such that each edge only connects a  $v \in V$  and a  $w \in W$ .

We have seen the notion of degree of a vertex  $v$  in an undirected graph, it is the number of edges incident with. We note that in this case, a loop at a vertex contributes twice. For directed graphs, you may like to know that the notion of degree is more precise, one distinguishes in-degree and out-degree, which we will not discussed here.

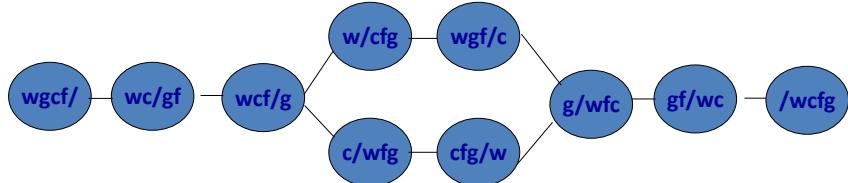
**Definition 81.** The [total degree](#) of an undirected graph  $G = (V, E)$  is the sum of the degrees of all the vertices of  $G$ :  $\sum_{v \in V} \deg(v)$ .

## Example: Wolf, Goat, Cabbage (IV)

Either he takes

- f = ferryman
- g = goat
- w = wolf
- c = cabbage

1. The cabbage, bring back the goat, leave the goat and take the wolf across, return, and take the goat across.
2. Or the wolf, bring back the goat, leave the goat and the cabbage across, return, and take the goat across.



## Complete & Bipartite

A **complete graph with  $n$  vertices** is a simple graph that has every vertex connected to every other distinct vertex.



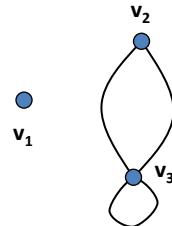
A **bipartite graph** is a graph whose vertices can be partitioned into 2 (disjoint) subsets  $V$  and  $W$  s.t. each edge only connects a  $v \in V$  and a  $w \in W$ .



## More on Node Degree

The **degree  $\deg(v)$**  of a vertex  $v$  in an undirected graph is the number of edges incident with it ( a loop at a vertex contributes twice) .  
In-degree and Out-degree are distinguished for directed graphs.

$$\text{total degree} = \deg(v_1) + \deg(v_2) + \deg(v_3) = 0 + 2 + 4 = 6.$$



The **total degree  $\deg(G)$**  of an undirected graph  $G$  is the sum of the degrees of all the vertices of  $G$ :  $\sum_{v \in V} \deg(v)$

## The Handshaking Theorem

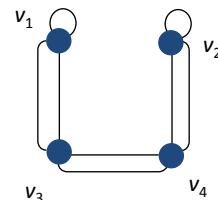
Let  $G = (V, E)$  be an undirected graph with  $e$  edges. Then

$$2e = \sum_{v \in V} \deg(v)$$

(Note that this even applies if multiple edges and loops are present.)

### Proof.

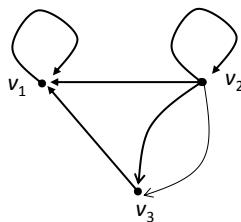
Choose an  $e \in E(G)$  with endpoints  $v, w \in V$ .  
 $e$  contributes 1 to  $\deg(v)$  and 1 to  $\deg(w)$ . True even when  $v = w$ . Thus each edge contributes 2 to the total degree.



$$\deg(v_1) = \deg(v_2) = \deg(v_3) = \deg(v_4) = 4$$

$$2e = \sum \deg(v) = 4 \times 4 = 16 \text{ and } e = 8$$

## Adjacency Matrix

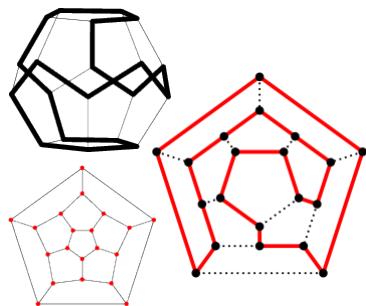


$$A = \begin{matrix} & v_1 & v_2 & v_3 \\ v_1 & 1 & 0 & 0 \\ v_2 & 1 & 1 & 2 \\ v_3 & 1 & 0 & 0 \end{matrix}$$

A graph can be represented by a matrix  $A = (a_{ij})$  called **adjacency matrix**, with  
 $a_{ij}$  = the number of arrows from  $v_i$  to  $v_j$

What is the adjacency matrix of a complete graph?

## Hamiltonian Circuit



- *The Icosian game* (1857): Along the edges of a dodecahedron, find a path such that every vertex is visited a single time, and the ending point is the same as the starting point.
- Hamilton sold it to a London game dealer in 1859 for 25 pounds.

A **Hamiltonian path** of a graph  $G$  is a walk such that every vertex is visited exactly once.

A **Hamiltonian circuit** of a graph  $G$  is a closed walk such that every vertex is visited exactly once (except the same start/end vertex).

The [Handshaking Theorem](#) links the number of edges in a graph to the numbers of vertices.

**Theorem 6.** *Let  $G = (V, E)$  be an undirected graph with  $|E|$  edges. Then*

$$2|E| = \sum_{v \in V} \deg(v).$$

*Proof.* The proof follows the name of the theorem. The idea of handshaking is that if two people shake hands, there must be...well...two persons involved. In a graph  $G$ , this becomes, if there is an edge  $e$  between  $v$  and  $w$ , then  $e$  contributes to 1 to the degree of  $v$  and to 1 to the degree of  $w$ . This is also true when  $v = w$ . Therefore each edge contributes 2 to the total degree.  $\square$

To represent a graph, a useful way to do so is to use a matrix.

**Definition 82.** The [adjacency matrix](#) of a graph  $G$  is a matrix  $A$  whose coefficients are denoted  $a_{ij}$ , where  $a_{ij}$ , the coefficient in the  $i$ th row and  $j$ th column, counts the number of arrows from  $v_i$  to  $v_j$ .

You may replace the term arrow by edge in this definition if your graph is undirected.

**Example 110.** The adjacency of a complete graph with 4 vertices is

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

The first row reads that  $v_1$  is connected to  $v_2, v_3, v_4$ , but not to itself (since it is a simple graph by definition).

We saw earlier Euler paths as walks going through exactly every edge of a graph. If you replace "exactly every edge" by "exactly every vertex", this becomes a Hamiltonian path!

**Definition 83.** A [Hamiltonian path](#) of a graph  $G$  is a walk such that every vertex is visited exactly one. A [Hamiltonian circuit](#) of a graph  $G$  is a closed walk such that every vertex is visited exactly one, except the same start/end vertex.

Hamiltonian paths are harder to characterize than Euler paths. In particular, finding an algorithm that will identify Hamiltonian paths in a graph is hard!

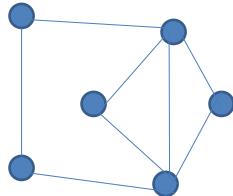


## Hamiltonian vs Eulerian

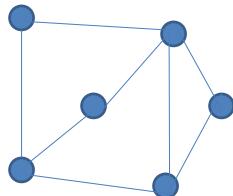


- Path (or trail) vs Circuit (or cycle): for circuits, the walk starts and finishes at the same vertex, not needed for path.
- Eulerian: walk through every edge exactly once.
- Hamiltonian: walk through every vertex exactly.

## Examples

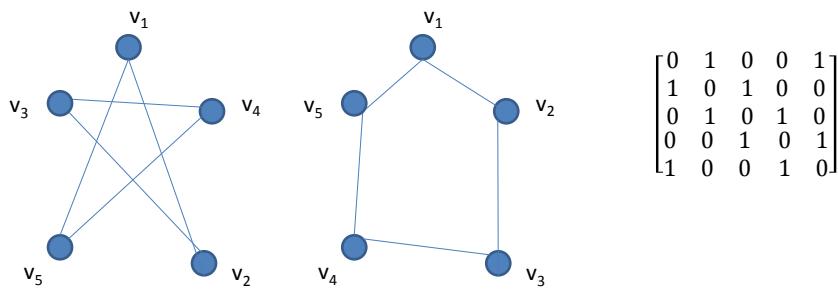


Euler circuit  
Hamiltonian path  
No Hamiltonian circuit



No Euler circuit, but Euler path  
Hamiltonian path  
No Hamiltonian circuit

## Graph Isomorphism



Finally, it is useful to pay attention that one draws a graph, it is just a visualization...and several visualizations may be different, giving the impression that the graphs are different, while they are actually the same. If two graphs differ by their labeling, but their adjacency structure is the same, we say that these graphs are isomorphic. More formally:

**Definition 84.** A [graph isomorphism](#) between two graphs  $G$  and  $H$  is a bijection  $f$  between the set of vertices of  $G$  and the set of vertices of  $H$  such that any two vertices  $u, v$  in  $G$  are adjacent if and only if  $f(u), f(v)$  are adjacent in  $H$ .

## Exercises for Chapter 11

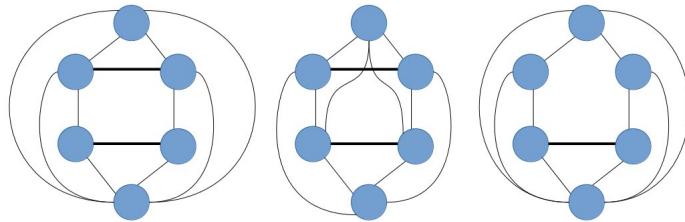
**Exercise 96.** Prove that if a connected graph  $G$  has exactly two vertices which have odd degree, then it contains an Euler path.

**Exercise 97.** Draw a complete graph with 5 vertices.

**Exercise 98.** Show that in every graph  $G$ , the number of vertices of odd degree is even.

**Exercise 99.** Show that in every simple graph (with at least two vertices), there must be two vertices that have the same degree.

**Exercise 100.** Decide whether the following graphs contain a Euler path/cycle.



# Chapter 12

## Solutions to the Exercises

*“Intuition comes from experience, experience from failure, and failure from trying.”*

### Exercises for Chapter 1

**Exercise 1.** Show that 2 is the only prime number which is even.

*Solution.* Take  $p$  a prime number. Then  $p$  has only 2 divisors, 1 and  $p$ . If  $p$  is also even, then one of its divisors has to be 2, thus  $p = 2$ .

**Exercise 2.** Show that if  $n^2$  is even, then  $n$  is even, for  $n$  an integer.

*Solution.* An integer  $n$  is either even, that is  $n = 2n'$ , for some integer  $n'$ , or odd, that is  $n = 2n' + 1$  for some integer  $n'$ . Thus  $n^2$  is either  $4(n')^2$  or  $4(n')^2 + 4n' + 1$ . The case where  $n^2$  is even is thus when  $n = 2n'$ .

**Exercise 3.** The goal of this exercise is to show that  $\sqrt{2}$  is irrational. We provide a step by step way of doing so.

1. Suppose by contradiction that  $\sqrt{2}$  is rational, that is  $\sqrt{2} = \frac{m}{n}$ , for  $m$  and  $n$  integers with no common factor. Show that  $m$  has to be even, that is  $m = 2k$ .
2. Compute  $m^2$ , and deduce that  $n$  has to be even too, a contradiction.

*Solution.* 1. Suppose by contradiction that  $\sqrt{2}$  is rational, that is  $\sqrt{2} = \frac{m}{n}$ , for  $m$  and  $n$  integers with no common factor. Then

$$2 = \frac{m^2}{n^2}$$

and thus  $m^2 = 2n^2$ , showing that  $m^2$  is even, that is, using Exercise 2,  $m$  has to be even, say  $m = 2k$  for  $k$  some integer.

2. Now  $m^2 = (2k)^2 = 4k^2$ . This tells us, combining with the first step of the exercise, that

$$m^2 = 4k^2 = 2n^2$$

which implies that  $2k^2 = n^2$ , that is  $n^2$  is even and by again by Exercise 2, it must be that  $n$  is even. This is a contradiction, since we assumed that  $m$  and  $n$  have no common factor.

**Exercise 4.** This exercise is optional, it requires to write things quite formally. Show the following two properties of integers modulo  $n$ :

1.  $(a \bmod n) + (b \bmod n) \equiv (a + b) \bmod n$ .
2.  $(a \bmod n)(b \bmod n) \equiv (a \cdot b) \bmod n$ .

*Solution.* 1. Suppose  $(a \bmod n) = a'$ , that is  $a = qn + a'$ , and  $(b \bmod n) = b'$ , that is  $b = rn + b'$ , for some integers  $q, r$ . Then

$$(a \bmod n) + (b \bmod n) \equiv a' + b' \bmod n$$

and

$$(a + b) \bmod n \equiv (qn + a' + rn + b') \bmod n \equiv (a' + b') \bmod n.$$

2. Similarly

$$(a \bmod n)(b \bmod n) \equiv a'b' \bmod n$$

and

$$(ab) \bmod n \equiv (qn + a')(rn + b') \equiv qrn^2 + qnb' + a'rn + a'b' \bmod n \equiv (a'b') \bmod n.$$

**Exercise 5.** Compute the addition table and the multiplication tables for integers modulo 4.

*Solution.* We represent integers modulo 4 by the set of integers  $\{0, 1, 2, 3\}$ . Then

+	0	1	2	3
0	0	1	2	3
1	1	2	3	0
2	2	3	0	1
3	3	0	1	2

Similarly

*	0	1	2	3
0	0	0	0	0
1	0	1	2	3
2	0	2	0	2
3	0	3	2	1

Note that these tables are great to observe the closure property! Elements computed in these tables are the same as those given as input.

**Exercise 6.** Show that  $\frac{p(p+1)}{2} \equiv 0 \pmod{p}$  for  $p$  an odd prime.

*Solution.* Suppose that  $p$  is an odd prime. Then  $p+1$  is even, thus divisible by 2, say  $p+1 = 2k$  for some  $k$ . Now

$$\frac{p(p+1)}{2} = pk \equiv 0 \pmod{p}.$$

Note that the critical part is that  $p$  is odd. If  $p = 2$ , this does not work, indeed  $2 \cdot 3/2 = 3$  which is not  $0 \pmod{2}$ .

**Exercise 7.** Consider the following sets  $S$ , with respective operator  $\Delta$ .

- Let  $S$  be the set of rational numbers, and  $\Delta$  be the multiplication. Is  $S$  closed under  $\Delta$ ? Justify your answer.
- Let  $S$  be the set of natural numbers, and  $\Delta$  be the subtraction. Is  $S$  closed under  $\Delta$ ? Justify your answer.
- Let  $S$  be the set of irrational numbers, and  $\Delta$  be the addition. Is  $S$  closed under  $\Delta$ ? Justify your answer.

*Solution.* • Take two rational numbers  $\frac{m}{n}$  and  $\frac{m'}{n'}$ . Then

$$\frac{\frac{m}{n} \frac{m'}{n'}}{n n'} = \frac{mm'}{nn'}$$

which is a rational number. Thus the answer is yes,  $S$  is closed under multiplication.

- The subtraction of two natural numbers does not always give a number natural, for example,

$$5 - 10 = -5.$$

Thus  $S$  is not closed under subtraction.

