Proofs and Concepts The Fundamentals of Abstract Mathematics

Dave Witte Morris and Joy Morris

Contribution:

Open Educational Resources
UKM Literasi Informasi & Perpustakaan Unsyiah

Proofs and Concepts

the fundamentals of abstract mathematics

by

Dave Witte Morris and Joy Morris

University of Lethbridge

incorporating material by

P. D. Magnus

University at Albany, State University of New York

Preliminary Version 0.92 of December 2016

The presentation of Logic in this textbook is adapted from

forallχ

An Introduction to Formal Logic

P. D. Magnus

University at Albany, State University of New York

The most recent version of forallχ is available on-line at
http://www.fecundity.com/logic

We thank Professor Magnus for making forallχ *freely available,*
and for authorizing derivative works such as this one.
He was not involved in the preparation of this manuscript,
so he is not responsible for any errors or other shortcomings.

Please send comments and corrections to:
Dave.Morris@uleth.ca or Joy.Morris@uleth.ca

Preface

Unlike in earlier courses, success in advanced undergraduate mathematics classes (and beyond) does not depend nearly so much on being able to find the right answer to a question as it does on being able to provide a convincing explanation that the answer is correct. (Mathematicians call this explanation a **proof**.) This textbook is designed to help students acquire this essential skill, by developing a working knowledge of:

1) proof techniques (and their basis in Logic), and

2) fundamental concepts of abstract mathematics.

We start with the language of Propositional Logic, where the rules for proofs are very straightforward. Adding sets and quantifiers to this yields First-Order Logic, which is the language of modern mathematics. Being able to do proofs in this setting is the main skill necessary for success in advanced mathematics.

It is also important to be familiar with (and be able to prove statements about) sets and functions, which are the building blocks of modern mathematics. In addition, a chapter on *Cardinality* provides an introduction to the surprising notion of "uncountable sets": infinite sets with so many elements that it is impossible to make a list $x_1, x_2, x_3, \ldots$ of all of them (even if the list is allowed to be infinitely long).

to Harmony

Contents

Part I. Introduction to Logic and Proofs

Part II. Sets and First-Order Logic

Part I

Introduction to Logic and Proofs

Chapter 1

Propositional Logic

The grand aim of all science is to cover the greatest number of empirical facts by logical deduction from the smallest number of hypotheses or axioms.
Albert Einstein (1879–1955), Nobel prize-winning physicist
in *Life* magazine

For our purposes, Logic is the business of deciding whether or not a deduction is valid; that is, deciding whether or not a particular conclusion is a consequence of particular assumptions. (The assumptions can also be called "**hypotheses**" or "**axioms**.")

1.1. Assertions, deductions, and validity

We will begin our discussion of Logic by introducing three basic ingredients: assertions, deductions, and validity.

Here is one possible deduction:

Hypotheses:

(1) It is raining heavily.

(2) If you do not take an umbrella, you will get soaked.

Conclusion: You should take an umbrella.

(The validity of this particular deduction will be analyzed in Example 1.1.10 below.)

In Logic, we are only interested in sentences that can be a hypothesis or conclusion of a deduction. These are called "assertions":

DEFINITION 1.1.1. An **assertion** is a sentence that is either true or false.

OTHER TERMINOLOGY. Some textbooks use the term *proposition* or *statement* or *sentence*, instead of *assertion*.

EXAMPLE 1.1.2.

- **Questions** The sentence "Are you sleepy yet?" is *not* an assertion. Although you might be sleepy or you might be alert, the question itself is neither true nor false. For this reason, questions do not count as assertions in Logic.
- **Imperatives** Commands are often phrased as imperatives like "Wake up!," "Sit up straight," and so on. Although it might be good for you to sit up straight or it might not, the command itself is neither true nor false.
- **Exclamations** "Ouch!" is sometimes called an exclamatory sentence, but it is neither true nor false. so it is another example of a sentence that is not an assertion.

Remark 1.1.3. Roughly speaking, an assertion is a statement of fact, such as "The earth is bigger than the moon" or "Edmonton is the capital of Alberta." However, it is important to remember that an assertion may be false, in which case it is a mistake (or perhaps a deliberate lie), such as "There are less than 1,000 automobiles in all of Canada." In many cases, the truth or falsity of an assertion depends on the situation. For example, the assertion "It is raining" is true in certain places at certain times, but is false at others.

In this chapter and the next, which are introductory, we will deal mostly with assertions about the real world, where facts are not always clear-cut. (For example, if Alice and Bob are almost the same height, it may be impossible to determine whether it is true that "Alice is taller than Bob.") We are taking a mathematical (or scientific) view toward Logic, not a philosophical one, so we will ignore the imperfections of these real-world assertions, which provide motivation and illustration, because our goal is to learn to use Logic to understand mathematical objects (not real-world objects), where there are no grey areas.

Throughout this text, you will find exercises that review and explore the material that has just been covered. There is no substitute for actually working through some problems, because this course, like most advanced mathematics, is more about a way of thinking than it is about memorizing facts.

EXERCISES 1.1.4. Which of the following are "assertions" in the logical sense?

1) England is smaller than China.
2) Greenland is south of Jerusalem.
3) Is New Jersey east of Wisconsin?
4) The atomic number of helium is 2.
5) The atomic number of helium is π.
6) Take your time.
7) This is the last question.
8) Rihanna was born in Barbados.

DEFINITION 1.1.5. A **deduction** is a series of hypotheses that is followed by a conclusion. (The conclusion and each of the hypotheses must be an assertion.)

If the hypotheses are true and the deduction is a good one, then you have a reason to accept the conclusion.

EXAMPLE 1.1.6. Here are two deductions.

1)
Hypotheses:
All men are mortal.
Socrates is a man.
Conclusion: Socrates is mortal.

2)
Hypotheses:
The Mona Lisa was painted by Leonardo da Vinci.
Neil Armstrong was the first man on the moon.
Conclusion: Justin Trudeau went swimming yesterday.

The first of these deductions is very famous (and was discussed by the ancient Greek philosopher Aristotle), but the second one is lame. It may seem odd to even call it a deduction, because the two hypotheses have nothing at all to do with the conclusion, but, given our definition, it does count as a deduction. However, it is is a very poor one, so it cannot be relied on as evidence that the conclusion is true.

We are interested in the deductions that *do* provide solid evidence for their conclusions:

DEFINITION 1.1.7. A deduction is **valid** if its conclusion is true whenever all of its hypotheses are true. In other words, it is impossible to have a situation in which all of the hypotheses are true, but the conclusion is false.

The task of Logic is to distinguish valid deductions from invalid ones.

EXAMPLE 1.1.8.

Hypotheses:

Oranges are either fruits or musical instruments.

Oranges are not fruits.

Conclusion: Oranges are musical instruments.

The conclusion is ridiculous. Nevertheless, the deduction is valid, because its conclusion follows validly from its hypotheses; that is, *if* both hypotheses were true, *then* the conclusion would necessarily be true. For example, you might be able to imagine that, in some remote river valley, there is a variety of orange that is not a fruit, because it is hollow inside, like a gourd. Well, if the other hypothesis is also true in that valley, then the residents must use the oranges to play music.

This shows that a logically valid deduction does not need to have true hypotheses or a true conclusion. Conversely, having true hypotheses and a true conclusion is not enough to make a deduction valid:

EXAMPLE 1.1.9.

Hypotheses:

London is in England.

Beijing is in China.

Conclusion: Paris is in France.

The hypotheses and conclusion of this deduction are, as a matter of fact, all true. This is a terrible deduction, however, because the hypotheses have nothing to do with the conclusion. For example, if Paris declared independence from the rest of France, then the conclusion would be false, even though the hypotheses would both still be true. Thus, it is *logically possible* to have a situation in which the hypotheses of this deduction are true and the conclusion is false. Therefore, the deduction is *not* valid.

EXAMPLE 1.1.10. Recall the deduction that you should take an umbrella (on p. 3, above), and suppose for a moment that both of its hypotheses are true. (Thus, you will get wet if you do not take an umbrella.) Now is it necessarily true that you should take an umbrella? No—perhaps you enjoy walking in the rain, and you would like to get soaked. In that case, even though the hypotheses were true, the conclusion would be false. Thus, the deduction is not valid.

EXAMPLE 1.1.11.

Hypotheses:

You are reading this book.

This is an undergraduate textbook.

Conclusion: You are an undergraduate student.

This is not a terrible deduction, because most people who read this book are undergraduate students. Yet, it is possible for someone besides an undergraduate to read this book. For example, if your mother or father picked up the book and thumbed through it, they would not

immediately become an undergraduate. So the hypotheses of this deduction, even though they are true, do not guarantee the truth of the conclusion. Thus, even though some people might say that the deduction has some value, it is certainly not valid.

Remark 1.1.12. It is important to remember that validity of a deduction is not about the truth or falsity of the deduction's assertions in the real world. Instead, it is about the form of the deduction, in that the truth of the hypotheses is incompatible with the falsity of the conclusion in every possible world (real or imaginary). Furthermore, a deduction gives you a reason to believe its conclusion only in situations where its hypotheses are true.

EXERCISES 1.1.13. Which of the following is possible? If it is possible, give an example. If it is not possible, explain why.

1) A valid deduction that has one false hypothesis and one true hypothesis.

2) A valid deduction that has a false conclusion.

3) A valid deduction that has at least one false hypothesis, and a true conclusion.

4) A valid deduction that has all true hypotheses, and a false conclusion.

5) An invalid deduction that has at least one false hypothesis, and a true conclusion.

1.2. Logic puzzles

Clear thinking (or logic) is important not only in mathematics, but in everyday life, and can also be fun; many logic puzzles (or brain teasers) can be found on the internet or in bookstores. Here are just a few. Solving problems like these provides good practice for some of the logic skills that will be needed later in this book.

EXERCISE 1.2.1 (found online at http://philosophy.hku.hk/think/logic/puzzles.php). There was a robbery in which a lot of goods were stolen. The robber(s) left in a truck. It is known that:

1) No one other than A, B and C was involved in the robbery.

2) C never commits a crime without including A as an accomplice.

3) B does not know how to drive.

So, can you tell whether A is innocent?

EXERCISES 1.2.2. On the island of Knights and Knaves,* every resident is either a Knight or a Knave (and they all know the status of everyone else). It is important to know that:

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

More precisely, every assertion spoken by a Knight is true, and every assertion spoken by a Knave is false.

You will meet some residents of the island, and your job is to figure out whether each of them is a Knight or a Knave.

1) You meet Alice and Bob on the island. Alice says "Bob and I are Knights." Bob says, "That's a lie! She's a Knave!" What are they?

2) You meet Charlie, Diane, and Ed on the island. Charlie says, "Be careful, not all three of us are Knights." Diane says, "But not all of us are Knaves, either." Ed says, "Don't listen to them, I'm the only Knight." What are they?

*http://en.wikipedia.org/wiki/Knights_and_knaves

3) You meet Frances and George on the island. Frances mumbles something, but you can't understand it. George says, "She said she's a Knave. And she sure is — don't trust her!" What are they?

EXERCISE 1.2.3. Complete each of these mini-Sudoku puzzles, by putting a number from 1 to 4 in each box, so that no number appears twice in any row or column, or in any of the four 2×2 boxes with dark outlines. (Each of the puzzles has a unique solution.)

a)

4			
		1	
		2	
3			

b)

		3	
			2
2			3
4			

c)

		4	1
			2
	1		

EXERCISE 1.2.4. In a game similar to *Mastermind*, one player chooses a secret 4-digit number, using only the digits 1–6, inclusive. (Repeated digits *are* allowed.) The other player makes a series of guesses. After each guess, the first player tells the other how many of the digits are in the guess are perfectly correct, and how many of the other digits are correct, but in the wrong place. In each of the following games, there is now enough information to determine the secret number. Figure out the secret number.

a)

guess	# correct	# in wrong spot
1234	0	3
2354	0	3
3642	1	1
5143	0	3
4512	1	2
	4	0

b)

guess	# correct	# in wrong spot
1234	0	2
4516	1	2
4621	1	1
6543	0	2
5411	0	3
	4	0

1.3. Using letters to symbolize assertions

In the remainder of this chapter, we will discuss a logical language called Propositional Logic. It provides a convenient way to describe the logical relationship between two (or more) assertions, by using capital letters to represent assertions. Considered only as a symbol of Propositional Logic, the letter A could mean any assertion. So, when translating from English into Propositional Logic, it is important to provide a **symbolization key** that specifies what assertion is represented by each letter.

For example, consider this deduction:

Hypotheses:

There is an apple on the desk.

If there is an apple on the desk, then Jenny made it to class.

Conclusion: Jenny made it to class.

This is obviously a valid deduction in English.

What happens if we replace each assertion with a letter? Our symbolization key would look like this:

A: There is an apple on the desk.

B: If there is an apple on the desk, then Jenny made it to class.

C: Jenny made it to class.

We would then symbolize the deduction in this way:

Hypotheses:

A

B

Conclusion: C

Unfortunately, there is no necessary connection between two assertions A and B, which could be any assertions, and a third assertion C, which could be any assertion, so this is not a valid deduction. Thus, the validity of the original deduction has been lost in this translation; we need to do something else in order to preserve the logical structure of the original deduction and obtain a valid translation.

The important thing about the original deduction is that the second hypothesis is not merely *any* assertion, logically divorced from the other assertions in the deduction. Instead, the second hypothesis contains the first hypothesis and the conclusion *as parts*. Our symbolization key for the deduction only needs to include meanings for A and C, and we can build the second hypothesis from those pieces. So we could symbolize the deduction this way:

Hypotheses:

A

If A, then C.

Conclusion: C

This preserves the structure of the deduction that makes it valid, but it still makes use of the English expression "If... then..." Eventually, we will replace all of the English expressions with mathematical notation, but this is a good start.

The assertions that are symbolized with a single letter are called *atomic assertions*, because they are the basic building blocks out of which more complex assertions are built. Whatever logical structure an assertion might have is lost when it is translated as an atomic assertion. From the point of view of Propositional Logic, the assertion is just a letter. It can be used to build more complex assertions, but it cannot be taken apart.

NOTATION 1.3.1. The symbol "$\therefore$" means "therefore," and we will often use

$$A, B, C, \ldots, \therefore Z$$

as an abbreviation for the deduction

Hypotheses:

A

B

C

$\vdots$

Conclusion: Z.

1.4. Connectives ($\neg$, $\&$, $\vee$, $\Rightarrow$, $\Leftrightarrow$)

Logical connectives are used to build complex assertions from simpler pieces. There are five logical connectives in Propositional Logic. This table summarizes them, and they are explained below.

symbol	nickname	what it means
$\neg$	not	"It is not the case that _____"
&	and	"Both _____ and _____"
$\vee$	or	"Either _____ or _____"
$\Rightarrow$	implies	"If _____ then _____"
$\Leftrightarrow$	iff	"_____ if and only if _____"

Remark 1.4.1. As we learn to write proofs, it will be important to be able to produce a deduction in Propositional Logic from a sequence of assertions in English. It will also be important to be able to retrieve the English meaning from a sequence of assertions in Propositional Logic, given a symbolization key. The above table should prove useful in both of these tasks.

1.4A. Not ($\neg$). As an example, consider how we might symbolize these assertions:

1. Mary is in Barcelona.
2. Mary is not in Barcelona.
3. Mary is somewhere other than Barcelona.

In order to symbolize Assertion 1, we will need one letter. We can provide a symbolization key:

B: Mary is in Barcelona.

Note that here we are giving B a different interpretation than we did in the previous section. The symbolization key only specifies what B means *in a specific context*. It is vital that we continue to use this meaning of B so long as we are talking about Mary and Barcelona. Later, when we are symbolizing different assertions, we can write a new symbolization key and use B to mean something else.

Now, Assertion 1 is simply B.

Since Assertion 2 is obviously related to Assertion 1, we do not want to introduce a different letter to represent it. To put it partly in English, the assertion means "It is not true that B." For short, logicians say "Not B." This is called the **logical negation** of B. In order to convert it entirely to symbols, we will use "$\neg$" to denote logical negation. Then we can symbolize "Not B" as $\neg B$.

Assertion 3 is about whether or not Mary is in Barcelona, but it does not contain the word "not." Nevertheless, they both mean "It is not the case that Mary is in Barcelona." As such, we can translate both Assertion 2 and Assertion 3 as $\neg B$.

An assertion can be symbolized as $\neg\mathcal{A}$ if
it can be paraphrased in English as "It is not the case that $\mathcal{A}$."

Consider these further examples:

4. The widget can be replaced if it breaks.
5. The widget is irreplaceable.
6. The widget is not irreplaceable.

If we let R mean "The widget is replaceable," then Assertion 4 can be translated as R.

What about Assertion 5? Saying the widget is irreplaceable means that it is not the case that the widget is replaceable. So even though Assertion 5 is not negative in English, we symbolize it using negation as $\neg R$.

Assertion 6 can be paraphrased as "It is not the case that the widget is irreplaceable." Now, as we have already discussed, "The widget is irreplaceable" can be symbolized as "$\neg R$." Therefore, Assertion 6 can be formulated as "it is not the case that $\neg R$." Hence, it is the negation of $\neg R$, so it can be symbolized as $\neg\neg R$. This is a *double negation*. (However, if you think about

the assertion in English, it is another way of saying the same thing as Assertion 4. In general, we will see that if A is any assertion, then A and $\neg\neg A$ are "logically equivalent.")

More examples:

7. Elliott is short.

8. Elliott is tall.

If we let S mean "Elliot is short," then we can symbolize Assertion 7 as S.

However, it would be a mistake to symbolize Assertion 8 as $\neg S$. If Elliott is tall, then he is not short—but Assertion 8 does not mean the same thing as "It is not the case that Elliott is short." It could be that he is not tall but that he is not short either: perhaps he is somewhere between the two (average height). In order to symbolize Assertion 8, we would need a new assertion letter.

For any assertion $\mathcal{A}$:

- If $\mathcal{A}$ is true, then $\neg\mathcal{A}$ is false.
- If $\neg\mathcal{A}$ is true, then $\mathcal{A}$ is false.

