Alexander Knop

Introduction to Discrete Mathematics

MARCH 2, 2020

Contents

Preface

> - Why is a math book so sad?
> - Because it's full of problems.
>
> Anonymous, Unknown

If you are reading this book, you probably have never studied proofs before. So let me give you some advice: mathematical books are very different from fiction, and even books in other sciences. Quite often you may see that some steps are missing, and some steps are not really explained and just claimed as obvious. The main reason behind this is to make the ideas of the proof more visible and to allow grasping the essence of proofs quickly.

Since the steps are skipped, you cannot just read the book and believe that you studied the topic; the best way to actually study the topic is to try to prove every statement before you read the actual proof in the book. In addition to this, I recommend trying to solve all the exercises in the book (you may find exercises in the middle and at the end of every chapter).

Additionally, many topics in this book have a corresponding five-minute video explaining the material of the chapter, it is useful to watch them before you go into the topic.

Organization

?? covers the basics of mathematics and provide the language we use in the next parts. We start from the explanation of what a mathematical proof is (in **??**). **??** shows how to prove theorems indirectly using proof by contradiction. **??** explains the most powerful method in our disposal, proof by induction. Finally, **????????** define several important objects such as sets, functions, and relations.

?? studies the basics of combinatorics, a branch of mathematics that answers the question "how many objects of this kind?". **??** gives a formal definition of "size" of a set and show how to compare sizes of two sets. **??** proves several simple principles that allow to find sizes of sets. In **??** we learn how to prove existence of an object with some properties using simple inequalities between sizes of sets. **??????** prove several properties of standard combinatorial objects. Finally, **??** provides a framework that helps to find sizes of sets in many cases.

?? returns back to proofs; however, instead of studying *how* to prove something we study what can we prove and how to define "proof" so that we can use computer to generate proofs and verify them.

In **??** we study basics of graph theory. **??** gives the definition of a graph and prove the one of the simplest and at the same time most important theorems in graph theory. In **??** we define what it means beeing connected and how to use this notion in real-life applications.

Finally, **??** defines a tree and show how to use these objects in computer networks.

Alexander Knop
San Diego, California, USA

Part I

Introduction to Mathematical Reasoning

1. Proofs

What is a Mathematical Proof:
Introduction to Mathematical Reasoning #1

`https://youtu.be/eJD0gGqveIE`

1.1 Direct Proofs

We start the discussion of the proofs in mathematics from an example of a proof in "everyday" life. Assume that we know that the following statements are true.

1. If a salmon has fins and scales it is kosher,
2. if a salmon has scales it has fins,
3. any salmon has scales.

Using these facts we may conclude that any salmon is kosher; indeed, any salmon has scales by the third statement, hence, by the second statement any salmon has fins, finally, by the first statement any salmon is kosher since it has fins and scales.

One may notice that this explanation is a sequence of conclusions such that each of them is true because the previous one is true. Mathematical proof is also a sequence of statements such that every statement is true if the previous statement is true. If P and Q are some statements and Q is always true when P is true, then we say that P implies Q. We denote the statement that P implies Q by $P \implies Q$.

In order to define the implication formally let us consider the following table.

P	Q	$P \implies Q$
T	T	T
T	F	F
F	T	T
F	F	T

Let P and Q be some statements. Then this table says that if P and Q are both false, then $P \implies Q$ is true etc.

Exercise 1.1. *Let n be an integer.*

1. *Is it always true that "n^2 is positive" implies "n is not equal to 0"?*

2. *Is it always true that "$n^2 - n - 2$ is equal to 0" implies "n is equal to 2"?*

In the example we gave at the beginning of the section we used some *known* facts. But what does it mean to know something? In math we typically say that we know a statement if we can prove it. But in order to prove this statement we need to know something again, which is a problem! In order to solve it, mathematicians introduced the notion of an *axiom*. An axiom is a statement that is believed to be true and when we prove a statement we prove it under the assumption that these axioms are true[1].

For example, we may consider axioms of inequalities for real numbers.

1. Let $a, b \in \mathbb{R}$. Only one of the following is true:
   - $a < b$,
   - $b < a$, or
   - $a = b$.
2. Let $a, b, c \in \mathbb{R}$. Then $a < b$ iff $a + c < b + c$ (iff is an abbreviation for "if and only if").
3. Let $a, b, c \in \mathbb{R}$. Then $a < b$ iff $ac < bc$ provided that $c > 0$ and $a < b$ iff $ac > bc$ if $c < 0$.
4. Let $a, b, c \in \mathbb{R}$. If $a < b$ and $b < c$, then $a < c$.

Let us now try to prove something using these axioms, we prove that if $a > 0$, then $a^2 > 0$. Note that $a > 0$, hence, by the third axiom $a^2 > 0$ (note that we also used an additional statement saying that $0 \cdot 0 = 0$).

Similarly, we may prove that if $a < 0$, then $a^2 > 0$. And combining these two statements together we may prove that if $a \neq 0$, then $a^2 > 0$.

Such a way of constructing proof is called direct proofs.

What We Know and How to Find a Proof: Introduction to Mathematical Reasoning #2

https://youtu.be/nBjJi6aTk2M

Exercise 1.2. *Axiomatic system for a four-point geometry.*

Undefined terms: point, line, is on.

Axioms:

- *For every pair of distinct points x and y, there is a unique line ℓ such that x is on l and y is on l.*
- *Given a line ℓ and a point x that is not on ℓ, there is a unique line m such that x is on m and no point on ℓ is also on m.*
- *There are exactly four points.*
- *It is impossible for three points to be on the same line.*

Prove that there are at least two distinct lines.

[1] Note that in different parts of math axioms may be different.

Let n and m be some integers. Using direct proofs we may prove the following two statements.

- if n is even, then nm is also even (a number ℓ is even if there is an integer k such that $\ell = 2k$),
- if n is even and m is even, then $n + m$ is also even.

We start from proving the first statement. There is an integer k such that $n = 2k$ since n is even. As a result, $nm = 2(nk)$ so nm is even.

Now we prove the second statement. Since n and m are even there are k and ℓ such that $n = 2k$ and $m = 2\ell$. Hence, $n + m = 2(k + \ell)$ so $n + m$ is even.

1.2 Constructing Proofs Backwards

However, sometimes it is not easy to find the proof. In this case one of the possible methods to deal with this problem is to try to prove starting from the end.

For example, we may consider the statement $(a + b)^2 = a^2 + 2ba + b^2$. Imagine, for a second, that you have not learned about axioms. In this case you would write something like this:

$$\begin{aligned}(a + b)^2 = (a + b) \cdot (a + b) = \\ a(a + b) + b(a + b) = \\ a^2 + ab + ba + b^2 = a^2 + 2ba + b^2.\end{aligned}$$

Let us try to prove it completely formally using the following axioms.

1. Let a, b, and c be reals. If $a = b$ and $b = c$, then $a = c$.
2. Let a, b, and c be reals. If $a = b$, then $a + c = b + c$ and $c + a = c + b$.
3. Let a, b, and c be reals. Then $a(b + c) = ab + ac$.
4. Let a and b be reals. Then $ab = ba$.
5. Let a and b be reals. Then $a + b = b + a$.
6. Let a be a real number. Then $a^2 = a \cdot a$ and $a \cdot a = a^2$.
7. Let a be a real number. Then $a + a = 2a$.

So the formal proof of the statement $(a + b)^2 = a^2 + 2ab + b^2$ is as follows. First note that $(a + b)^2 = (a + b) \cdot (a + b)$ (by axiom 6), hence, by axiom 1, it is enough to show that $(a + b) \cdot (a + b) = a^2 + 2ab + b^2$. By axiom 3, $(a + b) \cdot (a + b) = (a + b) \cdot a + (a + b) \cdot b$. Axiom 4 implies that $(a+b)\cdot a = a\cdot(a+b)$ and $(a+b)\cdot b = b\cdot(a+b)$ Hence, by axioms 1 and 2 applied twice

$$a\cdot(a+b)+b\cdot(a+b) = (a+b)\cdot a + b\cdot(a+b) = (a+b)\cdot a + (a+b)\cdot b.$$

As a result,

$$(a+b)\cdot(a+b) = (a+b)\cdot a + (a+b)\cdot b = \\ a\cdot(a+b)+b\cdot(a+b) = a\cdot a + a\cdot b + b\cdot a + b\cdot b;$$

so by axiom 1, it is enough to show that $a\cdot a + a\cdot b + b\cdot a + b\cdot b = a^2 + 2ab + b^2$. Additionally, by axiom 6, $a\cdot a = a^2$ and $b\cdot b = b^2$. Hence, by axiom 2, it is enough to show that $a^2 + a\cdot b + b\cdot a + b^2 = a^2 + 2ab + b^2$. By axiom 4, $a\cdot b = b\cdot a$, hence, by axiom 2, $a\cdot b + b\cdot a = b\cdot a + b\cdot a$. Therefore by axiom 7, $a\cdot b + b\cdot a = 2b\cdot a$. Finally, by axiom 2, $a\cdot b + b\cdot a + a^2 + b^2 = 2b\cdot a + a^2 + b^2$ and by axiom 5, $a\cdot b + b\cdot a + a^2 + b^2 = a^2 + a\cdot b + b\cdot a + b^2$ and $2b\cdot a + a^2 + b^2 = a^2 + 2b\cdot a + b^2$. Which finishes the proof by axiom 1.

1.3 Analysis of Simple Algorithms

We can use this knowledge to analyze simple algorithms. For example, let us consider the following algorithm. Let us prove that it is correct

Algorithm 1.1: The algorithm that finds the maximum element of a, b, c.

```
 1: function MAX($a, b, c$)
 2:     $r \leftarrow a$
 3:     if $b > r$ then
 4:         $r \leftarrow b$
 5:     end if
 6:     if $c > r$ then
 7:         $r \leftarrow c$
 8:     end if
 9:     return $r$
10: end function
```

i.e. it returns the maximum of a, b, and c. We need to consider the following cases.

- If the maximum is equal to a. In this case, at line 2, we set $r = a$, at line 3 the inequality $b > r$ is false (since $a = r$ is the maximum) and at line 6 the inequality $c > r$ is also false (since $a = r$ is the maximum). Hence, we do not change the value of r after line 2 and the returned value is a.

- If the maximum is equal to b. We set $r = a$ at line 2. The inequality $b > r$ at line 3 is true (since b is the maximum) and we set r to be equal to b. So at line 6, the inequality $c > r$ is false (since $b = r$ is the maximum). Hence, the returned value is b.

- If the maximum is equal to c. We set $r = a$ at line 2. If the inequality $b > r$ is true at line 3 we set r to be equal to b. So at line 6 the inequality $c > r$ is true (since c is the maximum). Hence, we set r being equal to c and the returned value is c.

1.4 Proofs in Real-life Mathematics

In this chapter we explicitly used axioms to prove statements. However, it leads us to really long and hard to understand proofs (the last example in the previous section is a good example of this phenomenon). Because of this mathematicians tend to skip steps in the proofs when they believe that they are clear. It is worth to mention a nice quotation of Scott Aaronson about this problem

> When mathematicians say that a theorem has been "proved," they still mean, as they always have, something more like: "we've reached a social consensus that all the ideas are now in place for a strictly formal proof that could be verified by a machine . . . with the only task remaining being massive rote coding work that none of us has any intention of ever doing!"

This is the reason why it is arduous to read mathematical texts and it is very different from reading non-mathematical books. A problem that arises because of this tendency is that some mistakes may happen if we skip way too many steps. In the last two centuries there were several attempts to solve this issue, one approach to this we are going to discuss in Part **??**.

Death of proof greatly exaggerated

`https://scottaaronson.com/blog/?p=4133`

End of The Chapter Exercises

1.3 Using the axioms of inequalities show that if a is a non-zero real number, then $a^2 > 0$.

1.4 Using the axioms of inequalities prove that for all real numbers a, b, and c,

$$bc + ac + ab \leq a^2 + b^2 + c^2.$$

1.5 *(recommended)* Prove that for all integers a, b, and c, If a divides b and b divides c, then a divides c. Recall that an integer m divides an integer n if there is an integer k such that $mk = n$.

1.6 *(recommended)* Show that square of an even integer is even.

1.7 Prove that 0 divides an integer a iff $a = 0$.

1.8 Using the axioms of inequalities, show that if $a > 0$, b, and c are real numbers, then $b \geq c$ implies that $ab \geq ac$.

1.9 Using the axioms of inequalities, show that if $a, b < 0$ are real numbers, then $a \leq b$ implies that $a^2 \geq b^2$.

2. Proofs by Contradiction

Proofs by Contradiction:
Introduction to Mathematical Reasoning #3

https://youtu.be/bWP0VYx75DI

2.1 Proving Negative Statements

The direct method is not very convenient when we need to prove a negation of some statement.

For example, we may try to prove that $78n + 102m = 11$ does not have integer solutions. It is not clear how to prove it directly since we can not consider all possible n and m. Hence, we need another approach. Let us assume that such a solution n, m exists. Note that $78n + 102m$ is even, but 11 is odd. In other words, an odd number is equal to an even number, it is impossible. Thus, the assumption was false.

Let us consider a more useful example, let us prove that if p^2 is even, then p is also even (p is an integer). Assume the opposite i.e. that p^2 is even but p is not. Let $p = 2b + 1$[1]. Note that $p^2 = (2b+1)^2 = 2(2b^2 + 2b) + 1$. Hence, p^2 is odd which contradicts to the assumption that p^2 is even.

Using this idea we may prove much more complicated results e.g. one may show that $\sqrt{2}$ is irrational. For the sake of contradiction, let us assume that it is not true. In other words there are p and q such that $\sqrt{2} = \frac{p}{q}$ and $\frac{p}{q}$ is an irreducible fraction.

Note that $\sqrt{2}q = p$, so $2q^2 = p^2$. Which implies that p is even and 4 divides p^2. Therefore 4 divides $2q^2$ and q is also even. As a result, we get a contradiction with the assumption that $\frac{p}{q}$ is an irreducible fraction.

Template for proving a statement by contradiction.

Assume, for the sake of contradiction, that *the statement* is false. Then *present some argument that leads to a contradiction*. Hence, the assumption is false and *the statement* is true.

Exercise 2.1. *Show that $\sqrt{3}$ is irrational.*

[1] Note that we use here the statement that an integer n is not even iff it is odd, which, formally speaking, should be proven.

2.2 *Proving Implications by Contradiction*

This method works especially well when we need to prove an implication. Since the implication $A \implies B$ is false only when A is true but B is false. Hence, you need to derive a contradiction from the fact that A is true and B is false.

We have already seen such examples in the previous section, we proved that p^2 is even implies p is even for any integer p. Let us consider another example. Let a and b be reals such that $a > b$. We need to show that $(ac < bc) \implies c < 0$. So we may assume that $ac < bc$ but $c \geq 0$. By the multiplicativity of the inequalities we know that if $(a > b)$ and $c > 0$, then $ac > bc$ which contradicts to $ac < bc$.

A special case of such a proof is when we need to prove the implication $A \implies B$, assume that B is false and derive that A is false which contradicts to A (such proofs are called proofs by contraposition); note that the previous proof is a proof of this form.

2.3 *Proof of "OR" Statements*

Another important case is when we need to prove that at least one of two statements is true. For example, let us prove that $ab = 0$ iff $a = 0$ or $b = 0$. We start from the implication from the right to the left. Since if $a = 0$, then $ab = 0$ and the same is true for $b = 0$ this implication is obvious.

The second part of the proof is the proof by contradiction. Assume $ab = 0$, $a \neq 0$, and $b \neq 0$. Note that $b = \frac{ab}{a} = 0$, hence $b = 0$ which is a contradiction to the assumption.

End of The Chapter Exercises

2.2 *(recommended)* Prove that if n^2 is odd, then n is odd.

2.3 In Euclidean (standard) geometry, prove: If two lines share a common perpendicular, then the lines are parallel.

2.4 *(recommended)* Let us consider four-lines geometry, it is a theory with undefined terms: point, line, is on, and axioms:

1. there exist exactly four lines,
2. any two distinct lines have exactly one point on both of them, and
3. each point is on exactly two lines.

Show that every line has exactly three points on it.

2.5 Let us consider group theory, it is a theory with undefined terms: group-element and times (if a and b are group elements, we denote a times b by $a \cdot b$), and axioms:

1. $(a \cdot b) \cdot c = a \cdot (b \cdot c)$ for every group-elements a, b, and c;
2. there is a unique group-element e such that $e \cdot a = a = a \cdot e$ for every group-element a (we say that such an element is the identity element);
3. for every group-element a there is a group-element b such that $a \cdot b = e$, where e is the identity element;
4. for every group-element a there is a group-element b such that $b \cdot a = e$, where e is the identity element.

Let e be the identity element. Show the following statements

- if $b_0 \cdot a = b_1 \cdot a = e$, then $b_0 = b_1$, for every group-elements a, b_0, and b_1.
- if $a \cdot b_0 = a \cdot b_1 = e$, then $b_0 = b_1$, for every group-elements a, b_0, and b_1.
- if $a \cdot b_0 = b_1 \cdot a = e$, then $b_0 = b_1$, for every group-elements a, b_0, and b_1.

2.6 Let us consider three-points geometry, it is a theory with undefined terms: point, line, is on, and axioms:

1. There exist exactly three points.
2. Two distinct points are on exactly one line.
3. Not all the three points are collinear i.e. they do not lay on the same line.
4. Two distinct lines are on at least one point i.e. there is at least one point such that it is on both lines.

Show that there are exactly three lines.

2.7 Show that there are irrational numbers a and b such that a^b is rational.

2.8 *(recommended)* Show that there does not exist the largest integer.

2.9 Let us consider Young's geometry, it is a theory with undefined terms: point, line, is on, and axioms:

1. there exists at least one line,
2. every line has exactly three points on it,
3. not all points are on the same line,

4. for two distinct points, there exists exactly one line on both of them,
5. if a point does not lie on a given line, then there exists exactly one line on that point that does not intersect the given line.

Show that for every point, there are exactly four lines on that point.

Solutions to The Exercises

2.6 Let us denote the points by p_1, p_2, and p_3 (they exist by Axiom 1). By Axiom 2, there are lines $l_{1,2}$, $l_{1,3}$, and $l_{2,3}$ such that p_i and p_j are on $l_{i,j}$ $(i \neq j)$.

Note that the lines $l_{1,2}$, $l_{1,3}$, and $l_{2,3}$ are different. Indeed, assume the opposite, i.e., without loss of generality that $l_{1,2} = l_{1,3}$. Note that p_1, p_2, and p_3 are on $l_{1,2}$ which contradicts Axiom 3.

Let us now prove that there are no other lines. Assume the opposite i.e. that there is another line l. There is a point that is on l and $l_{1,2}$. Without loss of generality, this point is p_1. Additionally there is a point p_i $(i \neq 1)$ that is on l and $l_{2,3}$. However, it means that p_1 and p_i are on l which contradicts Axiom 2.

3. Proofs by Induction

The Induction Principle:
Introduction to Mathematical Reasoning #4

https://youtu.be/jOnZTWGpX_I

3.1 Simple Induction

Let us consider a simple problem: what is bigger 2^n or n? In this chapter, we are going to study the simplest way to prove that $2^n > n$ for all positive integers n. First, let us check that it is true for small integers n.