- The addition of two irrational numbers does not always give an irrational number, for example,

$$(2 + \sqrt{2}) + (2 - \sqrt{2}) = 4$$

and 4 is not an irrational number. Thus  $S$  is not closed under addition. Note that we are using here the claim that  $2 + \sqrt{2}$  is irrational. Indeed, suppose that  $2 + \sqrt{2}$  were rational, that is  $2 + \sqrt{2} = \frac{m}{n}$  for  $m, n$  some integers. Then

$$\sqrt{2} = \frac{m}{n} - 2 = \frac{m - 2n}{n}$$

which is a contradiction to the fact that  $\sqrt{2}$  is irrational.

## Exercises for Chapter 2

**Exercise 8.** Decide whether the following statements are propositions. Justify your answer.

1.  $2 + 2 = 5$ .
2.  $2 + 2 = 4$ .
3.  $x = 3$ .
4. Every week has a Sunday.
5. Have you read “Catch 22”?

*Solution.* 1.  $2 + 2 = 5$ : this is a proposition, because it is a statement that always takes the truth value "false".

2.  $2 + 2 = 4$ : this is a proposition, because it is a statement that always takes the truth value "true".
3.  $x = 3$ : the statement depends on the value of  $x$ . Maybe it is true (if  $x$  was assigned the value 3), or maybe it is false (if  $x$  was assigned a different value). Thus this is not a proposition.
4. Every week has a Sunday: this is a proposition, because it is a statement that always takes the truth value "true".
5. Have you read "Catch 22"? this is a question, thus it is not a proposition.

**Exercise 9.** Show that

$$\neg(p \vee q) \equiv \neg p \wedge \neg q.$$

This is the second law of De Morgan.

*Solution.* We show the equivalence using truth tables:

$p$	$q$	$p \vee q$	$\neg(p \vee q)$	$p$	$q$	$\neg p$	$\neg q$	$\neg p \wedge \neg q$
T	T	T	F	T	T	F	F	F
T	F	T	F	T	F	F	T	F
F	T	T	F	F	T	T	F	F
F	F	F	T	F	F	T	T	T

Since both truth tables are the same, the two logical expressions are equivalent.

**Exercise 10.** Show that the second absorption law  $p \wedge (p \vee q) \equiv p$  holds.

*Solution.* We show the equivalence using a truth table:

$p$	$q$	$p \vee q$	$p \wedge (p \vee q)$
T	T	T	T
T	F	T	T
F	T	T	F
F	F	F	F

Since the column of  $p$  is the same as that of  $p \wedge (p \vee q)$ , both logical expressions are equivalent.

**Exercise 11.** These two laws are called distributivity laws. Show that they hold:

1. Show that  $(p \wedge q) \vee r \equiv (p \vee r) \wedge (q \vee r)$ .
2. Show that  $(p \vee q) \wedge r \equiv (p \wedge r) \vee (q \wedge r)$ .

*Solution.* We use truth tables:

$p$	$q$	$r$	$p \wedge q$	$(p \wedge q) \vee r$	$p \vee r$	$q \vee r$	$(p \vee r) \wedge (q \vee r)$
T	T	T	T	T	T	T	T
T	T	F	T	T	T	T	T
T	F	T	F	T	T	T	T
T	F	F	F	F	T	F	F
F	T	T	F	T	T	T	T
F	T	F	F	F	F	T	F
F	F	T	F	T	T	T	T
F	F	F	F	F	F	F	F

$p$	$q$	$r$	$p \vee q$	$(p \vee q) \wedge r$	$p \wedge r$	$q \wedge r$	$(p \wedge r) \vee (q \wedge r)$
T	T	T	T	T	T	T	T
T	T	F	T	F	F	F	F
T	F	T	T	T	F	F	T
T	F	F	T	F	F	F	F
F	T	T	T	T	F	T	T
F	T	F	T	F	F	F	F
F	F	T	F	F	F	F	F
F	F	F	F	F	F	F	F

**Exercise 12.** Verify  $\neg(p \vee \neg q) \vee (\neg p \wedge \neg q) \equiv \neg p$  by

- constructing a truth table,
- developing a series of logical equivalences.

*Solution.* We start with a truth table:

$p$	$q$	$\neg p$	$\neg q$	$p \vee \neg q$	$\neg(p \vee \neg q)$	$\neg p \wedge \neg q$	$\neg(p \vee \neg q) \vee (\neg p \wedge \neg q)$
$T$	$T$	$F$	$F$	$T$	$F$	$F$	$F$
$T$	$F$	$F$	$T$	$T$	$F$	$F$	$F$
$F$	$T$	$T$	$F$	$F$	$T$	$F$	$T$
$F$	$F$	$T$	$T$	$T$	$F$	$T$	$T$

Next we want to prove this result without using the truth table, but by developing logical equivalences:

$$\begin{aligned}
 \neg(p \vee \neg q) \vee (\neg p \wedge \neg q) &\equiv (\neg p \wedge q) \vee (\neg p \wedge \neg q) \text{ De Morgan} \\
 &\equiv \neg p \wedge (q \vee \neg q) \text{ Distributivity} \\
 &\equiv \neg p \wedge T \text{ since } (q \vee \neg q) \equiv T \\
 &\equiv \neg p.
 \end{aligned}$$

**Exercise 13.** Using a truth table, show that:

$$\neg q \rightarrow \neg p \equiv p \rightarrow q.$$

*Solution.* We compute the truth table:

$p$	$q$	$\neg p$	$\neg q$	$\neg q \rightarrow \neg p$	$p \rightarrow q$
$T$	$T$	$F$	$F$	$T$	$T$
$T$	$F$	$F$	$T$	$F$	$F$
$F$	$T$	$T$	$F$	$T$	$T$
$F$	$F$	$T$	$T$	$T$	$T$

**Exercise 14.** Show that  $p \vee q \rightarrow r \equiv (p \rightarrow r) \wedge (q \rightarrow r)$ .

*Solution.* We use logical equivalences:

$$\begin{aligned}
 p \vee q \rightarrow r &\equiv (p \vee q) \rightarrow r \text{ precedence} \\
 &\equiv \neg(p \vee q) \vee r \text{ conversion theorem} \\
 &\equiv (\neg p \wedge \neg q) \vee r \text{ De Morgan} \\
 &\equiv (\neg p \vee r) \wedge (\neg q \vee r) \text{ Distributivity} \\
 &\equiv (p \rightarrow r) \wedge (q \rightarrow r) \text{ conversion theorem}
 \end{aligned}$$

**Exercise 15.** Are  $(p \rightarrow q) \vee (q \rightarrow r)$  and  $p \rightarrow r$  equivalent statements?

*Solution.* They are not equivalent. Here is a proof using truth tables:

$p$	$q$	$r$	$p \rightarrow q$	$q \rightarrow r$	$(p \rightarrow q) \vee (q \rightarrow r)$	$p \rightarrow r$
$T$	$T$	$T$	$T$	$T$	$T$	$T$
$T$	$T$	$F$	$T$	$F$	$T$	$F$
$T$	$F$	$T$	$F$	$T$	$T$	$T$
$T$	$F$	$F$	$F$	$T$	$T$	$F$
$F$	$T$	$T$	$T$	$T$	$T$	$T$
$F$	$T$	$F$	$T$	$F$	$T$	$T$
$F$	$F$	$T$	$T$	$T$	$T$	$T$
$F$	$F$	$F$	$T$	$T$	$T$	$T$

We see that the first rows for example are giving different truth values. This can be done using equivalences as well:

$$\begin{aligned}
 (p \rightarrow q) \vee (q \rightarrow r) &\equiv (\neg p \vee q) \vee (\neg q \vee r) \text{ conversion theorem} \\
 &\equiv \neg p \vee r \vee T \text{ since } \neg q \vee q \equiv T \\
 &\equiv T.
 \end{aligned}$$

Since  $p \rightarrow q$  is not equivalent to  $T$ , both statements cannot be equivalent.

**Exercise 16.** Show that this argument is valid:

$$\boxed{\neg p \rightarrow F; \therefore p.}$$

*Solution.* The premise is  $\neg p \rightarrow F$ , which is true when  $\neg p$  is false, which is exactly when  $p$  is true.

**Exercise 17.** Show that this argument is valid, where  $C$  denotes a contradiction.

$$\boxed{\neg p \rightarrow C; \therefore p.}$$

*Solution.* The premise is  $\neg p \rightarrow C$ , which is true when  $\neg p$  is false, which is exactly when  $p$  is true.

**Exercise 18.** Determine whether the following argument is valid:

$$\begin{aligned}
 &\neg p \rightarrow r \wedge \neg s \\
 &t \rightarrow s \\
 &u \rightarrow \neg p \\
 &\neg w \\
 &u \vee w \\
 &\therefore t \rightarrow w.
 \end{aligned}$$

*Solution.* We start by noticing that we have

$$u \vee w; \neg w; \therefore u.$$

Indeed, if  $u \vee w$  and  $\neg w$  are both true, then  $w$  is false, and  $u$  must be true. Next

$$u \rightarrow \neg p; u; \therefore \neg p.$$

Indeed, if  $u \rightarrow \neg p$  is true, either  $u$  is true and  $\neg p$  is true, or  $u$  is false. But  $u$  is true, thus  $\neg p$  is true (Modus Ponens). Then

$$\neg p \rightarrow r \wedge \neg s; \neg p; \therefore r \wedge \neg s,$$

this is again Modus Ponens. Then

$$r \wedge \neg s; \therefore \neg s.$$

Indeed, for  $r \wedge \neg s$  to be true, it must be that  $\neg s$  is true. Finally

$$t \rightarrow s; \neg s; \therefore \neg t$$

since for  $t \rightarrow s$  to be true, we need either  $t$  to be false, or  $t$  and  $s$  to be true, but since  $s$  is false,  $t$  must be false (Modus Tollens), and

$$\neg t \therefore \neg t \vee w$$

or equivalently

$$\neg t \vee w \equiv t \rightarrow w$$

using the Conversion theorem, which shows that the argument is valid.

**Exercise 19.** Determine whether the following argument is valid:

$$\begin{aligned} & p \\ & p \vee q \\ & q \rightarrow (r \rightarrow s) \\ & t \rightarrow r \\ & \therefore \neg s \rightarrow \neg t. \end{aligned}$$

*Solution.* For this question, there is no obvious way to combine the known statements with inference rules. The only 2 related statements are  $p$  and  $p \vee q$ , and assuming that both are true, all can be deduced is that  $q$  is either true or false. Now if  $q$  is false,  $q \rightarrow (r \rightarrow s)$  is always true, while if  $q$  is true,  $q \rightarrow (r \rightarrow s)$  is true only if  $(r \rightarrow s)$  is true, which excludes the possibility  $r = T$  and  $s = F$ .

$q$	$r$	$s$	$t$
$F$	$T$	$T$	
$F$	$T$	$F$	
$F$	$F$	$T$	
$F$	$F$	$F$	
$F$	$T$	$T$	
$T$	$F$	$T$	
$T$	$F$	$F$	

Now we look at the last premise  $t \rightarrow r$ . For it to be true, we need  $t$  false, or  $t$  true and  $r$  true.

$q$	$r$	$s$	$t$
$F$	$T$	$T$	$T/F$
$F$	$T$	$F$	$T/F$
$F$	$F$	$T$	$F$
$F$	$F$	$F$	$F$
$T$	$T$	$T$	$T/F$
$T$	$F$	$T$	$F$
$T$	$F$	$F$	$F$

Now if  $s$  is true, then  $\neg s$  is always false, and the conclusion is always true. We thus focus on  $s$  is false, and  $\neg t$  is false, that is  $t$  is true. The second row gives a counter-example:

$$q = F, r = T, s = F, t = T.$$

## Exercises for Chapter 3

**Exercise 20.** Consider the predicates  $M(x, y) = “x$  has sent an email to  $y”$ , and  $T(x, y) = “x$  has called  $y”$ . The predicate variables  $x, y$  take values in the domain  $D = \{\text{students in the class}\}$ . Express these statements using symbolic logic.

1. There are at least two students in the class such that one student has sent the other an email, and the second student has called the first student.
2. There are some students in the class who have emailed everyone.

*Solution.* 1. We need two predicate variables since at least 2 students are involved, say  $x$  and  $y$ . There are at least two students in the class becomes

$$\exists x \in D, \exists y \in D.$$

Then  $x$  sent an email to  $y$ , that is  $M(x, y)$  and  $y$  has called  $x$ , that is  $T(y, x)$ , thus

$$M(x, y) \wedge T(y, x).$$

Furthermore, we need to take into account the fact that there are at least "two" students, so  $x$  and  $y$  have to be distinct! Thus the final answer is

$$\exists x \in D, \exists y \in D, ((x \neq y) \wedge M(x, y) \wedge T(y, x)).$$

2. There are students becomes

$$\exists x \in D,$$

then  $x$  has emailed everyone, that is

$$\exists x \in D, (\forall y \in D M(x, y)).$$

Note that the order of the quantifiers is important.

**Exercise 21.** Consider the predicate  $C(x, y) = "x \text{ is enrolled in the class } y"$ , where  $x$  takes values in the domain  $S = \{\text{students}\}$ , and  $y$  takes values in the domain  $D = \{\text{courses}\}$ . Express each statement by an English sentence.

1.  $\exists x \in S, C(x, \text{MH1812}).$
2.  $\exists y \in D, C(\text{Carol}, y).$
3.  $\exists x \in S, (C(x, \text{MH1812}) \wedge C(x, \text{CZ2002})).$
4.  $\exists x \in S, \exists x' \in S, \forall y \in D, ((x \neq x') \wedge (C(x, y) \leftrightarrow C(x', y))).$

*Solution.* 1. There exists a student such that this student is enrolled in the class MH1812, that is some student enrolled in the class MH1812.

2. There exists a course such that Carol is enrolled in this course, that is, Carol is enrolled in some course, or Carol is enrolled in at least one course.

3. There exists a student, such that this student is enrolled in MH1812 and in CZ2002, that is some student is enrolled in both MH1812 and CZ2002.
4. There exist two distinct students  $x$  and  $x'$ , such that for all courses,  $x$  is enrolled in the course if and only if  $x'$  is enrolled in the course. In other words, there exist two students which are enrolled in exactly the same courses.

**Exercise 22.** Consider the predicate  $P(x, y, z) = "xyz = 1"$ , for  $x, y, z \in \mathbb{R}$ ,  $x, y, z > 0$ . What are the truth values of these statements? Justify your answer.

1.  $\forall x, \forall y, \forall z, P(x, y, z)$ .
2.  $\exists x, \exists y, \exists z, P(x, y, z)$ .
3.  $\forall x, \forall y, \exists z, P(x, y, z)$ .
4.  $\exists x, \forall y, \forall z, P(x, y, z)$ .

*Solution.* 1.  $\forall x, \forall y, \forall z, P(x, y, z)$  is false: take  $x = 1$  and  $y = 1$ , then whenever  $z \neq 1$ ,  $xyz = z \neq 1$ .