Using "T" for true and "F" for false, we can summarize this in a *truth table* for negation:

$\mathcal{A}$	$\neg\mathcal{A}$
T	F
F	T

EXERCISES 1.4.2. Using the given symbolization key, translate each English-language assertion into Propositional Logic.

M: Those creatures are men in suits.

C: Those creatures are chimpanzees.

G: Those creatures are gorillas.

1) Those creatures are not men in suits.

2) It is not the case that those creatures are not gorillas.

3) Of course those creatures are not chimpanzees!

EXERCISES 1.4.3. Using the same symbolization key, translate each symbolic assertion into English.

1) G 2) $\neg M$ 3) $\neg\neg C$

1.4B. And (&). Consider these assertions:

9. Adam is athletic.

10. Barbara is athletic.

11. Adam is athletic, and Barbara is also athletic.

We will need separate assertion letters for Assertions 9 and 10, so we define this symbolization key:

A: Adam is athletic.

B: Barbara is athletic.

Assertion 9 can be symbolized as A.

Assertion 10 can be symbolized as B.

Assertion 11 can be paraphrased as "A and B." In order to fully symbolize this assertion, we need another symbol. We will use "&." We translate "A and B" as $A \,\&\, B$. We will call this connective "and" (but many logicians call it **conjunction**).

Notice that we make no attempt to symbolize "also" in Assertion 11. Words like "both" and "also" function to draw our attention to the fact that two things are being conjoined. They

are not doing any further logical work, so we do not need to represent them in Propositional Logic.

Some more examples:

12. Barbara is athletic and energetic.
13. Barbara and Adam are both athletic.
14. Although Barbara is energetic, she is not athletic.
15. Barbara is athletic, but Adam is more athletic than she is.

Assertion 12 is obviously a conjunction. The assertion says two things about Barbara, so in English it is permissible to refer to Barbara only once. It might be tempting to try this when translating the deduction: Since B means "Barbara is athletic," one might paraphrase the assertions as "B and energetic." This would be a mistake. Once we translate part of an assertion as B, any further structure is lost. B is an atomic assertion; it is nothing more than true or false. Conversely, "energetic" is not an assertion; on its own it is neither true nor false. We should instead paraphrase the assertion as "B and Barbara is energetic." Now we need to add an assertion letter to the symbolization key. Let E mean "Barbara is energetic." Now the assertion can be translated as $B \,\&\, E$.

> An assertion can be symbolized as $\mathcal{A} \,\&\, \mathcal{B}$ if
> it can be paraphrased in English as "Both $\mathcal{A}$, and $\mathcal{B}$."

Assertion 13 says one thing about two different subjects. It says of both Barbara and Adam that they are athletic, and in English we use the word "athletic" only once. In translating to Propositional Logic, it is important to realize that the assertion can be paraphrased as, "Barbara is athletic, and Adam is athletic." Thus, this translates as $B \,\&\, A$.

Assertion 14 is a bit more complicated. The word "although" sets up a contrast between the first part of the assertion and the second part. Nevertheless, the assertion says both that Barbara is energetic and that she is not athletic. In order to make the second part into an atomic assertion, we need to replace "she" with "Barbara."

So we can paraphrase Assertion 14 as, "*Both* Barbara is energetic, *and* Barbara is not athletic." The second part contains a negation, so we paraphrase further: "*Both* Barbara is energetic *and it is not the case that* Barbara is athletic." This translates as $E \,\&\, \neg B$.

Assertion 15 contains a similar contrastive structure. It is irrelevant for the purpose of translating to Propositional Logic, so we can paraphrase the assertion as "*Both* Barbara is athletic, *and* Adam is more athletic than Barbara." (Notice that we once again replace the pronoun "she" with her name.) How should we translate the second part? We already have the assertion letter A which is about Adam's being athletic and B which is about Barbara's being athletic, but neither is about one of them being more athletic than the other. We need a new assertion letter. Let M mean "Adam is more athletic than Barbara." Now the assertion translates as $B \,\&\, M$.

> Assertions that can be paraphrased "$\mathcal{A}$, but $\mathcal{B}$" or "Although $\mathcal{A}$, $\mathcal{B}$"
> are best symbolized using "and": $\mathcal{A} \,\&\, \mathcal{B}$.

It is important to keep in mind that the assertion letters A, B, and M are atomic assertions. Considered as symbols of Propositional Logic, they have no meaning beyond being true or false. We have used them to symbolize different English language assertions that are all about people being athletic, but this similarity is completely lost when we translate to Propositional Logic. No formal language can capture all the structure of the English language, but as long as this structure is not important to the deduction there is nothing lost by leaving it out.

For any assertions $\mathcal{A}$ and $\mathcal{B}$,

$\mathcal{A}$ & $\mathcal{B}$ is true if and only if both $\mathcal{A}$ and $\mathcal{B}$ are true.

We can summarize this in the truth table for "and":

$\mathcal{A}$	$\mathcal{B}$	$\mathcal{A}$ & $\mathcal{B}$
T	T	T
T	F	F
F	T	F
F	F	F

EXERCISES 1.4.4. Using the given symbolization key, translate each English-language assertion into Propositional Logic.

E_1: Ava is an electrician.

E_2: Harrison is an electrician.

F_1: Ava is a firefighter.

F_2: Harrison is a firefighter.

S_1: Ava is satisfied with her career.

S_2: Harrison is satisfied with his career.

1) Ava and Harrison are both electricians.

2) Harrison is an unsatisfied electrician.

3) Neither Ava nor Harrison is an electrician.

4) Both Ava and Harrison are electricians, but neither of them find it satisfying.

5) It cannot be that Harrison is both an electrician and a firefighter.

6) Ava is neither an electrician, nor a firefighter.

EXERCISES 1.4.5. Using the given symbolization key, translate each symbolic assertion into English.

J: Romeo likes Juliet.

M: Mercutio likes Juliet.

T: Romeo likes Tybalt.

1) M & J 2) J & $\neg T$ 3) $\neg M$ & J

1.4C. Or ($\vee$)**.** Consider these assertions:

16. Either Denison will play golf with me, or he will watch movies.

17. Either Denison or Ellery will play golf with me.

For these assertions we can use this symbolization key:

D: Denison will play golf with me.

E: Ellery will play golf with me.

M: Denison will watch movies.

Assertion 16 is "Either D or M." To fully symbolize this, we introduce a new symbol. The assertion becomes $D \vee M$. We will call this connective "or" (but many logicians call it **disjunction**).

Assertion 17 is only slightly more complicated. There are two subjects, but the English assertion only gives the verb once. In translating, we can paraphrase it as. "Either Denison will play golf with me, or Ellery will play golf with me." Now it obviously translates as $D \vee E$.

An assertion can be symbolized as $\mathcal{A} \vee \mathcal{B}$ if
it can be paraphrased in English as "Either $\mathcal{A}$, or $\mathcal{B}$."

Sometimes in English, the word "or" excludes the possibility that both disjuncts are true. This is called an **exclusive or**. An *exclusive or* is clearly intended when it says, on a restaurant menu, "Entrees come with either soup or salad." You may have soup; you may have salad; but, if you want *both* soup *and* salad, then you have to pay extra.

At other times, the word "or" allows for the possibility that both disjuncts might be true. This is probably the case with Assertion 17, above. I might play with Denison, with Ellery, or with both Denison and Ellery. Assertion 17 merely says that I will play with *at least* one of them. This is called an **inclusive or**.

The symbol "$\vee$" represents an *inclusive or*. So $D \vee E$ is true if D is true, if E is true, or if both D and E are true. It is false only if both D and E are false. We can summarize this with the truth table for "or":

$\mathcal{A}$	$\mathcal{B}$	$\mathcal{A} \vee \mathcal{B}$
T	T	T
T	F	T
F	T	T
F	F	F

Like "and," the connective "or" is commutative: $\mathcal{A} \vee \mathcal{B}$ is logically equivalent to $\mathcal{B} \vee \mathcal{A}$ (see Exercise 1.7.5(2)).

In mathematical writing, "or" *always* means **inclusive** or.

These assertions are somewhat more complicated:

18. Either you will not have soup, or you will not have salad.

19. You will have neither soup nor salad.

20. You get either soup or salad, but not both.

We let S_1 mean that you get soup and S_2 mean that you get salad.

Assertion 18 can be paraphrased in this way: "Either *it is not the case that* you get soup, or *it is not the case that* you get salad." Translating this requires both "or" and "not." It becomes $\neg S_1 \vee \neg S_2$.

Assertion 19 also requires negation. It can be paraphrased as, "*It is not the case that* either you get soup or you get salad." We use parentheses to indicate that "not" negates the entire assertion $S_1 \vee S_2$, not just S_1 or S_2: "It is not the case that $(S_1 \vee S_2)$." This becomes simply $\neg(S_1 \vee S_2)$.

Notice that the parentheses are doing important work here. The assertion $\neg S_1 \vee S_2$ would mean "Either you will not have soup, or you will have salad."

Assertion 20 is an *exclusive or*. We can break the assertion into two parts. The first part says that you get one or the other. We translate this as $(S_1 \vee S_2)$. The second part says that you do not get both. We can paraphrase this as, "It is not the case that both you get soup and you get salad." Using both "not" and "and," we translate this as $\neg(S_1 \,\&\, S_2)$. Now we just need to put the two parts together. As we saw above, "but" can usually be translated as "and." Assertion 20 can thus be translated as $(S_1 \vee S_2) \,\&\, \neg(S_1 \,\&\, S_2)$.

Although "$\vee$" is an *inclusive or*, the preceding paragraph illustrates that we can symbolize an *exclusive or* in Propositional Logic. We just need more than one connective to do it.

EXERCISES 1.4.6. Using the given symbolization key, translate each English-language assertion into Propositional Logic.

M: Those creatures are men in suits.

C: Those creatures are chimpanzees.

G: Those creatures are gorillas.

1) Those creatures are men in suits, or they are not.

2) Those creatures are either gorillas or chimpanzees.

3) Either those creatures are chimpanzees, or they are not gorillas.

EXERCISES 1.4.7. Give a symbolization key and symbolize the following assertions in Propositional Logic.

1) Either Alice or Bob is a spy, but not both.

2) Either Bob is a spy, or it is the case both that the code has been broken and the German embassy is in an uproar.

3) Either the code has been broken or it has not, but the German embassy is in an uproar regardless.

4) Alice may or may not be a spy, but the code has been broken in any case.

EXERCISES 1.4.8. Using the given symbolization key, translate each assertion into English.

J: Romeo likes Juliet.

M: Mercutio likes Juliet.

T: Romeo likes Tybalt.

1) $M \vee T$ 2) $T \vee (\neg J \,\&\, M)$ 3) $\neg(M \vee J) \,\&\, \neg T$

1.4D. Implies ($\Rightarrow$). For the following assertions, let R mean "You will cut the red wire" and B mean "The bomb will explode."

21. If you cut the red wire, then the bomb will explode.

22. The bomb will explode if you cut the red wire.

23. The bomb will explode only if you cut the red wire.

Assertion 21 can be translated partially as "If R, then B." We can rephrase this as "R implies B." We will use the symbol "$\Rightarrow$" to represent "implies": the assertion becomes $R \Rightarrow B$. We call this connective "implies" or "if-then" (but many logicians call it a **conditional**). The assertion on the left-hand side (R in this example) is called the **hypothesis**, and the assertion on the right-hand side (B) is called the **conclusion**.

Assertion 22 tells us that if you cut the red wire, then the bomb will explode. Thus, it is logically equivalent to Assertion 21, so it can be symbolized as $R \Rightarrow B$.

Assertion 23 is also a conditional assertion that tells us something must be true if some other thing is true. Since the word "if" appears in the second half of the assertion, it might be tempting to symbolize this in the same way as Assertions 21 and 22. That would be a mistake.

The implication $R \Rightarrow B$ says that *if* R were true, *then* B would also be true. It does not say that your cutting the red wire is the *only* way that the bomb could explode. Someone else might cut the wire, or the bomb might be on a timer. The assertion $R \Rightarrow B$ does not say anything about what to expect if R is false. Assertion 23 is different. It says that the only conditions under which the bomb will explode involve your having cut the red wire; i.e., if the

bomb explodes, then you must have cut the wire. As such, Assertion 23 should be symbolized as $B \Rightarrow R$.

Remark 1.4.9. The paraphrased assertion "$\mathcal{A}$ only if $\mathcal{B}$" is logically equivalent to "If $\mathcal{A}$, then $\mathcal{B}$."

"If $\mathcal{A}$, then $\mathcal{B}$" means that if $\mathcal{A}$ is true, then so is $\mathcal{B}$. So we know that if the hypothesis $\mathcal{A}$ is true, but the conclusion $\mathcal{B}$ is false, then the implication "If $\mathcal{A}$, then $\mathcal{B}$" is false. (For example, if you cut the red wire, but the bomb does not explode, then Assertion 21 is obviously false.) We now consider the other possible situations, and determine whether the assertion "If $\mathcal{A}$, then $\mathcal{B}$" is true or not.

- Suppose, for instance, that you do *not* cut the red wire. Then Assertion 21 is not a lie, whether the bomb explodes or not, because the assertion does not promise anything in this case. Thus, we consider Assertion 21 to be true in this case. In general, if $\mathcal{A}$ is false, then the implication "$\mathcal{A} \Rightarrow \mathcal{B}$" is true. (It does not matter whether $\mathcal{B}$ is true or not.)
- The only remaining case to consider is when you cut the red wire and the bomb does explode. In this case, Assertion 21 has told the truth. In general, if $\mathcal{A}$ and $\mathcal{B}$ are true, then the implication "$\mathcal{A} \Rightarrow \mathcal{B}$" is true.

$\mathcal{A} \Rightarrow \mathcal{B}$ is true unless $\mathcal{A}$ is true and $\mathcal{B}$ is false.
In that case, the implication is false.

We can summarize this with a truth table for "implies."

$\mathcal{A}$	$\mathcal{B}$	$\mathcal{A} \Rightarrow \mathcal{B}$
T	T	T
T	F	F
F	T	T
F	F	T

Remark 1.4.10. Logic students are sometimes confused by the fact that $\mathcal{A} \Rightarrow \mathcal{B}$ is true whenever $\mathcal{A}$ is false, but it is actually quite natural. For example, suppose a teacher promises, "If you do all of the homework, then you will pass the course." A student who fails to do all of the homework cannot accuse the teacher of a falsehood, whether he passes the course or not.

Also, people often use this principle when speaking sarcastically. An example is the assertion, "If Rudy is the best player on the team, then pigs can fly." We all know that pigs cannot fly, but, logically, the assertion is true as long as Rudy is *not* the best player on the team.

WARNING. The connective "implies" is *not* commutative: you cannot swap the hypothesis and the conclusion without changing the meaning of the assertion, because it is easy to find a situation in which $\mathcal{A} \Rightarrow \mathcal{B}$ is true, but $\mathcal{B} \Rightarrow \mathcal{A}$ is false. (Namely, suppose $\mathcal{A}$ is false and $\mathcal{B}$ is true.)

Let us go back to the example with which we started our discussion of "$\Rightarrow$," in which R is the assertion "You will cut the red wire," and B means "The bomb will explode." There are many different ways of saying $R \Rightarrow B$ in English. Here are some of the ways; all of these mean the same thing!

- If you cut the red wire, then the bomb will explode.
- You cutting the red wire implies that the bomb will explode.
- In any circumstances in which you cut the red wire, the bomb will explode.
- Whenever you cut the red wire, the bomb will explode.
- The bomb will explode whenever you cut the red wire.

- The bomb exploding is a necessary consequence of you cutting the red wire.
- You cutting the red wire is sufficient to ensure that the bomb will explode.
- You cutting the red wire guarantees that the bomb will explode.
- You cut the red wire only if the bomb will explode.
- If the bomb does not explode, you must not have cut the red wire.
- Either you will not cut the red wire, or the bomb will explode.

EXERCISES 1.4.11. Using the given symbolization key, translate each English-language assertion into Propositional Logic.

A: Mister Ace was murdered.

B: The butler committed the murder.

C: The cook committed the murder.

D: The Duchess is lying.

E: Mister Edge was murdered.

F: The murder weapon was a frying pan.

1) If Mister Ace was murdered, then the cook did it.
2) If Mister Edge was murdered, then the cook did not do it.
3) If the murder weapon was a frying pan, then the culprit must have been the cook.
4) If the murder weapon was not a frying pan, then the culprit was either the cook or the butler.
5) Either the Duchess is lying, or it was Mister Edge who was murdered.
6) If Mister Ace was murdered, he was done in with a frying pan.
7) The cook murdered Mister Edge, but she did not use the frying pan.

EXERCISES 1.4.12. Give a symbolization key and symbolize the following assertions in Propositional Logic.

1) If Gregor plays first base, then the team will lose.
2) If either Gregor or Evan plays first base, then there will not be a miracle.
3) If neither Gregor nor Evan plays first base, then there will be a miracle.
4) The team will lose if there is no miracle.
5) If there is a miracle, then Gregor's mom will not bake cookies.

EXERCISES 1.4.13. For each deduction, write a symbolization key and translate the deduction as well as possible into Propositional Logic.

1) If Dorothy plays the piano in the morning, then Roger wakes up cranky. Dorothy plays piano in the morning if she is not distracted. So if Roger does not wake up cranky, then Dorothy must be distracted.
2) It will either rain or snow on Tuesday. If it rains, Neville will be sad. If it snows, Neville will be cold. Therefore, Neville will either be sad or cold on Tuesday.
3) If Zoog remembered to do his chores, then things are clean but not neat. If he forgot, then things are neat but not clean. Therefore, things are either neat or clean—but not both.

EXERCISES 1.4.14. Using the given symbolization key, translate each assertion into English.

J: Romeo likes Juliet.

M: Mercutio likes Juliet.

T: Romeo likes Tybalt.