	1	2	3	4	5	6	7	8
n	1	2	3	4	5	6	7	8
2^n	2	4	8	16	32	64	128	256

We may also note that 2^n is growing faster than n, so we expect that if $2^n > n$ for small integers n, then it is true for all positive integers n.

In order to prove this statement formally, we use the following principle.

Principle 3.1 (The Induction Principle)**.** *Let $P(n)$ be some statement about a positive integer n. Hence, $P(n)$ is true for every positive integer n iff*

(the base case) $P(1)$ *is true and*

(the induction step) $P(k) \implies P(k+1)$ *is true for all positive integers k.*

Let us prove now the statement using this principle. We define $P(n)$ be the statement that "$2^n > n$". $P(1)$ is true since $2^1 > 1$. Let us assume now that $2^n > n$. Note that $2^{n+1} = 2 \cdot 2^n > 2n \geq n + 1$. Hence, we proved the induction step.

Exercise 3.1. *Prove that $(1 + x)^n \geq 1 + nx$ for all positive integers n and real numbers $x \geq -1$.*

3.2 Changing the Base Case

Let us consider functions n^2 and 2^n.

	1	2	3	4	5	6	7	8
n^2	1	4	9	16	25	36	49	64
2^n	2	4	8	16	32	64	128	256

Note that 2^n is greater than n^2 starting from 5. But without some trick we can not prove this using induction since for $n = 3$ it is not true!

The trick is to use the statement $P(n)$ stating that $(n+4)^2 < 2^{n+4}$. The base case when $n = 1$ is true. Let us now prove the induction step. Assume that $P(k)$ is true i.e. $(k+4)^2 < 2^{k+4}$. Note that $2(k+4)^2 < 2^{k+1+4}$ but $(k+5)^2 = k^2 + 10k + 25 \leq 2k^2 + 16k + 32 = 2(k+4)^2$. Which implies that $2^{k+1+4} > (k+5)^2$. So $P(k+1)$ is also true.

In order to avoid this strange $+4$ we may change the base case and use the following argument.

Theorem 3.1. *Let $P(n)$ be some statement about an integer n. Hence, $P(n)$ is true for every integer $n > n_0$ iff*

(the base case) $P(n_0 + 1)$ *is true and*

(the induction step) $P(k) \implies P(k+1)$ *is true for all integers $k > n_0$.*

Using this generalized induction principle we may prove that $2^n \geq n^2$ for $n \geq 4$. The base case for $n = 4$ is true. The induction step is also true; indeed let $P(k)$ be true i.e. $(k+4)^2 < 2^{k+4}$. Hence, $2(k+4)^2 < 2^{k+1+4}$ but $(k+5)^2 = k^2 + 10k + 25 \leq 2k^2 + 16k + 32 = 2(k+4)^2$.

Let us now prove the theorem. Note that the proof is based on an idea similar to the trick with $+4$, we just used.

Proof of Theorem **??**. $\Rightarrow$ If $P(n)$ is true for any $n > n_0$ it is also true for $n = n_0 + 1$ which implies the base case. Additionally, it true for $n = k + 1$ so the induction step is also true.

$\Leftarrow$ In this direction the proof is a bit harder. Let us consider a statement $Q(n)$ saying that $P(n + n_0)$ is true. Note that by the base case for P, $Q(1)$ is true; by the induction step for P we know that $Q(n)$ implies $P(n+1)$. As a result, by the induction principle $Q(n)$ is true for all positive integers n. Which implies that $P(n)$ is true for all integers $n > n_0$.

□

3.3 *Inductive Definitions*

We may also define objects inductively. Let us consider the sum $1 + 2 + \cdots + n$ a line of dots indicating "and so on" which indicates the definition by induction. In this case, a more precise notation is $\sum_{i=1}^{n} i$.

Definition 3.1. *Let $a(1), \ldots, a(n), \ldots$ be a sequence of integers. Then $\sum_{i=1}^{n} a(i)$ is defined inductively by the following statements:*

- $\sum_{i=1}^{1} a(i) = a(1)$, *and*
- $\sum_{i=1}^{k+1} a(i) = \sum_{i=1}^{k} a(i) + a(k+1)$.

Let us prove that $\sum_{i=1}^{n} i = \frac{n(n+1)}{2}$. Note that by definition $\sum_{i=1}^{1} i = 1$ and $\frac{1(1+1)}{2} = 1$; hence, the base case holds. Assume that $\sum_{i=1}^{n} i = \frac{n(n+1)}{2}$. Note that $\sum_{i=1}^{n+1} i = \sum_{i=1}^{n} i + (n+1)$ and by the induction hypothesis $\sum_{i=1}^{n} i = \frac{n(n+1)}{2}$. Hence, $\sum_{i=1}^{n+1} i = \frac{n(n+1)}{2} + (n+1) = \frac{(n+1)(n+2)}{2}$.

Exercise 3.2. *Prove that $\sum_{i=1}^{n} 2^i = 2^{n+1} - 2$.*

3.4 Analysis of Algorithms with Cycles

Induction is very useful for analysing algorithms using cycles. Let us extend the example we considered in Section **??**.

Let us consider the following algorithm. We prove that it is working correctly. First, we need to define $r_1, \ldots, r_n$ the value of r during the execution of the algorithm. It is easy to see that $r_1 = a_1$ and

Algorithm 3.1: The algorithm that finds the maximum element of $a_1, \ldots, a_n$.

```
function MAX($a_1, \ldots, a_n$)
    $r \leftarrow a_1$
    for $i$ from 2 to $n$ do
        if $a_i > r$ then
            $r \leftarrow a_i$
        end if
    end for
    return r
end function
```

$$r_{i+1} = \begin{cases} r_i & \text{if } r_i > a_{i+1} \\ a_{i+1} & \text{otherwise} \end{cases}.$$

Secondly, we prove by induction that r_i is the maximum of $a_1, \ldots, a_i$. It is clear that the base case for $i = 1$ is true. Let us prove the induction step from k to $k+1$. By the induction hypothesis, r_k is the maximum of $a_1, \ldots, a_k$. We may consider two following cases.

- If $r_k > a_{k+1}$, then $r_{k+1} = r_k$ is the maximum of $a_1, \ldots, a_{k+1}$ since r_k is the maximum of $a_1, \ldots, a_k$.

- Otherwise, a_{k+1} is greater than or equal to $a_1, \ldots, a_k$, hence, $r_{k+1} = a_{k+1}$.

Exercise 3.3. *Show that line 6 in the following sorting algorithm executes $\frac{n(n+1)}{2}$ times.*

```
function SelectionSort($a_1, \ldots, a_n$)
    for $i$ from 1 to $n$ do
        $r \leftarrow a_i$
        $\ell \leftarrow i$
        for $j$ from $i$ to $n$ do
            if $a_j > r$ then
                $r \leftarrow a_j$
                $\ell \leftarrow j$
            end if
        end for
        Swap $a_i$ and $a_\ell$.
    end for
end function
```

Algorithm 3.2: The algorithm is selection sort, it sorts $a_1, \ldots, a_n$.

3.5 *Strong Induction*

Sometimes $P(k)$ is not enough to prove $P(k+1)$ and we need all the statements $P(1), \ldots, P(k)$. In this case we may use the following induction principle.

Theorem 3.2 (The Strong Induction Principle)**.** *Let $P(n)$ be some statement about positive integer n. Hence, $P(n)$ is true for every integer $n > n_0$ iff*

(the base case) $P(n_0 + 1)$ *is true and*

(the induction step) If $P(n_0 + 1), \ldots, P(n_0 + k)$ are true, then $P(n_0 + k + 1)$ is also true for all positive integers k.

Before we prove this theorem let us present some applications of this principle.

The Fibonacci numbers are defined as follows: $f_0 = 0$, $f_1 = 1$, and $f_k = f_{k-1} + f_{k-2}$ for $k \geq 2$ (note that they are also defined using strong induction since we use not only f_{k-1} to define f_k).

Theorem 3.3 (The Binet formula)**.** *The Fibonacci numbers are given by the following formula*

$$f_n = \frac{\alpha^n - \beta^n}{\sqrt{5}},$$

where $\alpha = \frac{1+\sqrt{5}}{2}$ and $\beta = \frac{1-\sqrt{5}}{2}$.

Proof. We use the strong induction principle to prove this statement with $n_0 = -1$. Let us first prove the base case, $\frac{(\alpha^0 - \beta^0)}{\sqrt{5}} = 0 = f_0$. We also need to prove the induction step.

- If $k = 1$, then $\frac{(\alpha^1 - \beta^1)}{\sqrt{5}} = 1 = f_1$.

- Otherwise, by the induction hypothesis, $f_k = \frac{\alpha^k - \beta^k}{\sqrt{5}}$ and $f_{k-1} = \frac{\alpha^{k-1} - \beta^{k-1}}{\sqrt{5}}$. By the definition of the Fibonacci numbers $f_{k+1} = f_k + f_{k-1}$. Hence,
$$f_{k+1} = \frac{\alpha^k - \beta^k}{\sqrt{5}} + \frac{\alpha^{k-1} - \beta^{k-1}}{\sqrt{5}}.$$
Note that it is enough to show that
$$\frac{\alpha^{k+1} - \beta^{k+1}}{\sqrt{5}} = \frac{\alpha^k - \beta^k}{\sqrt{5}} + \frac{\alpha^{k-1} - \beta^{k-1}}{\sqrt{5}}. \tag{3.1}$$
Note that it is the same as
$$\frac{\alpha^{k+1} - \alpha^k - \alpha^{k-1}}{\sqrt{5}} = \frac{\beta^{k+1} - \beta^k - \beta^{k-1}}{\sqrt{5}}.$$
Additionally, note that α and β are roots of the equation $x^2 - x - 1 = 0$. Hence, $\alpha^{k+1} - \alpha^k - \alpha^{k-1} = \alpha^{k-1}(\alpha^2 - \alpha - 1) = 0$ and $\beta^{k+1} - \beta^k - \beta^{k-1} = \beta^{k-1}(\beta^2 - \beta - 1) = 0$. Which implies equality (**??**).

□

Another example of an application of the strong induction is the proof that any number can be written in digital numeral systems with any base.

Theorem 3.4. *Let $b > 1$ be an integer. Then there is a unique representation of any positive number in the base-b digital numeral system. In other words, for any positive integer n, there are unique $0 \le c_0, \dots, c_\ell < b$ such that $n = \sum_{i=0}^{\ell} b^i c_i$.*

Proof. We prove the statement using strong induction by n. The base case for $n < b$ is clear (we can choose $\ell = 1$ and $c_0 = n$). Let us now prove the induction step. Assume the statement is true for all $k < n$. Let n divided by b be equal to q with the reminder c_0. Note that $(n - c_0)/b < n$ is a positive integer. Hence, by the induction hypothesis, there are $0 \le c_1, \dots, c_\ell < b$ such that $(n - c_0)/b = \sum_{i=1}^{\ell} b^{i-1} c_i$. Hence, $n = \sum_{i=0}^{\ell} b^i c_i$. □

Now we are ready to prove the strong induction principle.

Proof of Theorem **??**. It is easy to see that if $P(n)$ is true for all $n > n_0$, then the base case and the induction steps are true. Let us prove that if the base case and the induction step are true, then $P(n)$ is true for all $n > n_0$.

Let $Q(k)$ be the statement that $P(n_0 + 1), \dots, P(n_0 + k)$ are true. Note that $Q(1)$ is true by the base case for P. Additionally, note that if $Q(k)$ is true, then $Q(k+1)$ is also true, by the induction step for P. Hence, by the induction principle, $Q(k)$ is true for all positive integers k. Which implies that $P(n_0 + k)$ is true for all positive integers k. □

3.6 *Analysis of Recursive Algorithms*

To illustrate the power of recursive definitions and strong induction, let us analyze Algorithm **??**. We prove that number of comparisons of this algorithm is bounded by $6 + 2\log_2(n)$. First step of the proof is to denote the worst number of comparisons when we run the algorithm on the list of length n by $C(n)$. It is easy to see that $C(n) = n$ for $n \leq 5$. Additionally, $C(n) \leq 1 + \max(C\left(\left\lfloor \frac{n}{2} \right\rfloor\right), C\left(n - \left\lfloor \frac{n}{2} \right\rfloor\right))$ for $n > 5$. As we mentioned we prove that $C(n) \leq 6 + 2\log_2(n)$, we prove it by induction. The base case is clear; let us now prove the induction step. By the induction hypothesis,

```
function BinarySearch($e, a_1, \dots, a_n$)
    if $n \leq 5$ then
        for $i$ from 1 to $n$ do
            if $a_i = e$ then
                return i
            end if
        end for
    else
        $\ell \leftarrow \lfloor \frac{n}{2} \rfloor$
        if $a_\ell \leq e$ then
            BinarySearch($e, a_1, \dots, a_\ell$)
        else
            BinarySearch($e, a_{\ell+1}, \dots, a_n$)
        end if
    end if
end function
```

Algorithm 3.3: The binary search algorithm that finds an element e in the sorted list $a_1, \dots, a_n$.

$$C\left(\left\lfloor \frac{n}{2} \right\rfloor\right) \leq 6 + 2\log_2\left(\left\lfloor \frac{n}{2} \right\rfloor\right)$$

and

$$C\left(n - \left\lfloor \frac{n}{2} \right\rfloor\right) \leq 6 + 2\log_2\left(n - \left\lfloor \frac{n}{2} \right\rfloor\right),$$

where $\lfloor \alpha \rfloor$ denotes the integer part of a real number α. Since $\left\lfloor \frac{n}{2} \right\rfloor \leq \frac{n}{2}$ and $n - \left\lfloor \frac{n}{2} \right\rfloor \leq \frac{n}{2} + 1$, $C(n) \leq 1 + 2\log_2(\frac{n}{2} + 1)$. However,

$$1 + 6 + 2\log_2\left(\frac{n}{2} + 1\right) \leq 6 + 2\log_2\left(\frac{n}{\sqrt{2}} + \sqrt{2}\right) \leq 6 + 2\log_2(n)$$

for $n \geq 5$. As a result, we proved the induction step.

End of The Chapter Exercises

3.4 Show that there does not exist the largest integer.

3.5 *(recommended)* Show that for any positive integer n, $n^2 + n$ is even.

3.6 Show that for any positive integer n, 3 divides $n^3 + 2n$.

3.7 Show that for any integer $n \geq 10$, $n^3 \leq 2^n$.

3.8 Show that for any positive integer n, $\sum_{i=0}^{n} x^i = \frac{1-x^{n+1}}{1-x}$.

3.9 Let $a_0 = 2$, $a_1 = 5$, and $a_n = 5a_{n-1} - 6a_{n-2}$ for all integers $n \geq 2$. Show that $a_n = 3^n + 2^n$ for all integers $n \geq 0$.

3.10 *(recommended)* Show that $\sum_{i=1}^{n} i^2 = \frac{n(n+1)(2n+1)}{6}$ for all integers $n \geq 1$.

3.11 Show that $\sum_{i=1}^{n} \frac{1}{i(i+1)} = \frac{n}{n+1}$ for all integers $n \geq 1$.

3.12 Show that $\sum_{i=1}^{n} \frac{1}{i^2} \leq 2 - \frac{1}{n}$ for all integers $n \geq 1$.

3.13 Show that $\sum_{i=1}^{n}(2i - 1) = n^2$ for any positive integer n.

3.14 Prove that $\sum_{i=1}^{n} \frac{1}{i(i+1)} = \frac{n}{n+1}$ for any positive integer n.

3.15 Prove that $\sum_{i=1}^{n}(i+1)2^i = n2^{n+1}$ for all integers $n > 2$.

3.16 Let $a_1, \ldots, a_n$ be a sequence of real numbers. We define inductively $\prod_{i=k}^{n} a_i$ as follows:

- $\prod_{i=1}^{1} a_i = a_1$ and
- $\prod_{i=1}^{k+1} a_i = \left(\prod_{i=1}^{k} a_i\right) \cdot a_{k+1}$.

Prove that $\prod_{i=1}^{n-1} \left(1 - \frac{1}{(i+1)^2}\right) = \frac{n+1}{2n}$ for all integers $n > 1$.

3.17 Let $f_0 = 1$, $f_1 = 1$, and $f_{n+2} = f_{n+1} + f_n$ for all integers $n \geq 0$. Show that $f_n \geq \left(\frac{3}{2}\right)^{n-2}$.

3.18 Show that $f_{n+m} = f_{n-1}f_{m-1} + f_n f_m$.

3.19 Show that two arithmetic formulas $(x_1 + x_2) \cdot x_3$ and $x_1 \cdot x_3 + x_2 \cdot x_3$ on the variables x_1, x_2, and x_3 have the same values.

3.20 Let us define $n!$ as follows: $1! = 1$ and $n! = (n-1)! \cdot n$. Show that $n! \geq 2^n$ for any $n \geq 4$.

3.21 *(open)* Fins all the natural numbers n such that $n! = m^2$ for some integer m.

3.22 Show that $\int\limits_{0}^{+\infty} x^n e^{-x}\, \mathrm{d}x = n!$ for all $n \geq 0$.

3.23 Prove that $\sum_{i=1}^{n}(i+1)2^i = n2^{n+1}$ for all integers $n \geq 1$.

3.24 Show that $\sum_{k=1}^{n} k \cdot k! = (n+1)! - 1$.

3.25 Show that the algorithm from **??** sorts the array.

Solutions to The Exercises

3.9 We prove this using induction by n. The base case for $n \leq 1$ is clear since $3^0 + 2^0 = 2$ and $3^1 + 2^1 = 5$.

Let us prove the induction step. Assume that $a_n = 3^n + 2^n$ and $a_{n-1} = 3^{n-1} + 2^{n-1}$, we need to prove that $a_{n+1} = 3^{n+1} + 2^{n+1}$. Note that

$$a_{n+1} = 5a_n - 6a_{n-1} = 5 \cdot 3^n + 5 \cdot 2^n - 6 \cdot 3^{n-1} - 6 \cdot 2^{n-1} = \\ 3^{n-1} \cdot 9 + 2^{n-1}4 = 3^{n+1} + 2^{n+1}.$$

4. Predicates and Connectives

Connectives and Propositions:
Introduction to Mathematical Reasoning #5

`https://youtu.be/0unvlq2OTaE`

4.1 Propositions and Predicates

In the previous chapters we used the word "statement" without any even relatively formal definition of what it means. In this chapter we are going to give a semi-formal definition and discuss how to create complicated statements from simple statements.

It is difficult to give a formal definition of what a mathematical statement is, hence, we are not going to do it in this book. The goal of this section is to enable the reader to recognize mathematical statements.

A *proposition* or a mathematical statement is a declarative sentence which is either true or false but not both. Consider the following list of sentences.

1. $2 \times 2 = 4$
2. $\pi = 4$
3. n is even
4. 32 is special
5. The square of any odd number is odd.
6. The sum of any even number and one is prime.

Of those, the first two are propositions; note that this says nothing about whether they are true or not. Actually, the first is true and the second is false. However, the third sentence becomes a proposition only when the value of n is fixed. The fourth is not a proposition. Finally, the last two are propositions (the fifth is true and the sixth is false).