2.  $\exists x, \exists y, \exists z, P(x, y, z)$  is true: take  $x = y = z = 1$ .
3.  $\forall x, \forall y, \exists z, P(x, y, z)$  is true: choose any  $x$  and any  $y$ , then there exists a  $z$ , namely  $z = \frac{1}{xy}$  such that  $xyz = 1$ .
4.  $\exists x, \forall y, \forall z, P(x, y, z)$  is false: one cannot find a single  $x$  such that  $xyz = 1$  no matter what are  $y$  and  $z$ . This is because once  $yz$  are chosen, then  $x$  is completely determined, so  $x$  changes whenever  $yz$  does.

**Exercise 23.** 1. Express

$$\neg(\forall x, \forall y, P(x, y))$$

in terms of existential quantification.

2. Express

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

in terms of universal quantification.

*Solution.* 1. We see that  $\neg(\forall x, \forall y, P(x, y))$  is a negation of two universal quantifications. Denote  $Q(x) = \text{"}\forall y, P(x, y)\text{"}$ , then  $\neg(\forall x, Q(x))$  is  $(\exists x, \neg Q(x))$ , thus

$$\neg(\forall x, \forall y, P(x, y)) \equiv \exists x, \neg(\forall y, P(x, y))$$

and now we iterate the same rule on the next negation, to get

$$\neg(\forall x, \forall y, P(x, y)) \equiv \exists x, \exists y, \neg P(x, y).$$

2. We repeat the same procedure with the negation of two existential quantifications, by setting this time  $Q(x) = \text{"}\exists y, P(x, y)\text{"}$ :

$$\begin{aligned} \neg(\exists x, \exists y, P(x, y)) &\equiv \neg(\exists x Q(x)) \\ &\equiv \forall x \neg Q(x) \\ &\equiv \forall x \neg(\exists y, P(x, y)) \\ &\equiv \forall x \forall y \neg P(x, y). \end{aligned}$$

**Exercise 24.** Consider the predicate  $C(x, y) = \text{"}x \text{ is enrolled in the class } y\text{"}$ , where  $x$  takes values in the domain  $S = \{\text{students}\}$ , and  $y$  takes values in the domain  $C = \{\text{courses}\}$ . Form the negation of these statements:

1.  $\exists x, (C(x, \text{MH1812}) \wedge C(x, \text{CZ2002}))$ .
2.  $\exists x \exists y, \forall z, ((x \neq y) \wedge (C(x, z) \leftrightarrow C(y, z)))$ .

*Solution.* 1. We have

$$\begin{aligned} \neg(\exists x, (C(x, \text{MH1812}) \wedge C(x, \text{CZ2002}))) &\equiv \forall x \neg(C(x, \text{MH1812}) \wedge C(x, \text{CZ2002})) \\ &\equiv \forall x \neg C(x, \text{MH1812}) \vee \neg C(x, \text{CZ2002}) \end{aligned}$$

where the first equivalence is the negation of quantification, and the second equivalence De Morgan's law.

2. We have

$$\begin{aligned} \neg(\exists x \exists y, \forall z, ((x \neq y) \wedge (C(x, z) \leftrightarrow C(y, z)))) &\equiv \forall x \neg(\exists y, \forall z, ((x \neq y) \wedge (C(x, z) \leftrightarrow C(y, z)))) \\ &\equiv \forall x \forall y \neg(\forall z, ((x \neq y) \wedge (C(x, z) \leftrightarrow C(y, z)))) \\ &\equiv \forall x \forall y \exists z \neg((x \neq y) \wedge (C(x, z) \leftrightarrow C(y, z))) \\ &\equiv \forall x \forall y \exists z \neg(x \neq y) \vee \neg(C(x, z) \leftrightarrow C(y, z)) \end{aligned}$$

using three times the negation of quantification, and lastly the Morgan's law. Next  $\neg(x \neq y) = (x = y)$  and using that

$$C(x, z) \leftrightarrow C(y, z) \equiv (C(x, z) \rightarrow C(y, z)) \wedge (C(y, z) \rightarrow C(x, z))$$

we get

$$\neg(C(x, z) \leftrightarrow C(y, z)) \equiv \neg(C(x, z) \rightarrow C(y, z)) \vee \neg(C(y, z) \rightarrow C(x, z))$$

so that, using the Conversion theorem to get  $\neg(\neg C(x, z) \vee C(y, z))$  and  $\neg(\neg C(y, z) \vee C(x, z))$

$$\begin{aligned} & \neg(\exists x \exists y, \forall z, ((x \neq y) \wedge (C(x, z) \leftrightarrow C(y, z)))) \\ & \equiv \forall x \forall y \exists z((x = y) \vee [(C(x, z) \wedge \neg C(y, z)) \vee (C(y, z) \wedge \neg C(x, z))]). \end{aligned}$$

The last term can be further modified using distributivity:

$$\begin{aligned} & (C(x, z) \wedge \neg C(y, z)) \vee (C(y, z) \wedge \neg C(x, z)) \\ & \equiv [(C(x, z) \wedge \neg C(y, z)) \vee C(y, z)] \wedge [(C(x, z) \wedge \neg C(y, z)) \vee \neg C(x, z)] \\ & \equiv (C(x, z) \vee C(y, z)) \wedge (\neg C(x, z) \vee \neg C(y, z)) \end{aligned}$$

to finally get

$$\begin{aligned} & \neg(\exists x \exists y, \forall z, ((x \neq y) \wedge (C(x, z) \leftrightarrow C(y, z)))) \\ & \equiv \forall x \forall y \exists z((x = y) \vee [(C(x, z) \vee C(y, z)) \wedge (\neg C(x, z) \vee \neg C(y, z))]). \end{aligned}$$

When many steps are involved, it is often a good idea to check the sanity of the answer. If we look at  $\neg(C(x, z) \leftrightarrow C(y, z))$ , it is false exactly when  $C(x, z)$  and  $C(y, z)$  are taking the same truth value (either both true or both false). Now we look at  $(C(x, z) \vee C(y, z)) \wedge (\neg C(x, z) \vee \neg C(y, z))$ : when  $C(x, z)$  and  $C(y, z)$  are taking the same value, we get false, and true otherwise. This makes sense!

**Exercise 25.** Show that  $\forall x \in D, P(x) \rightarrow Q(x)$  is equivalent to its contrapositive.

*Solution.* For every instantiation of  $x$ ,  $(\forall x \in D, P(x) \rightarrow Q(x))$  is a proposition, thus we can use the conversion theorem:

$$\begin{aligned} & (\forall x \in D, P(x) \rightarrow Q(x)) \\ & \equiv (\forall x \in D, \neg P(x) \vee Q(x)) \\ & \equiv (\forall x \in D, Q(x) \vee \neg P(x)) \\ & \equiv (\forall x \in D, \neg \neg Q(x) \vee \neg P(x)) \\ & \equiv (\forall x \in D, \neg Q(x) \rightarrow \neg P(x)). \end{aligned}$$

**Exercise 26.** Show that

$$\neg(\forall x, P(x) \rightarrow Q(x)) \equiv \exists x, P(x) \wedge \neg Q(x).$$

*Solution.*

$$\begin{aligned} & \neg(\forall x, P(x) \rightarrow Q(x)) \\ & \equiv \exists x, \neg(P(x) \rightarrow Q(x)) \\ & \equiv \exists x, \neg(\neg P(x) \vee Q(x)) \\ & \equiv \exists x, (P(x) \wedge \neg Q(x)) \end{aligned}$$

where the first equivalence is the negation of quantifications, the second equivalence is the conversion theorem, and the third one is De Morgan's law.

**Exercise 27.** Let  $y, z$  be positive integers. What is the truth value of “ $\exists y, \exists z, (y = 2z \wedge (y \text{ is prime}))$ ”.

*Solution.* The truth value is true, take  $y = 2$  and  $z = 1$ .

**Exercise 28.** Write in symbolic logic “Every SCE student studies discrete mathematics. Jackson is an SCE student. Therefore Jackson studies discrete mathematics”.

*Solution.* Consider the domain  $D = \{ \text{SCE students} \}$ . Set  $P(x) = "x \text{ studies discrete mathematics}"$ . Then every SCE student studies discrete mathematics becomes

$$\forall x \in D, P(x).$$

Now Jackson is a SCE student means Jackson belongs to  $D$ . This gives

$$\forall x \in D, P(x); \text{Jackson} \in D; \therefore P(\text{Jackson}).$$

**Exercise 29.** Here is an optional exercise about universal generalization. Consider the following two premises: (1) for any number  $x$ , if  $x > 1$  then  $x - 1 > 0$ , (2) every number in  $D$  is greater than 1. Show that therefore, for every number  $x$  in  $D$ ,  $x - 1 > 0$ .

*Solution.* Set  $P(x) = "x > 1"$  and  $Q(x) = "x - 1 > 0"$ . Let us formalize what we want to prove:

$$[\forall x (P(x) \rightarrow Q(x)) \wedge \forall x \in D P(x)] \rightarrow \forall x \in D, Q(x).$$

1.  $\forall x (P(x) \rightarrow Q(x))$ , by hypothesis
2.  $\forall x \in D P(x)$ , also by hypothesis
3.  $P(y) \rightarrow Q(y)$ , by universal instantiation on the first hypothesis
4.  $P(y)$ , by universal instantiation on  $D$  in the second hypothesis
5.  $Q(y)$ , using modus ponens
6.  $\forall x \in D, Q(x)$ , using universal generalization.

**Exercise 30.** Let  $q$  be a positive real number. Prove or disprove the following statement: if  $q$  is irrational, then  $\sqrt{q}$  is irrational.

*Solution.* We prove the contrapositive of this statement, namely: if  $\sqrt{q}$  is rational, then  $q$  is rational. But if  $\sqrt{q}$  is rational, then  $\sqrt{q} = \frac{a}{b}$ ,  $a, b$  integers,  $b \neq 0$ , and thus  $q = \frac{a^2}{b^2}$  which shows that  $q$  is rational.

**Exercise 31.** Prove using mathematical induction that the sum of the first  $n$  odd positive integers is  $n^2$ .

*Solution.* We want to prove that  $\forall n, P(n)$  where

$$P(n) = \left( \sum_{j=1}^n (2j - 1) = n^2 \right), \quad n \in \mathbb{Z}, \quad \geq 1.$$

- Basis Step: we need to show that  $P(1)$  is true.

$$P(1) = (2 - 1) = 1 = 1^2$$

which is true.

- Inductive Step: Assume  $P(k)$  is true, that is we assume that

$$\sum_{j=1}^k (2j - 1) = k^2$$

is true. We now need to prove that  $P(k + 1)$  is true.

$$\begin{aligned} & \sum_{j=1}^{k+1} (2j - 1) \\ &= \sum_{j=1}^k (2j - 1) + 2(k + 1) - 1 \\ &= k^2 + 2(k + 1) - 1 \text{ using } P(k) \\ &= k^2 + 2k + 1 = (k + 1)^2. \end{aligned}$$

This shows that  $P(k + 1)$  is true, therefore  $P(n)$  is true for all  $n$ .

**Exercise 32.** Prove using mathematical induction that  $n^3 - n$  is divisible by 3 whenever  $n$  is a positive integer.

*Solution.* We first set  $P(n) = "3 \mid n^3 - n"$ , that is 3 divides  $n^3 - n$ .

- Basis Step:  $P(1) = "3|0"$  which is true.
- Inductive Step: Assume  $P(k)$  is true, that is we assume that

$$3 \mid (k^3 - k).$$

is true. We now need to prove that  $P(k + 1)$  is true. When  $n = k + 1$ , we get

$$(k + 1)^3 - (k + 1) = (k^3 + 3k^2 + 3k + 1) - (k + 1) = (k^3 - k) + 3k^2 + 3k$$

which is divisible by 3, since  $(k^3 - k)$  is divisible by 3, and so is  $3(k^2 + k)$ . Therefore  $P(k + 1)$  is true, and we conclude that  $P(n)$  is true for all  $n$ .

## Exercises for Chapter 4

**Exercise 33.** 1. Show that

$$\binom{n}{k} + \binom{n}{k-1} = \binom{n+1}{k}$$

for  $1 \leq k \leq l$ , where by definition

$$\binom{n}{k} = \frac{n!}{k!(n-k)!}, \quad n! = n \cdot (n-1) \cdot (n-2) \cdots 2 \cdot 1.$$

2. Prove by mathematical induction that

$$(x+y)^n = \sum_{k=0}^n \binom{n}{k} x^{n-k} y^k.$$

You will need 1. for this!

3. Deduce that the cardinality of the power set  $P(S)$  of a finite set  $S$  with  $n$  elements is  $2^n$ .

*Solution.* To prove

$$\binom{n}{k} + \binom{n}{k-1} = \binom{n+1}{k},$$

we first expand the left hand side:

$$\frac{n!}{k!(n-k)!} + \frac{n!}{(k-1)!(n-k+1)!} = \frac{n!(n-k+1)}{k!(n-k+1)!} + \frac{n!k}{k!(n-k+1)!}$$

which is equal to

$$\frac{n!(n+1)}{k!(n-k+1)!} = \frac{(n+1)!}{k!(n-k+1)!} = \binom{n+1}{k}.$$

To prove the binomial theorem by mathematical induction, we set

$$P(n) = \text{"}(x+y)^n = \sum_{k=0}^n \binom{n}{k} x^{n-k} y^k\text{"},$$

and we want to prove that  $\forall n, P(n)$ . The basis step is to prove that  $P(1)$  holds, which is given by

$$(x+y) = \sum_{k=0}^1 \binom{1}{k} x^{1-k} y^k = x+y.$$

Next for the inductive step, we suppose that  $P(l)$  is true, namely

$$(x+y)^l = \sum_{k=0}^l \binom{l}{k} x^{l-k} y^k \tag{12.1}$$

and we want to prove  $P(l + 1)$ .

$$\begin{aligned}
 (x + y)^{l+1} &= (x + y)(x + y)^l \\
 &= (x + y) \sum_{k=0}^l \binom{l}{k} x^{l-k} y^k \text{ (using (12.1))} \\
 &= x \sum_{k=0}^l \binom{l}{k} x^{l-k} y^k + y \sum_{k=0}^l \binom{l}{k} x^{l-k} y^k \\
 &= \sum_{k=0}^l \binom{l}{k} x^{l-k+1} y^k + \sum_{k=0}^l \binom{l}{k} x^{l-k} y^{k+1}.
 \end{aligned}$$

At this point, it is probably a good idea to remember what we want to prove, namely

$$(x + y)^{l+1} = \sum_{k=0}^{l+1} \binom{l+1}{k} x^{(l+1)-k} y^k.$$

From our aim, we notice that the first sum has already right exponents, namely  $x^{l-k+1} y^k$  is a term we want. So we first work on the other sum to get a similar right term present, by doing a change of variable  $j = k + 1$ , to get

$$\sum_{k=0}^l \binom{l}{k} x^{l-k} y^{k+1} = \sum_{j=1}^{l+1} \binom{l}{j-1} x^{l-j+1} y^j.$$

We next combine this derivation:

$$\begin{aligned}
 &(x + y)^{l+1} \\
 &= \sum_{k=0}^l \binom{l}{k} x^{l-k+1} y^k + \sum_{j=1}^{l+1} \binom{l}{j-1} x^{l-j+1} y^j \\
 &= \sum_{k=1}^l \binom{l}{k} x^{l-k+1} y^k + \binom{l}{0} x^{l+1} + \sum_{j=1}^l \binom{l}{j-1} x^{l-j+1} y^j + \binom{l}{l} y^{l+1} \\
 &= \sum_{k=1}^l \left[ \binom{l}{k} + \binom{l}{k-1} \right] x^{l-k+1} y^k + y^{l+1} + x^{l+1}
 \end{aligned}$$

and now is the point where we recognize the formula that we derived in 1.,

thus

$$\begin{aligned}
 & (x+y)^{l+1} \\
 = & \sum_{k=1}^l \binom{l+1}{k} x^{l-k+1} y^k + y^{l+1} + x^{l+1} \\
 = & \sum_{k=0}^{l+1} \binom{l+1}{k} x^{l-k+1} y^k
 \end{aligned}$$

Finally, evaluate the binomial theorem in  $x = y = 1$ . The only thing left to be seen is the interpretation of  $\binom{n}{k}$  as “ $n$  choose  $k$ ”, which will be discussed into more details in the next chapter, namely  $\binom{n}{k}$  counts the number of ways of picking  $k$  elements out of  $n$ . Therefore to count the number of elements in  $P(S)$  we just count how many subsets we have with 1 element, with 2 elements, ..., and we sum these numbers up!