1) $M \Rightarrow J$ 2) $J \vee (M \Rightarrow \neg T)$ 3) $(T \Rightarrow J)\ \&\ (M \Rightarrow J)$

1.4E. Iff ($\Leftrightarrow$). Consider these assertions:

24. The figure on the board is a triangle if it has exactly three sides.

25. The figure on the board is a triangle only if it has exactly three sides.

26. The figure on the board is a triangle if and only if it has exactly three sides.

Let T mean "The figure is a triangle" and S mean "The figure has exactly three sides."

Assertion 24 can be rephrased as, "If the figure has exactly three sides, then it is a triangle." So it can be translated as $S \Rightarrow T$.

Assertion 25 is importantly different. Remark 1.4.9 tells us that it can be translated as $T \Rightarrow S$.

Assertion 26 says two things: that "T is true if S is true" *and* that "T is true only if S is true." The first half is Assertion 24, and the second half is Assertion 25; thus, it can be translated as

$$(S \Rightarrow T)\ \&\ (T \Rightarrow S).$$

However, this "if and only if" comes up so often that it has its own name. We call this connective "**iff**," which is short for "if and only if" (but many logicians call it a **biconditional**).

Because we could always write $(\mathcal{A} \Rightarrow \mathcal{B})\ \&\ (\mathcal{B} \Rightarrow \mathcal{A})$ instead of $\mathcal{A} \Leftrightarrow \mathcal{B}$, we do not strictly speaking *need* to introduce a new symbol for "iff." Nevertheless, it is useful so often that it is commonly accepted as one of the basic logical connectives.

> $\mathcal{A} \Leftrightarrow \mathcal{B}$ is true if and only if $\mathcal{A}$ and $\mathcal{B}$ have the same truth value (either both are true or both are false).

This is the truth table for "iff":

$\mathcal{A}$	$\mathcal{B}$	$\mathcal{A} \Leftrightarrow \mathcal{B}$
T	T	T
T	F	F
F	T	F
F	F	T

EXERCISES 1.4.15. Using the given symbolization key, translate each English-language assertion into Propositional Logic.

E_1: Ava is an electrician.

E_2: Harrison is an electrician.

F_1: Ava is a firefighter.

F_2: Harrison is a firefighter.

S_1: Ava is satisfied with her career.

S_2: Harrison is satisfied with his career.

1) If Ava is not an electrician, then neither is Harrison, but if she is, then he is too.

2) Ava is satisfied with her career if and only if Harrison is not satisfied with his.

3) Harrison and Ava are both firefighters if and only if neither of them is an electrician.

EXERCISES 1.4.16. Using the given symbolization key, translate each assertion into English.

J: Romeo likes Juliet.

M: Mercutio likes Juliet.

T: Romeo likes Tybalt.

Y: Romeo likes Yorick.

1) $T \Leftrightarrow Y$ 2) $M \Leftrightarrow (J \vee Y)$ 3) $(J \Leftrightarrow M)\ \&\ (T \Rightarrow Y)$

1.5. Determining whether an assertion is true

To put them all in one place, the truth tables for the connectives of Propositional Logic are repeated here:

$\mathcal{A}$	$\neg\mathcal{A}$
T	F
F	T

$\mathcal{A}$	$\mathcal{B}$	$\mathcal{A}\ \&\ \mathcal{B}$	$\mathcal{A} \vee \mathcal{B}$	$\mathcal{A} \Rightarrow \mathcal{B}$	$\mathcal{A} \Leftrightarrow \mathcal{B}$
T	T	T	T	T	T
T	F	F	T	F	F
F	T	F	T	T	F
F	F	F	F	T	T

Truth tables for the connectives of Propositional Logic.

Every advanced math student needs to be able to quickly reproduce *all* of these truth tables, without looking them up.

Using these tables, you should be able to decide whether any given assertion is true or false, for any particular values of its assertion letters. (In this section, we often refer to assertion letters as "**variables**.")

EXAMPLE 1.5.1. Assume A is true, B is false, and C is false. Is $(A \vee B) \Rightarrow (B\ \&\ \neg C)$ true?

SOLUTION. We have

$$\begin{aligned}(A \vee B) \Rightarrow (B\ \&\ \neg C) &= (\mathsf{T} \vee \mathsf{F}) \Rightarrow (\mathsf{F}\ \&\ \neg\mathsf{F}) \\ &= \mathsf{T} \Rightarrow (\mathsf{F}\ \&\ \mathsf{T}) \\ &= \mathsf{T} \Rightarrow \mathsf{F} \\ &= \mathsf{F}.\end{aligned}$$

The assertion is not true. □

What does this mean in English? Suppose, for example, that we have the symbolization key

A: Bill baked an apple pie,

B: Bill baked a banana pie,

C: Bill baked a cherry pie.

Also suppose Ellen tells us (maybe because she knows what ingredients Bill has):

If Bill baked either an apple pie or a banana pie,
then he baked a banana pie, but did not bake a cherry pie.

Now, it turns out that

Bill baked an apple pie, but did not bake a banana pie, and did not bake a cherry pie.

Then the above calculation shows that *Ellen was wrong*; her assertion is *false*.

EXAMPLE 1.5.2. Assume A is true, B is false, and C is true. Is $(A \vee C) \Rightarrow \neg(A \Rightarrow B)$ true?

SOLUTION. We have

$$\begin{aligned}(A \vee C) \Rightarrow \neg(A \Rightarrow B) &= (\mathsf{T} \vee \mathsf{T}) \Rightarrow \neg(\mathsf{T} \Rightarrow \mathsf{F}) \\ &= \mathsf{T} \Rightarrow \neg\mathsf{F} \\ &= \mathsf{T} \Rightarrow \mathsf{T} \\ &= \mathsf{T}.\end{aligned}$$

The assertion is true. □

EXERCISES 1.5.3. Determine whether each assertion is true for the given values of the variables.

1) $(A \vee C) \Rightarrow \neg(A \Rightarrow B)$
   (a) A is true, B is false, and C is false.
   (b) A is false, B is true, and C is false.

2) $\big(P \vee \neg(Q \Rightarrow R)\big) \Rightarrow \big((P \vee Q)\,\&\,R\big)$
   (a) P, Q, and R are all true.
   (b) P is true, Q is false, and R is true.
   (c) P is false, Q is true, and R is false.
   (d) P, Q, and R are all false.

3) $\big((U\,\&\,\neg V) \vee (V\,\&\,\neg W) \vee (W\,\&\,\neg U)\big) \Rightarrow \neg(U\,\&\,V\,\&\,W)$
   (a) U, V, and W are all true.
   (b) U is true, V is true, and W is false.
   (c) U is false, V is true, and W is false.
   (d) U, V, and W are all false.

4) $(X \vee \neg Y)\,\&\,(X \Rightarrow Y)$
   (a) X and Y are both true.
   (b) X is true and Y is false.
   (c) X is false and Y is true.
   (d) X and Y are both false.

1.6. Tautologies and contradictions

Most assertions are true in some situations, and false in others. But some assertions are true in all situations, and others are false in all situations.

DEFINITION 1.6.1.

- A **tautology** is an assertion of Propositional Logic that is true in all situations; that is, it is true for all possible values of its variables.
- A **contradiction** is an assertion of Propositional Logic that is false in all situations; that is, it is false for all possible values of its variables.

EXAMPLE 1.6.2. The assertion $A \vee B$ is true when A is true (or B is true), but it is false when A and B are both false. Thus, the assertion is sometimes true and sometimes false; it is neither a contradiction nor a tautology.

EXAMPLE 1.6.3. Show that the assertion $\big(P\&(\neg Q \vee \neg R)\big) \Rightarrow (P \Rightarrow \neg Q)$ is neither a tautology nor a contradiction.

Scratchwork. We need to find values of the variables that make the assertion true, and other values that make the assertion false.

It is easy to make the assertion true, because an implication is true whenever its conclusion is true, so we just need to make $P \Rightarrow \neg Q$ true. And we can make this true by making $\neg Q$ true.

So we let Q be false. Then we can let P and R be whatever we want: it's probably simplest to let them both be false (the same as Q).

To make the assertion false, we need to make its hypothesis true and its conclusion false.

- Let's start with the conclusion $P \Rightarrow \neg Q$. To make this false, we need to make P true and $\neg Q$ false. Thus, we let $P = \mathsf{T}$ and $Q = \mathsf{T}$.
- Now, we consider the hypothesis $P \,\&\, (\neg Q \vee \neg R)$. Fortunately, we already decided to make P true, but we also need to make $\neg Q \vee \neg R$ true. Since we already decided to make Q true, we need to make $\neg R$ true, so we let $R = \mathsf{F}$.

SOLUTION. If P, Q, and R are all false, then

$$\begin{aligned}\big(P \,\&\, (\neg Q \vee \neg R)\big) \Rightarrow (P \Rightarrow \neg Q) &= \big(\mathsf{F} \,\&\, (\neg\mathsf{F} \vee \neg\mathsf{F})\big) \Rightarrow (\mathsf{F} \Rightarrow \neg\mathsf{F}) \\ &= \big(\mathsf{F} \,\&\, (\mathsf{T} \vee \mathsf{T})\big) \Rightarrow (\mathsf{F} \Rightarrow \mathsf{T}) \\ &= \big(\mathsf{F} \,\&\, \mathsf{T}\big) \Rightarrow (\mathsf{T}) \\ &= \mathsf{F} \Rightarrow \mathsf{T} \\ &= \mathsf{T},\end{aligned}$$

whereas if P and Q are true, but R is false, then

$$\begin{aligned}\big(P \,\&\, (\neg Q \vee \neg R)\big) \Rightarrow (P \Rightarrow \neg Q) &= \big(\mathsf{T} \,\&\, (\neg\mathsf{T} \vee \neg\mathsf{F})\big) \Rightarrow (\mathsf{T} \Rightarrow \neg\mathsf{T}) \\ &= \big(\mathsf{T} \,\&\, (\mathsf{F} \vee \mathsf{T})\big) \Rightarrow (\mathsf{T} \Rightarrow \mathsf{F}) \\ &= \big(\mathsf{T} \,\&\, \mathsf{T}\big) \Rightarrow (\mathsf{F}) \\ &= \mathsf{T} \Rightarrow \mathsf{F} \\ &= \mathsf{F}.\end{aligned}$$

Thus, the assertion is sometimes true and sometimes false, so it is neither a tautology nor a contradiction. □

EXERCISES 1.6.4. Show that each of the following assertions is neither a tautology nor a contradiction.

1) $A \Rightarrow (A \,\&\, B)$ 2) $(A \vee B) \Rightarrow A$ 3) $(A \Leftrightarrow B) \vee (A \,\&\, \neg B)$
4) $(X \Rightarrow Z) \Rightarrow (Y \Rightarrow Z)$ 5) $\big(P \& \neg(Q \& R)\big) \vee (Q \Rightarrow R)$

EXAMPLE 1.6.5 (Law of Excluded Middle). It is easy to see that the assertion $A \vee \neg A$ is true when A is true, and also when A is false. Thus, the assertion is true for both possible values of the variable A, so it is a tautology:

$A \vee \neg A$ is a tautology.

Remark 1.6.6. The above tautology is called the "Law of Excluded Middle" because it says every assertion is either true or false: there is no middle ground where an assertion is partly true and partly false.

EXAMPLE 1.6.7. It is easy to see that the assertion $A \,\&\, \neg A$ is false when A is true, and also when A is false. Thus, the assertion is false for both possible values of the variable A, so it is a contradiction:

$A \,\&\, \neg A$ is a contradiction.

Remark 1.6.8. The assertions $A \vee \neg A$ and $A \,\&\, \neg A$ are the most important (and most common) examples of tautologies and contradictions. However, they will usually arise with some other expression plugged into the variable A. For example, by letting A be the assertion $(P \vee Q) \Rightarrow R$, we obtain the tautology

$$\big((P \vee Q) \Rightarrow R\big) \vee \neg\big((P \vee Q) \Rightarrow R\big),$$

which is a more complicated example of the Law of Excluded Middle, and we also obtain the contradiction

$$\big((P \vee Q) \Rightarrow R\big) \,\&\, \neg\big((P \vee Q) \Rightarrow R\big).$$

EXAMPLE 1.6.9. We can also give examples in English, rather than in symbols; consider these assertions:

27. It is raining.
28. Either it is raining, or it is not.
29. It is both raining and not raining.

In order to know whether Assertion 27 is true, you would need to check the weather. Logically speaking, it could be either true or false, so it is neither a tautology nor a contradiction.

Assertion 28 is different. You do not need to look outside to know that it is true, regardless of what the weather is like. So it is a tautology.

You do not need to check the weather to know about Assertion 29, either. It must be false, simply as a matter of logic. It might be raining here and not raining across town, or it might be raining now but stop raining even as you read this, but it is impossible for it to be both raining and not raining in any given situation (at any particular time and place). Thus, the third assertion is false in every possible situation; it is a contradiction.

EXERCISES 1.6.10. Which of the following are possible? For those that are possible, give an example. For those that are not, explain why.

1) A valid deduction whose conclusion is a contradiction.
2) A valid deduction whose conclusion is a tautology.
3) A valid deduction that has a tautology as one of its hypotheses.
4) A valid deduction that has a contradiction as one of its hypotheses.
5) An invalid deduction whose conclusion is a contradiction.
6) An invalid deduction whose conclusion is a tautology.
7) An invalid deduction that has a tautology as one of its hypotheses.
8) An invalid deduction that has a contradiction as one of its hypotheses.

1.7. Logical equivalence

In your previous mathematics classes (such as algebra and trigonometry), you encountered many examples where two different-looking formulas turned out to be equal. Analogously, in Logic, there can be two different assertions that happen to have the same truth value in all possible situations. (This means that, for every possible assignment of true or false to the variables, either both of the assertions are true, or both are false.) Such assertions are said to be **logically equivalent**.

NOTATION 1.7.1. We will write $\mathcal{A} \equiv \mathcal{B}$ to denote that $\mathcal{A}$ is logically equivalent to $\mathcal{B}$.

It can take a lot of work to verify that two assertions are logically equivalent. On the other hand, to show that two assertions are *not* logically equivalent, you only need to find one example of an assignment to the variables, such that one of the assertions is true and the other is false.

EXAMPLE 1.7.2. If A is true and B is false, then $A \vee B$ is true, but $A \Rightarrow B$ is false. Therefore, the assertions $A \vee B$ and $A \Rightarrow B$ are *not* logically equivalent.

EXERCISE 1.7.3. Show that each of the following pairs of sentences are *not* logically equivalent.

1) $A \vee B \vee \neg C$, $(A \vee B)\ \&\ (C \Rightarrow A)$

2) $(P \Rightarrow Q) \vee (Q \Rightarrow P)$, $P \vee Q$

3) $(X\ \&\ Y) \Rightarrow Z$, $X \vee (Y \Rightarrow Z)$

EXAMPLE 1.7.4. Show that $\neg(A \vee B) \equiv \neg A\ \&\ \neg B$.

SOLUTION. The variables A and B may each be either true or false, and we will evaluate both assertions for all possible combinations. To make it clear that none of the possibilities have been missed, we proceed systematically: for each value of A, we consider the two possible values for B.

Case 1: Assume A is true.

Subcase 1.1: Assume B is true. We have

$$\neg(A \vee B) \quad = \quad \neg(T \vee T) \quad = \quad \neg T \quad = \quad F$$

and

$$\neg A\ \&\ \neg B \quad = \quad \neg T\ \&\ \neg T \quad = \quad F\ \&\ F \quad = \quad F.$$

Both assertions are false.

Subcase 1.2: Assume B is false. We have

$$\neg(A \vee B) \quad = \quad \neg(T \vee F) \quad = \quad \neg T \quad = \quad F$$

and

$$\neg A\ \&\ \neg B \quad = \quad \neg T\ \&\ \neg F \quad = \quad T\ \&\ F \quad = \quad F.$$

Both assertions are false.

Case 2: Assume A is false.

Subcase 2.1: Assume B is true. We have

$$\neg(A \vee B) \quad = \quad \neg(F \vee T) \quad = \quad \neg T \quad = \quad F$$

and

$$\neg A\ \&\ \neg B \quad = \quad \neg F\ \&\ \neg T \quad = \quad T\ \&\ F \quad = \quad F.$$

Both assertions are false.

Subcase 2.2: Assume B is false. We have

$$\neg(A \vee B) \quad = \quad \neg(F \vee F) \quad = \quad \neg F \quad = \quad T$$

and

$$\neg A\ \&\ \neg B \quad = \quad \neg F\ \&\ \neg F \quad = \quad T\ \&\ T \quad = \quad T.$$

Both assertions are true.

In all cases, either both assertions are true, or both are false. Therefore, they are logically equivalent. □

We can also solve the problem without doing so much work:

EASIER SOLUTION. Note that the assertion $\neg(A \vee B)$ is true if and only if $A \vee B$ is false, which means that neither A nor B is true. Therefore,

$$\neg(A \vee B) \text{ is true if and only if } A \text{ and } B \text{ are both false.}$$

Also, $\neg A \,\&\, \neg B$ is true if and only if $\neg A$ and $\neg B$ are both true, which means that:

$$\neg A \,\&\, \neg B \text{ is true if and only if } A \text{ and } B \text{ are both false.}$$

So the two assertions $\neg(A \vee B)$ and $\neg A \,\&\, \neg B$ are true in exactly the same situation (namely, when A and B are both false); and they are both false in all other situations. Therefore, they are logically equivalent. □

EXERCISES 1.7.5. Verify each of the following important logical equivalences. For most of these, you should not need to evaluate the assertions for all possible values of the variables.