The third statement is somewhat special, because there is a simple way to make it a proposition: one just needs to fix the value of the variables. Such sentences are called predicates and the variables that need to be specified are called free variables of these predicates.

Note that the fourth sentence is also interesting, since if we define what it means to be special, the phrase became a proposition. Mathematicians tend to do such things to give mathematical meanings to everyday words.

4.2 Connectives

Mathematicians often need to decide whether a given proposition is true or false. Many statements are complicated and constructed from simpler statements using *logical connectives*. For example we may consider the following statements:

1. $3 > 4$ and $1 < 1$;
2. $1 \times 2 = 5$ or $6 > 1$.

Logical connective "OR". The second statement is an example of usage of this connective. The statement "P or Q" is true if and only if at least one of P and Q is true. We may define the connective using the truth table of it.

P	Q	P or Q
T	T	T
T	F	T
F	T	T
F	F	F

The or connective is also called *disjunction* and the disjunction of P and Q is often dented as $P \vee Q$.

Warning: Note that in everyday speech "or" is often used in the exclusive case, like in the sentence "we need to decide whether it is an insect or a spider". In this case the precise meaning of "or" is made clear by the context. However, mathematical language should be formal, hence, we always use "or" inclusively.

Logical connective "AND". The first statement is an example of this connective. The statement "P and Q" is true if and only if both P and Q are true. We may define the connective using the truth table of it.

P	Q	P and Q
T	T	T
T	F	F
F	T	F
F	F	F

The or connective is also called *conjunction* and the conjunction of P and Q is often dented as $P \wedge Q$.

Warning: Not all the properties of "and" from everyday speech are captured by logical conjunction. For example, "and" sometimes implies order. For example, "They got married and had a child" in common language means that the marriage came before the child. The word "and" can also imply a partition of a thing into parts, as "The American flag is red, white, and blue." Here it is not meant that the flag is at once red, white, and blue, but rather that it has a part of each color.

Logical connective "NOT". The last connective is called *negation* and examples of usage of it are the following:

1. 5 is not greater than 8;
2. Does not exist an integer n such that $n^2 = 2$.

Note that it is not straightforward where to put the negation in these sentences.

The negation of a statement P is denoted as $\neg P$ (sometimes it is also denoted as $\sim P$).

End of The Chapter Exercises

4.1 Construct truth tables for the statements

- not (P and Q);
- (not P) or (not Q);
- P and (not Q);
- (not P) or Q;

4.2 *(recommended)* Consider the statement "All gnomes like cookies". Which of the following statements is the negation of the above statement?

- All gnomes hate cookies.
- All gnomes do not like cookies.
- Some gnomes do not like cookies.
- Some gnomes hate cookies.
- All creatures who like cookies are gnomes.
- All creatures who do not like cookies are not gnomes.

4.3 Using truth tables show that the following statements are equivalent:

- $P \implies Q$,
- $(P \vee Q) \iff Q$ ($A \iff B$ is the same as $(A \implies B) \wedge (B \implies A)$),
- $(P \wedge Q) \iff P$

4.4 Prove that three connectives "or", "and", and "not" can all be written in terms of the single connective "notand" where "P notand Q" is interpreted as "not (P and Q)" (this operation is also known as Sheffer stroke or NAND).

4.5 Show the same statement about the connective "notor" where "P notor Q" is interpreted as "not (P or Q)" (this operation is also known as Peirce's arrow or NOR).

5. Sets

Sets:
Introduction to Mathematical Reasoning #6

`https://youtu.be/bshBV2H4Sqo`

5.1 The Intuitive Definition of a Set

A set is one of the two most important concepts in mathematics. Many mathematical statements involve "an integer n" or "a real number a". Set theory notation provides a simple way to express that a is a real number. However, this language is much more expressible and it is impossible to imagine modern mathematics without this notation.

As in the previous chapter it is difficult to define a set formally so we give a less formal definition which should be enough to use the notation. A *set* is a well-defined collection of objects. Important examples of sets are:

1. $\mathbb{R}$ a set of reals,
2. $\mathbb{Z}$ the set of integers[1],
3. $\mathbb{N}$ the set of natural numbers[2],
4. $\mathbb{Q}$ a set of rational numbers,
5. $\mathbb{C}$ a set of complex numbers.

Usually, sets are denoted by single letter.

Objects in a set are called *elements* of the set and we denote the statement "x is in the set E" by the formula $x \in E$ and the negation of this statement by $x \notin E$. For example, we proved that $\sqrt{2} \notin \mathbb{Q}$[3].

Exercise 5.1. *Which of the following sets are included in which? Recall that a number is prime iff it is an integer greater than* 1 *and divisible only by* 1 *and itself.*

1. *The set of all positive integers less than* 10.
2. *The set of all prime numbers less than* 11.
3. *The set of all odd numbers greater than* 1 *and less than* 6.
4. *The set of all positive integers less than* 10.
5. *The set whose only elements are* 1 *and* 2.

[1] "Z" stands for the German word Zahlen ("numbers").

[2] Note that in the literature there are two different traditions: in one 0 is a natural number, in another it is not; in this book we are going to assume that 0 is not a natural number.

[3] The symbol $\in$ was first used by Giuseppe Peano 1889 in his work "Arithmetices principia, nova methodo exposita". Here he wrote on page X: "The symbol $\in$ means is. So $a \in b$ is read as a *is* a b; . . . " The symbol itself is a stylized lowercase Greek letter epsilon ("ϵ"), the first letter of the word εστι, which means "is".

6. *The set whose only element is* 1.

7. *The set of all prime numbers less than* 11.

5.2 Basic Relations Between Sets

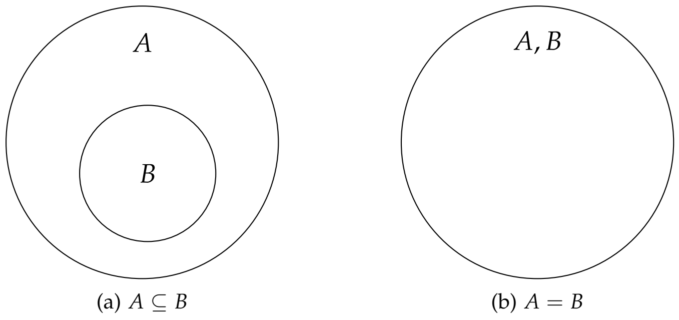
Many problems in mathematics are problems of determining whether two descriptions of sets are describing the same set or not. For example, when we learn how to solve quadratic equations of the form $ax^2 + bx + c = 0$ $(a, b, c \in \mathbb{R})$ we learn how to list the elements of the set $\left\{x \in \mathbb{R} \,:\, ax^2 + bx + c = 0\right\}$.

We say that two sets A and B are equal if they contain the same elements (we denote it by $A = B$). If all the elements of A belong to B we say that A is a subset of B and denote it by $A \subseteq B$[4].

For example, $\mathbb{Q} \subseteq \mathbb{R}$ since any rational number is also a real number. A special set is an empty set i.e. the set that does not have elements, we denote it $\varnothing$.

Diagrams

If we think of a set A as represented by all the points within a circle or any other closed figure, then it is easy to represent the notion of A being a subset of another set B also represented by all the points within a circle. We just put a circle labeled by A inside of the circle labeled by B. We can also diagram an equality by drawing a circle labeled by both A and B. (see fig. **??**). Such diagrams are called Euler diagrams and it is clear that one may draw Euler diagrams for more than two sets.



(a) $A \subseteq B$

(b) $A = B$

Figure 5.1: Euler diagrams for subset and equality relations

Descriptions of Sets

In this section we describe how to define new sets, this notation is also known as *set-builder notation*.

[4] In the literature there are three symbols for "subset": $\subseteq$, $\subsetneq$, and $\subset$. $A \subseteq B$ means that A is a subset of B and we allow $A = B$ and $A \subsetneq B$ means that A is a subset of B and we forbid $A = B$. However, there is a problem with the third symbol, some people use it as a synonym of $\subseteq$ and some use it as a synonym of $\subsetneq$. Due to this ambiguity we are going to avoid using it in this book.

Listing elements. The simplest way to define a set is just to list the elements. For example

1. $\{1, 2, \pi\}$ is the set consisting of three elements 1, 2, and π, and

2. $\{1, 2, 3, \ldots\}$ is the set of all positive integers i.e. it is the set $\mathbb{N}$.

Conditional definitions. We may also describe a set using some constraint e.g we may list all the even numbers using the following formula $\{n \in \mathbb{Z} \ : \ n \text{ is even}\}$ (we read it as "the set of all integers n such that n is even").

Using this we may also define the set of all integers from 1 to m, we denote it $[m]$; i.e. $[m] = \{n \in \mathbb{N} \ : \ 0 < n \leq m\}$.

Constructive definitions. Another way to construct a set of all even numbers is to use the constructive definition of a set: $\{2k \ : \ k \in \mathbb{Z}\}$.

We may also describe a set of rational numbers using this description: $\mathbb{Q} = \{a/b \ : \ a \in \mathbb{Z}, b \in \mathbb{N}\}$ (note that we may also use a mix of a conditional and constructive definitions, $\mathbb{Q} = \{a/b \ : \ a, b \in \mathbb{Z}, b \neq 0\}$).

Exercise 5.2. *Describe a set of perfect squares using constructive type of definition.*

Disjoint Sets

Two sets are *disjoint* iff they do not have common elements. We also say that two sets are *overlapping* iff they are not disjoint i.e. they share at least one element.

More generally, A_1, $\ldots$, A_ℓ are pairwise disjoint iff A_i is disjoint with A_j for all $i \neq j \in [\ell]$

Exercise 5.3. *Of the sets in Exercise* **??**, *which are disjoint from which?*

5.3 *Operations over Sets.*

Another way to describe a set is to apply operation to other sets. Let A and B be sets.

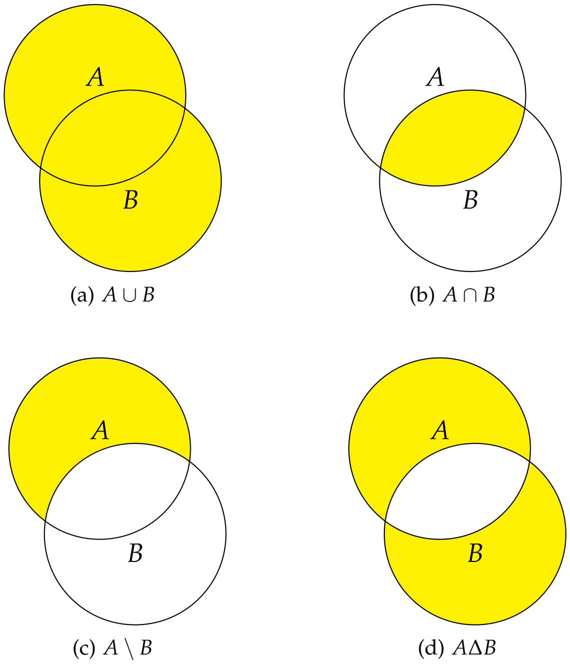
The first example of the operations on sets is the *union* operation. The union of A and B is the set containing all the elements of A and all the elements of B i.e. $A \cup B = \{x \ : \ x \in A \text{ or } x \in B\}$[5].

Another example of such an operation is *intersection*. The intersection of A and B is the set of all the elements belonging to both A and B i.e $A \cap B = \{x \ : \ x \in A \text{ and } x \in B\}$[6].

The third operation we are going to discuss this lecture is *set difference*. If A and B are some sets, then $A \setminus B = \{x \ : \ x \in A \text{ and } x \notin B\}$.

[5] Note that this definition is not correct since in the conditional definitions we have to specify the set x belongs to and we cannot do this here.

[6] You may notice that in the definition of the union we use disjunction and in the definition of intersection we use conjunction. Actually this is the reason the symbol of the conjunction is similar to the symbol of intersection and the symbol of the disjunction is similar to the symbol of union.



(a) $A \cup B$ (b) $A \cap B$

(c) $A \setminus B$ (d) $A \Delta B$

Figure 5.2: Euler diagrams for set operations

The last operation is *symmetric difference*. If A and B are some sets, then $A\Delta B = (A \setminus B) \cup (B \setminus A)$. Note that alternatively $A\Delta B = (A \cup B) \setminus (A \cap B)$

Exercise 5.4. *Describe the set* $\{n \in \mathbb{N} \ : \ n \text{ is even}\} \cap \{3n \ : \ n \in \mathbb{N}\}$.

Theorem 5.1. *Let A, B, and C be some sets. Then we have the following identities.*

(associativity) $A \cup (B \cup C) = (A \cup B) \cup C$ *and* $A \cap (B \cap C) = (A \cap B) \cap C$.

(commutativity) $A \cup B = B \cup A$ *and* $A \cap B = B \cap A$.

(distributivity) $A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$ *and* $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$.

Proof. One may prove these properties using the Euler diagrams. Alternatively they can be proven by definitions. Let us prove only the first part of the distributivity, the rest is Exercise **??**.

Our proof consists of two parts in the first part we prove that $A \cup (B \cap C) \subseteq (A \cup B) \cap (A \cup C)$. Suppose that $x \in A \cup (B \cap C)$. Then $x \in A$ or $x \in (B \cap C)$.

- If $x \in A$, then $x \in (A \cup B)$ and $x \in (A \cup C)$ i.e. $x \in ((A \cup B) \cap (A \cup C))$.

- If $x \in (B \cap C)$, then $x \in B$ and $x \in C$. Which implies that $x \in (A \cup B)$ and $x \in (A \cup C)$. As a result, $x \in ((A \cup B) \cap (A \cup C))$.

□

Exercise 5.5. *Prove the rest of the equalities in Theorem* **??**.

Probably the most difficult concept connected to sets is the concept of a power set. Let A be some set, then the set of all possible subsets of A is denoted by 2^A (sometimes this set is denoted by $\mathcal{P}(A)$) and called the power set of A. In other words $2^A = \{B : B \subseteq A\}$.

Warning: Please do not forget about two extremal elements of the power set 2^A: the empty set and A itself.

For example if $A = \{1,2,3\}$, then

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

5.4 The Well-ordering Principle

Using the set notation we may finally justify the proof of the statement that $2^n > n$ for all positive integers n from the video about mathematical induction. In order to do this let us first formulate the following theorem.

Theorem 5.2. *Let $A \subseteq \mathbb{Z}$ be a non-empty set. We say that $b \in \mathbb{Z}$ is a lower bound for the set A iff $b \leq a$ for all $a \in A$. Additionally, we say that the set A is bounded if there is a lower bound for A.*

Given this, if A is bounded, then there is a lower bound $a \in A$ for the set A (we say that a is the minimum of the set A).

Note that this theorem also states that any subset of natural numbers have a minimum.

Recall that we wish to prove that $2^n > n$ for all positive n. Assume that it is not true, in this case the set $A = \{n \in \mathbb{N} : 2^n < n\}$ is non-empty. Denote by n_0 the minimum of the set A, n_0 exists by Theorem **??**. We may consider the following two cases.

- If $n_0 = 1$, then it leads to a contradiction since $2 = 2^1 > 1$.
- Otherwise, note that $1 \leq n_0 - 1 < n_0$, hence, $2^{n_0 - 1} > n_0 - 1$. So $2^{n_0} > 2n_0 - 2 \geq n_0$. Which is a contradiction with the definition of n_0.

Finally, we prove Theorem **??**.

Proof of Theorem **??**. Let b be a lower bound for the set A. Assume that there is no minimum of the set A. Let $P(n)$ be the statement that $n \notin A$.

First, we are going to prove that $P(n)$ is true for all $n \geq b$. The base case is true since if $b \in A$, then b is the minimum of A which contradicts to the assumption that there is no minimum of A. The induction step is also clear, by the induction hypothesis we know that $P(b)$, $\ldots$, $P(k)$ are true, hence, $(k+1) \in A$ implies that $k+1$ is the minimum of A.

Now we prove that A is empty. Assume the opposite i.e. assume that there is $x \in A$. Note that $x \geq b$ since b is a lower bound of A. However, $P(x)$ is true which implies that $x \notin A$. Therefore the assumption was false and A is empty, but this contradicts to the fact that A is non- empty. □

End of The Chapter Exercises

5.6 Find the power sets of $\emptyset$, $\{1\}$, $\{1,2\}$, $\{1,2,3,4\}$. How many elements in each of this sets?

5.7 *(recommended)* Prove that

- $A \subseteq B \iff A \cup B = B$,
- $A \subseteq B \iff A \cap B = A$.

5.8 Let A be a subset of a set U we call this set a universe. We say that the set $\overline{A} = U \setminus A$ is a complement of A in U. Show the following equalities

- $\overline{\overline{A}} = A$.
- $\overline{A \cup B} = \overline{A} \cap \overline{B}$.
- $\overline{A \cap B} = \overline{A} \cup \overline{B}$.

5.9 *(recommended)* Let us define an intersection of more than two sets as follows. Let $A_1, \ldots, A_n$ be some sets. Then

- $\bigcap_{i=1}^{1} A_i = A_1$ and
- $\bigcap_{i=1}^{k+1} A_i = \left(\bigcap_{i=1}^{k} A_i\right) \cap A_{k+1}$.

Show that $\bigcap_{i=1}^{n} \{x \in \mathbb{N} \ : \ i \leq x \leq n\} = \{n\}$ for all integers $n > 0$.

5.10 Let us define a union of more than two sets as follows. Let A_1, $\ldots$, A_n be some sets. Then

- $\bigcup_{i=1}^{1} A_i = A_1$ and
- $\bigcup_{i=1}^{k+1} A_i = \left(\bigcup_{i=1}^{k} A_i\right) \cup A_{k+1}$.

Show that $\bigcup_{i=1}^{n} [i] = [n]$ for all integers $n > 0$.

5.11 *(recommended)* Let Ω be some set and $A_1, \ldots, A_n \subseteq \Omega$. Show that $\bigcup_{i=1}^{n} A_i = \{x \in \Omega \ : \ \exists i \in [n] \ x \in A_i\}$.

5.12 Let $A_1, \dots, A_n$ be some sets. Show that $\bigcup_{i=1}^{n}(A_i \cap B) = (\bigcup_{i=1}^{n} A_i) \cap B$.

5.13 Show that $A\Delta(B\Delta C) = (A\Delta B)\Delta C$.

5.14 *(recommended)* Let $\mathbb{R}^{m \times n}$ be the set of all matrices $m \times n$ and $\mathbb{R}^n$ be the set of n dimensional vectors. Show that for any matrix $A \in \mathbb{R}^{m \times n}$ $(n > m)$ there is a nonzero vector $x \in \mathbb{R}^n$ such that $Ax = 0$.

6. Functions

Functions and Quantifiers:
Introduction to Mathematical Reasoning #7

https://youtu.be/VHJeUrCedTU

Another important type of objects in mathematics are functions. Function f from a set X to a set Y (written as $f : X \to Y$) is a unique assignment of elements of Y to the elements of X (note that it is not necessary that all the elements of Y are used). In other words, for each element $x \in X$ there is one assigned element $f(x) \in Y$. We call such an element the *value* of f at x, we also say that $f(x)$ is an *image* of x.

Unfortunately, the definition is not formal. Through this chapter we will provide a more formal definition.