**Exercise 34.** Let  $P(C)$  denote the power set of  $C$ . Given  $A = \{1, 2\}$  and  $B = \{2, 3\}$ , determine:

$$P(A \cap B), P(A), P(A \cup B), P(A \times B).$$

*Solution.* •  $A \cap B = \{2\}$ , therefore  $P(A \cap B) = \{\emptyset, \{2\}\}$ .

- $P(A) = \{\emptyset, \{1\}, \{2\}, \{1, 2\}\}$ .
- $A \cup B = \{1, 2, 3\}$ , therefore  $P(A \cup B) = \{\emptyset, \{1\}, \{2\}, \{3\}, \{1, 2\}, \{1, 3\}, \{2, 3\}, \{1, 2, 3\}\}$ .
- $A \times B = \{(1, 2), (1, 3), (2, 2), (2, 3)\}$ , therefore  $P(A \times B)$  contains
  - $\emptyset, \{(1, 2)\}, \{(1, 3)\}, \{(2, 2)\}, \{(2, 3)\}$ ,
  - $\{(1, 2), (1, 3)\}, \{(1, 2), (2, 2)\}, \{(1, 2), (2, 3)\}, \{(1, 3), (2, 2)\}, \{(1, 3), (2, 3)\}, \{(2, 2), (2, 3)\}$ ,
  - $\{(1, 2), (1, 3), (2, 2)\}, \{(1, 2), (1, 3), (2, 3)\}, \{(1, 2), (2, 3), (2, 2)\}, \{(1, 3), (2, 2), (2, 3)\}$ ,
  - $\{(1, 2), (1, 3), (2, 2), (2, 3)\}$ .

**Exercise 35.** Prove by contradiction that for two sets  $A$  and  $B$

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

*Solution.* Suppose by contradiction that  $(A - B) \cap (B - A)$  is not empty. Then there exists an element  $x$  which belongs to both  $(A - B)$  and  $(B - A)$ . This means that  $x$  belongs to  $A$  (since  $x \in A - B$ ), and  $x$  does not belong to  $A$  (since  $x \in B - A$ ), which is a contradiction! Therefore the assumption was false, and  $(A - B) \cap (B - A)$  is empty.

Note that from a propositional logic point of view, what we did is set  $p = "A - B) \cap (B - A) = \emptyset"$ ,  $q = "x \in A"$ , and we prove that

$$\neg p \rightarrow (q \wedge \neg q)$$

which turns out to be equivalent to  $p$ .

**Exercise 36.** Let  $P(C)$  denote the power set of  $C$ . Prove that for two sets  $A$  and  $B$

$$P(A) = P(B) \iff A = B.$$

*Solution.* We need to show that  $P(A) = P(B) \rightarrow A = B$  and  $A = B \rightarrow P(A) = P(B)$ .

- suppose  $P(A) = P(B)$ : then all sets containing one element are the same for  $P(A)$  and  $P(B)$ , and  $A = B$ .
- suppose  $A = B$ : subsets of  $A$  and subsets of  $B$  are the same, and  $P(A) = P(B)$ .

**Exercise 37.** Let  $P(C)$  denote the power set of  $C$ . Prove that for two sets  $A$  and  $B$

$$P(A) \subseteq P(B) \iff A \subseteq B.$$

*Solution.* We need to show that  $P(A) \subseteq P(B) \rightarrow A \subseteq B$  and  $A \subseteq B \rightarrow P(A) \subseteq P(B)$ .

- suppose  $P(A) \subseteq P(B)$ : then  $A \subseteq P(B)$ , from which  $A \subseteq B$ .
- suppose  $A \subseteq B$ : then for any  $X \in P(A)$ ,  $X \subseteq A$ ,  $X \subseteq B$ , therefore  $X \in P(B)$ .

**Exercise 38.** Show that the empty set is a subset of all non-null sets.

*Solution.* Recall the definition of subset:  $Y \subseteq X$  means by definition that  $\forall x, (x \in Y \rightarrow x \in X)$ . Now take  $Y$  to be the empty set  $\emptyset$ . Since  $x \in Y$  is necessarily false (one cannot take any  $x$  in the empty set), then the conditional statement is vacuously true.

**Exercise 39.** Show that for two sets  $A$  and  $B$

$$A \neq B \equiv \exists x[(x \in A \wedge x \notin B) \vee (x \in B \wedge x \notin A)].$$

*Solution.*

$$\begin{aligned} A \neq B &= \neg \forall x(x \in A \leftrightarrow x \in B) \\ &\equiv \exists x \neg(x \in A \leftrightarrow x \in B) \text{ (negation of universal quantifier)} \\ &\equiv \exists x \neg[(x \in A \rightarrow x \in B) \wedge (x \in B \rightarrow x \in A)] \text{ (definition)} \\ &\equiv \exists x [\neg(x \in A \rightarrow x \in B) \vee \neg(x \in B \rightarrow x \in A)] \text{ (DeMorgan)} \\ &\equiv \exists x [\neg(x \notin A \vee x \in B) \vee \neg(x \notin B \vee x \in A)] \text{ (Conversion)} \\ &\equiv \exists x [(x \in A \wedge x \notin B) \vee (x \in B \wedge x \notin A)] \text{ (DeMorgan)} \end{aligned}$$

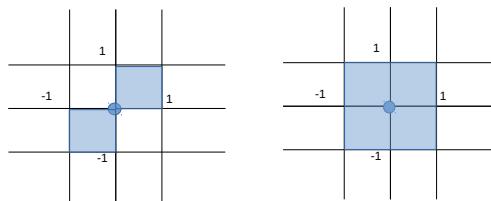
**Exercise 40.** Prove that for the sets  $A, B, C, D$

$$(A \times B) \cup (C \times D) \subseteq (A \cup C) \times (B \cup D).$$

Does equality hold?

*Solution.* Suppose  $x \in (A \times B) \cup (C \times D)$ , then  $x = (x_1, x_2)$  with  $x_1 \in A, x_2 \in B$  (or  $x_1 \in C, x_2 \in D$ ). But then  $x_1 \in A$  (or  $C$ ), and  $x_2 \in B$  (or  $D$ ), therefore  $x \in (A \cup C) \times (B \cup D)$ . The equality does not hold: take  $A = [-1, 0], B = [-1, 0], C = [0, 1], D = [0, 1]$  (all the set are intervals, that is  $[a, b]$  means the interval from  $a$  to  $b$ ). Then

$$([-1, 0] \times [-1, 0]) \cup ([0, 1] \times [0, 1]) \neq [-1, 1] \times [-1, 1].$$



**Exercise 41.** Does the equality

$$(A_1 \cup A_2) \times (B_1 \cup B_2) = (A_1 \times B_1) \cup (A_2 \times B_2)$$

hold?

*Solution.* No it does not. For example, take  $A_1 = \{0\}$ ,  $A_2 = \{1\}$ ,  $B_1 = \{0\}$ ,  $B_2 = \{1\}$ , then

$$\{0, 1\} \times \{0, 1\} \neq \{(0, 0), (1, 1)\}.$$

**Exercise 42.** For all sets  $A$ ,  $B$ ,  $C$ , prove that

$$\overline{(A - B) - (B - C)} = \bar{A} \cup B.$$

using set identities.

*Solution.* We have

$$\begin{aligned} \overline{(A - B) - (B - C)} &= \overline{(A \cap \bar{B}) - (B \cap \bar{C})} \\ &= \overline{(A \cap \bar{B})} \cap \overline{(B \cap \bar{C})} \\ &= \overline{(A \cap \bar{B})} \cup (B \cap \bar{C}) \\ &= (\bar{A} \cup B) \cup (B \cap \bar{C}) \\ &= \bar{A} \cup B \cup (B \cap \bar{C}) \\ &= \bar{A} \cup B \end{aligned}$$

Both the 3rd and 4th equality follows from De Morgan's Laws for sets, and  $\bar{\bar{S}} = S$  for any set  $S$ . The 5th equality is associativity, while the last equality is true because  $(B \cap \bar{C})$  is a subset of  $B$ .

**Exercise 43.** This exercise is more difficult. For all sets  $A$  and  $B$ , prove  $(A \cup B) \cap \overline{A \cap B} = (A - B) \cup (B - A)$  by showing that each side of the equation is a subset of the other.

*Solution.* We have to prove that (1)  $(A \cup B) \cap \overline{A \cap B} \subseteq (A - B) \cup (B - A)$  and (2)  $(A \cup B) \cap \overline{A \cap B} \supseteq (A - B) \cup (B - A)$ .

**Part (1).** Suppose that  $x \in (A \cup B) \cap \overline{A \cap B}$ , then

$$(x \in (A \cup B)) \wedge (x \in \overline{A \cap B}).$$

Now  $(x \in (A \cup B))$  means that  $(x \in A) \vee (x \in B)$ , that is

$$\begin{aligned} & [(x \in A) \vee (x \in B)] \wedge (x \in \overline{A \cap B}) \\ & \equiv [(x \in A) \wedge (x \in \overline{A \cap B})] \vee [(x \in B) \wedge (x \in \overline{A \cap B})] \end{aligned}$$

using distributivity.

Next  $(x \in \overline{A \cap B})$  means that  $x \notin A \cap B$  ( $x$  always lives in the universe  $U$ , so it is not repeated). Now the negation of  $x \in A \cap B$  is  $(x \notin A) \vee (x \notin B)$ . We thus get that  $(x \in A) \wedge (x \in \overline{A \cap B})$  becomes

$$(x \in A) \wedge [(x \notin A) \vee (x \notin B)] \equiv F \vee [(x \in A) \wedge (x \notin B)]$$

using distributivity. Repeating the same procedure by flipping the role of  $B$  and  $A$  in  $(x \in B) \wedge (x \in \overline{A \cap B})$ , we finally obtain that

$$[(x \in A) \wedge (x \notin B)] \vee [(x \in B) \wedge (x \notin A)].$$

We have thus shown that  $(A \cup B) \cap \overline{A \cap B} \subseteq (A - B) \cup (B - A)$ .

**Part (2).** For the second part, we need to show  $(A \cup B) \cap \overline{A \cap B} \supseteq (A - B) \cup (B - A)$ .

Suppose thus that  $x \in (A - B) \cup (B - A)$ , that is  $x \in (A - B)$  or  $x \in (B - A)$ . If  $x \in (A - B)$  then  $x \in A$  and  $x \notin B$  by definition. Therefore  $x \in A \cup B$  and  $x \notin A \cap B$ .

Similarly, if  $x \in (B - A)$  then  $x \in B$  and  $x \notin A$  by definition. Therefore  $x \in A \cup B$  and  $x \notin A \cap B$ .

**Exercise 44.** The symmetric difference of  $A$  and  $B$ , denoted by  $A \oplus B$ , is the set containing those elements in either  $A$  or  $B$ , but not in both  $A$  and  $B$ .

1. Prove that  $(A \oplus B) \oplus B = A$  by showing that each side of the equation is a subset of the other.
2. Prove that  $(A \oplus B) \oplus B = A$  using a membership table.

*Solution.* It is a good idea to draw a Venn diagram to visualize  $(A \oplus B)$ , which consists of  $A \cup B$  without the intersection  $A \cap B$ .

1. We have to show that (1)  $(A \oplus B) \oplus B \subseteq A$ , and (2)  $(A \oplus B) \oplus B \supseteq A$ .
  - $(A \oplus B) \oplus B \subseteq A$ : take  $x \in (A \oplus B) \oplus B$ . If  $x \in B$ , then  $x \notin A \oplus B$  by definition. But then, it must be that  $B \in A \cap B$  that is,  $x \in A$  as desired. Next if  $x \notin B$ , then  $x \in A \oplus B$  by definition. But then  $x$  is in the union  $A \cup B$  though not in the intersection  $A \cap B$ , and since it is not in  $B$ , it must be in  $A$ .

- $(A \oplus B) \oplus B \supseteq A$ : we now start with  $x \in A$ . If  $x \in B$  (that is,  $x \in A \cap B$ ), then  $x \notin A \oplus B$ , then  $x \in (A \oplus B) \oplus B$ . Next if  $x \notin B$ , then  $x \in A \oplus B$ , and thus  $x \in (A \oplus B) \oplus B$ .

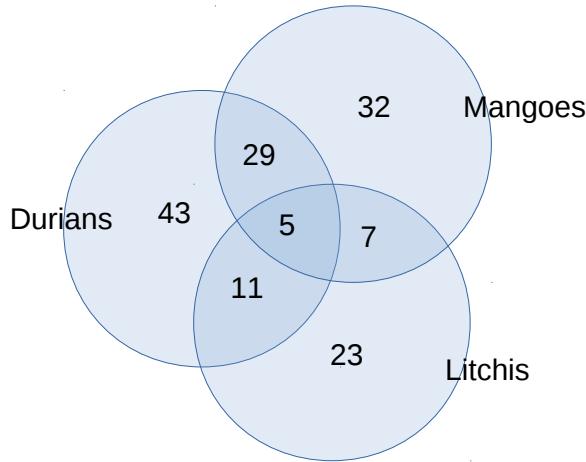
2. We construct a membership table as shown next:

$A$	$B$	$A \oplus B$	$(A \oplus B) \oplus B$
0	0	0	0
0	1	1	0
1	0	1	1
1	1	0	1

For example, for the second row,  $x \in B$  but not in  $A$ . Then  $x \in A \oplus B$ . But then it cannot be in  $(A \oplus B) \oplus B$  since  $x$  is in both  $B$  and  $A \oplus B$ . We conclude since both the first and last column are the same.