1) rules of negation:

$$\begin{aligned} \neg\neg A &\equiv A \\ \neg(A \,\&\, B) &\equiv \neg A \vee \neg B \\ \neg(A \vee B) &\equiv \neg A \,\&\, \neg B \\ \neg(A \Rightarrow B) &\equiv A \,\&\, \neg B \\ \neg(A \Leftrightarrow B) &\equiv A \Leftrightarrow \neg B \end{aligned}$$

2) commutativity of &, $\vee$, and $\Leftrightarrow$:

$$\begin{aligned} A \,\&\, B &\equiv B \,\&\, A \\ A \vee B &\equiv B \vee A \\ A \Leftrightarrow B &\equiv B \Leftrightarrow A \end{aligned}$$

3) associativity of & and $\vee$:

$$\begin{aligned} (A \,\&\, B) \,\&\, C &\equiv A \,\&\, (B \,\&\, C) \\ (A \vee B) \vee C &\equiv A \vee (B \vee C) \end{aligned}$$

Remark 1.7.6. The rules of negation for & and $\vee$ are often called *De Morgan's Laws*, in honour of the British mathematician Augustus De Morgan (1806–1871, `http://en.wikipedia.org/wiki/Augustus_De_Morgan`).

The rules of negation can be used to simplify the negation of any assertion.

EXAMPLE 1.7.7. Simplify $\neg\big((A \vee B) \Rightarrow (A \,\&\, \neg C)\big)$.

SOLUTION. We have

$$\begin{aligned} \neg\big((A \vee B) \Rightarrow (A \,\&\, \neg C)\big) &\equiv (A \vee B) \,\&\, \neg(A \,\&\, \neg C) \\ &\equiv (A \vee B) \,\&\, (\neg A \vee \neg\neg C) \\ &\equiv (A \vee B) \,\&\, (\neg A \vee C). \end{aligned}$$ □

Remark 1.7.8. If $\mathcal{A} \equiv \mathcal{B}$, then $\mathcal{A},\ \therefore \mathcal{B}$ is a valid deduction. For example, the above example shows that

$$\neg\big((A \vee B) \Rightarrow (A \,\&\, \neg C)\big),\ \therefore (A \vee B) \,\&\, (\neg A \vee C)$$

is a valid deduction.

EXERCISE 1.7.9. Use the rules of negation to simplify each of the following assertions (until negation is not applied to anything but variables).

1) $\neg\big((A \vee B) \Rightarrow (C \,\&\, D)\big)$
2) $\neg\big((A \Rightarrow B) \vee (C \,\&\, D)\big)$
3) $\neg\Big(A \Rightarrow \big(B \Rightarrow (C \Rightarrow D)\big)\Big)$
4) $\neg\Big(\big((A \Rightarrow B) \Rightarrow C\big) \Rightarrow D\Big)$
5) $\neg\big((P \vee \neg Q) \,\&\, R\big)$
6) $\neg(P \,\&\, Q \,\&\, R \,\&\, S)$
7) $\neg\Big(\big(P \Rightarrow (Q \,\&\, \neg R)\big) \vee (P \,\&\, \neg Q)\Big)$

EXERCISE 1.7.10. Use the rules of negation to simplify the *negation* of each of these assertions. Express your answers in English.

1) If it is raining, then the bus will not be on time.
2) I am sick, and I am tired.
3) Either the Pope is here, or the Queen and the Registrar are both here.
4) If Tom forgot his backpack, then Sam will eat either a pickle or a potato, and either Bob will not have lunch, or Alice will drive to the store.

1.8. Converse and contrapositive

The **converse** of an implication $\mathcal{A} \Rightarrow \mathcal{B}$ is the implication $\mathcal{B} \Rightarrow \mathcal{A}$. For example, the converse of "if Bob pays the cashier a dollar, then the server gives Bob an ice cream cone" is "if the server gives Bob an ice cream, then Bob pays the cashier a dollar." It should be clear that these are not saying the same thing. (For example, perhaps Bob has a coupon for a free cone.) This illustrates the fact that the converse of an assertion is usually not logically equivalent to the original assertion. In other words (as was mentioned in Section 1.4D), the connective $\Rightarrow$ is *not* commutative:

EXERCISE 1.8.1. Show that $A \Rightarrow B$ is *not* logically equivalent to its converse $B \Rightarrow A$.

The **inverse** of an implication $\mathcal{A} \Rightarrow \mathcal{B}$ is the implication $\neg\mathcal{A} \Rightarrow \neg\mathcal{B}$. For example, the inverse of "if Bob pays the cashier a dollar, then the server gives Bob an ice cream cone" is "if Bob does not pay the cashier a dollar, then the server does not give Bob an ice cream cone." It should be clear that these are not saying the same thing (because one assertion is about what happens if Bob pays a dollar, and the other is about the completely different situation in which Bob does not pay a dollar). This illustrates the fact that the inverse of an assertion is usually not logically equivalent to the original assertion:

EXERCISE 1.8.2. Show that $A \Rightarrow B$ is *not* logically equivalent to its inverse $\neg A \Rightarrow \neg B$.

The **contrapositive** of an implication is the converse of its inverse (or the inverse of its converse, which amounts to the same thing). That is,

the contrapositive of $\mathcal{A} \Rightarrow \mathcal{B}$ is the implication $\neg\mathcal{B} \Rightarrow \neg\mathcal{A}$.

For example, the contrapositive of "if Bob pays the cashier a dollar, then the server gives Bob an ice cream cone" is "if the server does not give Bob an ice cream cone, then Bob does not pay the cashier a dollar." A bit of thought should convince you that these *are* saying the same thing. This illustrates the following important fact:

Any implication is logically equivalent to its contrapositive.

EXERCISE 1.8.3. Show that $A \Rightarrow B$ is logically equivalent to its contrapositive $\neg B \Rightarrow \neg A$.

Remark 1.8.4. The inverse will not be important to us, although the converse and the contrapositive are fundamental. However, it may be worth mentioning that the inverse is the contrapositive of the converse, and therefore the inverse and the converse are logically equivalent to each other.

WARNING. Implications (that is, those of the form $\mathcal{A} \Rightarrow \mathcal{B}$) are the only assertions that have a converse or a contrapositive. For example, the converse of "I hate cheese" does not exist, because this assertion is not an if-then statement.

EXERCISES 1.8.5. State (a) the converse and (b) the contrapositive of each implication. (*You do not need to show your work.*)

1) If the students come to class, then the teacher lectures.
2) If it rains, then I carry my umbrella.
3) If I have to go to school this morning, then today is a weekday.
4) If you give me \$5, I can take you to the airport.
5) If the Mighty Ducks are the best hockey team, then pigs can fly.
6) Alberta is a province.
7) If you want to do well in your math class, then you need to do all of the homework.

1.9. Some valid deductions

Recall that a deduction is **valid** if its conclusion is true in all situations where all of its hypotheses are true. This means, for each and every possible assignment of true/false to the variables, if all of the hypotheses true, then the conclusion is also true.

EXAMPLE 1.9.1. Explain how you know that the following deduction is valid.

$$A \vee B, \quad \neg A, \quad \therefore \ B.$$

SOLUTION. Assume we are in a situation in which both hypotheses of the deduction are true. Then, from the first hypothesis, we know that either A is true or B is true. However, from the second hypothesis, we know that A is *not* true. Therefore, it must be B that is true. Hence, the conclusion of the deduction is true. □

EXERCISES 1.9.2. Answer each of the questions below and justify your answer.

1) Assume $(\mathcal{A} \,\&\, \mathcal{B}) \Rightarrow \mathcal{C}$ is neither a tautology nor a contradiction. What can you say about the deduction "$\mathcal{A}$, $\mathcal{B}$, $\therefore \mathcal{C}$"?
2) Assume $\mathcal{A}$ is a contradiction. What can you say about the deduction "$\mathcal{A}$, $\mathcal{B}$, $\therefore \mathcal{C}$"?
3) Assume $\mathcal{C}$ is a tautology. What can you say about the deduction "$\mathcal{A}$, $\mathcal{B}$, $\therefore \mathcal{C}$"?

TERMINOLOGY 1.9.3. Any valid deduction can be called a **theorem**.

EXERCISE 1.9.4 (Rules of Propositional Logic). It is not difficult to see that each of the following is a valid deduction. For each of them, either give a short explanation of how you know that it is valid, or verify the deduction by evaluating the conclusion for all possible values of the variables that make the hypotheses true.

1) **repeat**: $A,\ \therefore A$
2) **&-introduction**: $A,\ B,\ \therefore A \,\&\, B$
3) **&-elimination**: $A \,\&\, B,\ \therefore A \qquad A \,\&\, B,\ \therefore B$
4) **$\vee$-introduction**: $A,\ \therefore A \vee B \qquad B,\ \therefore A \vee B$
5) **$\vee$-elimination**: $A \vee B,\ \neg A,\ \therefore B \qquad A \vee B,\ \neg B,\ \therefore A$
6) **$\Rightarrow$-elimination**: $A \Rightarrow B,\ A,\ \therefore B$
7) **$\Leftrightarrow$-introduction**: $A \Rightarrow B,\ B \Rightarrow A,\ \therefore A \Leftrightarrow B$
8) **$\Leftrightarrow$-elimination**: $A \Leftrightarrow B,\ \therefore A \Rightarrow B \qquad A \Leftrightarrow B,\ \therefore B \Rightarrow A$
9) **proof by cases**: $A \vee B,\ A \Rightarrow C,\ B \Rightarrow C,\ \therefore C$

All of the theorems in Exercise 1.9.4 will be used on a regular basis in the following chapters (and in your later mathematics courses).

OTHER TERMINOLOGY. Most logicians call the $\Rightarrow$-elimination rule by its Latin name, which is **modus ponens**. (According to *Wikipedia*, this is short for *modus ponendo ponens*, which means "the way that affirms by affirming".)

Remark 1.9.5. A theorem remains valid if we change the names of the variables. For example, $P \vee Q,\ \neg P,\ \therefore Q$ is the same as $\vee$-elimination, but we have replaced A with P and B with Q. (In the language of high-school algebra, we have plugged in P for A, and plugged in Q for B.) Indeed, it should be clear that any theorem remains valid even if we substitute more complicated expressions into the variables.

EXAMPLE 1.9.6. The theorem

$$(X \vee Y) \Rightarrow (Y \vee Z),\ X \vee Y,\ \therefore Y \vee Z$$

is obtained from "$\Rightarrow$-elimination," by letting $A = (X \vee Y)$ and $B = (Y \vee Z)$.

EXERCISE 1.9.7. Each of the following is a valid theorem that is obtained from one of the basic theorems of Exercise 1.9.4, by substituting some expressions into the variables. Identify the theorem it is obtained from, and the expressions that were substituted into each variable.

1) $(A \vee B) \,\&\, (Y \Rightarrow Z),\ \therefore Y \Rightarrow Z$
2) $(A \vee B) \,\&\, (Y \Rightarrow Z),\ \therefore (A \vee B) \,\&\, (Y \Rightarrow Z)$
3) $A \vee B,\ \therefore (A \vee B) \vee (Y \Rightarrow Z)$
4) $(A \vee B), (Y \Rightarrow Z),\ \therefore (A \vee B) \,\&\, (Y \Rightarrow Z)$

EXERCISE 1.9.8. Each of the following is the English-language version of a valid theorem that is obtained from one of the basic theorems of Exercise 1.9.4, by substituting some expressions into the variables. Identify the theorem it is obtained from.

1) Susie will stop at either the grocery store or the drug store. If she stops at the grocery store, she will buy milk. If she stops at the drug store, she will buy milk. Therefore, I am sure that Susie will buy milk.

2) My opponent in this election is a liar! My opponent in this election is a cheat! Therefore, I say to you that my opponent is a liar and a cheat!

3) John went to the store. Therefore, as I already told you, John went to the store.

4) If I had $50, I would be able to buy a new coat. Hey, look! I found a $50 bill on the sidewalk! So I will be able to buy a new coat.

SUMMARY:

- Assertions stated in English can be translated into Propositional Logic (and vice-versa)
- In mathematics, "or" is inclusive.
- Notation:
  - $\neg$ (not; means "It is not the case that_____")
  - & (and; means "Both _____ and _____")
  - $\vee$ (or; means "Either _____ or _____")
  - $\Rightarrow$ (implies; means "If _____ then _____")
  - $\Leftrightarrow$ (iff; means "_____ if and only if _____")
- Important definitions:
  - assertion
  - deduction
  - valid, invalid
  - tautology
  - contradiction
  - logically equivalent
  - converse
  - contrapositive
- Determining whether an assertion is true (for particular values of its variables)
- An implication might *not* be equivalent to its converse.
- Every implication is logically equivalent to its contrapositive.
- Basic laws of Propositional Logic:
  - Law of Excluded Middle
  - rules of negation
  - commutativity of &, $\vee$, and $\Leftrightarrow$
  - associativity of & and $\vee$
- "Theorem" is another word for "valid deduction"
- Basic theorems of Propositional Logic:
  - repeat
  - introduction and elimination rules for &, $\vee$, and $\Leftrightarrow$
  - elimination rule for $\Rightarrow$
  - proof by cases

Chapter 2

Two-Column Proofs

No way of thinking or doing, however ancient, can be trusted without proof.
What everybody echoes or in silence passes by as true to-day
may turn out to be falsehood to-morrow, mere smoke of opinion...

Henry David Thoreau (1817–1862)), American author
Walden

The aim of a *proof* is to show that a deduction is valid, and it does this by putting together a number of simpler deductions that are already known to be valid. Ultimately, our goal is to teach you to write clear and correct proofs, in English, of claims stated in English. But we will start with the simpler situation of proofs written in the language of Propositional Logic. This has several advantages:

- it allows assertions to be written more concisely, because entire English phrases are abbreviated to a single letter,
- it avoids the difficulties caused by the fact that sentences written in English can be ambiguous, and
- it displays the logical structure of a proof in a way that makes it easier to decide whether or not each step in a proof is valid.

After you are familiar with proofs in this simpler setting, you will employ the same principles to write proofs in English.

2.1. First example of a two-column proof

Let us begin our exploration of proofs by looking at the following simple deduction.

Hypotheses:

1. $P \Rightarrow (Q \,\&\, R)$
2. P

Conclusion: R

We will prove that it is valid by showing it is a combination of deductions that are already known to be valid. Informally, we could try to convince someone that the deduction is valid by making the following explanation:

> Assume the Hypotheses (1) and (2) are true. Then applying $\Rightarrow$-elimination (with P in the role of A, and $Q \,\&\, R$ in the role of B) establishes that $Q \,\&\, R$ is true. (This is an *intermediate conclusion*. It follows logically from the hypotheses, and is helpful, but it is not the conclusion we want.) Now, applying &-elimination (with Q in the role of A, and R in the role of B) establishes that R is true. This is the conclusion of the deduction. Thus, we see that if

the hypotheses of this deduction are true, then the conclusion is also true. So the deduction is valid.

For emphasis, let us repeat that this explanation shows that the deduction is merely a combination of deductions that were already known to be valid.

Remark 2.1.1. Notice that we are using the fact that a valid deduction applies to all possible values of its variables, so we can plug any assertions we want into its variables. This is what allows us to talk about using (for example) "Q & R in the role of B." Another way of saying this, is that we are introducing a new symbolization key in which we let A stand for P, and let B stand for Q & R.

Formally, a **proof** is a sequence of assertions. The first assertions of the sequence are assumptions; these are the hypotheses of the deduction. It is required that every assertion later in the sequence is an immediate consequence of earlier assertions. (There are specific rules that determine which assertions are allowed to appear at each point in the proof.) The final assertion of the sequence is the conclusion of the deduction.

In this chapter, we use the format known as "**two-column proofs**" for writing our proofs. As indicated in the tableau below:

- Assertions appear in the left column.
- The reason (or "justification") for including each assertion appears in the right column. (The allowable justifications will be discussed in the later sections of this chapter.)

⟨*assertion*⟩	⟨*justification*⟩

Every assertion in a two-column proof needs to have a justification in the second column.

For clarity, we draw a dark horizontal line to separate the hypotheses from the rest of the proof. (In addition, we will number each row of the proof, for ease of reference, and we will make the left border of the figure a dark line.) For example, here is a two-column proof that justifies the deduction above. It starts by listing the hypotheses of the deduction, and ends with the correct conclusion.

1	$P \Rightarrow (Q \,\&\, R)$	hypothesis
2	P	hypothesis
3	$Q \,\&\, R$	$\Rightarrow$-elim (lines 1 and 2)
4	R	&-elim (line 3)

In this example, the assertions were written in the language of Propositional Logic, but sometimes we will write our proofs in English. For example, here is a symbolization key that allows us to translate P, Q, and R into English. For convenience, this same symbolization key will be used in many of the examples in this chapter.

P: The Pope is here.

Q: The Queen is here.

R: The Registrar is here.

Now, we can translate the deduction into English:

Hypotheses:

1. If the Pope is here, then the Queen and the Registrar are also here.
2. The Pope is here.

Conclusion: The Registrar is here.

And we can provide a two-column proof in English:

1	If the Pope is here, then the Queen and the Registrar are also here.	hypothesis
2	The Pope is here.	hypothesis
3	The Queen and the Registrar are both here.	⇒-elim (lines 1 and 2)
4	The Registrar is here.	&-elim (line 3)

While you are getting accustomed to two-column proofs, it will probably be helpful to see examples in *both* English *and* Propositional Logic. To save space, and make it easier to compare the two, the text will sometimes combine both proofs into one figure, by adding a third column at the right that states the English-language versions of the assertions:

⟨ *assertion in Propositional Logic* ⟩	⟨*justification*⟩	⟨ *English-language version of the assertion* ⟩

For example, here is what we get by combining the above two proofs:

1	$P \Rightarrow (Q \,\&\, R)$	hypothesis	If the Pope is here, then the Queen and the Registrar are also here.
2	P	hypothesis	The Pope is here.
3	$Q \,\&\, R$	⇒-elim (lines 1 and 2)	The Queen and the Registrar are both here.
4	R	&-elim (line 3)	The Registrar is here.

The next few sections will explain the justifications that are allowed in a two-column proof.

2.2. Hypotheses and theorems in two-column proofs

A two-column proof must start by listing all of the hypotheses of the deduction, and each hypothesis is justified by writing the word **hypothesis** in the second column. (This is the only rule that is allowed above the dark horizontal line, and it is not allowed below the dark horizontal line.) We saw this rule in the above examples of two-column proofs. As a synonym for "hypothesis," one sometimes says "given" or "assumption."