6.1 Quantifiers.

The first ingredient is called quantifiers. Very often we use phrases like "all the people in the class have smartphones." However, we still do not know how to write it using symbols.

The Universal Quantifier. In order to say "all" or "every" we use the symbol $\forall$[1]: if $P(a)$ is a predicate about $a \in A$, then $\forall a \in A\ P(a)$ is a statement saying that all the elements of A satisfy the predicate P. In other words it is the same as the statement $\{a \in A : P(a)\} = A$. For example, $\forall x \in \mathbb{R}\ x \cdot 0 = 0$ says that product of every real number and zero is equal to zero.

The Existential Quantifier. The second quantifier means "there is" and is denoted by the symbol $\exists$[2]: if $P(a)$ is a predicate about an element of A, then $\exists a \in A\ P(a)$ says that there is an element of A satisfying the predicate P i.e. $\{a \in A : P(a)\} \neq \emptyset$. For example, $\exists x \in \mathbb{R}\ x^2 - 1 = 0$ states that there is a real solution of the equation $x^2 - 1 = 0$.

[1] The symbol is a turned "A" symbol, the first letter of the word "all".

[2] The symbol is a turned "E" symbol, the first letter of the word "exists". It is also interesting that the symbol for the universal quantifier was introduced by Gerhard Gentzen in 1935 but the symbol for the existential quantifier was introduced, 38 years earlier, by Giuseppe Peano in 1897.

Warning: Note that the word "any" sometimes indicates a universal statement and sometimes an existential statement.
Standard meaning of "any" is "every" like in the statement "$a^2 \geq 0$ for any real number", therefore this statement can be rewritten as $\forall a \in \mathbb{R}\ a^2 \geq 0$. Nonetheless, in the negative and interrogative statements "any" is used to mean "some". For example, "There is not any real number a such that $a^2 < 0$" is asserting that the statement $\exists a \in \mathbb{R}\ a^2 < 0$ is false. And "Is there any real number a such that $a^2 = 1$?" is asking whether the existential statement $\exists a \in \mathbb{R}\ a^2 = 1$ is true.
Real care is required with questions involving "any": "Is there any integer a such that $a \geq 1$?" clearly is asking whether $\exists a \in \mathbb{R}\ a^2 \geq 1$ is true; however, "Is $a \geq 1$ for any integer a" is less clear and might be taken to asking about the same question as the first question, $\exists a \in \mathbb{Z}\ a \geq 1$ (which is true) but might also be taken to be asking about $\forall a \in \mathbb{Z}\ a \geq 1$ (which is false).

Proving Statements Involving Quantifiers

Most of the statements in mathematics involve quantifiers. This is one of the factors distinguishing advanced from elementary mathematics. In this section we give an overview of the main methods of proof. Though the whole book is about proving such results.

Proving statements of the form $\forall a \in A\ P(a)$. Such statements can be rewritten in the form $a \in A \implies P(a)$. For example, we proved earlier that $a^2 \geq 0$ for all real numbers a using this approach.

Proving statements of the form $\exists a \in A\ P(a)$. The easiest way to prove such a statement is by simply exhibiting an element a of A such that $P(a)$ is true. This method is called *proof by example*.

Let us prove the statement $\exists x \in \mathbb{N}\ x^2 = 4$ using this method. Observe that $2 \in N$ and $2^2 = 4$ so $x = 2$ provides an example proving this statement. There are, however, less direct methods such as use of the counting arguments.

Proving statements involving both quantifiers. To illustrate problems of this type let us prove that for any integer n, if n is even, then n^2 is also even.

This statement is a universal statement $\forall n \in \mathbb{Z}\ (n \text{ is even} \implies n^2 \text{ is even})$. However, the hypothesis that n is even is an existential statement $\exists q \in \mathbb{Z}\ n = 2q$. So we begin the proof as follows:

Suppose that n is an even integer. Then $n = 2q$ for some integer q.

The conclusion we wish to prove is that n^2 is even, which may be written as $\exists q \in \mathbb{Z}\ n^2 = 2q$. Note that q here is a dummy variable used to express the statement n^2 is a doubled integer. We may replace it by any other letter not already in use, for example $\exists p \in \mathbb{Z}\ n^2 = 2p$. Hence, if we present p such that $n^2 = 2p$, we finish the proof. As a result, we can complete the proof as follows.

Therefore, $n^2 = (2q)^2 = 4q^2$ and so, since $2q^2$ is an integer n^2 is even.

Disproving Statements Involving Quantifiers

Disproving something seems a bit off from the first glance, but to some extent it is the same as proving the negation.

Disproving statements of the form $\forall a \in A\ P(a)$. We may note that the negation of such a statement is the statement $\exists a \in A\ \neg P(a)$. So we can disprove it by giving a single example for which it is false. This is called *disproof by counterexample* to $P(a)$.

For example, we may disprove the statement $\forall x \in \mathbb{R}\ x^2 > 2$ by giving a counterexample $x = 1$ since $1^2 = 1 < 2$.

Disproving statements of the form $\exists a \in A\ P(a)$. The negation of this statement is the statement $\forall a \in A\ \neg P(a)$. Which gives one way of disproving the statement.

Let us prove that there does not exist a real number x such that $x^2 = -1$. We know that, for all $x \in \mathbb{R}$, we have the inequality $x^2 \geq 0$ and so $x^2 \neq -1$. Hence, there does not exist $x \in \mathbb{R}$ such that $x^2 = -1$.

6.2 *Cartesian product*

Another ingredient is the notion of Cartesian product. If X and Y are two sets, then $X \times Y = \{(x, y)\ :\ x \in X \text{ and } y \in Y\}$. We also denote $\underbrace{X \times X \times \cdots \times X}_{k \text{ times}}$ by X^k.

Consider the following example. If $X = \{a, b, c\}$ and $Y = \{a, b\}$, then

$$X \times Y = \{(a, a), (a, b), (b, a), (b, b), (c, a), (c, b)\}.$$

Additionally, $\mathbb{R}^2 = \mathbb{R} \times \mathbb{R}$ is the familiar 2-dimensional Euclidean plane.

Exercise 6.1. *Find the set $\{a, b\} \times \{a, b\} \setminus \{(x, x)\ :\ x \in \{a, b\}\}$*

Theorem 6.1. *For all sets A, B, C, and D the following hold:*

- $A \times (B \cup C) = (A \times B) \cup (A \times C)$;
- $A \times (B \cap C) = (A \times B) \cap (A \times C)$;

- $(A \times B) \cup (C \times D) \subseteq (A \cup C) \times (B \cup D)$;
- $(A \times B) \cap (C \times D) = (A \cap C) \times (B \cap D)$.

Proof. It is easy to prove this statement by the definitions. Let us prove only the second equality, the rest is Exercise **??**.

Note that $(x, y) \in A \times (B \cap C)$ iff $x \in A$ and $y \in (B \cap C)$. Hence, $(x, y) \in A \times (B \cap C)$ iff $x \in A$, $y \in B$, and $y \in C$. Thus $(x, y) \in A \times (B \cap C)$ iff $(x, y) \in (A \times B)$ and $(x, y) \in (A \times C)$. As a result, $(x, y) \in A \times (B \cap C)$ iff $(x, y) \in (A \times B) \cap (A \times C)$ as required. □

Exercise 6.2. *Prove the rest of the equalities in Theorem* **??**.

6.3 *Graphs of Functions*

Now we have all the components to define a function. Mathematicians think about the functions in the way we defined them at the beginning of the chapter, however formally in order to define a function $f : X \to Y$ one need to define a set $D \subseteq X \times Y$ (such a set is called the *graph of the function* f) such that

- $\forall x \in X\ \exists y \in Y\ (x, y) \in D$ and
- $\forall x \in X, y_1, y_2 \in Y\ ((x, y_1) \in D \wedge (x, y_2) \in D \implies y_1 = y_2)$.

We say that $y \in Y$ is the value $f(x)$ of the function described by D at $x \in X$ iff $(x, y) \in D$.

The simplest way to think about the functions is in the terms of tables. Let us use this idea to list all the functions $\{a, b, c\}$ to $\{d, e\}$.

x	$f_1(x)$	$f_2(x)$	$f_3(x)$	$f_4(x)$	$f_5(x)$	$f_6(x)$	$f_7(x)$	$f_8(x)$
a	d	d	d	d	e	e	e	e
b	d	d	e	e	d	d	e	e
c	d	e	d	e	d	e	d	e

Exercise 6.3. *List all the functions from* $\{a, b\}$ *to* $\{a, b\}$.

However, listing all the values of a function is only possible when the domain of the function is finite. Thus the most common way to describe a function is using a formula which provides a way to find the value of a function. When the function is defined as a formula it is important to be clear which sets are the domain and the codomain of the function.

Let $\mathbb{R}_+ = \{x \in \mathbb{R} : x \geq 0\}$. Consider the following functions.

- $g_1 : \mathbb{R} \to \mathbb{R}$ such that $g_1(x) = x^2$;
- $g_2 : \mathbb{R}_+ \to \mathbb{R}$ such that $g_2(x) = x^2$;

- $g_3 : \mathbb{R} \to \mathbb{R}_+$ such that $g_3(x) = x^2$;
- $g_4 : \mathbb{R}_+ \to \mathbb{R}_+$ such that $g_4(x) = x^2$;

Nonetheless that all these functions are defined using the same formula x^2, we will see in the next chapters that these four functions have different properties.

Exercise 6.4. *Find the graph of the function $f : \mathbb{Z} \to \mathbb{Z}$ such that $f(x) = 3x$.*

Note that when you define the function you need to define it such that the definition makes sense for all the elements of the domain. For example, the formula $g(x) = \frac{x^2-3x+2}{x-1}$ does not define a function from $\mathbb{R}$ to $\mathbb{R}$ since it is not defined for $x = 1$. It is typical to define a function from real numbers to real numbers by a formula and the convention is that the domain is the set of all numbers for which the formula makes sense (unless the domain is specified explicitly). Using this convention the formula g defines a function from $\mathbb{R} \setminus \{1\}$ to $\mathbb{R}$.

If we really need a function from $\mathbb{R}$ there are two possible approaches for extending g.

Rewriting the formula. We can rewrite the formula such that it makes sense for all the real numbers. Note that for all $x \in \mathbb{R} \setminus \{1\}$,

$$\frac{x^2 - 3x + 2}{x - 1} = \frac{(x-2)(x-1)}{x-1} = x - 2.$$

Then $g_1(x) = x - 2$ defines a function on $\mathbb{R}$ extending the function g.

Explicit definition. Alternatively we can explicitly specify the value of g at 1. So

$$g_2(x) = \begin{cases} \frac{x^2-3x+2}{x-1} & \text{if } x \neq 1 \\ -1 & \text{if } x = 1 \end{cases}$$

defines a function from $\mathbb{R}$ to $\mathbb{R}$. Note that we can specify the values at individual points any way we want.

Similarly to sets we may define the equality between functions. We say that two functions $f, g : X \to Y$ are equal ($f = g$) iff $f(x) = g(x)$ for all $x \in X$ i.e. their graphs are equal. Note that two functions are equal only if they have the same domains and codomains. For example, g_1 and g_2 we just defined are equal to each other nonetheless that we defined them in two different ways.

We defined g_1 and g_2 to extend g to a bigger domain, similarly we can make a domain smaller.

Definition 6.1. *Let $f : X \to Y$ and $A \subseteq X$. Then $f|_A : A \to Y$ is a function such that $\forall x \in A\ f|_A(x) = f(x)$ (we say that $f|_A$ is the restriction of f to the set A).*

6.4 Composition of Functions

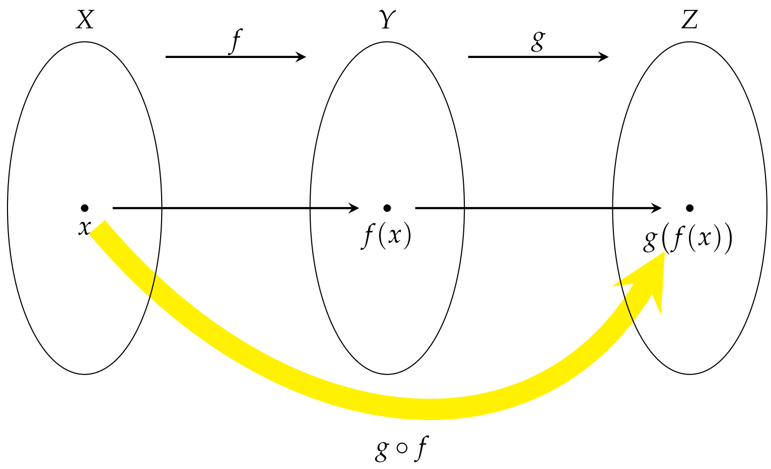


Figure 6.1: Composition of functions

Suppose $f : X \to Y$ and $g : Y \to Z$ be some function. Then, given an element $x \in X$, the function f assigns $y = f(x) \in Y$, and the function g assigns $z = g(y) = g(f(x)) \in Z$. Thus using f and g an element of Z can be assigned to x. This operation defines a function from X to Z and the result of this operation is called the *composition* of f and g.

Definition 6.2. *If $f : X \to Y$ and $g : Y \to Z$, then $h = g \circ f$ is a function from X to Z such that $g(f(x)) = h(x)$ for all $x \in X$.*

Let us consider an example. Let $f : \mathbb{R} \to \mathbb{R}$ such that $f(x) = x + 1$ and $g : \mathbb{R} \to \mathbb{R}$ such that $g(x) = x^2$. Then $(g \circ f) : \mathbb{R} \to \mathbb{R}$ and $(g \circ f)(x) = (x+1)^2$ for all $x \in \mathbb{R}$. Note that the order of f and g is important since $(f \circ g)(x) = x^2 + 1$. Thus composition is not *commutative*.

There are two special type functions.

- Let $A \subseteq X$, then $i : A \to X$ such that $i(a) = a$ for all $a \in A$ is called the *inclusion* function of A into X. Observe that $(f \circ i) : A \to Y$ and $(f \circ i) = f|_A$ for any function $f : X \to Y$.

- Another important function is called the *identity* function. Let X be some set. Then $I_X : X \to X$ is the identity function on X iff $I_X(x) = x$.

Theorem 6.2. *Let $f : X \to Y$, $g : Y \to Z$, and $h : Z \to W$. Then*

- $h \circ (g \circ f) = (h \circ g) \circ f$.

- $f \circ I_X = f = I_Y \circ f$.

Proof. These results can be proven simply by evaluating the functions. For example, both functions in the first equality assign $h(g(f(x)))$ for any $x \in X$ and so functions are equal. □

Notice that this theorem states that we may write $f \circ g \circ h$ without ambiguity.

6.5 The Image of a Function

Given a function $f : X \to Y$, it is not necessary that every element of Y is an image of some $x \in X$. For example, the function $\mathbb{R} \to \mathbb{R}$ defined by the formula x^2 does not have -1 as a value.

Thus we may give the following definition.

Definition 6.3. *The image of the function f is defined as follows*

$$\operatorname{Im} f = \{y \in Y \;:\; \exists x \in X \; f(x) = y\} = \{f(x) \;:\; x \in X\}$$

(in other words it is the projection of the graph D of f on the second coordinate: $\operatorname{Im} f = \{y \;:\; (x, y) \in D\}$*).*

End of The Chapter Exercises

6.5 Is there $x, y, z \in \mathbb{N}$ such that $29x + 30y + 31z = 366$.

6.6 Find then image of the function $f : \mathbb{Z} \to \mathbb{Z}$ such that $f(x) = 3x$.

6.7 *(recommended)* Determine the following sets:

- $\{m \in \mathbb{N} \;:\; \exists n \in \mathbb{N} \; m \le n\}$;
- $\{m \in \mathbb{N} \;:\; \forall n \in \mathbb{N} \; m \le n\}$;
- $\{n \in \mathbb{N} \;:\; \exists m \in \mathbb{N} \; m \le n\}$;
- $\{n \in \mathbb{N} \;:\; \forall m \in \mathbb{N} \; m \le n\}$.

6.8 Prove or disprove the following statements.

- $\forall m, n \in \mathbb{N} \; m \le n$.
- $\exists m, n \in \mathbb{N} \; m \le n$.
- $\exists m \in \mathbb{N} \; \forall n \in \mathbb{N} \; m \le n$.
- $\forall m \in \mathbb{N} \; \exists n \in \mathbb{N} \; m \le n$.
- $\exists n \in \mathbb{N} \; \forall m \in \mathbb{N} \; m \le n$.
- $\forall n \in \mathbb{N} \; \exists m \in \mathbb{N} \; m \le n$.

6.9 *(recommended)* We call elements of the set $\{0,1\}^n$ Binary strings of length n. Moreover, instead of $(c_1,\ldots,c_n)$ we write $c_1,\ldots,c_n$ and we call c_is characters. Show that all Binary strings of length n may be ordered such that every successive strings in this order are different only in one character. (For example, for $n = 2$ the order may be 00, 01, 11, 10.)

7. Relations

Nonetheless that function are used almost everywhere in mathematics, many relations are not functional by their nature. For example, for any real a, there are two solutions of $x^2 = a$ and there are zero solutions for $a < 0$. To work with such situations, relations are used.

In order to define a relation we need to relax the definition of the graph of a function (Section **??**) by allowing more than one "result" and by allowing zero "results". In other words we just say that any set $R \subseteq X_1 \times \cdots \times X_k$ is a k-ary *relation* on $X_1, \ldots, X_k$. We also say that $x_1 \in X_1, \ldots, x_k \in X_k$ are in the relation R iff $(x_1, \ldots, x_k) \in R$. If $k = 2$ such a relation is called a *binary relation* and we write xRy if x and y are in the relation R. If $X_1 = \cdots = X_k = X$, we say that R is a k-ary relation on X.

Note that $=$, $\leq$, $\geq$, $<$, and $>$ define relations on $\mathbb{R}$ (or any subset S of $\mathbb{R}$). For example, if $S = \{0, 1, 2\}$, then $<$ defines the relation $R = \{(0,1), (0,2), (1,2)\}$.

Another widely used family of relations on $\mathbb{Z}$ can be defined as follows. Let $n, a, b \in \mathbb{Z}$. If n divides $a - b$, we say that "a equivalent to b modulo n" and denote it as $a \equiv b \pmod{n}$. For example, 1 and 4 are equivalent modulo 3 since 3 divides $1 - 4 = -3$.

7.1 Equivalence Relations

The definition of a relation is way too broad. Hence, quite often we consider some types of relation. Probably the most interesting type of the relations is equivalence relations.

Definition 7.1. *Let R be a binary relation on a set X. We say that R is an* equivalence relation *if it satisfies the following conditions:*

(reflexivity) xRx *for any* $x \in X$*;*

(symmetry) xRy *iff* yRx *for any* $x, y \in X$*;*

(transitivity) for any $x, y, z \in X$*, if* xRy *and* yRZ*, then* xRz*.*

One may guess that the equivalence relation are mimicking $=$, so it is not a surprise that $=$ is an equivalence relation.

The definition seems quite bizarre, however, all of you are already familiar with another important example: you know that equivalent fractions represent the same number. For example, $\frac{2}{4}$ is the same as $\frac{1}{2}$. Let us consider this example more thorough, let S be a set of symbols of the form $\frac{x}{y}$ (note that it is not a set of numbers) where $x, y \in Z$ and $y \neq 0$. We define a binary relation R on S such that $\frac{x}{y}$ and $\frac{z}{w}$ are in the relation R iff $xw = zy$. It is easy to prove that this relation is an equivalence relation.