**Exercise 45.** In a fruit feast among 200 students, 88 chose to eat durians, 73 ate mangoes, and 46 ate litchis. 34 of them had eaten both durians and mangoes, 16 had eaten durians and litchis, and 12 had eaten mangoes and litchis, while 5 had eaten all 3 fruits. Determine, how many of the 200 students ate none of the 3 fruits, and how many ate only mangoes?

*Solution.* Let us draw a Venn diagram with 3 sets (one for each of the fruits) and start by identifying the numbers of students who ate all the 3 fruits, namely 5 of them. Then we identify the number of students who ate two types of fruits (for example, 34 ate durians and mangoes, so  $34 - 5 = 29$ ), and finally only one type of fruit.



We thus get a total of 150, meaning that 50 students ate nothing. 32 students ate only mangoes.

## Exercises for Chapter 5

**Exercise 46.** A set menu proposes 2 choices of starters, 3 choices of main dishes, and 2 choices of desserts. How many possible set menus are available?

*Solution.* You have 2 choices of starters, then for any choice, you get 3 choices of main dishes, or for each of them you get 2 choices of desserts. Therefore the total is

$$2 \cdot 3 \cdot 2 = 12.$$

**Exercise 47.** Consider the set  $A = \{1, 2, 3\}$ ,  $P(A)$  = power set of  $A$ .

- Compute the cardinality of  $P(A)$  using the binomial theorem approach.
- Compute the cardinality of  $P(A)$  using the counting approach.

*Solution.* • To compute the cardinality of  $P(A)$ , we need to count the empty set (1), the number of subset of size 1 ( $\binom{3}{1}$ ), the number of sets

of size 2 ( $\binom{3}{2}$ ) and the whole set (1), therefore:

$$|P(A)| = 1 + \binom{3}{1} + \binom{3}{2} + 1 = 2^3$$

where the last equality follows from the binomial theorem.

- In a counting approach, we write binary strings to identify whether a given element belongs to a subset, for example 000 corresponds to the empty set, and 000 is read as 1 is not in this subset, 2 is not, and 3 is not either. Now for every subset, each element either belongs, or does not belong, therefore we get the 8 possible binary strings, and the cardinality is  $2^3$ .

**Exercise 48.** • Two fair coins are tossed. What is the probability of getting 2 heads?

- Three fair coins are tossed. What is the probability of getting exactly 2 heads?

*Solution.* • Two fair coins are tossed, therefore the sample space is

$$\{HH, HT, TH, TT\}.$$

The probability of getting  $HH$  is therefore  $1/4$ .

- Three fair coins are tossed, therefore the sample space is

$$\{HHH, HTH, THH, TTH, HHT, HTT, THT, TTT\}.$$

The event getting exactly 2 heads is  $\{HHT, HTH, THH\}$ . Therefore the probability is  $3/8$ .

**Exercise 49.** Ten fair coins are tossed together. What is the probability that there were at least seven heads?

*Solution.* To have at least seven heads means that the number of heads is either 7, 8, 9 or 10. The number of patterns for 7 heads is  $C(10, 7)$ , the number of patterns for 8 heads is  $C(10, 8)$ , the number of patterns for 9 heads is  $C(10, 9)$ , and finally  $C(10, 10)$  for 10 heads. The total number of outcomes is  $2^{10}$ , thus we get

$$\frac{\sum_{i=7}^{10} C(10, i)}{2^{10}} = \frac{120 + 45 + 10 + 1}{2^{10}} = \frac{176}{2^{10}}$$

since

$$C(10, 7) = \frac{10!}{7!3!} = 10 \cdot 3 \cdot 4, \quad C(10, 8) = \frac{10!}{8!2!} = 5 \cdot 9, \quad C(10, 9) = \frac{10!}{9!} = 10.$$

**Exercise 50.** Snow white is going to a party with the seven dwarves. Each of the eight of them owns a red dress and a blue dress. If each of them is likely to choose either colored dress randomly and independently of the other's choices, what is the chance that all of them go to the party wearing the same colored dress?

*Solution.* Either they all dress in blue, or they all dress in blue, therefore

$$\frac{2}{2^8} = \frac{1}{2^7}.$$

## Exercises for Chapter 6

**Exercise 51.** Consider the linear recurrence  $a_n = 2a_{n-1} - a_{n-2}$  with initial conditions  $a_1 = 3, a_0 = 0$ .

- Solve it using the backtracking method.
- Solve it using the characteristic equation.

*Solution.* • We have  $a_n = 2a_{n-1} - a_{n-2}$ , thus  $a_{n-1} = 2a_{n-2} - a_{n-3}, a_{n-2} = 2a_{n-3} - a_{n-4}, a_{n-3} = 2a_{n-4} - a_{n-5}$ , etc therefore

$$\begin{aligned} a_n &= 2a_{n-1} - a_{n-2} \\ &= 2(2a_{n-2} - a_{n-3}) - a_{n-2} = 3a_{n-2} - 2a_{n-3} \\ &= 3(2a_{n-3} - a_{n-4}) - 2a_{n-3} = 4a_{n-3} - 3a_{n-4} \\ &= 4(2a_{n-4} - a_{n-5}) - 3a_{n-4} = 5a_{n-4} - 4a_{n-5} \\ &= \dots \end{aligned}$$

We see that a general term is  $(i+1)a_{n-i} - ia_{n-(i+1)}$ . Therefore the last term is when  $n - i - 1 = 0$  that is  $i = n - 1$ , for which we have  $na_1 - (n-1)a_0$ , therefore with initial condition  $a_0 = 0$  and  $a_1 = 3$ , we get

$$a_n = 3n.$$

*Optional.* Now if one wants to be sure that this is indeed the right answer, this can be checked using a proof by mathematical induction! However here, the mathematical induction is slightly different from our usual one! We have

$$P(n) = "a_n = 3n",$$

so the basis step which is  $P(0) = "a_0 = 0"$  holds. However we will also need a second basis step, which is  $P(1) = "a_1 = 3"$ , which still holds. Now suppose  $P(k) = "a_k = 3k"$  and  $P(k-1) = "a_{k-1} = 3(k-1)"$  are both true. Then

$$\begin{aligned} a_{k+1} &= 2a_k - a_{k-1} \\ &= 6k - 3(k-1) \\ &= 6k - 3k + 3 = 3k + 3 = 3(k+1) \end{aligned}$$

as needed, where we used both our induction hypotheses!

- Suppose now we want to solve the same recurrence using a characteristic equation. We have  $x^n = 2x^{n-1} - x^{n-2}$  that is

$$x^n - 2x^{n-1} + x^{n-2} = 0 \iff x^{n-2}(x^2 - 2x + 1) = 0.$$

We have  $x^2 - 2x + 1 = (x-1)^2$ , therefore

$$a_n = u + vn.$$

Then

$$a_0 = u = 0, \quad a_1 = u + v = 3$$

thus  $v = 3$ , yielding

$$a_n = 3n.$$

**Exercise 52.** What is the solution of the recurrence relation

$$a_n = a_{n-1} + 2a_{n-2}$$

with  $a_0 = 2$  and  $a_1 = 7$ ?

*Solution.* The characteristic equation is  $x^2 - x - 2 = 0$ . Its roots are  $x = -1$  and  $x = 2$  since  $(x+1)(x-2) = 0$ . Therefore  $a_n = u2^n + v(-1)^n$  is a solution. We are left with identifying  $u, v$  using the initial conditions.

$$a_0 = 2 = u + v, \quad a_1 = 2u - v = 7.$$

So  $u = 3, v = -1$ , therefore

$$a_n = 3 \cdot 2^n - (-1)^n.$$

**Exercise 53.** Let  $a_n = c_1 a_{n-1} + c_2 a_{n-2} + \dots + c_k a_{n-k}$  be a linear homogeneous recurrence. Assume both sequences  $a_n, a'_n$  satisfy this linear homogeneous recurrence. Show that  $a_n + a'_n$  and  $\alpha a_n$  also satisfy it, for  $\alpha$  some constant.

*Solution.* We have

$$\begin{aligned} a_n + a'_n &= (c_1 a_{n-1} + c_2 a_{n-2} + \dots + c_k a_{n-k}) + (c_1 a'_{n-1} + c_2 a'_{n-2} + \dots + c_k a'_{n-k}) \\ &= c_1(a_{n-1} + a'_{n-1}) + c_2(a_{n-2} + a'_{n-2}) + \dots + c_k(a_{n-k} + a'_{n-k}). \end{aligned}$$

Thus  $a_n + a'_n$  is a solution of the recurrence. Similarly

$$\begin{aligned} \alpha a_n &= \alpha(c_1 a_{n-1} + c_2 a_{n-2} + \dots + c_k a_{n-k}) \\ &= c_1 \alpha a_{n-1} + c_2 \alpha a_{n-2} + \dots + c_k \alpha a_{n-k}. \end{aligned}$$

Therefore  $\alpha a_n$  is a solution of the recurrence.

## Exercises for Chapter 7

**Exercise 54.** Set  $i = \sqrt{-1}$ . Compute

$$i^5, \frac{1}{i^2}, \frac{1}{i^3}.$$

*Solution.* First

$$i^5 = (i^4)i = i.$$

Note that  $i^4 = 1$  because  $i^2 = -1$ .

Then since  $i^2 = -1$ , we have

$$\frac{1}{i^2} = -1.$$

Finally since  $i^3 = -i$ , we have

$$\frac{1}{i^3} = \frac{1}{-i} = i.$$

Indeed  $-i \cdot i = 1$ .

**Exercise 55.** Set  $i = \sqrt{-1}$ . Compute the real part and the imaginary part of

$$\frac{(1+2i)-(2+i)}{(2-i)(3+i)}.$$

*Solution.* We have

$$\begin{aligned} \frac{(1+2i)-(2+i)}{(2-i)(3+i)} &= \frac{-1+i}{7-i} \\ &= \frac{(-1+i)(7+i)}{(7-i)(7+i)} \\ &= \frac{-4}{25} + i\frac{3}{25}. \end{aligned}$$

**Exercise 56.** Set  $i = \sqrt{-1}$ . Compute  $d, e \in \mathbb{R}$  such that

$$4 - 6i + d = \frac{7}{i} + ei.$$

*Solution.* Note that for a complex number to be zero, we need both its real and its imaginary parts to be 0. We thus need  $4+d=0$ , that is  $d=-4$ , and

$$\frac{7}{i} + ei + 6i = 0,$$

for which we need  $e=1$ .

**Exercise 57.** For  $z_1, z_2 \in \mathbb{C}$ , prove that

- $\overline{z_1 + z_2} = \bar{z}_1 + \bar{z}_2$ .
- $\overline{z_1 \cdot z_2} = \bar{z}_1 \cdot \bar{z}_2$ .

*Solution.* Write  $z_1 = a_1 + ib_1$ ,  $z_2 = a_2 + ib_2$ .

- We have

$$\begin{aligned} \overline{z_1 + z_2} &= \overline{(a_1 + a_2) + i(b_1 + b_2)} \\ &= (a_1 + a_2) - i(b_1 + b_2) \\ &= a_1 - ib_1 + a_2 - ib_2 \\ &= \bar{z}_1 + \bar{z}_2. \end{aligned}$$

- Similarly:

$$\begin{aligned}
 \overline{z_1 \cdot z_2} &= \overline{(a_1 + ib_1)(a_2 + ib_2)} \\
 &= a_1 a_2 - i(a_1 b_2 + b_1 a_2) - b_1 b_2 \\
 &= (a_1 - ib_1)(a_2 - ib_2) \\
 &= \bar{z}_1 \cdot \bar{z}_2.
 \end{aligned}$$

**Exercise 58.** Consider the complex number  $z$  in polar form:  $z = re^{i\theta}$ . Express  $re^{-i\theta}$  as a function of  $z$ .

*Solution.* Since

$$z = r(\cos \theta + i \sin \theta),$$

we have

$$re^{-i\theta} = r(\cos \theta - i \sin \theta) = \bar{z}.$$

**Exercise 59.** Prove that

$$(\cos x + i \sin x)^n = \cos nx + i \sin nx$$

for  $n$  an integer.

*Solution.* We have by Euler formula that

$$e^{ix} = \cos x + i \sin x,$$

thus

$$(e^{ix})^n = e^{inx} \cos nx + i \sin nx.$$

**Exercise 60.** Compute  $|e^{i\theta}|$ ,  $\theta \in \mathbb{R}$ .

*Solution.* We have

$$|e^{i\theta}|^2 = |\cos \theta + i \sin \theta|^2 = (\cos \theta)^2 + (\sin \theta)^2 = 1.$$

Therefore  $|e^{i\theta}| = 1$ .

**Exercise 61.** Prove the so-called triangle inequality:

$$|a + b| \leq |a| + |b|, \quad a, b \in \mathbb{C}.$$

*Solution.* Write  $a = a_1 + ia_2$ ,  $b = b_1 + ib_2$ . Then

$$|a + b|^2 = (a + b)\overline{(a + b)} = |a|^2 + |b|^2 + a\bar{b} + \bar{a}b$$

while

$$(|a| + |b|)^2 = |a|^2 + 2|a||b| + |b|^2$$

so we are left to show that  $a\bar{b} + \bar{a}b \leq 2|a||b|$ . We have

$$a\bar{b} = (a_1 + ia_2)\overline{(b_1 + ib_2)} = a_1b_1 - a_1ib_2 + ia_2b_1 + a_2b_2$$

and

$$\bar{a}b = \overline{(a_1 + ia_2)}(b_1 + ib_2) = a_1b_1 + a_1ib_2 - ia_2b_1 + a_2b_2$$

thus we are left to show that

$$a\bar{b} + \bar{a}b = 2a_1b_1 + 2a_2b_2 \leq 2|a||b|,$$

or equivalently that  $(a_1b_1 + a_2b_2)^2 \leq |ab|^2$ . But  $|ab|^2 = (a_1^2 + a_2^2)(b_1^2 + b_2^2)$  and we get

$$2a_1b_1a_2b_2 \leq a_1^2b_2^2 + a_2^2b_1^2 \iff 0 \leq (a_1b_2 - a_2b_1)^2$$

which is true.

**Exercise 62.** Compute the two roots of  $4i$ , that is

$$\sqrt{4i}.$$

*Solution.* We have

$$\sqrt{4i} = 2\sqrt{i} = \pm 2e^{2\pi i/8} = \pm 2(\cos 2\pi/8 + i \sin 2\pi/8).$$

One may further compute that  $\cos 2\pi/8 = \sin 2\pi/8 = 1/\sqrt{2}$  (this is for example seen by considering a square with edges of size 1). Therefore the two roots are

$$\pm 2(1/\sqrt{2} + i/\sqrt{2}).$$

## Exercises for Chapter 8

**Exercise 63.** Compute the sum  $A + B$  of the matrices  $A$  and  $B$ , where  $A$  and  $B$  are as follows:

1.

$$A = \begin{pmatrix} 2 & \sqrt{2} \\ -1 & 3 \end{pmatrix}, \quad B = \begin{pmatrix} 0 & \sqrt{2} \\ 4 & 2 \end{pmatrix}$$

where  $A, B$  are matrices with coefficients in  $\mathbb{R}$ .