Any deduction that is *already known to be valid* can be used as a justification *if* its hypotheses have been verified earlier in the proof. (And the lines where the hypotheses appear are written in parentheses after the name of the theorem.) For example, the theorems "⇒-elim" and "∨-elim were used in our first examples of two-column proofs. These and several other very useful theorems were given in Exercise 1.9.4. You will be expected to be familiar with all of them.

EXAMPLE 2.2.1. Here is a proof of the deduction

$$P \vee Q, \quad Q \Rightarrow R, \quad \neg P, \quad \therefore R.$$

We provide an English translation by using the symbolization key on page 30.

1	$P \vee Q$	hypothesis	Either the Pope is here, or the Queen is here.
2	$Q \Rightarrow R$	hypothesis	If the Queen is here, then the Registrar is also here.
3	$\neg P$	hypothesis	The Pope is not here.
4	Q	$\vee$-elim (lines 1 and 3)	The Queen is here.
5	R	$\Rightarrow$-elim (lines 2 and 4)	The Registrar is here.

EXAMPLE 2.2.2. Here is a proof of the deduction $\neg L \Rightarrow (J \vee L)$, $\neg L$, $\therefore J$.

1	$\neg L \Rightarrow (J \vee L)$	hypothesis
2	$\neg L$	hypothesis
3	$J \vee L$	$\Rightarrow$-elim (lines 1 and 2)
4	J	$\vee$-elim (lines 3 and 2)

EXERCISE 2.2.3. Provide a justification (rule and line numbers) for each line of this proof.

1	$W \Rightarrow \neg B$	
2	$A \,\&\, W$	
3	$\neg B \Rightarrow (J \,\&\, K)$	
4	W	
5	$\neg B$	
6	$J \,\&\, K$	
7	K	

EXERCISE 2.2.4. Provide a justification (rule and line numbers) for each line of this proof.

1	$A \Rightarrow B$	
2	$\neg A \Rightarrow B$	
3	$A \vee \neg A$	
4	B	

EXERCISES 2.2.5. Write a two-column proof of each of the following deductions:

1) $P \vee Q$, $Q \vee R$, $\neg Q$, $\therefore P \,\&\, R$.

2) $(E \vee F) \vee G$, $\neg F \,\&\, \neg G$, $\therefore E$.

EXERCISES 2.2.6. Write a two-column proof of each of the following deductions. (Write the assertions in English.)

1) *Hypotheses:*
The Pope and the Queen are here.
Conclusion: The Queen is here.

2) *Hypotheses:*
The Pope is here.
The Registrar and the Queen are here.
Conclusion: The Queen and the Pope are here.

3) *Hypotheses:*
If the Pope is here, then the Queen is here.
If the Queen is here, then the Registrar is here.
The Pope is here.
Conclusion: The Registrar is here.

4) Grace is sick.
Frank is sick.
$\therefore$ Either Grace and Frank are both sick, or Ellen is sick.

EXAMPLE 2.2.7. Many proofs use the Rules of Negation or the fact that any statement is logically equivalent to its contrapositive. Here is an example.

1	$\neg P \Rightarrow (Q \,\&\, R)$	hypothesis	If the Pope is not here, then the Queen and the Registrar are here.
2	$\neg Q \vee \neg R$	hypothesis	Either the Queen is not here, or the Registrar is not here.
3	$\neg(Q \,\&\, R) \Rightarrow \neg\neg P$	contrapositive of line 1	If it is not the case that both the Queen and the Registrar are here, then it is not the case that the Pope is not here.
4	$(\neg Q \vee \neg R) \Rightarrow P$	Rules of Negation applied to line 3	If it is the case either that the Queen is not here, or that the Registrar is not here, then the Pope is here.
5	P	$\Rightarrow$-elim (lines 4 and 2)	The Pope is here.

EXERCISE 2.2.8. Provide a justification (rule and line numbers) for each line of these proofs.

1)

1	$H \Rightarrow F$	
2	$H \Rightarrow G$	
3	$(F \,\&\, G) \Rightarrow I$	
4	$\neg I$	
5	$\neg I \Rightarrow \neg(F \,\&\, G)$	
6	$\neg(F \,\&\, G)$	
7	$\neg F \vee \neg G$	
8	$\neg F \Rightarrow \neg H$	
9	$\neg G \Rightarrow \neg H$	
10	$\neg H$	

2)

1	$(W \vee X) \Rightarrow (Y \,\&\, Z)$	
2	$\neg Y$	
3	$\neg Y \vee \neg Z$	
4	$\neg(Y \,\&\, Z) \Rightarrow \neg(W \vee X)$	
5	$(\neg Y \vee \neg Z) \Rightarrow (\neg W \,\&\, \neg X)$	
6	$\neg W \,\&\, \neg X$	
7	$\neg X$	

EXERCISES 2.2.9. Give a two-column proof of each of these deductions.

1) $A \Rightarrow B$, $\neg B$, $\therefore \neg A$

2) $(L \vee M) \Rightarrow (N \,\&\, O)$, M, $\therefore O$

3)
Hypotheses:
Either the Pope is not here, or the Queen is here.
The Pope is here.
Conclusion: Either the Queen is here, or else the Registrar and the Pope are both here.

4)
Hypotheses:
If Sammy is not tired, then she does not need a nap.
Sammy needs a nap.
Conclusion: Sammy is tired.

Remark 2.2.10. You should also remember that & and $\vee$ are commutative, so, for example,

$$F \Rightarrow \big((E \,\&\, D \,\&\, C \,\&\, B) \vee (A \,\&\, B)\big) \quad \equiv \quad F \Rightarrow \big((A \,\&\, B) \vee (B \,\&\, C \,\&\, D \,\&\, E)\big).$$

2.3. Subproofs for ⇒-introduction

Consider this deduction:

	$P \Rightarrow R$	If the Pope is here, then the Registrar is here.
$\therefore$	$(P \,\&\, Q) \Rightarrow R$	If the Pope and the Queen are both here, then the Registrar is here.

The deduction is a valid one. Intuitively, we can justify it by noting that if $P \,\&\, Q$ is true, then P is certainly true, so the hypothesis implies R is true. Thus, we have verified that $(P \,\&\, Q) \Rightarrow R$. The ⇒-introduction rule will allow us to turn this intuitive justification into an official proof.

We begin the proof by writing down the hypothesis of the deduction and drawing a dark horizontal line, like this:

1	$P \Rightarrow R$	hypothesis	If the Pope is here, then the Registrar is here.

The conclusion of the deduction is an assertion about what happens when $P \,\&\, Q$ is true. That is, we want to see what happens if we assume, for the sake of argument, that the assertion $P \,\&\, Q$ is true. To accomplish this, what we will do is start a **subproof**, a proof within the main proof, where we assume that $P \,\&\, Q$ is true. When we start a subproof, we start a new set of double columns, and indent them from the left margin. Then we write in an assumption for the subproof. This can be anything we want. In the case at hand, we want to assume $P \,\&\, Q$. Our proof now looks like this:

1	$P \Rightarrow R$	hypothesis	If the Pope is here, then the Registrar is here.
2	$P \,\&\, Q$	assumption	Suppose the Pope and the Queen are both here.

It is important to notice that we are not claiming to have proven $P \,\&\, Q$ (that the Pope and the Queen are here). You can think of the subproof as posing the question: What could we show *if* $P \,\&\, Q$ were true? For one thing, we can derive P. So we do:

1	$P \Rightarrow R$	hypothesis	If the Pope is here, then the Registrar is here.
2	$P \,\&\, Q$	assumption	Suppose the Pope and the Queen are both here.
3	P	&-elim (line 2)	The Pope is here.

And now, since the Pope is here, we can derive R, from the hypothesis that $P \Rightarrow R$:

1	$P \Rightarrow R$	hypothesis	If the Pope is here, then the Registrar is here.
2	$P \,\&\, Q$	assumption	Suppose the Pope and the Queen are both here.
3	P	&-elim (line 2)	The Pope is here.
4	R	⇒-elim (lines 1 and 3)	The Registrar is here.

This has shown that *if* we had $P \,\&\, Q$ as a hypothesis, *then* we could prove R. In effect, we have proven $(P \,\&\, Q) \Rightarrow R$: that if the Pope and the Queen are here, then the Registrar is here. In recognition of this, the if-introduction rule (⇒-intro) will allow us to close the subproof and derive $(P \,\&\, Q) \Rightarrow R$ in the main proof. Our final proof looks like this:

1	$P \Rightarrow R$	hypothesis	If the Pope is here, then the Registrar is here.
2	$P \,\&\, Q$	assumption	Suppose the Pope and the Queen are both here.
3	P	&-elim (line 2)	The Pope is here.
4	R	$\Rightarrow$-elim (lines 1 and 3)	The Registrar is here.
5	$(P \,\&\, Q) \Rightarrow R$	$\Rightarrow$-intro (lines 2–4)	If the Pope and the Queen are both here, then the Registrar is here.

Notice that the justification for applying the $\Rightarrow$-intro rule is the entire subproof. Usually that will be more than just three lines.

It may seem as if the ability to assume anything at all in a subproof would lead to chaos: does it allow you to prove any conclusion from any hypotheses? The answer is no, it does not. Consider this proof:

1	P	hypothesis	The Pope is here.
2	Q	assumption	Suppose that the Queen is here.
3	Q	repeat (line 2)	As mentioned previously, the Queen is here.

It may seem as if this is a proof that you can derive any conclusion Q (such as the conclusion that the Queen is here) from any hypothesis P (such as the hypothesis that the Pope is here). When the vertical line for the subproof ends, the subproof is *closed*. In order to complete a proof, you must close all of the subproofs. And you cannot close the subproof and use the repeat rule again on line 4 to derive Q in the main proof. Once you close a subproof, you cannot refer back to individual lines inside it.

You **cannot** use a line from a subproof as a hypothesis for a theorem that is being applied in the main proof. Lines in a subproof stay in the subproof.

In particular, you **cannot** use the repeat theorem to copy a line from a subproof into the main proof.

Of course, it is legitimate to do this:

1	P	hypothesis	The Pope is here.
2	Q	assumption	Suppose that the Queen is here.
3	Q	repeat (line 2)	As mentioned previously, the Queen is here.
4	$Q \Rightarrow Q$	$\Rightarrow$-intro (lines 2 and 3)	If the Queen is here, then the Queen is here.

This should not seem so strange, though. Since $Q \Rightarrow Q$ is a tautology, it follows validly from any hypotheses.

Put in a general form, the **⇒-introduction** rule looks like this:

m		$\mathcal{A}$	assumption (want $\mathcal{B}$)	Suppose that Alberta is big.
		⋮	⋮	⋮
n		$\mathcal{B}$	⟨*whatever reason*⟩	Then BC is big.
	$\mathcal{A} \Rightarrow \mathcal{B}$		⇒-intro (lines m–n)	If Alberta is big, then BC is big.

When we introduce a subproof, it is helpful to make a note of what we want to derive (and add it to the justification). This is so that anyone reading the proof will find it easier to understand why we are doing what we are doing (and also so that we do not forget why we started the subproof if it goes on for five or ten lines). There is no "want" rule. It is a note to ourselves and to the reader; it is not formally part of the proof.

Although it is legal to open a subproof with any assumption you please, there is some strategy involved in picking a useful assumption. Starting a subproof with an arbitrary, wacky assumption would just waste lines of the proof. In order to derive an if-then statement by using ⇒-intro, for instance, you must assume the hypothesis of the statement in a subproof.

Now that we have both the introduction rule and the elimination rule for "implies," we can prove that the following deduction is valid:

$P \Rightarrow Q$ If the Pope is here, then so is the Queen.
$Q \Rightarrow R$ If the Queen is here, then so is the Registrar.
$\therefore$ $P \Rightarrow R$ Therefore, if the Pope is here, then so is the Registrar.

We begin the proof by writing the two hypotheses as assumptions. Since the main logical operator in the conclusion is ⇒, we can expect to use the ⇒-introduction rule. For that, we need a subproof—so we write in the hypothesis of the if-then statement as the assumption of a subproof:

1		$P \Rightarrow Q$	hypothesis	If the Pope is here, then so is the Queen.
2		$Q \Rightarrow R$	hypothesis	If the Queen is here, then so is the Registrar.
3		P	assumption (want R)	Suppose that the Pope is here.

We made P available by assuming it in a subproof, allowing us to apply ⇒-elim to line 1. This gives us Q, which allows us to apply ⇒-elim to line 2. Having derived R, we close the subproof. By assuming P we were able to prove R, so ⇒-elim completes the proof. Here it is written out:

1		$P \Rightarrow Q$	hypothesis	If the Pope is here, then so is the Queen.
2		$Q \Rightarrow R$	hypothesis	If the Queen is here, then so is the Registrar.
3		P	assumption (want R)	Suppose that the Pope is here.
4		Q	⇒-elim (lines 1 and 3)	The Queen is here.
5		R	⇒-elim (lines 2 and 4)	The Registrar is here.
6	$P \Rightarrow R$		⇒-intro (lines 3–5)	If the Pope is here, then so is the Registrar.

EXERCISE 2.3.1. Provide a justification (rule and line numbers) for each line of these proofs.

1)

1	$A \Rightarrow B$	
2	$\neg A \Rightarrow C$	
3	$A \vee \neg A$	
4	A	
5	B	
6	$B \vee C$	
7	$A \Rightarrow (B \vee C)$	
8	$\neg A$	
9	C	
10	$B \vee C$	
11	$\neg A \Rightarrow (B \vee C)$	
12	$B \vee C$	

2)

1	$L \Leftrightarrow \neg O$	
2	$L \vee \neg O$	
3	L	
4	L	
5	$L \Rightarrow L$	
6	$\neg O \Rightarrow L$	
7	L	

3)

1	$F \Rightarrow \big((G \,\&\, H) \vee I\big)$	
2	$\neg I$	
3	$\neg G$	
4	$\neg G \vee \neg H$	
5	$(\neg G \vee \neg H) \,\&\, \neg I$	
6	$\neg\big((G \,\&\, H) \vee I\big)$	
7	$\neg\big((G \,\&\, H) \vee I\big) \Rightarrow \neg F$	
8	$\neg F$	
9	$\neg G \Rightarrow \neg F$	
10	$\neg\neg F \Rightarrow \neg\neg G$	
11	$F \Rightarrow G$	

4)

1	$\neg C \Rightarrow B \vee C$	
2	$C \vee \neg C$	
3	C	
4	$B \vee C$	
5	$C \Rightarrow (B \vee C)$	
6	$B \vee C$	
7	$A \,\&\, \neg B$	
8	$\neg B$	
9	C	
10	$(A \,\&\, \neg B) \Rightarrow C$	

5)

1	$Z \Rightarrow (C \,\&\, \neg N)$	
2	$\neg Z \Rightarrow (N \,\&\, \neg C)$	
3	$Z \vee \neg Z$	
4	Z	
5	$C \,\&\, \neg N$	
6	C	
7	$N \vee C$	
8	$Z \Rightarrow (N \vee C)$	
9	$\neg Z$	
10	$N \,\&\, \neg C$	
11	N	
12	$N \vee C$	
13	$\neg Z \Rightarrow (N \vee C)$	
14	$N \vee C$	

6)

1	$A \Rightarrow E$	
2	$C \Rightarrow E$	
3	$A \vee C$	
4	E	
5	$(A \vee C) \Rightarrow E$	

EXERCISES 2.3.2. Write a two-column proof of each of the following deductions:

1) $P \Rightarrow (Q \vee R)$, $Q \Rightarrow S$, $R \Rightarrow S$, $\quad \therefore\ P \Rightarrow S$

2) $Q \Rightarrow (Q \Rightarrow P)$, $\quad \therefore\ Q \Rightarrow P$

3) $M \vee (N \Rightarrow M)$, $\quad \therefore\ \neg M \Rightarrow \neg N$

4) $R \Rightarrow \Big(R \Rightarrow \big(R \Rightarrow (R \Rightarrow Q)\big)\Big)$, $\quad \therefore R \Rightarrow (Q \vee P)$

5) $A \vee B$, $A \Rightarrow C$, $B \Rightarrow D$, $C \Rightarrow E$, $D \Rightarrow E$, $\quad \therefore E$

6) *Hypotheses:*

The Pope is here if and only if the Queen is here.

The Queen is here if and only if the Registrar is here.

Conclusion: The Pope is here if and only if the Registrar is here.

7)

Hypotheses:

If Jim is sick, he should stay in bed.

If Jim is not sick, he should go outside to play.

Conclusion: Jim should either stay in bed or go outside to play.

8)

Hypotheses:

If the King will sing, then the Queen will sing.

If the King and the Queen will both sing, then the Prince and the Princess will also sing.

If the King and the Queen and the Prince and the Princess will all sing, then the party will be fun.

Conclusion: If the King will sing, then the party will be fun.

9)

Hypotheses:

If the Pope is here, then either the Queen is here or the Registrar is here.

If the Queen is here, then the Spy is here.

If the Registrar is here, then the Spy is here.

Conclusion: If the Pope is here, then the Spy is here.

2.4. Proof by contradiction

How often have I said to you that when you have eliminated the impossible, whatever remains, however improbable, *must be the truth?*

Sherlock Holmes, fictional British detective in *The Sign of the Four*

The usual way to prove that an assertion is false is to show that it cannot be true. We do this by considering what would happen if it were indeed true. That is, we assume, for the sake of argument, that the assertion is true. If, by using logic, we can show that this assumption leads to a contradiction, then we can conclude that the hypothesis was wrong: the assertion we are interested in must be false. This is known as **proof by contradiction**.

Proofs by contradiction are used quite commonly in everyday life. Here is a contrived example:

EXAMPLE 2.4.1. Suppose someone has stolen a bracelet from Ms. Haslot's parlour. Could the butler have done it? Inspector Thinkright might say: "Let us suppose that Jeeves, the butler, took the bracelet. We know he was in the kitchen until 8pm, so the theft must have taken place at a later time. However, Jeeves is highly allergic to Fifi, and that dog was in the parlour from 7:30pm until the theft was discovered at midnight, so Jeeves must have sneezed continuously while he was in the parlour. But the security guard absolutely refutes this: he states unequivocally that there were precisely 2 coughs, and no sneezes at all, coming from the parlour last evening. Thus, Jeeves could *not* have taken the bracelet; we eliminate him as a suspect."