(reflexivity) Let $\frac{a}{b} \in S$. Since $ab = ab$, we have that $\frac{a}{b} R \frac{a}{b}$.

(symmetry) Let $\frac{a}{b}, \frac{c}{d} \in S$. Suppose that $\frac{a}{b} R \frac{c}{d}$, by the definition of R, it implies that $ac = db$. As a result, $\frac{c}{d} R \frac{a}{b}$.

(transitivity) Let $\frac{a}{b}, \frac{c}{d}, \frac{e}{f} \in S$ with $\frac{a}{b} R \frac{c}{d}$ and $\frac{c}{d} R \frac{e}{f}$. Then $ad = cb$ and $cf = ed$. The first equality can be rewritten as $c = ad/b$. Hence, $adf/b = ed$ and $af = eb$ since $d \neq 0$. So $\frac{a}{b} R \frac{e}{f}$.

Partitions

Let S be some set. We say that $\{P_1, \ldots, P_k\}$ form a partition of S iff P_1, $\ldots$, P_k are pairwise disjoint and $P_1 \cup \cdots \cup P_k = S$; in other words, a partition is a way of dividing a set into overlapping pieces.

Exercise 7.1. *Let $\{P_1, \ldots, P_k\}$ be a partition of a set S and R be a binary relation of S such that aRb iff $a, b \in P_i$ for some $i \in [k]$. Show that R is an equivalence relation.*

This exercise shows that one may transform a partition of the set S into an equivalence relation on S. However, it is possible to do the opposite.

Theorem 7.1. *Let R be a binary equivalence relation on a set S. For any element $x \in S$, define $R_x = \{y \in S : xRy\}$ (the set of all the elements of S related to x) we call such a set the equivalence class of x. Then $\{R_x : x \in S\}$ is a partition of S.*

Exercise 7.2. *Prove Theorem* **??**.

Modular Arithmetic

The relation "$\equiv \pmod{n}$" is actively used in the number theory. One of the important properties of this relation is that it is an equivalence relation.

Theorem 7.2. *The relation $\equiv \pmod{n}$ is an equivalence relation.*

Proof. To prove this statement we need to prove all three properties: reflexivity, symmetry, and transitivity.

(reflexivity) Note that for any integer x, $x - x = 0$ is divisible by any integer including n. Hence, $x \equiv x \pmod{n}$.

(symmetry) Let us assume that $x \equiv y \pmod{n}$; i.e., $x - y = kn$ for some integer k. Note that $y - x = (-k)n$, so $y \equiv x \pmod{n}$.

(transitivity) finally, assume that $x \equiv y \pmod{n}$ and $y \equiv z \pmod{n}$; i.e. $x - y = kn$ and $y - z = \ell n$ for some integers k and ℓ. It is easy to note that $x - z = (x - y) + (y - z) = (k + \ell)n$. As a result, $x \equiv z \pmod{n}$.

Thus, we proved that $\equiv \pmod{n}$ is an equivalence relation. □

Let $x \in \mathbb{Z}$; we denote by $r_{x,n}$ the equivalence class of x with respect to the relation $\equiv \pmod{n}$, we also denote by $\mathbb{Z}/n\mathbb{Z}$ the set of all the equivalence classes with respect to the relation $\equiv \pmod{n}$.

Another important property of these relations is that they behave well with respect to the arithmetic operations.

Theorem 7.3. *Let $x, y \in Z$ and $n \in \mathbb{N}$. Suppose that $a \in r_{x,n}$ and $b \in r_{y,n}$, then $(a + b) \in r_{x+y,n}$ and $ab \in r_{xy,n}$.*

Using this theorem we may define arithmetic operations on the equivalence classes with respect to the relation $\equiv \pmod{n}$. Let $x, y \in \mathbb{Z}$ and $n \in \mathbb{N}$. Then $r_{x,n} + r_{y,n} = \{a + b \ : \ a \in r_{x,n}, b \in r_{y,n}\} = r_{x+y,n}$ and $r_{x,n}r_{y,n} = \{ab \ : \ a \in r_{x,n}, b \in r_{y,n}\} = r_{xy,n}$. Moreover, these operations have plenty of good properties.

Exercise 7.3. *Let $a, b, c \in \mathbb{Z}/n\mathbb{Z}$. Show that the following equalities are true:*

- $a + (b + c) = (a + b) + c$,
- $a + r_{0,n} = a$ *(thus we denote $r_{0,n}$ as 0)*,
- $ar_{1,n} = a$ *(thus we denote $r_{1,n}$ as 1)*,
- *there is a class $d \in \mathbb{Z}/n\mathbb{Z}$ such that $a + d = r_{0,n}$ (thus we denote this d as $-a$)*,
- $a + b = b + a$,
- $ab = ba$,
- $a(b + c) = ab + ac$,

7.2 *Partial Orderings*

In the previous section we discussed a generalization of "$=$". In this section we are going to give a way to analyze relations similar to "$<$".

Definition 7.2. *A binary relation R on S is a partial ordering if it satisfies the following constraints.*

(reflexivity) xRx *for any* $x \in S$*;*

(antisymmetry) *if* xRy *and* yRx*, then* $x = y$ *for all* $x, y \in S$*;*

(transitivity) *for any* $x, y, z \in S$*, if* xRy *and* yRZ*, then* xRz*;*

We say that a partial ordering R on a set S is total iff for any $x, y \in S$, either xRy or yRx.

Note that $\leq$ defines a partial ordering on any $S \subseteq \mathbb{R}$; moreover, it defines a total order.

Typically we use symbols similar to $\preceq$ to denote partial orderings and we write $a \prec b$ to express that $a \preceq b$ and $a \neq b$.

Let $|$ be the relation on $\mathbb{Z}$ such that $d \mid n$ iff d divides n.

Theorem 7.4. *The relation $|$ is a partial ordering of the set $\mathbb{N}$.*

Proof. To prove that this relation is a partial ordering we need to check all three properties.

(reflexivity) Note that $x = 1 \cdot x$ for any integer x; hence, $x \mid x$ for any integer x.

(antisymmetry) Assume that $x \mid y$ and $y \mid x$. Note that it means that $kx = y$ and $\ell y = x$ for some integers k and ℓ. Hence, $y = (k \cdot \ell)y$ which implies that $k \cdot \ell = 1$ and $k = \ell = 1$. Thus, $x = y$.

(transitivity) finally, assume that $x \mid y$ and $y \mid z$; i.e., $kx = y$ and $\ell y = z$. As a result, $(k \cdot \ell)x = z$ and $x \mid z$.

□

Exercise 7.4. *Let S be some set, show that $\subseteq$ defines a partial ordering on the set 2^S.*

Topological Sorting

Partial orderings are very useful for describing complex processes. Suppose that some process consists of several tasks, T denotes the set of these tasks. Some tasks can be done only after some others e.g. when you cooking a salad you need to wash vegetables before you chop them. If $x, y \in T$ be some tasks, $x \preceq y$ if x should be done before y and this is a partial ordering.

In the applications this order is not a total order because some steps do not depend on other steps being done first (you can chop tomatoes and chop cucumbers in any order). However, if we need to create a schedule in which the tasks should be done, we need to create a total ordering on T. Moreover, this order should be compatible with the partial ordering. In other words, if $x \preceq y$, then $x \preceq_t y$ for all $x, y \in T$, where $\preceq_t$ is the total order. The technique of finding such a total ordering is called *topological sorting*.

Theorem 7.5. *Let S be a finite set and $\preceq$ be a partial order on S. Then there is a total order $\preceq_t$ on S such that if $x \preceq y$, then $x \preceq_t y$ for all $x, y \in S$*

This sorting can be done using the following procedure.

- Initiate the set S being equal to T

- Choose the minimal element of the set S with respect to the ordering $\preceq$ (such an element exists since S is a finite set, see Chapter **??**). Add this element to the list, remove it from the set S, and repeat this step if $S \neq \emptyset$.

Let us consider the following example. In the left column we list the classes and in the right column the prerequisite.

Courses	Prerequisite
Math 20A	
Math 20B	Math 20A
Math 20C	Math 20B
Math 18	
Math 109	Math 20C, Math 18
Math 184A	Math 109

We need to find an order to take the courses.

1. We start with

$$S = \{\text{Math 20A}, \text{Math 20B}, \text{Math 20C}, \text{Math 18}, \text{Math 109}, \text{Math 184}\}.$$

There are two minimal elements: Math 20A and Math 18. Let us remove Math 18 from S and add it to the resulting list R.

2. Now we have

$$R = \text{Math 18}$$

and

$$S = \{\text{Math 20A}, \text{Math 20B}, \text{Math 20C}, \text{Math 109}, \text{Math 184}\}.$$

There is only one minimal element Math 20A. We remove it and add it to the list R.

3. On this step
$$R = \text{Math 18}, \text{Math 20A}$$
and
$$S = \{\text{Math 20B}, \text{Math 20C}, \text{Math 109}, \text{Math 184}\}.$$
Again there is only one minimal element: Math 20B.

4.
$$R = \text{Math 18}, \text{Math 20A}, \text{Math 20B}$$
and
$$S = \{\text{Math 20C}, \text{Math 109}, \text{Math 184}\}.$$
There is only one minimal element: Math 20C.

5.
$$R = \text{Math 18}, \text{Math 20A}, \text{Math 20B}, \text{Math 20C}$$
and
$$S = \{\text{Math 109}, \text{Math 184}\}.$$
There is only one minimal element: Math 109.

6. Finally,
$$R = \text{Math 18}, \text{Math 20A}, \text{Math 20B}, \text{Math 20C}, \text{Math 109}$$
and
$$S = \{\text{Math 184}\}.$$
There is only one minimal element: Math 184A.

As a result, the final list is

$$R = \text{Math 18}, \text{Math 20A}, \text{Math 20B}, \text{Math 20C}, \text{Math 109}, \text{Math 184A}.$$

End of The Chapter Exercises

7.5 *(recommended)* Show that the relation $|$ does not define a partial ordering on $\mathbb{Z}$.

7.6 Let a relation R be defined on the set of real numbers as follows: xRy iff $2x + y = 3$. Show that it is antisymmetric.

7.7 Are there any minimal elements in $\mathbb{N}$ with respect to $|$? Are there any maximal elements?

8. Structural Induction

To illustrate the notions we introduce in this chapter let us consider the following game: Alice have chosen a number from 1 to 1000. Bob wants to guess the number so he is asking Alice "yes" or "no" questions. How many questions does Bob need to ask to determine the number in the worst-case scenario?

The following simple algorithm allows Bob to learn the number using 10 questions.[1]

1. Bob start with two numbers $\ell = 0$ and $u = 1000$.

2. Bob asks whether the Alice's number is at most $(\ell + u)/2$. If the answer is yes, then Bob replace u by $(\ell + u)/2$; otherwise Bob replace ℓ by $(\ell + u)/2$.

3. If there is only one integer x between ℓ and u, Bob says that Alice's number is x. Otherwise Bob goes to step **??**

Now we have two problems:

- we need to prove that the algorithm is correct and

- we would like to prove that it is impossible to guess the number using less questions.

To study both these problems we need a formal definition of an algorithm that Bob may use. However, the first problem is somewhat less demanding and the solution can be explained without a precise definition.

First note that the Alice's number is always between ℓ and u. Hence, the answer given at step **??** is always correct. You may also notice that on each step $u - \ell$ decreases by 2; hence, after $\lceil \log_2 1000 \rceil = 10$ questions $u - \ell < 1$. As a result, after 10 questions, there is only one integer between ℓ and u.

Exercise 8.1. *Give a nonadaptive algorithm for Bob that allows him to guess the number using* 10 *queries. In other words, write* 10 *questions such that answers to these questions allow Bob to guess the number.*

[1] This algorithm is based on the same ideas as **??**.

8.1 Recursive Definitions

First note that any question for Alice can be formulated as follows: "Is the value of a function f at your number equal to T?", where f is a function from $\mathbb{Z}$ to $\{0,1\}$ (here and in the sequel we interpret 1 as "yes" and 0 as "no").

Hence, there are two possible behaviours of any algorithm for Bob.

- The algorithm prescribes Bob to just say that the answer is some number $x \in \mathbb{Z}$, or
- The algorithm prescribes Bob to ask whether f at Alice's number is equal to T. If the answer is yes, then the algorithm prescribes Bob to proceed according to an algorithm A_1, otherwise the algorithm prescribe Bob to proceed according to an algorithm A_0.

Hence, any algorithm for Bob can be described using the following object.

Definition 8.1. *We say that T is a B-decision tree if*

(base case) either T is equal to an integer, or

(recursion step) T is equal to (f, T_0, T_1), where $f : \mathbb{Z} \to \{0,1\}$, and T_0 and T_1 are B-decision trees.

Note that this definition is not quite formal since it is recursive and we usually do not allow recursive definitions. So we will need to give a more formal way to define B-decision trees. However, this definition allows us to prove that $(f_1, (f_2, 1, 2), (f_3, 3, 4))$ is a B-decision tree, where

$$f_1(x) = \begin{cases} 1 & \text{if } x \leq 2 \\ 0 & \text{otherwise} \end{cases},$$

$$f_2(x) = \begin{cases} 1 & \text{if } x = 1 \\ 0 & \text{otherwise} \end{cases},$$

and

$$f_3(x) = \begin{cases} 1 & \text{if } x = 3 \\ 0 & \text{otherwise} \end{cases}.$$

This can be explained as follows.

- It is clear that 1, 2, 3, and 4 are B-decision trees by the base case.
- Hence, by the recursion step case, $(f_2, 1, 2)$ and $(f_3, 3, 4)$ are B-decision trees.

- Finally, by the recursion step case, $(f_1, (f_2, 1, 2), (f_3, 3, 4))$ is a B-decision tree.

In other words we proved that $(f_1, (f_2, 1, 2), (f_3, 3, 4))$ is a B-decision tree by providing $T_1, \dots, T_7$ such that $T_7 = (f_1, (f_2, 1, 2), (f_3, 3, 4))$ and for each $i \in [7]$, T_i is a number from $\mathbb{Z}$ or $T_i = (f, T_j, T_k)$ for $j, k < i$ and $f : \mathbb{Z} \to \{0, 1\}$.

This idea leads to the framework that would allows us to give a formal definition of B-decision trees.

Definition 8.2. *Let U be a set, let $S_0 \subseteq U$, and let*

$$\mathcal{F} = \left\{ F_1 : U^{\ell_1} \to U, \dots, F_n : U^{\ell_n} \to U, \dots \right\}.$$

Then we say that the set S is generated by $\mathcal{F}$ from S_0 if it is the set of all $u \in U$ such that there is a sequence $u_1, \dots, u_m$ satisfying the following constraints: $u_m = u$ and for each $i \in [m]$,

- *either $u_i \in S_0$, or*
- *$u_i = F(u_{k_1}, \dots, u_{k_\ell})$ for $F \in \mathcal{F}$ and $k_1, \dots, k_\ell < i$.*

In case of B-decision trees U is the set of all sequences of numbers, parentheses, commas, and functions from $\mathbb{Z}$ to $\{0, 1\}$, S_0 is the set of all sequences consisting of one number, and

$$\mathcal{F} = \left\{ F_f \ : \ f \text{ is a function from } \mathbb{Z} \text{ to } \{0, 1\} \right\},$$

where $F_f(T_0, T_1) = (f, T_0, T_1)$.

Remark 8.1. *Let U be a set, let $S_0 \subseteq U$, and let*

$$\mathcal{F} = \left\{ F_1 : U^{\ell_1} \to U, \dots, F_n : U^{\ell_n} \to U, \dots \right\}.$$

Let us consider the sets $S_0, \dots, S_n, \dots \subseteq U$ such that

$$S_{i+1} = S_i \cup \{ F(u_1, \dots, u_\ell) \ : \ u_1, \dots, u_\ell \in S_i, F \in \mathcal{F} \}.$$

It is clear that $\bigcup_{i \geq 0} S_i$ is the set generated by $\mathcal{F}$ from S_0.

Using similar ideas we may define some functions on the objects defined recursively.

Definition 8.3. *Let T be a B-decision tree. Then the height $h(T)$ of T can be defined as follows.*

(base case) *If T is equal to a number $x \in \mathbb{Z}$, then $h(T) = 0$.*

(recursion step) *If T_1 and T_2 are B-decision trees, then $h\big((f, T_1, T_2)\big) = \max(h(T_1), h(T_2)) + 1$.*

Note that $h(T)$ correpsonds to the worst-case number of queries made by Bob if we interpret T as a description of an algorithm.

However, before we explain how to formalize such a definition, we need to note that in the general case such definition may be contradictory. Consider $U = \mathbb{R}$, $S_0 = \{0\}$, and $\mathcal{F} = \{f, g\}$, where $f(x, y) = xy$ and $g(x) = x + 1$. We define $v : U \to \mathbb{R}$ as follows.

(base case) $v(0) = 0$.

(recursion step) $v(f(x, y)) = f(v(x), v(y))$ and $v(g(x)) = v(x) + 2$.

Note that $v(f(g(0), g(0))) = f(v(g(0)), v(g(0))) = \big(v(g(0))\big)^2 = 4$ and $v(g(0)) = v(0) + 2 = 2$. However, $g(0) = 1$ and $f(g(0), g(0)) = 1$.

Therefore handle such an issue, we consider S that is *freely* generated from S_0 by by $\mathcal{F}$.

Definition 8.4. *The set S is freely generated by $\mathcal{F}$ from S_0 iff it is generated by $\mathcal{F}$ from S_0, $S_0 \cap \operatorname{Im} F = \varnothing$, and $\operatorname{Im} F \cap \operatorname{Im} G = \varnothing$ for any $F, G \in \mathcal{F}$.*

The following theorem claims existence of functions defined recursively.

Theorem 8.1. *Let $S \subseteq U$ be the set freely generated from $S_0 \subseteq U$ by $\mathcal{F} = \Big\{F_1 : U^{\ell_1} \to U, \dots, F_n : U^{\ell_n} \to U, \dots\Big\}$. In addition, let $G_0 : S_0 \to V$ and $G_1 : V^{\ell_1} \to V, \dots, G_n : V^{\ell_n} \to V, \dots$ be some functions.*

Then there is a function $h : S \to V$ such that

(base case) $h(u) = G_0(u)$ *for any* $u \in S_0$.

(recursion step) $h(F_i(u_1, \dots, u_{\ell_i})) = G_i(h(u_1), \dots, h(u_{\ell_i}))$ *for any i and $u_1, \dots, u_{\ell_i} \in S$.*

Exercise 8.2. *Prove* **??**.

B-decision tree are used in this chapter to represent Bob's algorithms; hence, we need to define values of B-decision trees.

Definition 8.5. *The value $\operatorname{val}(T, n)$ of a B-decision tree T at an integer n can be defined as follows.*

(base case) If T is just a number $x \in [1000]$, then the result $\operatorname{val}(T, n)$ is equal to x.