2.

$$A = \begin{pmatrix} 2+i & -1 \\ -1+i & 3 \end{pmatrix}, \quad B = \begin{pmatrix} -i & 1 \\ -1 & 2 \end{pmatrix}$$

where  $A, B$  are matrices with coefficients in  $\mathbb{C}$ , and  $i = \sqrt{-1}$ .

3.

$$A = \begin{pmatrix} 0 & 1 & 2 \\ 2 & 0 & 2 \end{pmatrix}, \quad B = \begin{pmatrix} 1 & 0 & 2 \\ 0 & 1 & 1 \end{pmatrix}$$

where  $A, B$  are matrices with coefficients that are integers  $\pmod{3}$ .

What are the dimensions of the matrices involved?

*Solution.* 1. Matrices are  $2 \times 2$ .

$$A + B = \begin{pmatrix} 2 & \sqrt{2} \\ -1 & 3 \end{pmatrix} + \begin{pmatrix} 0 & \sqrt{2} \\ 4 & 2 \end{pmatrix} = \begin{pmatrix} 2 & 2\sqrt{2} \\ 3 & 5 \end{pmatrix}$$

2. Matrices are  $2 \times 2$ .

$$A + B = \begin{pmatrix} 2+i & -1 \\ -1+i & 3 \end{pmatrix} + \begin{pmatrix} -i & 1 \\ -1 & 2 \end{pmatrix} = \begin{pmatrix} 2 & 0 \\ -2+i & 5 \end{pmatrix}$$

3. Matrices are  $2 \times 3$ .

$$A + B = \begin{pmatrix} 0 & 1 & 2 \\ 2 & 0 & 2 \end{pmatrix} + \begin{pmatrix} 1 & 0 & 2 \\ 0 & 1 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 1 & 1 \\ 2 & 1 & 0 \end{pmatrix}$$

**Exercise 64.** 1. Compute the transpose  $A^T$  of  $A$  for

$$A = \begin{pmatrix} 1 & 0 & 2 \\ 0 & 1 & 1 \end{pmatrix}.$$

2. Show that  $(A + B)^T = A^T + B^T$ .

*Solution.* 1. We have

$$A^T = \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 2 & 1 \end{pmatrix}.$$

2. If  $A$  has coefficients  $a_{i,j}$ , then the coefficients of  $A^T$  are  $a_{j,i}$ , thus  $(A + B)^T$  has coefficients  $a_{j,i} + b_{j,i}$  and  $(A + B)^T = A^T + B^T$  as needed.

**Exercise 65.** Compute

$$2A + BC + B^2 + AD$$

where

$$A = \begin{pmatrix} 2 & 0 \\ -1 & 1 \end{pmatrix}, B = \begin{pmatrix} 1 & -1 \\ 0 & 2 \end{pmatrix}, C = \begin{pmatrix} -1 & 0 \\ 2 & 1 \end{pmatrix}$$

are real matrices and  $D = I_2$  is the 2-dimensional identity matrix.

*Solution.* First we notice that  $AD = A$ , thus  $2A + AD = 3A$ , and we are left to compute

$$3A + BC + B^2.$$

Then

$$3A = \begin{pmatrix} 6 & 0 \\ -3 & 3 \end{pmatrix},$$

and

$$BC = \begin{pmatrix} 1 & -1 \\ 0 & 2 \end{pmatrix} \begin{pmatrix} -1 & 0 \\ 2 & 1 \end{pmatrix} = \begin{pmatrix} -3 & -1 \\ 4 & 2 \end{pmatrix}$$

and finally

$$B^2 = \begin{pmatrix} 1 & -1 \\ 0 & 2 \end{pmatrix} \begin{pmatrix} 1 & -1 \\ 0 & 2 \end{pmatrix} = \begin{pmatrix} 1 & -3 \\ 0 & 4 \end{pmatrix}$$

therefore

$$3A + BC + B^2 = \begin{pmatrix} 6 & 0 \\ -3 & 3 \end{pmatrix} + \begin{pmatrix} -3 & -1 \\ 4 & 2 \end{pmatrix} + \begin{pmatrix} 1 & -3 \\ 0 & 4 \end{pmatrix} = \begin{pmatrix} 4 & -4 \\ 1 & 9 \end{pmatrix}$$

**Exercise 66.** Consider the complex matrix

$$A = \begin{pmatrix} 0 & i \\ -1 & 0 \end{pmatrix},$$

where  $i = \sqrt{-1}$ . What is  $A^l$ , for  $l \geq 1$ .

*Solution.* We have

$$A^2 = \begin{pmatrix} 0 & i \\ -1 & 0 \end{pmatrix} \begin{pmatrix} 0 & i \\ -1 & 0 \end{pmatrix} = \begin{pmatrix} -i & 0 \\ 0 & -i \end{pmatrix} = -iI_2$$

where  $I_2$  is the identity matrix. Thus

$$A^3 = -i \begin{pmatrix} 0 & i \\ -1 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ i & 0 \end{pmatrix}$$

and

$$A^4 = (A^2)^2 = (-iI_2)^2 = -I_2.$$

Next

$$A^5 = -A, \quad A^6 = -A^2 = iI_2, \quad A^7 = -A^3 = iA, \quad A^8 = (A^4)^2 = I_2.$$

This shows that  $A^l$  is decided by  $l \pmod{8}$ , since we may write  $l = l' + 8k$  for some integers  $l', k$ , and

$$A^l = A^{l'+8k} = A^{l'}(A^8)^k = A^{l'}$$

so we conclude that

$$A^l = \begin{cases} A & l \equiv 1 \pmod{8} \\ -iI_2 & l \equiv 2 \pmod{8} \\ -iA & l \equiv 3 \pmod{8} \\ -I_2 & l \equiv 4 \pmod{8} \\ -A & l \equiv 5 \pmod{8} \\ iI_2 & l \equiv 6 \pmod{8} \\ iA & l \equiv 7 \pmod{8} \\ I_2 & l \equiv 0 \pmod{8} \end{cases}$$

**Exercise 67.** 1. Let  $S$  be the set of  $3 \times 3$  diagonal real matrices. Is  $S$  closed under matrix addition?

2. Consider the real matrix

$$A = \begin{pmatrix} 2 & 3 \\ -1 & 1 \end{pmatrix}.$$

Compute a matrix  $B$  such that  $A + B$  is a diagonal, and a matrix  $C$  such that  $AC$  is diagonal.

*Solution.* 1. Take two diagonal matrices  $D$  and  $E$  in  $S$

$$D = \begin{pmatrix} d_{1,1} & 0 & 0 \\ 0 & d_{2,2} & 0 \\ 0 & 0 & d_{3,3} \end{pmatrix}, \quad E = \begin{pmatrix} e_{1,1} & 0 & 0 \\ 0 & e_{2,2} & 0 \\ 0 & 0 & e_{3,3} \end{pmatrix}$$

and compute their sum:

$$D + E = \begin{pmatrix} d_{1,1} + e_{1,1} & 0 & 0 \\ 0 & d_{2,2} + e_{2,2} & 0 \\ 0 & 0 & d_{3,3} + e_{3,3} \end{pmatrix}$$

thus the sum  $D + E$  belongs to  $S$  and  $S$  is closed under matrix addition.

2. We want a matrix  $B$  such that

$$A + B = \begin{pmatrix} 2 & 3 \\ -1 & 1 \end{pmatrix} + \begin{pmatrix} b_{1,1} & b_{1,2} \\ b_{2,1} & b_{2,2} \end{pmatrix} = \begin{pmatrix} 2 + b_{1,1} & 3 + b_{1,2} \\ -1 + b_{2,1} & 1 + b_{2,2} \end{pmatrix}$$

is diagonal. Therefore we need  $b_{1,2} = -3$  and  $b_{2,1} = 1$ , then such a matrix  $B$  will work, independently of the choice of  $b_{1,1}$  and  $b_{2,2}$ . Then we want a matrix  $C$  such that

$$AC = \begin{pmatrix} 2 & 3 \\ -1 & 1 \end{pmatrix} \begin{pmatrix} c_{1,1} & c_{1,2} \\ c_{2,1} & c_{2,2} \end{pmatrix} = \begin{pmatrix} 2c_{1,1} + 3c_{2,1} & 2c_{1,2} + 3c_{2,2} \\ -c_{1,1} + c_{2,1} & -c_{1,2} + c_{2,2} \end{pmatrix}$$

is diagonal. For example, we can take  $c_{1,2} = -3$  and  $c_{2,2} = 2$ , and  $c_{2,1} = c_{1,1}$ .

**Exercise 68.** Let  $A$  and  $B$  be  $n \times n$  matrices which satisfy

$$A^2 + AB + A - I_n = 0,$$

where  $I_n$  means the  $n \times n$  identity matrix, and  $0$  the  $n \times n$  zero matrix. Show that  $A$  is invertible.

*Solution.* We can rewrite  $A^2 + AB + A - I_n = 0$  as

$$A(A + B + I_n) = I_n$$

therefore  $A$  is invertible with inverse  $A + B + I_n$ .

**Exercise 69.** Compute, if it exists, the inverse  $A^{-1}$  of the matrix  $A$ , where  $A$  is given by

•

$$A = \begin{pmatrix} 2 & 3 & -2 \\ -1 & 1 & 2 \\ 3 & 7 & 2 \end{pmatrix}$$

for  $A$  a real matrix.

•

$$A = \begin{pmatrix} 1 & 1+i \\ 1-i & 1 \end{pmatrix}$$

for  $A$  a complex matrix and  $i = \sqrt{-1}$ .

•

$$A = \begin{pmatrix} 2 & 3 \\ 1 & 1 \end{pmatrix}$$

for  $A$  a matrix with coefficients modulo 5.

*Solution.*

- We compute the row echelon form of the augmented matrix

$$A = \left( \begin{array}{cccccc} 2 & 3 & -2 & 1 & 0 & 0 \\ -1 & 1 & 2 & 0 & 1 & 0 \\ 3 & 7 & 2 & 0 & 0 & 1 \end{array} \right).$$

Replace the 3rd row by (row 3)+3(row 2)=(3, 7, 2, 0, 0, 1)+3(-1, 1, 2, 0, 1, 0) which is(3, 7, 2, 0, 0, 1)+(-3, 3, 6, 0, 3, 0)=(0, 10, 8, 0, 3, 1), to get

$$\left( \begin{array}{cccccc} 2 & 3 & -2 & 1 & 0 & 0 \\ -1 & 1 & 2 & 0 & 1 & 0 \\ 0 & 10 & 8 & 0 & 3 & 1 \end{array} \right).$$

Next replace the 1st row by (row 1)+2(row 2)=(2, 3, -2, 1, 0, 0)+2(-1, 1, 2, 0, 1, 0) which is(2, 3, -2, 1, 0, 0)+(-2, 2, 4, 0, 2, 0)=(0, 5, 2, 1, 2, 0), after which we switch row 1 and 2 to get

$$\left( \begin{array}{cccccc} -1 & 1 & 2 & 0 & 1 & 0 \\ 0 & 5 & 2 & 1 & 2 & 0 \\ 0 & 10 & 8 & 0 & 3 & 1 \end{array} \right).$$

Next replace the 3rd row by  $-2(\text{row 2}) + (\text{row 3}) = (0, -10, -4, -2, -4, 0) + (0, 10, 8, 0, 3, 1)$  which is  $(0, 0, 4, -2, -1, 1)$ , to get

$$\begin{pmatrix} -1 & 1 & 2 & 0 & 1 & 0 \\ 0 & 5 & 2 & 1 & 2 & 0 \\ 0 & 0 & 4 & -2 & -1 & 1 \end{pmatrix}.$$

Therefore we can already tell that this matrix is invertible. Next to find its inverse, let us compute the reduced row echelon form. Replace row 2 by  $2(\text{row 2}) - (\text{row 3}) = (0, 10, 4, 2, 4, 0) - (0, 0, 4, -2, -1, 1)$  which is  $(0, 10, 0, 4, 5, -1)$ . We get

$$\begin{pmatrix} -1 & 1 & 2 & 0 & 1 & 0 \\ 0 & 10 & 0 & 4 & 5 & -1 \\ 0 & 0 & 4 & -2 & -1 & 1 \end{pmatrix}.$$

Next replace (row 1) by  $2(\text{row 1}) - (\text{row 3}) = (-2, 2, 4, 0, 2, 0) + (0, 0, -4, 2, 1, -1) = (-2, 2, 0, 2, 3, -1)$  to get

$$\begin{pmatrix} -2 & 2 & 0 & 2 & 3 & -1 \\ 0 & 10 & 0 & 4 & 5 & -1 \\ 0 & 0 & 4 & -2 & -1 & 1 \end{pmatrix}.$$

Next we replace (row 1) by  $-5(\text{row 1}) + (\text{row 2}) = (10, -10, 0, -10, -15, 5) + (0, 10, 0, 4, 5, -1)$  which is equal to  $(10, 0, 0, -6, -10, 4)$

$$\begin{pmatrix} 10 & 0 & 0 & -6 & -10 & 4 \\ 0 & 10 & 0 & 4 & 5 & -1 \\ 0 & 0 & 4 & -2 & -1 & 1 \end{pmatrix}.$$

We are now left by normalizing the diagonal coefficients to get 1:

$$\begin{pmatrix} 1 & 0 & 0 & -6/10 & -1 & 4/10 \\ 0 & 1 & 0 & 4/10 & 5/10 & -1/10 \\ 0 & 0 & 1 & -1/2 & -1/4 & 1/4 \end{pmatrix}.$$

This gives us  $A^{-1}$ :

$$A^{-1} = \begin{pmatrix} -6/10 & -1 & 4/10 \\ 4/10 & 5/10 & -1/10 \\ -1/2 & -1/4 & 1/4 \end{pmatrix}.$$

It is always good to compute  $AA^{-1}$  to make sure the answer is correct!