They also arise in mathematics:

EXAMPLE 2.4.2. Here is an argument in English that shows there is no greatest (i.e., largest) natural number:

Suppose there is some greatest natural number. Call it n.

Then $n+1$ is also a natural number. And, obviously, $n+1 > n$.

So there is a natural number (namely, $n+1$) that is greater than n.

This contradicts the fact that n is the greatest natural number.

So our hypothesis leads to consequences that are impossible.

Conclusion: Our hypothesis cannot be true: there is no greatest natural number.

The "Proof by Contradiction" rule allows us to turn an explanation of this type into an official proof. If we assume that a particular assertion is true and show that this leads to something impossible (namely, a contradiction), then we have proven that our assumption is wrong; the assumption must be false, so its negation must be true:

$\mathcal{A}$: Alberta is big.

$\mathcal{B}$: BC is big.

m		$\mathcal{A}$	assumption (for contradiction)	Suppose that Alberta is big.
		⋮	⋮	⋮
n		$\mathcal{B} \,\&\, \neg\mathcal{B}$	⟨*whatever reason*⟩	Then BC is big and BC is not big.
	$\neg\mathcal{A}$	Proof by Contradiction (lines m–n)		Alberta must not be big.

For this rule to apply, the last line of the subproof must be an explicit contradiction of the form $\mathcal{B} \,\&\, \neg\mathcal{B}$: some assertion and its negation. We write "(will lead to a contradiction)" or "(for contradiction)" as a note to ourselves and the reader. It is an explanation of why we started the subproof, and is not formally part of the proof.

OTHER TERMINOLOGY. Proof by Contradiction allows us to put $\neg$ onto an assertion, so some logicians call it **¬-introduction**, but we use the terminology of mathematicians, who always refer to it as "Proof by Contradiction." (And the **¬-elimination** rule is the fact that $\neg\neg A$ is logically equivalent to A, which is one of the rules of negation in Exercise 1.7.5(1).)

EXERCISES 2.4.3. Provide a justification (rule and line numbers) for each line of these proofs.

1)

1		$\neg C \Rightarrow B \vee C$	
2		$A \,\&\, \neg B$	
3		$\neg C$	
4		$B \vee C$	
5		B	
6		$\neg B$	
7		$B \,\&\, \neg B$	
8		$\neg\neg C$	
9		C	
10		$(A \,\&\, \neg B) \Rightarrow C$	

2)

1	$(A \vee B) \Rightarrow \neg B$	
2	B	
3	$A \vee B$	
4	$\neg B$	
5	$B \,\&\, \neg B$	
6	$\neg B$	

3)

1	$P \Rightarrow Q$	
2	$Q \Rightarrow R$	
3	$R \Rightarrow \neg P$	
4	P	
5	Q	
6	R	
7	$\neg P$	
8	$P \,\&\, \neg P$	
9	$\neg P$	

4)

1	$(P \vee \neg Q) \Rightarrow \neg R$	
2	$Q \Rightarrow P$	
3	R	
4	Q	
5	P	
6	$P \vee \neg Q$	
7	$\neg R$	
8	$R \,\&\, \neg R$	
9	$\neg Q$	
10	$P \vee \neg Q$	
11	$\neg R$	
12	$R \,\&\, \neg R$	
13	$\neg R$	

5)

1	$Z \Rightarrow (C \,\&\, \neg N)$	
2	$\neg Z \Rightarrow (N \,\&\, \neg C)$	
3	$\neg(N \vee C)$	
4	N	
5	$N \vee C$	
6	$(N \vee C) \,\&\, \neg(N \vee C)$	
7	$\neg N$	
8	C	
9	$N \vee C$	
10	$(N \vee C) \,\&\, \neg(N \vee C)$	
11	$\neg C$	
12	Z	
13	$C \,\&\, \neg N$	
14	C	
15	$C \,\&\, \neg C$	
16	$\neg Z$	
17	$N \,\&\, \neg C$	
18	N	
19	$N \,\&\, \neg N$	
20	$\neg\neg(N \vee C)$	
21	$N \vee C$	

EXERCISES 2.4.4. Give a two-column proof of each of these deductions.

1) $Q \Rightarrow (Q \,\&\, \neg Q), \quad \therefore\ \neg Q$

2) $J \Rightarrow \neg J, \quad \therefore\ \neg J$

3) $U \Rightarrow X$, $V \Rightarrow \neg X, \quad \therefore\ \neg(U \,\&\, V)$

4) $(M \vee N) \Rightarrow \neg T$, $\neg T \Rightarrow \neg M, \quad \therefore\ \neg M$

5)

Hypotheses:

If Alice is here, then Bob is here.

If Bob is here, then Carol is here.

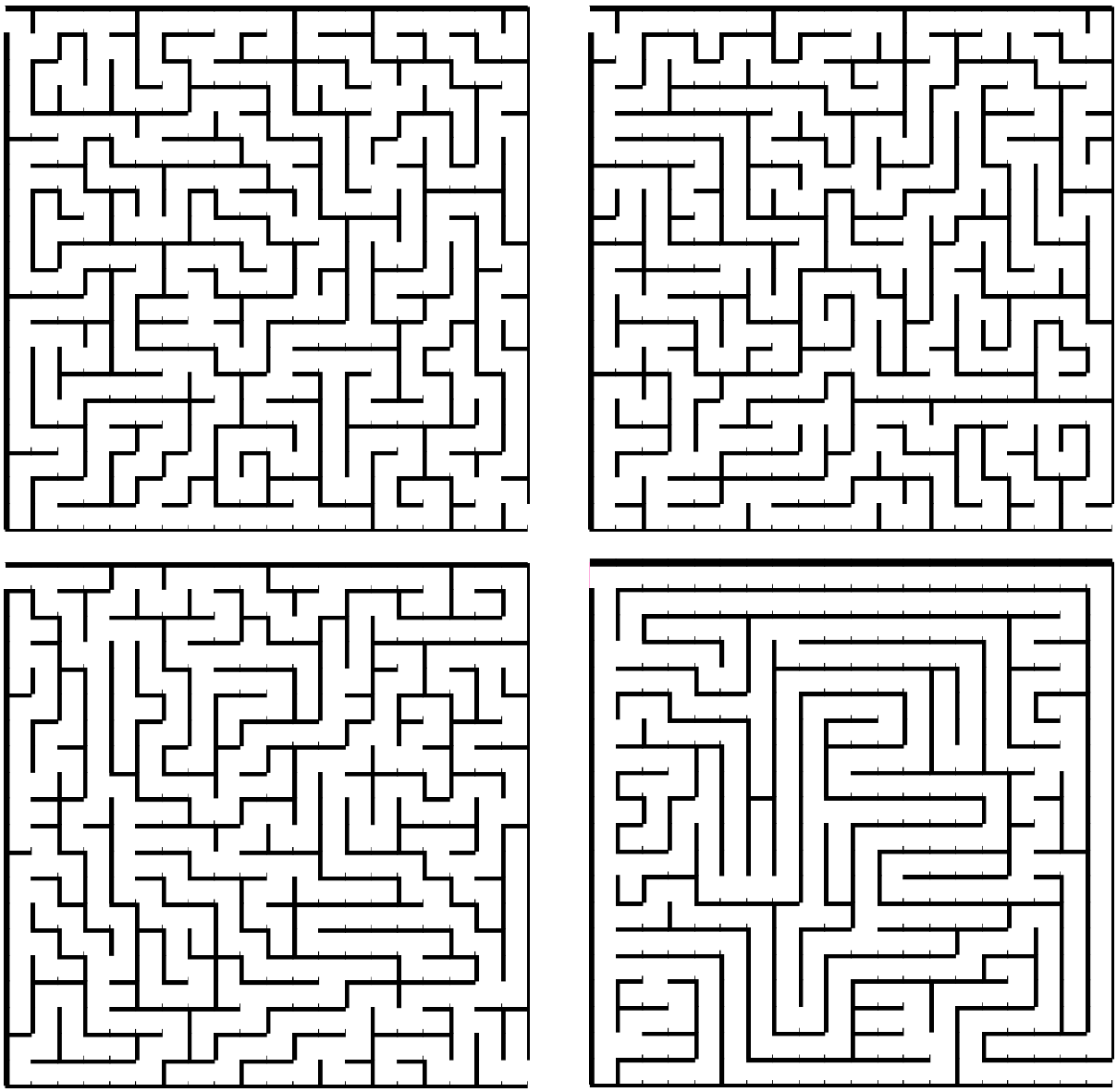
If Carol is here, then Bob is not here.

Conclusion: Alice is not here.

2.5. Proof strategies

Finding a proof is very much the same as solving a puzzle that tells you a starting position and a goal, and gives specific rules about exactly what steps you are allowed to take to complete the task. At each step (of a proof or a puzzle) there are several options for what to do, and you need to choose the right one. There is no simple recipe that will always tell you what to do; being successful requires the same skills that you need to solve a maze.

EXERCISE 2.5.1. Find a path through each maze from the top left corner to the bottom right corner.



There is no substitute for practice, but here are some suggestions and strategies to keep in mind when doing proofs.

Work forwards from what you have. Look at the hypotheses (and any other assertions that you have derived so far). Think about the elimination rules for the main operators in these assertions. These will tell you what your options are. For example:

- If you have $P \,\&\, Q$, you can immediately obtain both P and Q.

- If you have both P and $P \Rightarrow Q$, you can use $\Rightarrow$-elimination to obtain Q.
- If you have $P \vee Q$, you should consider using a proof by cases.

EXERCISE 2.5.2. Give a two-column proof of the deduction

$$A \Rightarrow (B \,\&\, C),\ A,\ \ \therefore C$$

Work backwards from what you want. The ultimate goal is to derive the conclusion. Look at the conclusion and ask what the introduction rule is for its main logical operator. This gives you an idea of where you want to be *just before* the last line of the proof. Then you can treat this line as if it were your goal; we call it a **subgoal** because it represents partial progress toward the true goal. Ask what you could do to derive the subgoal. For example:

- If your conclusion is $\mathcal{A} \,\&\, \mathcal{B}$, then you need to figure out a way to prove $\mathcal{A}$ and a way to prove $\mathcal{B}$.
- If your conclusion is a conditional $\mathcal{A} \Rightarrow \mathcal{B}$, plan to use the $\Rightarrow$-intro rule. This requires starting a subproof in which you assume $\mathcal{A}$. In the subproof, you want to derive $\mathcal{B}$.
- The last of the four mazes in Exercise 2.5.1 is easy if you work backwards from the finish, instead of forward from the start.

EXERCISE 2.5.3. Give a two-column proof of the deduction

$$(P \vee Q) \Rightarrow (R \,\&\, S),\ (R \vee S) \Rightarrow (P \,\&\, Q),\ \therefore P \Rightarrow Q$$

Try breaking the proof down into cases. If it looks like you need an additional hypothesis (A) to prove what you want, try considering two cases: since $A \vee \neg A$ is a tautology ("Law of Excluded Middle"), it suffices to prove that A and $\neg A$ each yield the desired conclusion.

EXERCISE 2.5.4. Give a two-column proof of the deduction

$$P \Rightarrow Q,\ \neg P \Rightarrow R,\ (Q \vee R) \Rightarrow S,\ \therefore S$$

Look for useful subgoals. Working backwards is one way to identify a worthwhile subgoal, but there are others. For example, if you have $\mathcal{A} \Rightarrow \mathcal{B}$, you should think about whether you can obtain $\mathcal{A}$ somehow, so that you can apply $\Rightarrow$-elimination.

EXERCISE 2.5.5. Give a two-column proof of the deduction

$$(R \vee S) \Rightarrow (P \vee Q),\ \neg Q,\ \ \therefore R \Rightarrow P$$

Change what you are looking at. Replacement rules can often make your life easier; if a proof seems impossible, try out some different substitutions. For example:

- the Rules of Negation should become second nature; they can often transform an assertion into a more useful form.
- Remember that every implication is logically equivalent to its contrapositive. The contrapositive may be easier to prove as a conclusion, and it might be more useful as a hypothesis.

EXERCISE 2.5.6. Give a two-column proof of the deduction

$$P,\ \neg(P \,\&\, Q),\ Q \vee R,\ \ \therefore R$$

Do not forget proof by contradiction. If you cannot find a way to show something directly, try assuming its negation, and then look for a contradiction. For example, instead of proving $\mathcal{A} \vee \mathcal{B}$ directly, you can *assume* both $\neg\mathcal{A}$ and $\neg\mathcal{B}$, which is likely to make the work easier.

EXERCISE 2.5.7. Give a two-column proof of the deduction

$$P \Rightarrow Q,\ Q \Rightarrow \neg P,\ \ \therefore \neg P$$

Repeat as necessary. After you have made some progress, by either deriving some new assertions or deciding on a new goal that would represent substantial progress, see what the above strategies suggest in your new situation.

Persist. Try different things. If one approach fails, try something else. When solving a difficult maze, you should expect to have to backtrack several times, and the same is true when doing proofs.

EXERCISES 2.5.8. Give a two-column proof of each of these deductions.

1) $(P \,\&\, \neg Q) \Rightarrow (Q \vee R),\ \ \therefore (P \,\&\, \neg Q) \Rightarrow (R \vee S)$

2) $P \Rightarrow (Q \vee R),\ Q \Rightarrow \neg P,\ R \Rightarrow S,\ \ \therefore P \Rightarrow S$

2.6. What is a proof?

> *I don't know — a proof is a proof. What kind of a proof? It's a proof. A proof is a proof, and when you have a good proof, it's because it's proven.*
>
> Jean Chrétien (b. 1934), Prime Minister of Canada

The goal of a mathematical proof is to provide a completely convincing explanation that a deduction is valid. It needs to be so carefully written that it would hold up in court forever, even against your worst enemy, in any country of the world, and without any further explanation required. Fortunately, the rules of logic are accepted worldwide, so, if applied properly, they create an irrefutable case.

In the previous sections of this chapter, we wrote our proofs in two-column format. We will now start the transition to writing proofs in English prose; our ideas will be expressed in sentences and paragraphs, using correct grammar, combining words with appropriate mathematical notation. A proof written in prose needs to convey the same information as would be found in a two-column proof, so essentially the same rules and strategies will still apply, but writing in ordinary English provides more freedom, and often leads to shorter proofs that are more reader-friendly.

Remark 2.6.1. The big advantage of two-column proofs is that the rules are very clear, so there are no ambiguities that require good judgment to resolve. This makes them easier for beginners who may have difficulty deciding what they are allowed to do. The disadvantage is that being required to record every detail of every step makes the proofs very verbose, so they cannot reasonably be used in the complicated situations that arise in the study of advanced mathematics.

Just as when using the two-column format, our proofs will be a sequence of assertions that lead from the hypotheses to the desired conclusion. Each assertion must have a logical justification based on assertions that were stated earlier in the proof. Any subproof (for $\Rightarrow$-introduction or Proof by Contradiction) will form a paragraph of its own within the proof.

Before the proof begins, we always provide a statement of the theorem that will be proved.

- The statement is preceded by the label "Theorem" (or a suitable substitute).

- The statement of the result begins with a list all of the hypotheses. To make it clear that they are assumptions, not conclusions, this list of assertions is introduced by an appropriate word or phrase such as "Assume...," or "Suppose that ...," or "If ...," or "Let ..."
- The statement of the result ends with a statement of the desired conclusion, introduced by an appropriate word or phrase such as "Then ...," or "Therefore, ..."

Following the statement of the result, we begin our proof in a new paragraph.

- The proof is labelled with the single word: "Proof."
- We then proceed to give a well-organized series of assertions that logically lead from our hypotheses to the desired conclusion.
- A small square is drawn at the right margin at the end of the proof to signify that the proof is complete.

For example, here is how the chapter's first deduction could be treated:

THEOREM. *Assume:*

a) *if the Pope is here, then the Queen and the Registrar are both here, and*

b) *the Pope is here.*

Then the Registrar is here.

PROOF. From Assumption (b), we know that the Pope is here. Therefore, Assumption (a) tells us that the Queen and the Registrar are both here. In particular, the Registrar is here. □

Here is another example:

EXAMPLE 2.6.2.

Hypotheses:

1. If the Pope is here, and the Queen is not here, then the Registrar is here.

Conclusion: If the Pope is here, then either the Queen or the Registrar is also here.

PROOF BY CONTRADICTION. Suppose the conclusion is false. (This will lead to a contradiction.) This means that the Pope is here, but neither the Queen nor the Registrar is here. In particular, the Pope is here and the Queen is not here, so Hypothesis (1) tells us that the Registrar is here. However, since neither the Queen nor the Registrar is here, we also know that the Registrar is not here. Therefore, the Registrar is both here and not here. This is a contradiction. □

ALTERNATE PROOF. Suppose the Pope is here. (We wish to show that either the Queen or the Registrar is also here.) From the Law of Excluded Middle, we know that the Queen is either here or not here, and we consider these two possibilities as separate cases.

Case 1. Assume the Queen is here. Then it is true that either the Queen or the Registrar is here, as desired.

Case 2. Assume the Queen is not here. Then the Pope is here, and the Queen is not here. From Hypothesis (1), we conclude that the Registrar is here. Therefore, either the Queen or the Registrar is here, as desired. □

Remark 2.6.3. Note that some of the rules of the two-column format are relaxed for proofs written in prose:

1) We will no longer list all of the hypotheses at the start of our proof. Instead, we refer to the list that is in the statement of the theorem.
2) We will no longer make a practice of numbering all of the assertions in our proofs. However, if there is a particular assertion that will be used repeatedly, we may label it with a number for easy reference.
3) We will usually not cite the basic rules of Propositional Logic by name every time they are used. However, we should be able to justify any assertion with a rule, if called upon to do so.

EXERCISE 2.6.4. Translate both proofs of Example 2.6.2 into two-column format (using our usual symbolization key).

EXERCISES 2.6.5. Write a proof of each of these theorems in English prose.