(recursion step) If T_1 and T_2 are B-decision trees, then

$$\operatorname{val}\Big((f, T_0, T_1), n\Big) = \begin{cases} \operatorname{val}(T_0, n) & \text{if } f(x) = 0 \\ \operatorname{val}(T_1, n) & \text{otherwise} \end{cases}.$$

Using all these notions we can reformulate the results about Alice and Bob's game.

Theorem 8.2. *1. There is a B-decision tree T such that*

- $h(T) \leq 10$ *and*
- $\operatorname{val}(T, n) = n$ *for any* $n \in [1000]$.

2. Let T be a B-decision tree such that $\operatorname{val}(T, n) = n$ *for any* $n \in [1000]$. *Then* $h(T) \geq 10$.

Exercise 8.3. *Prove the first part of* **??**.

8.2 Structural Induction Theorem

To prove the second part of **??** we need to introduce the notion of structural induction.

Theorem 8.3 (The Structural Induction Principle). *Let* $S \subseteq U$ *be the set freely generated from* $S_0 \subseteq U$ *by* $\mathcal{F} = \left\{F_1 : U^{\ell_1} \to U, \ldots, F_n : U^{\ell_n} \to U, \ldots\right\}$.

Assume that $S' \subseteq U$ *is a set such that the following constraints are true.*

(base case) $S_0 \subseteq S'$

(induction step) $F_i(u_1, \ldots, u_{\ell_i}) \in S'$ *for any* $u_1, \ldots, u_{\ell_i} \in S'$ *and* $i \in \mathbb{N}$.

Then $S \subseteq S'$.

Using this result, we may prove **??**.

Proof of **??**. We need to prove only the second part of the statement. Let $V(T) = \{\operatorname{val}(T, n) : n \in \mathbb{Z}\}$, where T is a B-decision tree. This proof is based on the following observation.

Claim 8.3.1. *For any B-decision tree T, size of* $V(T)$ *is at most* $2^{h(T)}$.

Assume that T is a B-decision tree such that $\operatorname{val}(T, n) = n$ for any $n \in [1000]$. Whence $[1000] \subseteq V(T)$. Therefore $V(T)$ has at least 1000 elements. As a result, $2^{h(T)} \geq 1000$; which implies that $h(T) \geq 10$.

So we just need to prove **??**. We prove it using structural induction. Let S' be the set of decision trees T such that size of $V(T)$ is at most $2^{h(T)}$.

(base case) If T is equal to an integer x, then $\operatorname{val}(T, n) = x$ for any $n \in \mathbb{Z}$. Hence, the size of $V(T)$ is equal to $1 = 2^0 = 2^{h(T)}$.

(induction step) Assume that $T = (f, T_0, T_1)$ for some $T_0, T_1 \in S'$. We know that the size of $V(T_0)$ is at most $2^{h(T_0)}$ and the size of $V(T_0)$ is at most $2^{h(T_1)}$. In addtion, it is clear that $V(T) \subseteq V(T_0) \cup V(T_1)$. Therefore, the size of $V(T)$ is at most[2]

$$2^{h(T_0)} + 2^{h(T_1)} \leq 2^{\max(h(T_0), h(T_1))+1} = 2^{h(T)}.$$

[2] Formally speaking, we use **??** to prove this; i.e. we use the fact that the size of a set $A \cup B$ is at most the size of A plus the size of B.

Hence, by **??**, S' is equal to the set of all B-decision trees. □

Now we are ready to prove **??**.

Proof of **??**. We prove the statement using induction. More precisely, we prove using induction by m that if there is a sequence $u_1, \dots, u_m$ such that for each $i \in [m]$, $u_i \in S_0$ or $u_i = F(u_{k_1}, \dots, u_{k_\ell})$ for $F \in \mathcal{F}$ and $k_1, \dots, k_\ell < i$, then $u_m \in S'$.

The case when $m = 1$ is clear since in this case $u_1 \in S_0$ which implies that it is in S'.

Let us now prove the induction step. Assume that the statement is true for any $k \leq m$. Consider a sequence $u_1, \dots, u_{m+1}$ such that for each $i \in [m+1]$, $u_i \in S_0$ or $u_i = F(u_{k_1}, \dots, u_{k_\ell})$ for $F \in \mathcal{F}$ and $k_1, \dots, k_\ell < i$. Let us consider $F \in \mathcal{F}$ and $k_1, \dots, k_\ell < m+1$ such that $u_{m+1} = f(u_{k_1}, \dots, u_{k_\ell})$. By the induction hypothesis, $u_{k_1}, \dots, u_{k_\ell} \in S'$. Therefore, by the properties of S', $u_{m+1} \in S'$. □

End of The Chapter Exercises

8.4 *(recommended)* Let $S \subseteq \mathbb{Z}$ be a set of size at least 2^k, and let T be a decision tree such that $\mathrm{val}(T, n) = n$ for any $n \in S$. Show that $h(T) \geq k$.

8.5 *(recommended)* Let S be a set of integers defined recursively such that

- $3 \in S$, and
- if $x, y \in S$, then $x + y \in S$.

Show that $S = \{3k : k \in \mathbb{N}\}$.

8.6 *(recommended)* Using recursive definition we can define an arithmetic formula on the variables $x_1, \dots, x_n$,

(base case) x_i is an arithmetic formula on the variables $x_1, \dots, x_n$ for all i; if c is a real number, then c is also an arithmetic formula on the variables $x_1, \dots, x_n$.

(recursion step) If P and Q are arithmetic formulas on the variables $x_1, \dots, x_n$, then $(P + Q)$ and $P \cdot Q$ are arithmetic formulas on the variables $x_1, \dots, x_n$.

We can also define recursively the value of such a formula. Let $v_1, \dots, v_n$ be some integers.

(base cases) $x_i\big|_{x_1=v_1,\dots,x_n=v_n} = v_i$; in other words, the value of the arithmetic formula x_i is equal to v_i when $x_1 = v_1, \dots, x_n = v_n$; if c is a real number, then $c|_{x_1=v_1,\dots,x_n=v_n} = c$.

(recursion steps) If P and Q are arithmetic formulas on the variables $x_1, \ldots, x_n$, then

$$(P+Q)\big|_{x_1=v_1,\ldots,x_n=v_n} = P\big|_{x_1=v_1,\ldots,x_n=v_n} + Q\big|_{x_1=v_1,\ldots,x_n=v_n}$$

and

$$(P \cdot Q)\big|_{x_1=v_1,\ldots,x_n=v_n} = P\big|_{x_1=v_1,\ldots,x_n=v_n} \cdot Q\big|_{x_1=v_1,\ldots,x_n=v_n}.$$

Prove that for any arithmetic formula A on x, there is a polynomial p such that $p(v) = A\big|_{x=v}$ for any $v \in \mathbb{R}$.

8.7
- Define arithmetic formulas with division and define their value (make sure that you handled divisions by 0).
- Show that for any arithmetic formula with division A on x, there are polynomials p and q such that $\frac{p(v)}{q(v)} = A\big|_{x=v}$ or $A\big|_{x=v}$ is not defined for any real value v.

8.8 *(recommended)* We say that L is a B-decision list

(base case) if either L is a number $y \in \mathbb{Z}$, or

(recursion step) L is equal to (f, v, L') where $f : \mathbb{Z} \to \{0,1\}$, $v \in \mathbb{Z}$, and L is a B-decision list.

We can also define the value $\mathrm{val}(L, x)$ of a B-decision list L at $x \in \mathbb{Z}$.

(base case) If L is a number y, then $\mathrm{val}(L, x) = y$, and

(recursion step) if $L = (f, v, L')$, then

$$\mathrm{val}(L, x) = \begin{cases} v & \text{if } f(x) = 1 \\ \mathrm{val}(L', x) & \text{otherwise} \end{cases}.$$

Similarly one may define the length $\ell(L)$ of a B-decition list L.

(base case) If L is a number y, then $\ell(L) = 1$, and

(recursion step) if $L = (f, v, L')$, then $\ell(L) = \ell(L') + 1$.

Assume that $\mathrm{val}(L, x) = x$ for any $x \in [1000]$ show that $\ell(L) \geq 1000$.

8.9 Let us consider the following game modification of the game studied in this chapter. Alice have chosen a number from 1 to 1000. Bob wants to guess the number so he is asking Alice "yes" or "no" questions. However, he cannot get more than ℓ "yes" answers; i.e., as soon as Alice says ℓth "yes", Bob is supposed to be able to guess her number.

1. Show that if $\ell = 1$, then Bob needs at least 1000 quesitons.
2. Show that if $\ell = 2$, then Bob needs at least $\sqrt{1000}$ quesions.

3. Show that if Bob needs at least $\sqrt[\ell]{1000}$ questions.
4. Show that if $\ell = 2$, there is an algorithm for Bob such that he is able to guess the number using $2\left\lceil\sqrt{1000}\right\rceil$ questions.

8.10 You have a 100-story building and two eggs. When you drop an egg from any floor of the building, the egg will either break or survive the fall. If the egg survives, then it would have survived any lesser fall. If the egg breaks, then any greater fall would have broken it as well. The eggs are all identical and interchangeable. You'd like to find the minimum height that will break an egg. What is the fewest number of drops in which you are guaranteed to find the right floor?

Part II

Introduction to Combinatorial Game Theory

9. *P-positions and N-positions*

In this part we use our knowledge about basics of mathematical reasoning to study games similar to checkers, chess, shogi, and tic tac toe. The games we are going to study are called combinatorial games. In these games there are two players, each know all the information, there are no chance moves, and when the game ends there is always a winner. (The last condition implies that among the beforementioned games only checkers are combinatorial since all of them allow draws; however, we may change the rules to disallow the draws and this change would make all of them combinatorial.) Such a game is determined by a set of positions, and possible moves from each position for each player. Usually, players are taking turns until they reach a position such that no moves are possible and one of the player is declared a winner.

This part is based on Part I of "Game Theory" by Ferguson.

Combinatorial Games:
Introduction to Combinatorial Game Theory #1

`https://youtu.be/DbCKHPlMN2c`

9.1 *Take-Away Game*

Since chess, shogi and even tic tac toe are relatively complicated, we are going to start from much simpler example of combinatorial games.

Game 9.1 (Take-Away Game). *In this game there are two players.*

- *They have a pile of* 21 *chips.*
- *They make moves in turns with player I starting, each move consists of moving one, two or three chips out of the pile.*
- *The player that removes the last chip wins.*

The question we would like to answer is there a strategy for one of the players to always win? So in the rest of this part we assume that both players are playing optimally; i.e., if there is a winning strategy they follow the strategy.

To analyze this game we need the following two observations:

1. the game is symmetric and the only difference between the players is who makes the first move, and
2. if at some point the players have n chips it does not matter how they achieved this, it will not affect the rest of the game.

Using this remarks and induction (this style of induction is sometimes referred as *backward induction*) we are able to analyze the game.

Let us consider some certain states of the game. Assume that they have at most 3 chips left, in this case the player that make the move wins. However, if there are 4 chips, the player that makes the first move should always take at least 1 chip so she loses since after her turn there are at most 3 chip. Similarly, if there are 5 chips, the first player to move wins since she can take a chip and make the second player to start with 4.

So we can formulate the following conjecture. Assume that n chips left in the pile. Let r be the reminder of n modulo 4. Then if $r = 0$, the first player to move loses, otherwise, the other player loses.

Let us prove this using induction. We already proved the base case so we need to prove the induction step from n to $n + 1$.

- If $n \equiv 0 \pmod 4$, then the first player to move can remove one chip and the other player will start with n chips so by the induction hypothesis he/she loses.

- If $n \equiv 1 \pmod 4$. then the first player to move can remove two chips and the other player will start with n chips so by the induction hypothesis he/she loses.

- If $n \equiv 2 \pmod 4$. then the first player to move can remove three chips and the other player will start with n chips so by the induction hypothesis he/she loses.

- If $n \equiv 3 \pmod 4$. then after the current player move the other player will start with either n, or $n - 1$, or $n - 2$ chips. But all these numbers have non-zero reminders modulo 4. So the other player can win in any case.

To study combinatorial games we need to give a formal definition of them.

Definition 9.1. *A game is combinatorial if*

- *there are two players,*

- *there is a set of possible positions in the game,*

- *for each position and each player, there is a fixed set of possible legal moves,*

- *players alternate moving,*

- *the game ends when no moves are possible for the player whose turn is to move.*

There are possible winning conditions, *normal play rule: the player that made the last move wins, and*

misère play rule: the player that made the last move loses.

If the game never ends, we declare a draw. If the game always ends, we say that the game satisfies the ending condition*.*

If the possible moves are the same for both players the game is called impartial *otherwise it is called* partisan*.*

Note that these games do not allow random moves, hidden information, simultaneous moves, and a draw in a finite number of steps so poker, battleships, rock-paper-scissors, and tick tack toe are not combinatorial games.

P-positions and N-positions: Introduction to Combinatorial Game Theory #2

`https://youtu.be/YV_oWBi1_ck`

Since we gave a formal definition of combinatorial games we can give a framework that allows to analyze these games.

Definition 9.2. *We say that a position in a combinatorial game is* terminal *if there are no legal moves.*

All terminal positions are P-positions*. Every position that allows for the current player to move to a P-position is an* N-position*. If all possible moves lead to N-positions, then the position is a P-position.*

For the game using the Misère rule, the definition is the same except the terminal positions are N-positions.

Using this definition, one may create the following procedure that would allow to determine which positions are P-positions and which are N-positions.

Steps necessary to find P- and N-positions

1. Label all terminal positions as P-positions.
2. If some position is not labeled but all the moves lead to labeled positions, then label the position using the definition; i.e., if there are moves leading to P-position, it is an N-position, otherwise it is a P-position.
3. If not all the positions are labeled, go to Step 2.

Note that P- and N-positions are defined recursively so in some games not all the positions are either P- or N-positions. (For example, if there are no terminal positions.) However, **??** proves that if the game satisfies the ending condition, then all the positions are either P- or N-positions.

0	1	2	3	4	5	6	7	8
P	N	N	N	P	N	N	N	P

Table 9.1: P-positions and N-positions for **??**

So in **??** the only terminal position is 0; hence, 0 is a P-position. Similarly we can go to 0 from 1, 2, and 3 so they are N-positions. Hence, 3 is a P-position, since all the moves from 4 lead to N-positions.

Exercise 9.1. *Show that a position n is a P-position if 4 divides n, and it is an N-position otherwise.*

In other words, in this game, P-positions coincide with the positions where the current player loses. However, it is not a coincidence.

Theorem 9.1. *If some position in a combinatorial game is an N-position, then the player to move has a winning strategy if we start from this position. If the position is a P-position, then the other player has a winning strategy.*

Subtractraction Games

Let us define a big class of games that generalizes the take-away game we discussed at the beginning of the chapter.

Game 9.2. *Let $S \subseteq \mathbb{N}$ be some set. The subtraction game with the subtraction set S is the following combinatorial game. Two players start with a pile of n chips. On each move they remove $s \in S$ chips out of the pile.*

So **??** is the subtraction game with the subtraction set $\{1, 2, 3\}$.

Let us analyze the subtraction game with the subtraction set $\{1, 3, 4\}$.

0	1	2	3	4	5	6	7	8
P	N	P	N	N	N	N	P	N

Table 9.2: P-positions and N-positions for **??**

Clearly 0 is a P-position since it is the only terminal position in the game. We can go to 0 from 1 so 1 is an N-position. The only possible move from 2 is to 1 so 2 is a P-position. From 3 and 4 we can go to 0 so they are N-positions. From 5 and 6 one may go to 2 so they are a N-positions as well. Hence, 7 is a P-position.

Now we may notice the pattern: n is a P-position iff $n \equiv 0 \pmod 7$ or $n \equiv 2 \pmod 7$. We prove is using induction. The base case for $n < 8$ we already proved. Let us now prove the induction step. Assume that the statement is true for all $k < n$. Consider the following cases.

1. If $n \equiv 0 \pmod 7$, the current player can move to $n - 1 \equiv 5 \pmod 7$, $n - 3 \equiv 4 \pmod 7$, or $n - 3 \equiv 5 \pmod 7$ which are all N-positions so n is a P-position.

2. If $n \equiv 1 \pmod 7$, the current player can move to $n-1$ which is a P-position so n is an N-position.

3. If $n \equiv 2 \pmod 7$, the current player can move to $n-1 \equiv 1 \pmod 7$, $n-3 \equiv 6 \pmod 7$, or $n-4 \equiv 5 \pmod 7$ which are all N-positions so n is a P-position.

4. If $n \equiv 3 \pmod 7$, the current player can move to $n-1$ which is a P-position so n is an N-position.

5. If $n \equiv 4 \pmod 7$, the current player can move to $n-4$ which is a P-position so n is an N-position.

6. If $n \equiv 5 \pmod 7$, the current player can move to $n-3$ which is a P-position so n is an N-position.

7. If $n \equiv 6 \pmod 7$, the current player can move to $n-4$ which is a P-position so n is an N-position.

End of The Chapter Exercises

9.2 Two players I and II are playing the following game.

- They start with a number 0 written on a blackboard.
- On each step one of the players replace a number n on the blackboard by either $n+1$ or by $n+2$.
- Player I makes the first move and players do moves one after another.
- The player who writes 20 wins.

Who has a winning strategy? (Note that the game is not a combinatorial game).

9.3 Two players I and II are playing the following game.

- Initially, there are 20 numbers written on a blackboard: 10 numbers 1 and 10 numbers 2.
- On each step one of the players select two numbers; and if they were the same, replace them by 2; otherwise, replace them by 1.
- Player I makes the first move and players do moves one after another.

Who is the winner? (Note that the game is not a combinatorial game).

9.4 Consider the subtraction game where players may subtract 2 and 3 chips on their turn, is 5 an N-position?

9.5 Consider the Misère subtraction game where players may subtract 1, 2 or 5 chips on their turn, identify N-positions and P-positions.

9.6 Consider the Misère subtraction game where players may subtract 1, 5 or 6 chips on their turn, identify N-positions and P-positions.

9.7 In the subtraction game where players may subtract 1, 2, or 5 chips on their turn, identify N-positions and P-positions.

9.8 Two players one by one put bishops on the chessboard such that none of the bishops attack each other. Determine the winning strategy.

9.9 Consider the following game: two players I and II are writing an 11-digit number from left to right, one digit after another. Player I wins if 7 divides the number and player II wins otherwise. Determine who is the winner if player I makes the first move.

10. The Game of Nim

This chapter discusses probably the most famous combinatorial game, the game of *Nim*. In this game there are several piles of chips on the table. On each turn the current player may remove some number of chips from *one* of the piles; however, the player should remove *at least one chip*. We say that a game of Nim is a k-pile game of Nim if there are k piles.

You can play Nim on this website

`https://dotsphinx.com/games/nim/`

We start from analysis of the game when we have one pile of chips. It is clear that the first player to move wins since he/she may remove all the chips.

Consider a more complicated case when we have two piles of size n and m respectively. We need to consider two cases:

1. If $n = m$, then the second player to move wins. Indeed, we can use the symmetric strategy; i.e., if the first player removes s chips from one pile we also remove s chips from the other pile. It is clear that we can always make a move as long as the first player can.

2. Otherwise, the first player wins because it can move to the state with two equal piles.

The case of three piles is even more complicated. So we spend the rest of the chapter studying it.