- We compute the row echelon form of

$$A = \begin{pmatrix} 1 & 1+i & 1 & 0 \\ 1-i & 1 & 0 & 1 \end{pmatrix}.$$

We replace (row 2) by (row 2)-(1-i)(row 1)=(1-i, 1, 0, 1)-(1-i, 2, 1-i, 0)=(0, -1, -1+i, 1), to get

$$\begin{pmatrix} 1 & 1+i & 1 & 0 \\ 0 & -1 & -1+i & 1 \end{pmatrix}.$$

Thus this matrix is invertible, and we compute its reduced row echelon form. We replace (row 1) by (row 1)+(1+i)(row 2)=(1, 1+i, 1, 0)+(0, -(1+i), -(1+i)(1-i), 1+i) to get

$$\begin{pmatrix} 1 & 0 & -1 & 1+i \\ 0 & -1 & -1+i & 1 \end{pmatrix}.$$

Finally we multiply the second row by -1:

$$\begin{pmatrix} 1 & 0 & -1 & 1+i \\ 0 & 1 & 1-i & -1 \end{pmatrix}.$$

This gives

$$A^{-1} = \begin{pmatrix} -1 & 1+i \\ 1-i & -1 \end{pmatrix}.$$

- We compute the row echelon form of

$$A = \begin{pmatrix} 2 & 3 & 1 & 0 \\ 1 & 1 & 0 & 1 \end{pmatrix}.$$

We replace (row 1) by (row 1)-2(row 2)=(2, 3, 1, 0)-2(1, 1, 0, 1)=(0, 1, 1, -2), and switch rows to get

$$A = \begin{pmatrix} 1 & 1 & 0 & 1 \\ 0 & 1 & 1 & 3 \end{pmatrix}.$$

This matrix is thus invertible, and we compute its reduced row echelon form. Replace row 1 by (row 1)-(row 2)=(1, 0, -1, -2) to get

$$A = \begin{pmatrix} 1 & 0 & 4 & 3 \\ 0 & 1 & 1 & 3 \end{pmatrix}.$$

Therefore

$$A^{-1} = \begin{pmatrix} 4 & 3 \\ 1 & 3 \end{pmatrix}.$$

**Exercise 70.** Write the following system of linear equations in a matrix form and solve it.

$$\begin{cases} x_1 + x_2 - 2x_3 = 1 \\ 2x_1 - 3x_2 + x_3 = -8 \\ 3x_1 + x_2 + 4x_3 = 7 \end{cases}$$

*Solution.* In matrix form, we get

$$\begin{pmatrix} 1 & 1 & -2 \\ 2 & -3 & 1 \\ 3 & 1 & 4 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 1 \\ -8 \\ 7 \end{pmatrix}.$$

We compute the row echelon form of the matrix of the system, augmented by the vector  $(1, -8, 7)$ : Replace row 2 by (row 2)-2(row 1) and row 3 by (row 3)-3(row 1):

$$\begin{pmatrix} 1 & 1 & -2 & 1 \\ 0 & -5 & 5 & -10 \\ 0 & -2 & 10 & 4 \end{pmatrix}$$

Replace row 3 by 5(row 3)-2(row 2):

$$\begin{pmatrix} 1 & 1 & -2 & 1 \\ 0 & -5 & 5 & -10 \\ 0 & 0 & 40 & 40 \end{pmatrix}$$

Divide the last row by 40, we already deduce that  $x_3 = 1$ , and after dividing the second row by -5, we get

$$\begin{pmatrix} 1 & 1 & -2 & 1 \\ 0 & 1 & -1 & 2 \\ 0 & 0 & 1 & 1 \end{pmatrix}$$

where we replace row 2 by (row 2)+(row 3):

$$\begin{pmatrix} 1 & 1 & -2 & 1 \\ 0 & 1 & 0 & 3 \\ 0 & 0 & 1 & 1 \end{pmatrix}$$

from which we get that  $x_2 = 3$ . Finally, by replacing row 1 by (row 1)+2(row 3), and then (row 1) again by (row 1)-(row 2) we get

$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 3 \\ 0 & 0 & 1 & 1 \end{pmatrix}$$

and  $x_1 = 0$ . In summary, this system has a unique solution, given by

$$x_1 = 0, \quad x_2 = 3, \quad x_3 = 1.$$

**Exercise 71.** Write the following system of linear equations in a matrix form and solve it.

$$\begin{cases} x_1 - x_2 + x_3 - x_4 &= 2 \\ x_1 - x_2 + x_3 + x_4 &= 0 \\ 4x_1 - 4x_2 + 4x_3 &= 4 \\ -2x_1 + 2x_2 - 2x_3 + x_4 &= -3 \end{cases}$$

*Solution.* In matrix form, we have

$$\begin{pmatrix} 1 & -1 & 1 & -1 \\ 1 & -1 & 1 & 1 \\ 4 & -4 & 4 & 0 \\ -2 & 2 & -2 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} 2 \\ 0 \\ 4 \\ -3 \end{pmatrix}.$$

Next we compute the row echelon form of

$$\begin{pmatrix} 1 & -1 & 1 & -1 & 2 \\ 1 & -1 & 1 & 1 & 0 \\ 4 & -4 & 4 & 0 & 4 \\ -2 & 2 & -2 & 1 & -3 \end{pmatrix}$$

Replace (row 4) by (row 4)+2(row 1) and (row 3) by (row 3)-4(row 1):

$$\begin{pmatrix} 1 & -1 & 1 & -1 & 2 \\ 1 & -1 & 1 & 1 & 0 \\ 0 & 0 & 0 & 4 & -4 \\ 0 & 0 & 0 & -1 & 1 \end{pmatrix}$$

Then row 3 is a multiple of row 4, and replace row 2 by (row 2)-(row 1):

$$\begin{pmatrix} 1 & -1 & 1 & -1 & 2 \\ 0 & 0 & 0 & 2 & -2 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 & 1 \end{pmatrix}$$

showing the row 2 is a multiple of row 4, and thus our system reduces to

$$\begin{pmatrix} 1 & -1 & 1 & -1 & 2 \\ 0 & 0 & 0 & -1 & 1 \end{pmatrix}$$

where we replace row 1 by (row 1)-(row 2):

$$\begin{pmatrix} 1 & -1 & 1 & 0 & 1 \\ 0 & 0 & 0 & -1 & 1 \end{pmatrix}.$$

This tells us that

$$x_4 = -1, \quad x_1 = 1 + x_2 - x_3.$$

There are infinitely many solutions.

## Exercises for Chapter 9

**Exercise 72.** Consider the sets  $A = \{1, 2\}$ ,  $B = \{1, 2, 3\}$  and the relation  $(x, y) \in R \iff (x - y)$  is even. Compute the inverse relation  $R^{-1}$ . Compute its matrix representation.

*Solution.* The relation  $R$  is

$$(1, 1), (1, 3), (2, 2),$$

therefore the relation  $R^{-1}$  is

$$(1, 1), (3, 1), (2, 2).$$

Its matrix representation is obtained by representing  $B$  as rows, that is row 1 is  $b_1 = 1$ , row 2 is  $b_2 = 2$ , row 3 is  $b_3 = 3$ , while column 1 is  $a_1 = 1$  and column 2 is  $a_2 = 2$ :

$$\begin{pmatrix} T & F \\ F & T \\ T & F \end{pmatrix}$$

**Exercise 73.** Consider the sets  $A = \{2, 3, 4\}$ ,  $B = \{2, 6, 8\}$  and the relation  $(x, y) \in R \iff x \mid y$ . Compute the matrix of the inverse relation  $R^{-1}$ .

*Solution.* The relation  $R$  is

$$(2, 2), (2, 6), (2, 8), (3, 6), (4, 8)$$

thus the inverse relation  $R^{-1}$  is

$$(2, 2), (6, 2), (8, 2), (6, 3), (8, 4)$$

that is  $(x, y) \in R^{-1} \iff x$  is a multiple of  $y$ , and the corresponding matrix is

$$\begin{pmatrix} T & F & F \\ T & T & F \\ T & F & T \end{pmatrix}$$

**Exercise 74.** Let  $R$  be a relation from  $\mathbb{Z}$  to  $\mathbb{Z}$  defined by  $xRy \leftrightarrow 2|(x - y)$ . Show that if  $n$  is odd, then  $n$  is related to 1.

*Solution.* Any odd number  $n$  can be written of the form  $n = 2m + 1$  for some integer  $m$ . Therefore  $n - 1 = 2m$  which is divisible by 2 and  $n$  is related to 1.

**Exercise 75.** This exercise is about composing relations.

1. Consider the sets  $A = \{a_1, a_2\}$ ,  $B = \{b_1, b_2\}$ ,  $C = \{c_1, c_2, c_3\}$  with the following relations  $R$  from  $A$  to  $B$ , and  $S$  from  $B$  to  $C$ :

$$R = \{(a_1, b_1), (a_1, b_2)\}, \quad S = \{(b_1, c_1), (b_2, c_1), (b_1, c_3), (b_2, c_2)\}.$$

What is the matrix of  $R \circ S$ ?

2. In general, what is the matrix of  $R \circ S$ ?

*Solution.* 1. Let us write the matrices of  $R$  and  $S$  first:

$$\begin{pmatrix} T & T \\ F & F \end{pmatrix} \circ \begin{pmatrix} T & F & T \\ T & T & F \end{pmatrix}.$$

Next we have that  $(a, c) \in R \circ S$  whenever  $aRb \wedge bSc$  for some  $b \in B$ . So to know, for example, whether  $(a_1, c_1)$  is in  $R \circ S$ , we have to check if we can find a  $b_i$  such that  $(a_1, b_i) \wedge (b_i, c_1)$ , that is whether

$$[(a_1, b_1) \wedge (b_1, c_1)] \vee [(a_1, b_2) \wedge (b_2, c_1)]$$

is true. But  $(a_i, b_j)$  means that the coefficient  $r_{ij}$  of the matrix  $R$  is true, and similarly  $(b_i, a_j)$  means that the coefficient  $s_{ij}$  of the matrix  $S$  is true. So we may rephrase the coefficient of the 1st row, 1st column of the matrix of  $R \circ S$  as

$$(r_{11} \wedge s_{11}) \vee (r_{12} \wedge s_{21}).$$

Notice that this is almost like doing the scalar product of the first row of  $R$  with the first column of  $S$ , except that multiplication is replaced by  $\wedge$ , and addition by  $\vee$ . Therefore, we have that the matrix of  $R \circ S$  is

$$\begin{pmatrix} T & T \\ F & F \end{pmatrix} \circ \begin{pmatrix} T & F & T \\ T & T & F \end{pmatrix} = \begin{pmatrix} T & T & T \\ F & F & F \end{pmatrix}$$

2. In general, we have a relation  $R$  from  $A$  to  $B$ , and a relation  $S$  from  $B$  to  $C$ , where the size of the set  $B$  is  $n$ . Denote the coefficients of the matrix of the relation  $R$  by  $r_{ij}$ , and that of the matrix of the relation  $S$  by  $s_{ij}$ . Then the matrix of  $R \circ S$  will have coefficients  $t_{ij}$  given by

$$t_{ij} = (r_{i1} \wedge s_{1j}) \vee (r_{i2} \wedge s_{2j}) \vee \dots \vee (r_{in} \wedge s_{nj}).$$

**Exercise 76.** Consider the relation  $R$  on  $\mathbb{Z}$ , given by  $aRb \iff a - b$  divisible by  $n$ . Is it symmetric?

*Solution.* Yes it is symmetric. Suppose  $aRb$ , then  $a - b$  is divisible by  $n$ . Thus  $-(a - b) = b - a$  is divisible by  $n$ , and  $bRa$  holds.

**Exercise 77.** Consider a relation  $R$  on any set  $A$ . Show that  $R$  symmetric if and only if  $R = R^{-1}$ .

*Solution.* Consider a relation  $R$ . The relation  $R^{-1}$  is defined by pairs  $(y, x)$  such that  $(x, y) \in R$ . If  $R$  is symmetric, it has the property that  $(x, y) \Rightarrow (y, x)$ , therefore  $(y, x) \in R$  and  $R = R^{-1}$ . Conversely, if  $R = R^{-1}$ , then if  $(x, y) \in R$ , it must be that  $(y, x) \in R$  and  $R$  is symmetric.

**Exercise 78.** Consider the set  $A = \{a, b, c, d\}$  and the relation

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

Is this relation reflexive? symmetric? transitive?

*Solution.* It is reflexive since  $(a, a), (b, b), (c, c), (d, d) \in R$ . It is symmetric since  $(a, b), (b, a), (a, d), (d, a) \in R$ . It is not transitive, indeed,  $(b, a), (a, d) \in R$  but  $(b, d) \notin R$ .

**Exercise 79.** Consider the set  $A = \{0, 1, 2\}$  and the relation  $R = \{(0, 2), (1, 2), (2, 0)\}$ . Is  $R$  antisymmetric?

*Solution.* No, since  $(0, 2)$  and  $(2, 0)$  are in  $R$ , but  $2 \neq 0$ .

**Exercise 80.** Are symmetry and antisymmetry mutually exclusive?

*Solution.* There is no connection between symmetry and antisymmetry, so no they are not mutually exclusive. For example, the relation  $A = B$  is both symmetric and antisymmetric. Then the relation “ $A$  is brother of  $B$ ” is neither symmetric (if  $A$  is a brother of  $B$ , it could be that  $B$  is a sister of  $A$ ) nor antisymmetric.

**Exercise 81.** Consider the relation  $R$  given by divisibility on positive integers, that is  $xRy \leftrightarrow x|y$ . Is this relation reflexive? symmetric? antisymmetric? transitive? What if the relation  $R$  is now defined over non-zero integers instead?

*Solution.* It is reflexive since  $x|x$  always. It is not symmetric, since for example  $1|y$  but  $y$  will never divide 1 if  $y > 1$ . It is antisymmetric, since if  $x|y$  then  $y = ax$  while if  $y|x$  then  $x = by$  and it must be that  $y = ax = a(by) = aby$  and  $a = b = 1$ . It is transitive, since  $x|y$  and  $y|z$  imply  $y = ax$ ,  $z = by$  thus  $z = by = b(ax)$  and  $x|z$ .

If we consider instead non-zero integers, the relation is not antisymmetric, indeed  $y = ax = a(by) = aby$  could imply  $a = b = -1$  in which case  $x|y$  and  $y|x$  when  $y = -x$  is possible.

**Exercise 82.** Consider the set  $A = \{0, 1, 2, 3, 4, 5, 6, 7, 8\}$ . Show that the relation  $xRy \leftrightarrow 2|(x - y)$  is an equivalence relation.

*Solution.* It is reflexive:  $2|(x - x)$ . It is symmetric: if  $2|(x - y)$  then  $(x - y) = 2n$  for some integer  $n$ , and thus  $(y - x) = -2n$  showing that  $2|(y - x)$ . It is transitive: if  $2|(x - y)$  and  $2|(y - z)$ , then  $(x - y) = 2n$ , and  $(y - z) = 2m$ , for some integers  $m, n$ . Therefore  $x - z = (x - y) + (y - z) = 2n + 2m = 2(n + m)$  and  $2|(x - z)$ .

**Exercise 83.** Show that given a set  $A$  and an equivalence relation  $R$  on  $A$ , then the equivalence classes of  $R$  partition  $A$ .

*Solution.* Let  $a, b \in A$ , and  $[a], [b]$  denote their equivalence classes. It is possible that  $[a] = [b]$ . Suppose that this is not the case. Then we will show that  $[a]$  and  $[b]$  are disjoint. Suppose by contradiction that there exists one element  $c \in [a] \cap [b]$ . Then  $aRc$  and  $bRc$ . But  $R$  is an equivalence relation, therefore it is symmetric (and  $cRb$ ) and transitive, implying that  $aRb$ . But then  $b \in [a]$  and  $a \in [b]$  by symmetry, and it must be that  $[a] = [b]$ . Indeed:

we show that  $[a] \subseteq [b]$  and  $[b] \subseteq [a]$ . Take an element  $x$  of  $[a]$ , then  $aRx$ , that is  $xRa$  (symmetry), and since  $aRb$  ( $b$  is in  $[a]$ ), it must be that  $xRb$  (transitivity) and thus  $bRx$  (symmetry again), which shows that  $x$  is in  $[b]$ . The same reasoning will show that  $[b]$  belongs to  $[a]$ .