1) *Hypotheses:*
   1. If the Pope is here, then the Queen is here.
   2. If the Queen is here, then the Registrar is here.

   Conclusion: If the Pope is here, then the Registrar is here.

2) **THEOREM.** *Assume:*
   (a) *If the Pope is here, then the Registrar is here.*
   (b) *If the Queen is here, then the Spy is here.*
   (c) *The Pope and the Queen are both here.*

   Then the Registrar and the Spy are both here.

3) **THEOREM.** *Assume:*
   (a) *If Adam is here, then Betty is here.*
   (b) *If Betty is not here, then Charlie is here.*
   (c) *Either Adam is here, or Charlie is not here.*

   Then Betty is here.

4) **THEOREM.** *Assume:*
   (a) *If Jack and Jill went up the hill, then something will go wrong.*
   (b) *If Jack went up the hill, then Jill went up the hill.*
   (c) *Nothing will go wrong.*

   Then Jack did not go up the hill.

2.7. Counterexamples

Not all deductions are valid. To show that a particular deduction is *not* valid, you need to show that it is possible for its conclusion to be false at the same time that all of its hypotheses are true. To do this, you should find an assignment to the variables that makes all of the hypotheses true, but makes the conclusion false.

EXAMPLE 2.7.1. Show that the deduction

$$A \vee B, \quad A \Rightarrow B, \quad \therefore \ A$$

is not valid.

Scratchwork. To make the conclusion false, we let A be false. Then, to make the first hypothesis true, we must let B be true. Fortunately, this also makes the second hypothesis true.

SOLUTION. Let A be false, and let B be true. Then

$$A \vee B \quad = \quad \mathsf{F} \vee \mathsf{T} \quad = \quad \mathsf{T}$$

and

$$A \Rightarrow B \quad = \quad \mathsf{F} \Rightarrow \mathsf{T} \quad = \quad \mathsf{T},$$

so both hypotheses of the deduction are true. However, the conclusion of the deduction (namely, A) is false.

Since we have a situation in which both hypotheses of the deduction are true, but the conclusion of the deduction is false, the deduction is not valid. □

DEFINITION 2.7.2. Any situation in which all of the hypotheses of a deduction are true, but the conclusion is false, is called a **counterexample** to the deduction.

To show that a deduction is *not* valid, find a counterexample.

EXERCISE 2.7.3. Show that each of these deductions is invalid, by finding a counterexample.

1) $A \vee B, \therefore A \Rightarrow B$
2) $P \vee Q, \ \therefore P \,\&\, Q$
3) $A \Rightarrow (B \,\&\, C), \ \neg A \Rightarrow (B \vee C), \ \therefore C$
4) $P \Rightarrow Q, \neg P \Rightarrow R, \therefore Q \,\&\, (P \vee R)$

SUMMARY:

- A "two-column proof" is a tool that we use to learn techniques for writing proofs.
  - The left-hand column contains a sequence of assertions.
  - The right-hand column contains a justification for each assertion.
  - Each row of the proof is numbered (in the left margin) for easy reference.
  - A dark horizontal line is drawn to indicate the end of the hypotheses.
  - A dark horizontal line is drawn along the left edge of the proof, and of each subproof.
- In addition to the basic theorems of Propositional Logic, we have two rules that use subproofs:
  - $\Rightarrow$-introduction
  - proof by contradiction
- Proofs often use the Law of Excluded Middle, the Rules of Negation, and contrapositives.
- Assertions that are in a subproof cannot be used as justification for lines that are not in that same subproof.
- Writing proofs takes practice, but there are some strategies that can help.
- Proofs can also be written in English prose, using sentences and paragraphs.
- To show that a deduction is *not* valid, find a counterexample.

Part II

Sets and First-Order Logic

Chapter 3

Sets

In mathematics these days, essentially everything is a set. Some knowledge of set theory is a necessary part of the background everyone needs for further study of mathematics.

Herbert B. Enderton (1936–2010), American mathematician
Elements of Set Theory

3.1. Propositional Logic is not enough

Consider the following deduction:

> Merlin is a wizard. All wizards wear funny hats.
> Therefore, Merlin wears a funny hat.

To symbolize it in Propositional Logic, we define a symbolization key:

W: Merlin is a wizard.

A: All wizards are wearing funny hats.

H: Merlin is wearing a funny hat.

Now we symbolize the deduction:

Hypotheses:

W

A

Conclusion: H

This is *not* valid in Propositional Logic. (If W and A are true, but H is false, then it is obvious that both hypotheses are true, but the conclusion is false.) There is something very wrong here, because the deduction that was written in English is clearly valid.

The problem is that symbolizing this deduction in Propositional Logic leaves out some of the important structure: The assertion "All wizards are wearing funny hats" is about both wizards and hat-wearing, but Propositional Logic is not able to capture this information: it loses the connection between Merlin's being a wizard and Merlin's wearing a hat. However, the problem is not that we have made a mistake while symbolizing the deduction; it is the best symbolization we can give for this deduction *in Propositional Logic*.

In order to symbolize this deduction properly, we need to use a more powerful logical language. This language is called **First-Order Logic**, and its assertions are built from "predicates" and "quantifiers."

A predicate is an expression like "_____ is wearing a funny hat." This is not an assertion on its own, because it is neither true nor false until we fill in the blank, to specify who it is that we claim is wearing a funny hat.

The details of this will be explained in Section 3.2D, but here is the basic idea: In First-Order Logic, we will represent predicates with capital letters. For instance, we could let H stand for "_____ is wearing a funny hat." However, we will use variables instead of blanks; so "x is wearing a funny hat" is a predicate, and we could represent it as $H(x)$.

The words "all" and "some" are *quantifiers*, and we will have symbols that represent them. For instance, "$\exists$" will mean "There exists some_____, such that." Thus, to say that someone is wearing a funny hat, we can write $\exists x, H(x)$; that is: There exists some x, such that x is wearing a funny hat. Quantifiers will be dealt with in Chapter 4, when First-Order Logic is fully explained.

With predicates and quantifiers, we will be talking about many people (or other things) all at once, instead of one at a time. For example, we may wish to talk about "the people who are wearing hats," or "the mammals that lay eggs." These are examples of *sets*.

3.2. Sets, subsets, and predicates

3.2A. Sets and their elements. In mathematics, a **set** is a collection of objects. The objects in the collection are called "**elements**" (or "**members**") of the set. If someone has a particular set in mind, they may wish to tell other people which set it is. One good way to do this is to list its elements. The list needs to be surrounded with curly braces ("{ }") to indicate that it represents a set, rather than some other type of object.

OTHER TERMINOLOGY. In mathematics, the term **collection** is a synonym for "set."

EXAMPLE 3.2.1.

1) $\{1, 2, 3, 4, 5\}$ is the set of natural numbers from 1 to 5.

2) $\{1, 2, 3, \ldots, 100\}$ is the set of natural numbers from 1 to 100.

3) $\{\clubsuit, \diamondsuit, \heartsuit, \spadesuit\}$ is the set of suits in a standard deck of cards.

4) The set of provinces in Canada is

$$\left\{\begin{array}{c}\text{British Columbia, Alberta, Saskatchewan, Manitoba,}\\ \text{Ontario, Quebec, Newfoundland and Labrador,}\\ \text{New Brunswick, Prince Edward Island, Nova Scotia}\end{array}\right\}.$$

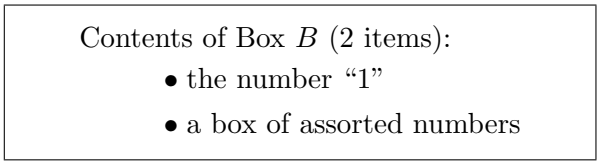
Remark 3.2.2. In everyday life, when you have a bunch of things that you want to keep together, you might look for a box to put them in. (The box itself probably has no value — you are interested only in the stuff that is in the box.) In mathematics, you should put the things into a set, not a box. If you think of a set as being a box of stuff, then the elements of the set are the things you see when you open the box.

EXAMPLE 3.2.3.

1) If $A = \{1, 2, 3\}$, then the elements of A are the numbers 1, 2, and 3.

2) If $B = \big\{1, \{2, 3\}\big\}$, then the elements of B are the number 1 and the set $\{2, 3\}$. It is important to note that the numbers 2 and 3 are *not* elements of B.

   (a) To understand this, it may help to consider the analogy with boxes: if we open the box B, we will see the number 1 and a box, but we will not see the number 2 or the number 3.



We would need to open up the box that is inside of B in order to see those extra numbers. So 2 and 3 are not elements of the set B — they are elements of a set that is an element of B.

(b) As another illustration of this same phenomenon, suppose we make a list of the teams in a chess tournament. The list might be:

(i) U of Lethbridge, (ii) U of Alberta, (iii) U of Calgary.

And maybe the members of the Lethbridge team are Alice, Bob, and Cindy. Then Alice is *not* on the list of teams; she is a *member* of one of the teams on the list.

NOTATION 3.2.4. We use

- "$\in$" as an abbreviation for "is an element of," and
- "$\notin$" as an abbreviation for "is *not* an element of."

For example, if $A = \{1, 2, 3, 4, 5\}$, then we have $3 \in A$ and $7 \notin A$, because 3 is an element of A, but 7 is not an element of A.

DEFINITION 3.2.5. The set with no elements can be denoted { }. (It is like an empty box.) It is called the **empty set**, and it comes up so often that is named by a special symbol:

$\varnothing$ denotes the empty set.

Remark 3.2.6. Because the empty set has no elements,

for all x, we have $x \notin \varnothing$.

EXERCISE 3.2.7. Fill in each blank with $\in$ or $\notin$.

1) t_____$\{\mathsf{t}, \mathsf{i}, \mathsf{m}, \mathsf{e}\}$	2) i_____$\{\mathsf{t}, \mathsf{i}, \mathsf{m}, \mathsf{e}\}$	3) m_____$\{\mathsf{t}, \mathsf{i}, \mathsf{m}, \mathsf{e}\}$
4) $\{\mathsf{t}\}$_____$\{\mathsf{t}, \mathsf{i}, \mathsf{m}, \mathsf{e}\}$	5) $\{\mathsf{i}\}$_____$\{\mathsf{t}, \mathsf{i}, \mathsf{m}, \mathsf{e}\}$	6) $\{\mathsf{m}\}$_____$\{\mathsf{t}, \mathsf{i}, \mathsf{m}, \mathsf{e}\}$
7) $\{\mathsf{t}, \mathsf{i}\}$_____$\{\mathsf{t}, \mathsf{i}, \mathsf{m}, \mathsf{e}\}$	8) $\{\mathsf{m}, \mathsf{e}\}$_____$\{\mathsf{t}, \mathsf{i}, \mathsf{m}, \mathsf{e}\}$	9) t_____$\{\mathsf{t}, \{\mathsf{i}\}, \{\mathsf{m}, \mathsf{e}\}\}$
10) i_____$\{\mathsf{t}, \{\mathsf{i}\}, \{\mathsf{m}, \mathsf{e}\}\}$	11) m_____$\{\mathsf{t}, \{\mathsf{i}\}, \{\mathsf{m}, \mathsf{e}\}\}$	12) $\{\mathsf{t}\}$_____$\{\mathsf{t}, \{\mathsf{i}\}, \{\mathsf{m}, \mathsf{e}\}\}$
13) $\{\mathsf{i}\}$_____$\{\mathsf{t}, \{\mathsf{i}\}, \{\mathsf{m}, \mathsf{e}\}\}$	14) $\{\mathsf{m}\}$_____$\{\mathsf{t}, \{\mathsf{i}\}, \{\mathsf{m}, \mathsf{e}\}\}$	15) $\{\mathsf{t}, \mathsf{i}\}$_____$\{\mathsf{t}, \{\mathsf{i}\}, \{\mathsf{m}, \mathsf{e}\}\}$
16) $\{\mathsf{m}, \mathsf{e}\}$_____$\{\mathsf{t}, \{\mathsf{i}\}, \{\mathsf{m}, \mathsf{e}\}\}$	17) $\varnothing$_____$\varnothing$	18) $\varnothing$_____$\{\varnothing\}$
19) $\{\varnothing\}$_____$\varnothing$	20) $\{\varnothing\}$_____$\{\varnothing\}$	21) $\{\varnothing\}$_____$\{\{\varnothing\}\}$

We said that a set is a collection of objects, but this needs a bit of elaboration:

1) *A set is **determined by its elements**.* This means that there cannot be two different sets that have exactly the same elements. (Or, in other words, if two sets have the same elements, then the two sets are equal.) For example, suppose:

(a) H is the set of students who had a perfect score on last week's history quiz,

(b) M is the set of students who had a perfect score on last week's math quiz.

(c) Alice and Bob are the only two students who had a perfect score on last week's history quiz, and

(d) Alice and Bob are also the only two students who had a perfect score on last week's math quiz.

Then H and M have exactly the same elements, so H and M are just different names for the same set: namely, both represent $\{\mathsf{Alice}, \mathsf{Bob}\}$. So the sets are equal: we have $H = M$.

Suppose A and B are two sets.
We have $A = B$ if and only if
for every x, $\big((x \in A) \Leftrightarrow (x \in B)\big)$.

2) *A set is an **unordered** collection.* This means that listing the elements of a set in a different order does not give a different set. For example, $\{1, 2, 3\}$ and $\{1, 3, 2\}$ are the same set. We write $\{1, 2, 3\} = \{1, 3, 2\}$. Both of them are the set whose elements are 1, 2, and 3.

3) *A set is a collection **without repetition**.* This means that repeating something in the list of elements does not change the set. For example, $\{1, 2, 2, 3, 3\}$ is the same set as $\{1, 2, 3\}$. We write $\{1, 2, 2, 3, 3\} = \{1, 2, 3\}$.

EXERCISES 3.2.8. Fill in the blank with $=$ or $\neq$.

1) $\{\mathsf{t}, \mathsf{i}, \mathsf{m}, \mathsf{e}\}$_____$\{\mathsf{t}, \mathsf{m}, \mathsf{i}, \mathsf{e}\}$
2) $\{\mathsf{t}, \mathsf{i}, \mathsf{m}\}$_____$\{\mathsf{t}, \mathsf{m}, \mathsf{i}, \mathsf{e}\}$
3) $\{\mathsf{t}, \mathsf{i}, \mathsf{m}\}$_____$\{\mathsf{t}, \mathsf{m}, \mathsf{i}, \mathsf{m}\}$
4) $\{\mathsf{t}, \mathsf{i}, \mathsf{m}\}$_____$\{\mathsf{t}, \mathsf{m}, \mathsf{i}\}$
5) $\{\mathsf{t}, \mathsf{i}, \mathsf{m}\}$_____$\{\mathsf{t}, \mathsf{i}, \mathsf{m}, \mathsf{i}\}$
6) $\{\mathsf{t}, \mathsf{i}, \mathsf{m}\}$_____$\{\mathsf{t}, \mathsf{t}, \mathsf{t}, \mathsf{i}, \mathsf{i}, \mathsf{m}\}$
7) $\{\mathsf{t}, \mathsf{t}\}$_____$\{\mathsf{t}\}$
8) $\{\mathsf{t}, \mathsf{t}\}$_____$\{\mathsf{i}, \mathsf{i}\}$
9) $\{\mathsf{t}, \mathsf{t}\}$_____$\{\mathsf{t}, \mathsf{i}\}$

EXERCISES 3.2.9. Provide a 2-column proof of each deduction.

1) $(a \in A) \Rightarrow (a \notin B)$, $(b \in B) \Rightarrow (a \in B)$, $\therefore (b \in B) \Rightarrow (a \notin A)$
2) $(p \in X)$ & $(q \in X)$, $(p \in X) \Rightarrow \big((q \notin X) \vee (Y = \varnothing)\big)$, $\therefore Y = \varnothing$.

EXERCISES 3.2.10. Write your proofs in English.

1) Assume:
   (a) If $p \in H$ and $q \in H$, then $r \in H$.
   (b) $q \in H$.
   Show that if $p \in H$, then $r \in H$.

2) Assume:
   (a) If $X \neq \varnothing$, then $a \in Y$.
   (b) If $X = \varnothing$, then $b \in Y$.
   (c) If either $a \in Y$ or $b \in Y$, then $Y \neq \varnothing$.
   Show $Y \neq \varnothing$.

NOTATION 3.2.11. It is traditional to use:

- capital letters (such as A, B, C, X, Y, Z) to represent sets, and
- lower-case letters (such as a, b, c, x, y, z) to represent numbers and other objects that are individual elements (or "atoms"), rather than sets.

Furthermore, it is a good idea to maintain a correspondence between the upper-case letters and lower-case letters: when feasible, use a to represent an element of A and b to represent an element of B, for example.

NOTATION 3.2.12. A few particularly important sets of numbers have been given names that every mathematician needs to know:

- $\mathbb{N} = \{0, 1, 2, ...\}$ is the set of *natural numbers*.
  (*Warning:* Some textbooks do not consider 0 to be a natural number.)
- $\mathbb{N}^+ = \{1, 2, ...\}$ is the set of *positive natural numbers*.
- $\mathbb{Z} = \{\ldots, -3, -2, -1, 0, 1, 2, 3, \ldots\}$ is the set of *integers*. A number is an integer if and only if it is either a natural number or the negative of a natural number.

- $\mathbb{Q} = \left\{ \dfrac{p}{q} \;\middle|\; \begin{matrix} p, q \in \mathbb{Z} \\ q \neq 0 \end{matrix} \right\}$ is the set of *rational numbers*. (This notation means that a number x is an element of $\mathbb{Q}$ if and only if there exist integers p and q, with $q \neq 0$, such that $x = p/q$ (cf. §3.2E).) For example, $1/2$, $7/5$, and $-32/9$ are elements of $\mathbb{Q}$.
- $\mathbb{R}$ is the set of all *real numbers*. (That is, the set of all numbers that are either positive or negative or 0. Unless you have learned about "complex numbers" or "imaginary numbers," it is probably the case that all the numbers you know are real numbers.) For example, $\sqrt[3]{n}$ is a real number whenever $n \in \mathbb{Z}$; and $\sqrt{n}$ is a real number whenever $n \in \mathbb{N}$. ("You can't take the square root of a negative number.")