10.1 Nim Sum

We start from a definition of the XOR operation $\oplus : \{0, 1\} \times \{0, 1\} \to \{0, 1\}$, also known as "exclusive or"), this operation is defined as follows: $a \oplus b = 1$ iff $a \neq b$.

It is well-known that any number $n \in \mathbb{N}_0$ can be represented as a binary number ($\mathbb{N}_0$ denotes nonnegative integers); we write $n = (a_\ell, \ldots, a_0)_2$ if $n = \sum_{i=0}^{\ell} a_i 2^i$. For example, $5 = 4 + 1 = 1 \cdot 2^2 + 0 \cdot 2^1 + 1 \cdot 2^0 = (1, 0, 1)_2$ and $6 = 4 + 2 = 1 \cdot 2^2 + 1 \cdot 2^1 + 0 \cdot 2^0 = (1, 1, 0)_2$. So we can define the Nim sum $\oplus : \mathbb{N}_0 \times \mathbb{N}_0 \to \mathbb{N}_0$, also known as bitwise xor, as follows: $(a_\ell, \ldots, a_0)_2 \oplus (b_\ell, \ldots, b_0)_2 = (a_\ell \oplus b_\ell, \ldots, a_0 \oplus b_0)_2$. For example, $5 \oplus 6 = (1, 0, 1)_2 \oplus (1, 1, 0)_2 = (1 \oplus 1, 0 \oplus 1, 1 \oplus 0)_2 = (0, 1, 1)_2$.

Exercise 10.1. *Show that $a \oplus (b \oplus c) = (a \oplus b) \oplus c$ for any $a, b, c \in \mathbb{N}_0$.*

Hence, we are going to write $a \oplus b \oplus c$ instead of $a \oplus (b \oplus c)$ and $(a \oplus b) \oplus c$.

10.2 Bouton's Theorem

Now we may notice that $a \oplus b = 0$ iff $a = b$. So our result about 2-pile Nim can be rephrased: a position (a, b) in the 2-pile Nim is a P-position iff $a \oplus b = 0$. Which leads us to the next theorem.

Theorem 10.1 (Bouton). *A position (a, b, c) in the 3-pile Nim is a P-position iff $a \oplus b \oplus c = 0$*

Proof. We prove the statement using structural induction. First note that the only terminal position the 3-pile Nim is $(0, 0, 0)$ and $(0, 0, 0)$ and $0 \oplus 0 \oplus 0 = 0$.

Let us consider some (a, b, c) such that $a \oplus b \oplus c \neq 0$. We need to show that there is a move from this position to a P-position. Let $a \oplus b \oplus c = (0, \dots, 0, 1, r_{k-1}, \dots, r_0)_2$. So among a, b, and c there is a number that has 1 in the kth position. Note that without loss of generality $a = (p_\ell, \dots, p_{k+1}, 1, p_{k-1}, \dots, p_0)_2$. Consider $a' = (p_\ell, \dots, p_{k+1}, 0, p_{k-1} \oplus r_{k-1}, \dots, r_0 \oplus p_0)_2$. It is clear that $a' < a$ and $a' \oplus b \oplus c = 0$. Hence, (a', b, c) is a P-position and therefore, (a, b, c) is an N-position.

Finally, let us consider (a, b, c) such that $a \oplus b \oplus c = 0$. Assume that there is a move to a position (a', b, c) such that $a' \oplus b \oplus c = 0$. This implies that $(a' \oplus b \oplus c) \oplus (a \oplus b \oplus c) = a \oplus a' = 0$, whence $a = a'$. □

11. Graph Games

This section gives an alternative definition of a combinatorial game. This definition allows us to study general combinatorial games.

Definition 11.1. *A* directed graph *G is a pair (V, N) such that V is a non-empty set and $N : V \to 2^V$.*[1]

We say that a game on G is the impartial game where elements of V are positions and each player can move from $x \in V$ to any $y \in N(x)$. (Elements of $N(x)$ are called followers of x.)

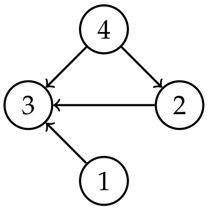
Remark 11.1. *It is also easy to see that any impartial game can be transformed into a graph G such that the game on G and the impartial game are equivalent.*

For example, the take-away game from **??** can be considered as a graph on a graph $G = (\mathbb{N}_0, N)$, where $N(0) = \emptyset$, $N(1) = \{0\}$, $N(2) = \{0, 1\}$, and $N(n+3) = \{n, n+1, n+2\}$ for any $n \in \mathbb{N}_0$.

The key ingredient for the analysis of games based on graphs was proposed by Sprague and Grundy. They proposed to consider the following function:

Definition 11.2. *Let $G = (V, N)$ be a directed graph. A function $g : V \to \mathbb{N}$ is a* Sprague–Grundy function *for G iff $g(x) = \text{mex } \{g(y) \ : \ y \in N(x)\}$, where* $\text{mex } S = \min \{n \in \mathbb{N}_0 \ : \ n \notin S\}$.

Consider the following graph (arrows depict possible moves).



Let us assume that g is a Sprague–Grundy function for this graph. Note that 3 is a terminal position so $g(3) = \text{mex } \emptyset = 0$. Since from 1 and 2 there are only moves to 3, it is clear that $g(1) = g(2) = \text{mex } \{0\} = 1$. Finally, $g(4) = \text{mex } \{0, 1\} = 2$.

Note that the Sprague–Grundy function is recursively defined so it may not exist or not to be unique if graph violets the ending condi-

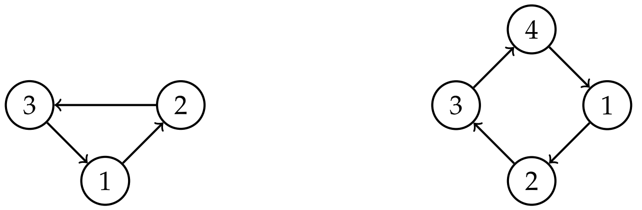
[1] We are going to have a more in-depth discussion of graphs in **??**.

tion. For example, the graph depicted on **??** does not have a Sprague–Grundy function. Indeed, assume that such a function g exists. Consider two following cases.

- First case is when $g(3) = 0$. Note that $g(2) = \text{mex }\{0\} = 1$. Hence, $g(1) = \text{mex }\{1\} = 0$ which contradicts the assumption that $g(3) = 0$ since $g(3) = \text{mex }\{g(0)\} = 1$.

- Second case is when $g(3) \neq 0$. Note that $g(2) = \text{mex }\{g(3)\} = 0$. Therefore $g(1) = \text{mex }\{0\} = 1$ and $g(3) = \text{mex }\{1\} = 0$ which is a contradiction.

Note that the graph depicted on **??** has several Sprague–Grundy functions. We may consider functions g_1 and g_2 such that $g_1(1) = g_1(3) = g_2(2) = g_2(4) = 0$ and $g_2(1) = g_2(3) = g_1(2) = g_1(4) = 1$. It is clear that they are Sprague–Grundy functions for the graph from **??**.



(a) A graph without a Sprague–Grundy function

(b) A graph with several Sprague–Grundy functions

Figure 11.1: Graphs where Sprague–Grundy function is either not unique or does not exist.

Unfortunately, even if there are no cycles, a graph my not have a Sprague–Grundy function or have several Sprague–Grundy functions. Indeed, consider the graph $G = (\mathbb{Z}, N)$ such that $N(x) = \{x - 1\}$. It is clear that the functions g_1 and g_2 such that

$$g_1(x) = \begin{cases} 0 & \text{if } x \text{ is even} \\ 1 & \text{if } x \text{ is odd} \end{cases}$$

and

$$g_2(x) = \begin{cases} 0 & \text{if } x \text{ is even} \\ 1 & \text{if } x \text{ is odd} \end{cases}$$

are Sprague–Grundy functions for G.

Exercise 11.1. *Let $G = (\mathbb{N} \cup \{\infty\}, N)$ such that $N(x) = \{y \in \mathbb{N}_0 : y < x\}$, and $N(\infty) = \mathbb{N}$. Show that G does not have a Sprague–Grundy function.*

However, all the combinatorial games we are going to consider the Sprague–Grundy function exists and is unique.

Theorem 11.1. *Let $G = (V, N)$ be a graph such that $N(v)$ is finite and G satisfies the ending condition. Then Sprague–Grundy function for G exists and it is unique.*

Before we prove this theorem, let us illustrate by proving that the Sprague–Grundy function for **??** is unique. Assume that a Sprague–Grundy function g for **??** exists. We are going to show that it is unique. Note that if x is a terminal position, then $g(x) = 0$. Hence, $g(0) = 0$. There is only one move from 1 so $g(1) = \text{mex } \{0\} = 1$. Similarly there are two moves from 2: one to 1 and one to 0 so $g(2) = \text{mex } \{0, 1\} = 2$. In the same way $g(3) = \text{mex } \{0, 1, 2\} = 3$ and $g(4) = \text{mex } \{1, 2, 3\} = 0$. One may notice that there is a pattern and conjecture that

$$g(x) = \begin{cases} 0 & \text{if } x \equiv 0 \pmod 4 \\ 1 & \text{if } x \equiv 1 \pmod 4 \\ 2 & \text{if } x \equiv 2 \pmod 4 \\ 3 & \text{if } x \equiv 3 \pmod 4 \end{cases}.$$

We already proved the base case, let us now prove the induction step. Assume the equality is true for all $y < x$ and consider the following cases.

- If $x \equiv 0 \pmod 4$, then $x - 1 \equiv 3 \pmod 4$, $x - 2 \equiv 2 \pmod 4$, and $x - 3 \equiv 1 \pmod 4$. Hence, $g(x) = \text{mex } \{1, 2, 3\} = 0$.

- If $x \equiv 1 \pmod 4$, similarly $g(x) = \text{mex } \{2, 3, 0\} = 1$.

- If $x \equiv 2 \pmod 4$, $g(x) = \text{mex } \{3, 0, 1\} = 2$.

- If $x \equiv 3 \pmod 4$, $g(x) = \text{mex } \{0, 1, 2\} = 3$.

It is also clear that the constructed function is indeed a Sprague–Grundy function for **??**. Therefore we proved that existence and uniqueness.

0	1	2	3	4	5	6	7	8
0	1	2	3	0	1	2	3	0

Table 11.1: The Sprague–Grundy function for **??**

Note that in this proof we used a procedure very similar to the one we used to find P- and N-positions.

Steps necessary to find a Sparague–Grundy function

Assume we are trying to construct a a Sparague–Grundy function g.

1. Set $g(x) = 0$ for all terminal positions x.
2. If $g(x)$ is not defined but $g(y)$ is defined for all $y \in N(x)$, then set $g(x) = \text{mex } \{g(y) \ : \ y \in N(x)\}$.
3. If $g(x)$ is not defined for some x, go to Step 2.

Let V_0 be the set of terminal positions, and let V_i be the set of x such that we define $g(x)$ on the ith iteration of the procedure. The following lemma eseentially says that this procedure defines g everywhere.

Lemma 11.1 (Kőnig)**.** *Let $G = (V, N)$ be a graph such that $N(v)$ is finite for all $v \in V$ and G satisfies the ending condition. Let V_0 be the set of terminal positions, and let $V_{n+1} = \{v \in V \ : \ N(v) \subseteq V_n\}$. Then $V = \bigcup_{n \in \mathbb{N}_0} V_n$.*

Proof of **??**. Let V_0 be the set of terminal positions, and let $V_{n+1} = \{v \in V \ : \ N(v) \subseteq V_n\}$. **??** claims that $V = \bigcup_{n \in \mathbb{N}_0} V_n$.

For each $n \in \mathbb{N}_0$, we define $g_n : V_n \to \mathbb{N}_0$ such that $g_{n+1}(v) = \text{mex } \{g_n(u) \ : \ u \in N(v)\}$ for $v \in V_{n+1}$ and $g_0(v) = 0$ for $v \in V_0$. It is easy to see that the function $g : V \to \mathbb{N}_0$ such that $g(v) = g_n(v)$ for $v \in V_n$ is a Sprague–Grundy function for G.

To finish the proof we need to prove that there are no other Sprague–Grundy functions for G. Assume that g' is a Sprague–Grundy function for G; we prove using induction by $n \in \mathbb{N}_0$ that $g(v) = g'(v)$ for $v \in V_n$. The base case for $v = 0$ is clear since $N(v) = \varnothing$ for $v \in V_0$. Let us consider the induction step from n to $n + 1$. By the induction hypothesis, $g(u) = g'(u)$ for $u \in V_n$. Let us consider some $v \in V_{n+1}$. Note that

$$g'(v) = \text{mex } \{g(u) \ : \ u \in N(v)\} = \text{mex } \{g(u) \ : \ u \in N(v)\} = g(v).$$

Therefore $g(v) = g'(v)$ for $v \in V_{n+1}$. As a result, $g(v) = g'(v)$ for $v \in V$; i.e., $g = g'$. □

One may note that in **??** P-positions are the positions where the Sprague–Grundy function is zero. In fact, this is not a coincidence.

Theorem 11.2. *Let $G = (V, N)$ be a graph such that $N(v)$ is finite for all $v \in V$ and G satisfies the ending condition. Then all the vertices of G can be labeled as either P- or N-positions. Moreover, $v \in V$ is a P-position iff $g(v) = 0$, where g is the Sprague–Grundy function for G.*

End of The Chapter Exercises

11.2 Prove **??**.

11.3 Show that there are only two Sprague–Grundy functions for the graph depicted on **??**.

11.4 Prove that there is unique Sprague–Grundy function for the one pile Nim game.

11.5 Prove that there is unique Sprague–Grundy function for the subtraction game where players may subtract 2 and 3 chips on their turn.

11.6 Prove that there is unique Sprague–Grundy function for the subtraction game where players may subtract 1, 2, or 5 chips on their turn.

12. Sums of Combinatorial Games

Let us consider the following game.

Game 12.1 (Take-Away Game)**.** *In this game there are two players.*

- *They have two piles of n and m chips, respectively.*
- *They make moves in turns with player I starting, each move consists of moving one or two chips out of the pile.*
- *The player that removes the last chip wins.*

It is not hard to draw the graph corresponding to the game for small n and m. Using this graph it is easy to see that all drawn positions, except $(1, 1)$ and $(0, 0)$ are N-positions.

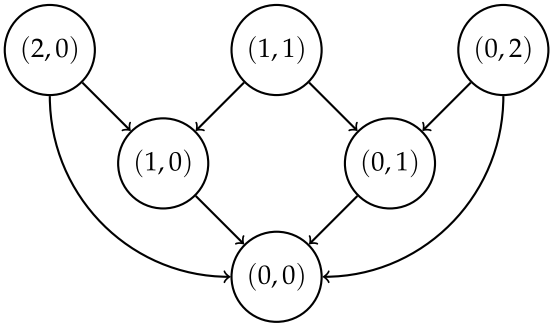


Figure 12.1: Part of the graph of **??**

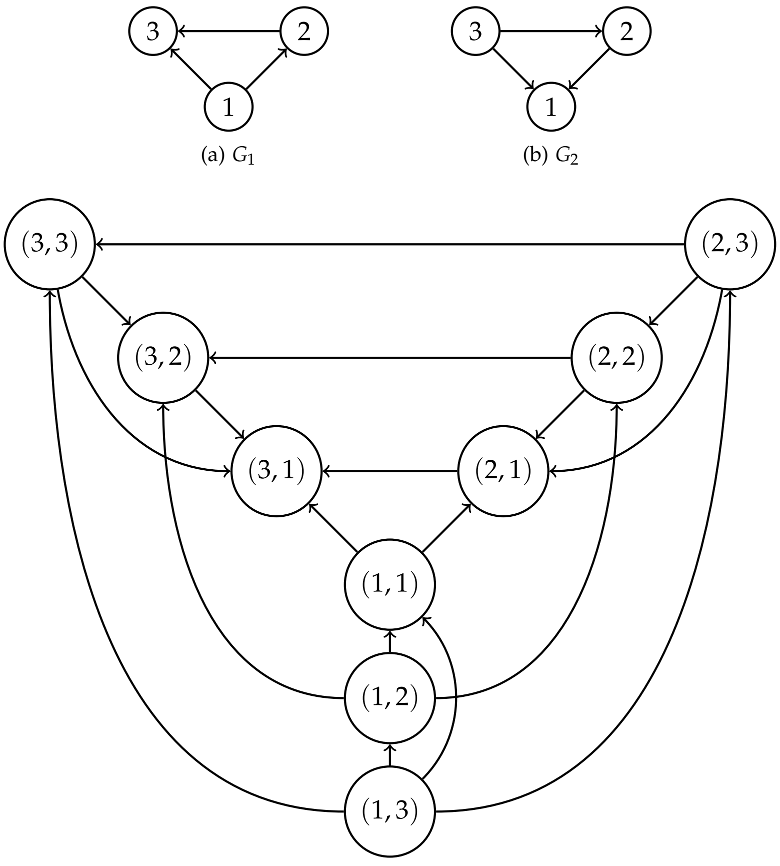
In the rest of the chapter we will discuss a method to study similar games. Assume we have two combinatorial games $\mathcal{G}_1$ and $\mathcal{G}_2$. One may form another game played as follows: the initial position of the new game consists of the pair of initial positions of $\mathcal{G}_1$ and $\mathcal{G}_2$, players alternate moves, and on each turn a player make a move in one of the game leaving the position in the second untouched. The new game is called *sum of $\mathcal{G}_1$ and $\mathcal{G}_2$*.

Let us give a formal definition.

Definition 12.1. *Let $G_1 = (V_1, F_1)$ and $G_2 = (V_1, F_2)$ be directed graphs. We say that G is the sum of G_1 and G_2, denoted $G_1 + G_2$, is a graph $(V_1 \times V_2, F)$ such that*

$$F(x_1, x_2) = \{(y_1, x_2) \ : \ y_1 \in F_1(x_1)\} \cup \{(x_1, y_2) \ : \ y_2 \in F_2(x_2)\}.$$

?? gives an example of this operation.



(a) G_1

(b) G_2

(c) $G_1 + G_2$

Figure 12.2: **??** depicts sum of graphs from **??** and **??**.

Another example is given by the game of Nim; it is easy to see that 2-pile Nim is a sum of two 1-pile Nims. This observation leads to a generalization of Bouton's Theorem (**??**).

Theorem 12.1 (The Sprague–Grundy Theorem)**.** *Let G_1 and G_2 be some graphs and g_1 and g_2 be corresponding Sprague–Grundy functions. Then the graph G_1 and G_2 has a Sprague–Grundy function g such that $g(x_1, x_2) = g_1(x_1) \oplus g_2(x_2)$.*

Proof. Let $G_1 = (V_1, F_1)$, $G_2 = (V_2, F_2)$, and $G = G_1 + G_2$. Consider some $x_1 \in V_1$ and $x_2 \in V_2$. Let $a = g_1(x_1) \oplus g_2(x_2)$. To prove the statement we need to show that

1. for any $0 \leq b < a$, there is $(y_1, y_2) \in F(x_1, x_2)$ such that $g(y_1, y_2) = b$;

2. for any $(y_1, y_2) \in F(x_1, x_2)$, $g(y_1, y_2) \neq a$.