Since either  $[a] = [b]$  or they are disjoint, take the union of the classes  $[a]$  that give distinct classes, and this gives a partition of  $A$ .

**Exercise 84.** Consider the set  $A = \{2, 3, 4, 5, 6, 7, 8, 9, 10\}$  and the relation

$$xRy \leftrightarrow \exists c \in \mathbb{Z}, y = cx.$$

Is  $R$  an equivalence relation? is  $R$  a partial order?

*Solution.*  $R$  is reflexive:  $xRx \leftrightarrow \exists c \in \mathbb{Z}, x = cx$ , take  $c = 1$ .  $R$  is not symmetric:  $xRy$  means  $x = cy$ , but then  $y = \frac{x}{c}$  so apart if  $c = \pm 1$ ,  $\frac{1}{c}$  will not be in  $\mathbb{Z}$ . For example,  $2R4$  since  $4 = c2$  with  $c = 2$ , but  $2 = c4$  means that  $c$  cannot be an integer. We conclude that  $R$  cannot be an equivalence relation.

Let us check antisymmetry and transitivity. Suppose  $y = cx$  and  $x = c'y$ , then  $x = c'cx$  and  $c'c = 1$ . So either  $c = c' = -1$ , which cannot happen because all elements of  $A$  are positive, or  $c = c' = 1$ , and the relation is antisymmetric. For transitivity, suppose  $xRy \iff y = cx$ ,  $yRz \iff z = c'y$ . Then  $z = c'y = c'cx$  with  $cc' \in \mathbb{Z}$  thus  $xRz$  as needed. We conclude that  $R$  is a partial order.

## Exercises for Chapter 10

**Exercise 85.** Consider the set  $A = \{a, b, c\}$  with power set  $P(A)$  and  $\cap : P(A) \times P(A) \rightarrow P(A)$ . What is its domain? its co-domain? its range? What is the cardinality of the pre-image of  $\{a\}$ ?

*Solution.* Its domain is the cartesian product  $P(A) \times P(A)$ , its co-domain is  $P(A)$ . Its range is  $P(A)$ : indeed, for any subset  $X$  of  $A$ ,  $X \cap X = X$ , therefore every element of  $P(A)$  has a pre-image. The pre-image of  $\{a\}$  is the set of elements in  $P(A) \times P(A)$  which are mapped to  $\{a\}$ , that is, pairs  $(X, Y)$  of subsets of  $A$  whose intersection is  $\{a\}$ . Now  $\{\{a\}\}, \{\{a\}, \{a, b\}\}, \{\{a\}, \{a, c\}\}, \{\{a\}, \{a, b, c\}\}$  are all the subsets containing  $\{a\}$ , so this gives  $2^4$  possible pairs, but among them, not all are suitable: we have to remove those with bigger intersection. So we can intersect  $\{a\}$  with all of them:

$$(\{\{a\}\}, \{\{a\}\}), (\{\{a\}\}, \{\{a, b\}\}), (\{\{a\}\}, \{\{a, c\}\}), (\{\{a\}\}, \{\{a, b, c\}\}),$$

or  $\{\{a, c\}\}$  with  $\{\{a, b\}\}$ . Note that the ordering of a pair matters, thus all those pairs give rise to another pair, apart for  $(\{\{a\}\}, \{\{a\}\})$  thus a total of 9.

**Exercise 86.** Show that  $\sin : \mathbb{R} \rightarrow \mathbb{R}$  is not one-to-one.

*Solution.* We have that  $\sin(0) = \sin(\pi) = 0$  but  $\pi \neq 0$ , which contradicts the definition of one-to-one, since there exist  $x_1 = 0, x_2 = \pi$  such that  $\sin(x_1) = \sin(x_2)$  but  $x_1 \neq x_2$ .

**Exercise 87.** Show that  $\sin : \mathbb{R} \rightarrow \mathbb{R}$  is not onto, but  $\sin : \mathbb{R} \rightarrow [-1, 1]$  is.

*Solution.* It is not onto because  $\exists y \in \mathbb{R}$ , say  $y = 2$ , such that for all  $x \in \mathbb{R}$ ,  $f(x) \neq 2$ .

**Exercise 88.** Is  $h : \mathbb{Z} \rightarrow \mathbb{Z}$ ,  $h(n) = 4n - 1$ , onto (surjective)?

*Solution.* No, it is not. For example, take  $y = 1$ . Then it is not possible that  $1 = 4n - 1$  for  $n$  an integer, because this equation means that  $n = 1/2$ .

**Exercise 89.** Is  $f : \mathbb{R} \rightarrow \mathbb{R}$ ,  $f(x) = x^3$ , a bijection (one-to-one correspondence)?

*Solution.* Injectivity: suppose  $f(x_1) = f(x_2)$ , then  $x_1^3 = x_2^3$  and it must be that  $x_1 = x_2$ . Surjectivity: take  $y \in \mathbb{R}$ , and  $x = \sqrt[3]{y} \in \mathbb{R}$ , then  $f(x) = y$ , so surjectivity holds. Therefore it is a bijection.

**Exercise 90.** Consider  $f : \mathbb{R} \rightarrow \mathbb{R}$ ,  $f(x) = x^2$  and  $g : \mathbb{R} \rightarrow \mathbb{R}$ ,  $g(x) = x + 5$ . What is  $g \circ f$ ? What is  $f \circ g$ ?

*Solution.* We have

$$g(f(x)) = g(x^2) = x^2 + 5, \quad f(g(x)) = f(x + 5) = (x + 5)^2.$$

**Exercise 91.** Consider  $f : \mathbb{Z} \rightarrow \mathbb{Z}$ ,  $f(n) = n + 1$  and  $g : \mathbb{Z} \rightarrow \mathbb{Z}$ ,  $g(n) = n^2$ . What is  $g \circ f$ ? What is  $f \circ g$ ?

*Solution.* We have

$$g(f(n)) = g(n + 1) = (n + 1)^2, \quad f(g(n)) = f(n^2) = n^2 + 1.$$

**Exercise 92.** Given two functions  $f : X \rightarrow Y$ ,  $g : Y \rightarrow Z$ . If  $g \circ f : X \rightarrow Z$  is one-to-one, must both  $f$  and  $g$  be one-to-one? Prove or give a counter-example.

*Solution.* It is not true. For example, take  $f : X \rightarrow Y$  and  $g : Y \rightarrow Z$  as follows,  $X = \{a, b, c\}$ ,  $Y = \{w, x, y, z\}$ ,  $Z = \{1, 2, 3\}$ :

$$f(a) = x, \quad f(b) = y, \quad f(c) = z, \quad g(w) = 1, \quad g(x) = 1, \quad g(y) = 2, \quad g(z) = 3.$$

Then  $g \circ f$  is one-to-one, but  $g$  is not.

**Exercise 93.** Show that if  $f : X \rightarrow Y$  is invertible with inverse function  $f^{-1} : Y \rightarrow X$ , then  $f^{-1} \circ f = i_X$  and  $f \circ f^{-1} = i_Y$ .

*Solution.* Take  $x \in X$ , with  $y = f(x)$ . Then  $f^{-1}(f(x)) = f^{-1}(y) = x$  by definition of inverse, and  $x = i_X(x)$  for all  $x \in X$  therefore  $f^{-1} \circ f = i_X$ . Similarly take  $y \in Y$  and  $x = f^{-1}(y)$ . Then  $f(f^{-1}(y)) = f(x) = y$  by definition of inverse, and  $y = i_Y(y)$  for all  $y \in Y$  therefore  $f \circ f^{-1} = i_Y$ .

**Exercise 94.** If you pick five cards from a deck of 52 cards, prove that at least two will be of the same suit.

*Solution.* If you pick 5 cards, then the first one will be of a given suit (say heart), if the second is also heart, then you got two of the same suit. If the second is not heart (say diamond), then take a 3rd. If it is either heart or diamond, then you got at least two of the same suit, if not, say it is club, pick a 4th card. Again, if the 4th card is heart, diamond or club, you got at least two of the same suit, if not, it must be that this 4th card is spade.

But now all the 4 possible choices of suits are picked, so no matter what is the next suit of the card, it will be one that has already appeared. This shows that you will always get at least two cards of the same suit. This is an application of the pigeonhole principle: you have 4 suits, and 5 cards, therefore 2 cards must be of the same suit.

**Exercise 95.** If you have 10 black socks and 10 white socks, and you are picking socks randomly, you will only need to pick three to find a matching pair.

*Solution.* Pick the first sock, it is say white. Pick the second sock, if it white, then you got a matching pair. If not, pick a third one. But by now, you have already one white and one black sock, so no matter which is the color of the third one, ou will have a matching pair. This is an application of the pigeonhole principle: you have 2 colors, and 3 socks, therefore 2 socks must be of the same color.

## Exercises for Chapter 11

**Exercise 96.** Prove that if a connected graph  $G$  has exactly two vertices which have odd degree, then it contains an Euler path.

*Solution.* Suppose that  $v$  and  $w$  are the vertices of  $G$  which have odd degrees, while all the other vertices have an even degree. Create a new graph  $G'$ , formed by  $G$ , with one more edge  $e$ , which connects  $v$  and  $w$ . Every vertex in  $G'$  has even degree, so by the theorem on Euler cycles, there is an Euler cycle. Say this Euler cycle is

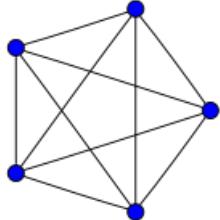
$$v, e_1, v_2, e_2, \dots, w, e, v$$

then

$$v, e_1, v_2, e_2, \dots, w$$

is an Euler path.

**Exercise 97.** Draw a complete graph with 5 vertices.



*Solution.*

**Exercise 98.** Show that in every graph  $G$ , the number of vertices of odd degree is even.

*Solution.* Let  $E$  denote the set of edges, and write the set  $V$  of vertices as  $V' \cup V''$  where  $V'$  is the set of nodes with odd degrees, and  $V''$  is the set of nodes with even degrees. Suppose that the number of vertices of odd degree is odd, then

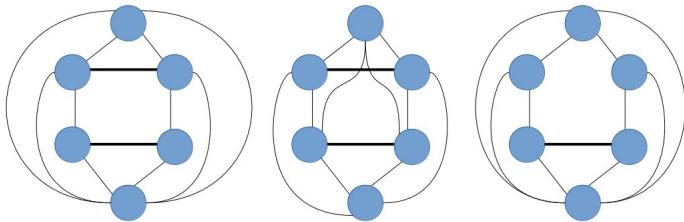
$$2|E| = \sum_{v \in V} \deg(v) = \sum_{v \in V'} + \sum_{v \in V''}$$

where the first sum (over  $V'$ ) is odd and the second sum is even, a contradiction.

**Exercise 99.** Show that in every simple graph (with at least two vertices), there must be two vertices that have the same degree.

*Solution.* Suppose there are  $n$  nodes. If all degrees are different, they must be exactly  $0, 1, \dots, n - 1$ , which is impossible: one cannot have one node of degree 0, yet another one with degree  $n - 1$ !

**Exercise 100.** Decide whether the following graphs contain a Euler path/cycle.



*Solution.* The first graph (left hand side) contains a Euler path and no Euler circuit, the middle graph contains a Euler circuit, the third one contains none!

# Index

- $n$  divides  $m$ , 9
- $n$  is divisible by  $m$ , 9
- addition table, 15
- adjacency matrix, 263
- adjacent, 251
- antisymmetric, 223
- argument, 53, 165
- assumption, hypothesis, 53
- biconditional, 49
- bijection, 237
- binary relation, 209
- bipartite graph, 259
- cardinality, 111
- Cartesian product, 115
- characteristic equation, 149
- closed under  $\Delta$ , 17
- closure, 17
- codomain, 231
- column vector, 175
- combination, 131
- commutative, 37
- complement of  $A$ , 115
- complement of  $B$  with respect to  $A$ , 115
- complete graph, 259
- complex number, 163
- composition, 215, 241
- compound proposition, 35
- conclusion, 53
- conditional statement, 45
- congruent  $(\text{mod } n)$ , 13
- conjugate, 163
- contradiction, 41
- contrapositive, 47
- converse, 47
- converse error, 59
- counter-example, 57
- critical rows, 55
- De Morgan's Law, 37
- degree of a node, 257
- diagonal, 181
- difference of  $A$  and  $B$ , 115
- directed graph, 253
- directed multigraph, 253
- disjoint, 115
- domain, 79, 231
- elementary row operations, 185
- empty set, 109
- equivalence class, 219
- equivalence relation, 219
- equivalent proposition, 37
- Euclidean division, 9
- Euler circuit/cycle, 253
- Euler Formula, 167
- Euler identity, 169
- Euler path/trail, 253
- even number, 11

- event, 135
- existential quantification, 83
- fallacy, 59
- function, 231
- graph, 251
- graph isomorphism, 265
- Hamiltonian circuit, 263
- Hamiltonian path, 263
- Handshaking Theorem, 263
- homogeneous, 191
- identity function, 237
- identity matrix , 181
- if and only if, 49
- image, 231
- image of a subset, 233
- imaginary number  $z$ , 161
- imaginary part, 163
- imaginary unit  $i$ , 161
- incident, 251
- injective, 233
- integer number, 5
- integers  $(\text{mod } n)$ , 13
- intersection, 115
- inverse, 47
- inverse error, 59
- inverse function, 237
- inverse relations, 211
- invertible, 183
- irrational numbers, 7
- liar paradox, 35
- linear homogeneous relation, 149
- logical equivalence (law), 43
- loop, 251
- matrix, 175
- modulus, 165
- multigraph, 253
- multiplication table, 15
- natural number, 5
- necessary condition, 49
- odd number, 11
- one-to-one, 233
- one-to-one correspondence, 237
- only if, 47
- onto, 237
- paradox, 35
- partial order, 225
- partition, 117
- permutation, 131
- pigeonhole principle, 247
- polar coordinates, 165
- power set, 111
- Predicate, 77
- predicate logic, 77
- preimage, 231
- premise, 53
- prime number, 11
- probability space, 137
- Proposition, 33
- proposition logic, 33
- quotient, 9
- range, 231
- rank, 195
- rational numbers, 7
- real numbers, 7
- real part, 163
- reduced row echelon form, 187
- reflexive, 217
- remainder, 9

root of unity, 171  
row echelon form, 183  
row vector, 175  
rule of inference, 59  
  
sample space, 133  
scalar product, 179  
set, 107  
simple, 251  
subset, 109  
sufficient condition, 49  
surjective, 237  
symmetric, 217  
system of linear equations, 191  
  
tautology, 41  
total degree, 259  
transitive, 217  
transitive closure, 225  
transpose matrix, 177  
true by default, 45  
truth table, 35  
  
union, 115  
universal quantification, 79  
  
vacuously true, 45  
valid argument, 53  
  
whole numbers, 5