We start from proving the first statement. Let us fix some $0 \leq b < a$ and let $c = a \oplus b$. Let $g_i(x_i) = (p_{i,\ell}, \ldots, p_{i,0})$ for each $i\,\{1,2\}$

and $c = (1, q_{k-1}, \ldots, q_0)$ where $k \leq \ell$. For some $j \in \{1,2\}$, $p_{j,k} = 1$ since $a = g_1(x_1) \oplus g_2(x_2)$. Without loss of generality $j = 1$. Hence, $c \oplus g_1(x_1) < g_1(x_1)$, whence there is x_1' such that $g_1(x_1') = c \oplus g_1(x_1)$. As a result, there is a move in G from (x_1, x_2) to (x_1', x_2) and $g(x_1', x_2) = g_1(x_1') \oplus g_2(x_2) = c \oplus g_1(x_1) \oplus g_2(x_2) = c \oplus a = b$.

To prove the second statement, assume that there is $(y_1, y_2) \in F(x_1, x_2)$ so that $g(y_1, y_2) = a$. Without loss of generality we may assume that $x_2 = y_2$. Hence, $0 = g(y_1, x_2) \oplus g(x_1, x_2) = g_1(y_1) \oplus g_1(x_1)$. However, $g_1(y_1) \neq g_1(x_1)$ since there is a move from x_1 to y_1. Therefore $g_1(y_1) \oplus g_1(x_1) \neq 0$ which is a contradiction. □

It is also easy to see that if G_1 and G_2 satisfy the ending condition, then $G_1 + G_2$ also satisfies the ending condition. Therefore, if G_1 and G_2 satisfy the ending condition and g_1, g_2 are Sprague–Grundy functions of them, $G_1 + G_2$ has unique Sprague–Grundy function g such that $g(x_1, x_2) = g_1(x_1) \oplus g_2(x_2)$.

The simple example of an application of this theorem is the analysis of the following game.

Game 12.2. *Alice and Bob have two piles with* 10 *and* 11 *chips respectively. They take turns and remove* 1, 2, *or* 3 *chips from one of the piles. If one of them cannot make a move he/she loses.*

To determine who is the winner in this game, we start with a subtraction game G with the subtraction set $\{1,2,3\}$. It is easy to see that $g : \mathbb{N}_0 \to \mathbb{N}_0$ such that

$$g(x) = \begin{cases} 0 & \text{if } x \equiv 0 \pmod 3 \\ 1 & \text{if } x \equiv 1 \pmod 3 \\ 2 & x\text{if } \equiv 2 \pmod 3 \end{cases}$$

is the Sprague–Grundy function for G. It is also clear that **??** is equal to $G + G$. Therefore the function $g' : \mathbb{N}_0^2 \to \mathbb{N}_0$ such that

$$g(x,y) = \begin{cases} 0 & \text{if } x \equiv 0 \pmod 3 \text{ and } y \equiv 0 \pmod 3 \\ 1 & \text{if } x \equiv 0 \pmod 3 \text{ and } y \equiv 1 \pmod 3 \\ 2 & \text{if } x \equiv 0 \pmod 3 \text{ and } y \equiv 2 \pmod 3 \\ 1 & \text{if } x \equiv 1 \pmod 3 \text{ and } y \equiv 0 \pmod 3 \\ 0 & \text{if } x \equiv 1 \pmod 3 \text{ and } y \equiv 1 \pmod 3 \\ 3 & \text{if } x \equiv 1 \pmod 3 \text{ and } y \equiv 2 \pmod 3 \\ 2 & \text{if } x \equiv 2 \pmod 3 \text{ and } y \equiv 0 \pmod 3 \\ 3 & \text{if } x \equiv 2 \pmod 3 \text{ and } y \equiv 1 \pmod 3 \\ 0 & \text{if } x \equiv 2 \pmod 3 \text{ and } y \equiv 2 \pmod 3 \end{cases}$$

A surprising example of the application of this theorem is the following game.

Game 12.3. *In this game position is described by a polygon with several diagonals. The game starts with a polygon with n sides. On each turn a player draw a new diagonal so that it does not intersect with previously drawn diagonals. Players take turns and the one who cannot make a move loses.*

It is easy to see that we do not care about the shape of the polygon and diagonals, the only important information is the number of nodes and which nodes are connected by a diagonal.

Let $g(n)$ be the value of the Sprague–Grundy function at the polygon with n sides. It is easy to see that if we split the polygon, by a diagonal, into two parts with ℓ and m sides, then the resulting position is essentially a position in the sum of two games; hence, $g(n) = \text{mex } \{g(\ell) \oplus g(m) \ : \ \ell, m \geq 3 \text{ and } \ell + m = n + 2\}$. Using this observation it is easy to compute the value of $g(n)$ for small n. It is easy to see to make a conjecture that $g(n) = 0$ for odd n and $g(n) = 1$ for even n. Let us prove this using induction. The base case follows from the computation necessary to write the table. Let us prove the induction step from $1, \ldots, n$ to $n + 1$.

3	4	5	6	7	8
0	1	0	1	0	1

Table 12.1: Sprague–Grundy function for **??**

- Let n be odd. It is clear that if $\ell + m = n + 3$, then ℓ and m have the same reminder modulo 2. Therefore, $g(m) = g(\ell)$ by the induction hypothesis. As a result,

  $$\{g(\ell) \oplus g(m) \ : \ \ell, m \geq 3 \text{ and } \ell + m = n + 2\} = \{0\}$$

  and $g(n + 1) = 1$.

- Let n be even. It is clear that if $\ell + m = n + 3$, then ℓ and m have different reminders modulo 2. Therefore, $g(m) \neq g(\ell)$ by the induction hypothesis. As a result,

  $$\{g(\ell) \oplus g(m) \ : \ \ell, m \geq 3 \text{ and } \ell + m = n + 2\} = \{1\}$$

  and $g(n + 1) = 0$.

End of The Chapter Exercises

12.1 Compute the Sprague–Grundy function for states of the subtraction game with two piles of chips where players and the subtraction set $\{1, 2, 5\}$.

12.2 Let G_1 be the subtraction game with the subtraction set $\{1, 2\}$ Let G_2 be the game of Nim with three piles. Find all the moves from $(11, (1, 6, 7))$ to P-positions in $G1 + G2$.

Part III

Introduction to Discrete Probability

13. Sample Spaces and Events

In the previous chapter we discussed games that do not involve move by chance. A theory that studies experiments, processes and interactions that are subject to chance is called *probability theory*; we are going to study this theory in this part.

The most important assumption of probability theory is that nonetheless the outcome of an experiment is not known in advance — the set of all possible outcomes is known. This set is called the *sample space* or the *probability space*.

For example, if our experiment consists of tossing a coin, then the sample space consists of two outcomes $\{H, T\}$, where H stands for heads and T stands for tails.

Exercise 13.1. *Write the sample space for the experiment consisting of tossing two coins.*

Each element of the sample space is called an *outcome* or an *elementary event*. Typically, we are interested in observing several outcomes; e.g., in the experiment consisting of tossing five coins, the set $\{HTTTT, THTTT, TTHTT, TTTHT, TTTTH\}$ describe all possible outcomes when exactly one coin shows heads. We say that a set of outcomes is an *event*.

Another assumption of probability theory is that every outcome ω of a sample space Ω is assigned some probability $\Pr(\omega)$; intuitively, $\Pr(\omega)$ is the likelihood that the outcome ω occur in the experiment. It is convenient to normalize probabilities, so we require that $0 \leq \Pr(\omega) \leq 1$ for all $\omega \in \Omega$ and $\sum_{\omega \in \Omega} \Pr(\omega) = 1$. The function $\Pr$ is called a *probability distribution* on Ω.

The pair of a sample space and a probability distribution on the space is called a *finite discrete probability space*.

We can extend the notion of probability from elementary events to all events as follows. Let $E \subseteq \Omega$ be an event in the finite discrete probability space $(\Omega, \Pr)$. Then $\Pr(E)$ denote the sum of all $\Pr(\omega)$ for $\omega \in \Omega$. In the sequel, we use $\sum_{\omega \in E} \Pr(\omega)$ to describe similar sums (see **??** for a formal definition).

Exercise 13.2. *In many cases we consider* uniform distribution *on a set Ω, the distribution* $\Pr$ *such that all the outcomes are equally likely. Let* $\Omega =$

$\{HH, HT, TH, TT\}$ *and* $\Pr$ *be the uniform distribution on* Ω. *Find the probability of the event* $\{HT, TH, TT\}$, *and give an informal interpretation of the answer.*

13.1 Basic Principles

Events are subsets of the sample space; hence, they can be combined using the standard set operations. So it is natural to ask whether the probabilities $\Pr(A \cup B)$ and $\Pr(A \cap B)$ can be expressed in terms of $\Pr(A)$ and $\Pr(B)$.

Theorem 13.1 (The Additive Principle). *Let* $(\Omega, \Pr)$ *be a finite discrete probability space, and let* $A, B \subseteq \Omega$ *be two disjoint events. Then* $\Pr(A \cup B) = \Pr(A) + \Pr(B)$.

Proof Sketch, see **??**. Let $A = \{\omega_{1,1}, \ldots, \omega_{1,k}\}$ and let $B = \{\omega_{2,1}, \ldots, \omega_{2,\ell}\}$. Then $A \cup B = \{\omega_{1,1}, \ldots, \omega_{1,k}, \omega_{2,1}, \ldots, \omega_{2,\ell}\}$. Therefore $\Pr(A \cup B) = \Pr(\omega_{1,1}) + \cdots + \Pr(\omega_{2,k}) + \Pr(\omega_{2,1}) + \cdots + \Pr(\omega_{2,\ell}) = \Pr(A) + \Pr(B)$. □

Exercise 13.3. *Let* $(\Omega, \Pr)$ *be a finite discrete probability space, and let* $A \subseteq A$ *be an event. Show that* $\Pr(A) = 1 - \Pr(A)$.

This result can be easily extended to the cases when A and B are not disjoint.

Corollary 13.1 (The Inclusion-exclusion Principle). *Let* $(\Omega, \Pr)$ *be a finite discrete probability space, and let* $A, B \subseteq \Omega$ *be two events. Then* $\Pr(A \cup B) = \Pr(A) + \Pr(B) - \Pr(A \cap B)$.

Unfortunately, $\Pr(A \cap B)$ cannot be expressed via $\Pr(A)$ and $\Pr(B)$. However, in many cases $\Pr(A \cap B) = \Pr(A)\Pr(B)$; if this equality holds, we say that A and B are *independent*.

For example, let us consider an experiment where we toss two fair coins; i.e., let us consider $\Omega = \{HH, HT, TH, TT\}$ and let Pr be a uniform distribution on Ω. It is easy to see that $\Pr(\{HH, HT\}) = 1/2$, $\Pr(\{HH, TH\}) = 1/2$, and $\Pr(\{HH, HT\} \cap \{HH, TH\}) = \Pr(\{HH\}) = 1/4$. Hence, these two events are independent.

To analyze experiments consisting of tossing several coins, we need to be able to study products of finite discrete probability spaces.

Theorem 13.2 (The Multiplicative Principle). *Let* $\Omega = \Omega_1 \times \Omega_2$ *and let* $(\Omega_1, \Pr_1)$ *and* $(\Omega_2, \Pr_2)$ *be finite discrete probability spaces. Then* $\Pr : \Omega \to \mathbb{R}$ *such that* $\Pr(\omega_1, \omega_2) = \Pr_1(\omega_1) \cdot \Pr_2(\omega_2)$ *is a probability distribution on* Ω. *Moreover,* $\Pr(E_1 \times E_2) = \Pr_1(E_1) \cdot \Pr_2(E_2)$ *for all* $E_1 \subseteq \Omega_1$ *and* $E_2 \subseteq \Omega_2$.

Using this principle we can show that in the experiment consisting of tossing five coins, the event where the first flip is H and the event where the second flip is H are independent. Indeed, let $\Omega = \{H,T\}^5$; Pr be the uniform distribution on Ω; and $E_1 = \left\{t \in \{H,T\}^5 \ : \ t_1 = H\right\}$ and $E_2 = \left\{t \in \{H,T\}^5 \ : \ t_2 = H\right\}$. By **??**, $\Pr(E_1) = \Pr(E_2) = 1/2$. Moreover, $E_1 \cap E_2 = \left\{t \in \{H,T\}^5 \ : \ t_1 = H \text{ and } t_2 = H\right\}$; therefore, $\Pr(E_1 \cap E_2) = \frac{1}{4}$.

13.2 Random Variables

Sometimes we are more interested in some function of the result of the experiment rather than the result itself. For example, Sasha may play Dungeons and Dragons and be interested in his chances to roill 7 on two dices together. Let us formalize the question. Let $\Omega = [6]^2$ and Pr be a the uniform distribution on Ω. Sasha is iterested in the probability of the event $\left\{(x,y) \in [6]^2 \ : \ x + y = 7\right\}$.

More generally, let $(\Omega, \Pr)$ be a finite discrete probability space. Then a function $\chi : \Omega \to \mathbb{R}$ is called a *random variable* and $\Pr(\chi = a)$ denotes $\Pr(\{\omega \in \Omega \ : \ \chi(\omega) = a\})$.

In the example about Dangeons and Dragons, $\chi(x,y) = x + y$ and we are interested in $\Pr(\chi = 7) = \Pr(\{(1,6),(2,5),\dots,(6,1)\} = 1/6$.

Exercise 13.4. *Let* $\Omega = [6]^2$ *and* Pr *be a the uniform distribution on* Ω*. Let* $\chi : \Omega \to \mathbb{R}$ *be the random variable such that* $\chi(x,y) = x + y$*. Find* $\Pr(\chi = 1), \dots, \Pr(\chi = 12)$*.*

We are going to adopt some simple additional notation, if $\chi_1, \chi_2 : \Omega \to \mathbb{R}$ are random variables then $(\chi_1 + \chi_2), (\chi_1 \cdot \chi_2) : \Omega \to \mathbb{R}$ are the random variables such that $(\chi_1 + \chi_2)(\omega) = \chi_1(\omega) + \chi_2(\omega)$ and $(\chi_1 \cdot \chi_2)(\omega) = \chi_1(\omega) \cdot \chi_2(\omega)$.

End of The Chapter Exercises

13.5 Prove **??**.

13.6 Let $\Omega = \{HH, HT, TH, TT\}$ and let Pr be a uniform distribution on Ω. Show that $\{HH\}$ and $\{TT\}$ are not independent.

13.7 Alice is rolling a dice n times, compute the probability that Alice sees 6, 6, and 6 in three consecutive rolls.

13.8 Alice is rolling a dice n times, compute the probability that Alice sees 4, 5, and 6 in three consecutive rolls.

14. Conditional Probability

In general, the occurrence of an event B changes the probability that another event A occurs, $\Pr(A \mid B)$ denotes the latter probability. More formally, $\Pr(A \mid B) = \Pr(A \cap B) / \Pr(B)$, where $(\Omega, \Pr)$ is finite discrete probability space, $A, B \subseteq \Omega$, and $\Pr(B) \neq 0$. We say that $\Pr(A \mid B)$ is the conditional probability that A occurs given that B occurs.

For example, let us consider $\Omega = [6]^2$ and uniform distribution $\Pr$ on Ω; i.e., we consider an experiment consisting of rolling two dices. Let us compute the probability that the sum of numbers on the dices exceeds 6 given that the first dice's number is 3. In other words we need to compute $\Pr(A \mid B)$, where $A = \{(i,j) \in [6]^2 \,:\, i+j > 6\}$ and $B = \{(3,j) \,:\, j \in [6]\}$. It is clear that $\Pr(B) = 1/6$ and $\Pr(A \cap B) = \{(3,4),(3,5),(3,6)\} = 1/12$. Hence, $\Pr(A \mid B) = 1/2$.

Exercise 14.1. *A family has two children. What is the probability that both are boys, given at least one is a boy? What if it is given that* the first child *is a boy. (You assume that the probability distribution of families is uniform.)*

Let us consider another example known as "Monty Hall Problem". On the television game *Let's make a deal*, a contestant is presented with a choice of three closed doors. Behind exactly one door is a prize; the other doors conceal cheap items. First, the contestant is asked to choose a door. Then Monty Hall, the host of the show, shows the contestant one of the worthless prizes behind one of the other doors. At this point, there are two closed doors, and the contestant is given the opportunity to switch from his original choice to the other closed door. The question is, is it better for the contestant to stick to his original choice or to switch doors?

Let us analyze this question using the conditional probabilities. Without loss of generality, we may assume that the contestant chooses door 1. Note that the sample space is equal to $\{(1,2),(1,3),(2,3),(3,2)\}$, where the first numbder denotes the door with the prize and the second number denotes the door opened by the host. The probability distribution is equal to

$$\Pr(x) = \begin{cases} 1/6 & x = (1,2) \\ 1/6 & x = (1,3) \\ 1/3 & x = (2,3) \\ 1/3 & x = (3,2) \end{cases}$$

since in the first two cases Monty has two possible choices to show a door without the prize. Suppose the host reveiled the door number 2 (the probability of this is 3/6). Then the probability that we win the price if we stick to the original choice is $(1/6)/(3/6) = 1/3$. However, the probability to win the prize in case of us swithcing the door is $(1/3)/(3/6) = 2/3$. Which implies, paradozicaly, that it is beneficial to switch the door!

Theorem 14.1 (Bayes' Rules). *Let* $(\Omega, \Pr)$ *be a finite discrete probability space.*

- *Let* $A, B \subseteq \Omega$ *be two events such that* $\Pr(A) > 0$ *and* $\Pr(B) > 0$. *Then* $\Pr(A \mid B) = \frac{\Pr(B \mid A)\Pr(B)}{\Pr(A)}$.
- *Let* $A, B \subseteq \Omega$ *be two events such that* $\Pr(A) < 1$ *and* $\Pr(B) < 1$; *i.e.,* $\Pr(\bar{A}) > 0$ *and* $\Pr(\bar{A}) > 0$, *where* $\bar{A} = \Omega \setminus A$ *and* $\bar{B} = \Omega \setminus B$. *Then* $\Pr(A) = \Pr(A \mid B)\Pr(B) + \Pr(A \mid \bar{B})\Pr(\bar{B})$.

Usefulness of this result can be illustrated with the following example. Assume that there is a rare disease that has the property that if a patient is affected by the desease, then the test is positive in 99% of the cases. However, it happens in 2% of the cases that a healthy patient tests positive. Statistical data shows that one person out of 1000 has the desease. What is the probability for a patient with a positive test to be affected by the desease?

Let S be the event that the patient has the desease, and P and N the events that the test is positive or negative. We know that $\Pr(S) = 0.001$, $\Pr(P \mid S) = 0.99$, and $\Pr(P \mid S) = 0.02$, where $\bar{S}$ is the event that the patient does not have the desease. Therefore $\Pr(S \mid P) = \frac{\Pr(P \mid S)\Pr(S)}{\Pr(P)}$ and $\Pr(P) = \Pr(P \mid S)\Pr(S) + \Pr(P \mid \bar{S})\Pr(\bar{S})$. As a result $\Pr(S \mid P) = \frac{0.99 \cdot 0.001}{0.99 \cdot 0.001 + 0.02 \cdot 0.999} \approx \frac{1}{20}$.