Remark 3.2.13. Sets are the most fundamental objects in mathematics. Indeed, modern mathematicians consider every object everywhere to be a set, but we will not be quite this extreme. Namely, in addition to sets, we will consider two additional types of objects: numbers and ordered pairs. (It is assumed that you already have a lot of experience with numbers, and know how to deal with them. You have probably also seen ordered pairs (x, y), but their basic properties will be reviewed in Notation 6.1.1.) Functions are another very important class of mathematical objects, but, as will be seen in Section 6.2, we can think of them as being a particular type of set.

3.2B. Cardinality.

NOTATION 3.2.14. We use $\#A$ to denote the number of elements in the set A. (Thus, for example, $\#\{\mathsf{a}, \mathsf{e}, \mathsf{i}, \mathsf{o}, \mathsf{u}\} = 5$.) Mathematicians call $\#A$ the **cardinality** of A. This seemingly simple notion actually has some complicated implications, and will be discussed in more detail in Chapter 9.

OTHER NOTATION. Many mathematicians use the notation $|A|$ instead of $\#A$.

Remark 3.2.15. You probably already know that some sets are finite and some (such as $\mathbb{N}$) are infinite. We will discuss this in more detail in Chapter 9. For now, we remind you that a set A is **finite** iff the elements of A can be counted (and the answer is some number n); that is, if $\#A = n$, for some $n \in \mathbb{N}$.

EXERCISES 3.2.16. What is the cardinality of each set? (*You do not need to show your work.*)

1) $\#\{\mathsf{a}, \mathsf{b}, \mathsf{c}, \mathsf{d}\} =$
2) $\#\{\mathsf{a}, \mathsf{a}, \mathsf{b}, \mathsf{c}, \mathsf{c}, \mathsf{d}\} =$
3) $\#\big\{\mathsf{a}, \{\mathsf{b}, \mathsf{c}\}\big\} =$
4) $\#\big\{\mathsf{a}, \mathsf{a}, \{\mathsf{b}, \mathsf{c}\}, \{\mathsf{b}, \mathsf{c}, \mathsf{d}\}\big\} =$
5) $\#\varnothing =$
6) $\#\{\varnothing\} =$

3.2C. Subsets. Geometry students are taught that every square is a rectangle. Translating this into the terms of set theory, we can say that if

- S is the set of all squares, and
- R is the set of all rectangles,

then every element of the set S is also an element of R. For short, we say that S is a *subset* of R, and we may write $S \subset R$.

DEFINITION 3.2.17. Suppose A and B are two sets. We say that B is a **subset** of A iff every element of B is an element of A.

When B is a subset of A:

- In symbols, we write $B \subset A$.
- We may say that B is **contained in** A or that A **contains** B.
- We may also write $A \supset B$ (and call A a **superset** of B).

EXAMPLE 3.2.18.

1) $\{1, 2, 3\}$ is a subset of $\{1, 2, 3, 4\}$, because the elements of $\{1, 2, 3\}$ are 1, 2, and 3, and every one of those numbers is an element of $\{1, 2, 3, 4\}$.
2) $\{1, 3, 5\}$ is *not* a subset of $\{1, 2, 3, 4\}$, because there is an element of $\{1, 3, 5\}$ (namely, 5) that is not an element of $\{1, 2, 3, 4\}$.
3) We have $\mathbb{N}^+ \subset \mathbb{N} \subset \mathbb{Z} \subset \mathbb{Q} \subset \mathbb{R}$.

Remark 3.2.19.

1) We write $B \not\subset A$ to denote that B is *not* a subset of A.
2) We have $B \not\subset A$ iff there is at least one element of B that is *not* an element of A.

Remark 3.2.20.

1) In the language of everyday life, suppose someone gives you a box A that has some stuff in it. You are allowed to take some of the things from the box and put them into a new box B. But you are not allowed to put anything into B if it was not in box A. Then B will be a subset of A.
2) If you decide to take all of the things that were in box A, then box B will end up being exactly the same as A; that is $B = A$. This illustrates the fact that every set is a subset of itself.

   For every set A, we have $A \subset A$.

3) If you decide not to take anything at all from box A, then box B will be empty. This illustrates the important fact that the empty set is a subset of every set.

   For every set A, we have $\varnothing \subset A$.

DEFINITION 3.2.21. Suppose A and B are sets. We say B is a **proper subset** of A iff $B \subset A$ *and* $B \neq A$.

OTHER NOTATION. Many mathematicians use a slightly different notation: they define $A \subset B$ to mean that A is a *proper* subset of B. Then, to say that A is a subset of B, they write $A \subseteq B$.

EXERCISE 3.2.22. Fill each blank with $\subset$ or $\not\subset$, as appropriate.

1) $\{\mathsf{s}\}$_____$\{\mathsf{h}, \mathsf{o}, \mathsf{r}, \mathsf{n}, \mathsf{s}\}$
2) $\{\mathsf{o}, \mathsf{r}\}$_____$\{\mathsf{h}, \mathsf{o}, \mathsf{r}, \mathsf{n}, \mathsf{s}\}$
3) $\{\mathsf{n}, \mathsf{o}, \mathsf{r}\}$_____$\{\mathsf{h}, \mathsf{o}, \mathsf{r}, \mathsf{n}, \mathsf{s}\}$
4) $\{\mathsf{p}, \mathsf{r}, \mathsf{o}, \mathsf{n}, \mathsf{g}\}$_____$\{\mathsf{h}, \mathsf{o}, \mathsf{r}, \mathsf{n}, \mathsf{s}\}$
5) $\{\mathsf{s}, \mathsf{h}, \mathsf{o}, \mathsf{r}, \mathsf{n}\}$_____$\{\mathsf{h}, \mathsf{o}, \mathsf{r}, \mathsf{n}, \mathsf{s}\}$
6) $\varnothing$_____$\{\mathsf{h}, \mathsf{o}, \mathsf{r}, \mathsf{n}, \mathsf{s}\}$
7) $\{\varnothing\}$_____$\{\mathsf{h}, \mathsf{o}, \mathsf{r}, \mathsf{n}, \mathsf{s}\}$
8) $\{\mathsf{h}, \mathsf{o}, \mathsf{r}, \mathsf{n}, \mathsf{s}\}$_____$\varnothing$

It is intuitively clear that a subset of a set cannot have more elements than the original set. That is:

If $B \subset A$, then $\#B \leq \#A$.

We will prove this fact in Exercise 9.1.15(2).

In Example 4.5.1, we will prove that two sets are equal if and only if they are subsets of each other. This is a basic principle that will be very important in later chapters when we are doing proofs with sets:

To show two sets A and B are equal, prove $A \subset B$ and $B \subset A$.

EXERCISES 3.2.23. Write a two-column proof to justify each assertion.

1) $X \subset Y \Rightarrow X \subset Z$, $X \subset Z \Rightarrow x \in Z$, $x \notin Z$, $\therefore X \not\subset Y$.

2) $(x \in Y) \Rightarrow (X \subset Y)$, $(x \in Y) \vee (Y \subset X)$, $\therefore (X \subset Y) \vee (Y \subset X)$.

3.2D. Predicates. The simplest predicates are things you can say about a single object; they are properties of individuals. For example, "x is a dog" and "x is a *Harry Potter* fan" are both predicates. In First-Order Logic, we symbolize predicates with capital letters A through Z (with or without subscripts). Thus, our symbolization key might include:

$D(x)$: x is a dog. $H(x)$: x is a *Harry Potter* fan.

Predicates like these are called **one-place** or **unary**, because there is only one variable. Assigning a value to this variable yields an assertion. For example, letting $x =$ "Lassie" in the first predicate yields the assertion "Lassie is a dog." Note that in translating English assertions, the variable will not always come at the beginning of the assertion: "the Louvre owns at least one watercolour painted by x" is also a predicate.

Other predicates are about the *relation* between two things. For instance, in algebra, we have the relations "x is equal to y," symbolized as $x = y$, and "x is greater than y," symbolized as $x > y$. These are **two-place** or **binary** predicates, because values need to be assigned to two variables in order to make an assertion. Our symbolization key might include:

$x\ F\ y$: x is a friend of y.

$x\ L\ y$: x is to the left of y.

$x\ M\ y$: x owes money to y.

In general, we can have predicates with as many variables as we need. Predicates with n variables, for some number n, are called ***n*-place** or ***n*-ary**. However, in practice, predicates almost always have only one or two variables.

A symbolization key may also include **constants** (that is, the names of specific objects). For example, we might have a symbolization key that looks like this:

$H(x)$: x is happy. $S(x)$: x is singing. $x\ T\ y$: x is teaching y.

g: Greg m: Mary v: Vikki

This allows us to symbolize assertions that use any combination of these predicates and terms. For example:

assertion	*symbolization*
Greg is happy.	$H(g)$
If Mary is singing, then Vikki is happy.	$S(m) \Rightarrow H(v)$
Greg and Mary are singing.	$S(g)\ \&\ S(m)$
If either Greg or Mary is singing, then Vikki is not happy.	$\big(S(g) \vee S(m)\big) \Rightarrow \neg H(v)$
Mary is teaching Greg.	$m\ T\ g$
Mary is not teaching Vikki.	$\neg(m\ T\ v)$
Vikki is teaching either Mary or Greg.	$(v\ T\ m) \vee (v\ T\ g)$
If Mary is teaching Greg, then Mary and Vikki are happy.	$(m\ T\ g) \Rightarrow \big(H(m)\ \&\ H(v)\big)$
Either Vikki is not singing, or she is not teaching Mary.	$\big(\neg S(v)\big) \vee \neg(v\ T\ m)$
If Mary is singing, then Greg is not teaching Vikki.	$S(m) \Rightarrow \neg(g\ T\ v)$

WARNING. Whenever you have a predicate with two (or more) variables, it is important to be careful about the order in which the variables occur. (For example, saying $x < y$ is certainly not the same as saying $y < x$.) Some special choices of predicates are "symmetric," which means

that if the predicate is true with variables in one order, then it is true for the same variables in a different order, but this should *never* be assumed. The order of the variables should always represent exactly what we know.

EXERCISES 3.2.24. Using the symbolization key given below, give an English version of each assertion.

$x\ O\ y$: x is older than y. r: Roger

$x\ F\ y$: x is a friend of y. s: Sam

S: the set of all students. t: Tess

1) $r\ O\ s$
2) $t\ O\ s$
3) $(r\ F\ t) \Rightarrow (t \in S)$
4) $\big((s \in S)\ \&\ (r \in S)\big) \Rightarrow (s\ F\ r)$
5) $(t \in S) \vee (r\ O\ t)$
6) $(r\ F\ s) \Leftrightarrow (t \notin S)$

EXERCISES 3.2.25. Using the same symbolization key, write these English assertions using predicates and logical connectives.

1) Tess is older than Roger.
2) Roger is a friend of Sam.
3) If Tess is a student then Tess is a friend of Sam.
4) Either Sam is a student, or Roger is not a student.
5) If Roger is a friend of Sam, then Sam is a student.
6) Sam is older than Roger if and only if Roger is a student.
7) If Sam and Roger both are students, then Sam is not a friend of Roger.

EXERCISES 3.2.26. Using the same symbolization key, write a two-column proof to justify each of the following deductions.

1) $(r \in S) \Rightarrow \big((r\ O\ s) \vee (r \notin S)\big), \quad \therefore\ \big((t \in S)\ \&\ \neg(r\ O\ s)\big) \Rightarrow (r \notin S)$
2) If either Roger is a student, or Tess is *not* a student, then Sam is older than Tess.
   If Tess is a student, then Roger is also a student.
   $\therefore$ Sam is older than Tess.

3.2E. Using predicates to specify subsets. Subsets arise in everyday life whenever you want only *part* of something. For example, suppose you are in a kitchen with a lot of plates. If you are washing dishes, then you do not want to be given *all* of the plates, but only the ones that are dirty. In mathematical terms, you do not want the set of all plates, but only want a subset, those that are dirty. That is, if P represents the set of all plates, and D represents the set of all dirty plates, then $D \subset P$.

This type of situation is handled by the following useful notation:

> Suppose A is a set and $P(x)$ is a predicate.
> Then $\{\, a \in A \mid P(a) \,\}$ denotes
> the set of all elements a of A, such that $P(a)$ is true.
> It is a subset of A.

In the example above, you are interested in the subset

$$\{\, p \in P \mid p \text{ is dirty} \,\},$$

because this is the set of plates that are dirty. The notation tells us to look through all of the plates in P, and check each one to see whether it is dirty. If it is, we put it in the subset. If it is not dirty, then we do not put it in the subset.

EXAMPLE 3.2.27.

1) Suppose $B = \{1, 2, 3, \ldots, 10\}$. Then:
   (a) $\{\, b \in B \mid b \text{ is odd} \,\} = \{1, 3, 5, 7, 9\}$.
   (b) $\{\, b \in B \mid b \text{ is even} \,\} = \{2, 4, 6, 8, 10\}$.
   (c) $\{\, b \in B \mid b \text{ is prime} \,\} = \{2, 3, 5, 7\}$.
   (d) $\{\, b \in B \mid b^2 - 1 \text{ is divisible by } 3 \,\} = \{1, 2, 4, 5, 7, 8, 10\}$.
   (e) $\{\, b \in B \mid (b-5)^2 > 4 \,\} = \{1, 2, 8, 9, 10\}$.
   (f) $\{\, b \in B \mid 3 \leq b \leq 8 \text{ and } b \text{ is even} \,\} = \{4, 6, 8\}$.
2) For any $n \in \mathbb{N}$, we have $\{\, i \in \mathbb{N} \mid 1 \leq i \leq n \,\} = \{1, 2, 3, \ldots, n\}$.

EXERCISE 3.2.28. Let $A = \{1, 2, 3, 4, 5\}$ and $B = \{1, 3, 5, 7, 9\}$. Specify each set by listing its elements.

1) $\{\, a \in A \mid a \text{ is even} \,\} =$
2) $\{\, b \in B \mid b \text{ is even} \,\} =$
3) $\{\, a \in A \mid a \text{ is odd} \,\} =$
4) $\{\, b \in B \mid b \text{ is odd} \,\} =$
5) $\{\, a \in A \mid a < 4 \,\} =$
6) $\{\, b \in B \mid b < 4 \,\} =$
7) $\{\, a \in A \mid (a-3)^2 = 9 \,\} =$
8) $\{\, b \in B \mid (b-3)^2 = 9 \,\} =$
9) $\{\, a \in A \mid a \in B \,\} =$
10) $\{\, b \in B \mid b \in A \,\} =$
11) $\{\, a \in A \mid a \notin B \,\} =$
12) $\{\, b \in B \mid b \notin A \,\} =$
13) $\{\, a \in A \mid 2a \in B \,\} =$
14) $\{\, b \in B \mid 2b \in A \,\} =$
15) $\{\, a \in A \mid a^2 \in B \,\} =$
16) $\{\, a \in A \mid a^2 < 0 \,\} =$

Students of mathematics are expected to be able to *prove* assertions about sets and their subsets. Before giving an example of this, let us point out that if A is a set, $P(x)$ is a predicate, and

$$B = \{\, a \in A \mid P(a) \,\},$$

then the assertion "$b \in B$" is logically equivalent to the assertion "$(b \in A)$ & $P(b)$." Thus, in a proof:

- if the assertion $b \in B$ is known to be true, then we also know that $b \in A$ and $P(b)$ are also true; and
- if the assertions $b \in A$ and $P(b)$ are both known to be true, then we also know that $b \in B$ is true.

In a 2-column proof, the justification for these assertions is "definition of B" or "because $B = \{\, a \in A \mid P(a) \,\}$."

Of course, not all sets are called A and B, but the same principles hold for sets named with other letters.

EXAMPLE 3.2.29. Suppose X and Y are sets, and let $Z = \{\, y \in Y \mid y \notin X \,\}$. Show that if $z \in Z$, then $z \notin X$.

SOLUTION. Assume $z \in Z$. Then, from the definition of Z, we know $z \in Y$ and $z \notin X$. In particular, $z \notin X$, as desired. □

EXERCISES 3.2.30.

1) Assume $A = \{\, d \in D \mid d \text{ is hungry} \,\}$.
   (a) If we *know* that $p \in A$, then what do we know about p?
   (b) If we want to *prove* that $q \in A$, then what do we need to show?
2) Assume $f(x) = x^2$ and $K = \{\, y \in \mathbb{R} \mid f(y) > 2 \,\}$.
   (a) If we *know* that $s \in K$, then what do we know about s?
   (b) If we want to *prove* that $t \in K$, then what do we need to show?
3) Assume $G = \{\, i \in I \mid M(i) \,\}$, where $M(x)$ is a predicate.
   (a) If we *know* that $v \in G$, then what do we know about v?
   (b) If we want to *prove* that $w \in G$, then what do we need to show?

EXERCISES 3.2.31. Write your proofs in English.

1) Assume A and B are sets, and let $C = \{\, a \in A \mid a \in B \,\}$. Show that if $c \in C$, then $c \in B$.
2) Let $A = \{\, x \in \mathbb{R} \mid x^2 - 5x = 14 \,\}$. Show that if $a \in A$, then $a < 10$.

NOTATION 3.2.32. When talking about sets or using predicates, we usually assume that a "**universe of discourse**" $\mathcal{U}$ has been agreed on. This means that all the elements of all of the sets under discussion are assumed to be members of $\mathcal{U}$. Then

$$\{\, x \mid P(x) \,\}$$

can be used as an abbreviation for $\{\, x \in \mathcal{U} \mid P(x) \,\}$.

The universe of discourse is sometimes assumed to be understood from the context, but it is an important concept, and it is best to specify it so that there is no room for confusion. For example, if we say "Everyone is happy," who is included in this *everyone*? We usually do not mean everyone now alive on the Earth. We certainly do not mean everyone who was ever alive or who will ever live. We mean something more modest: perhaps we mean everyone in the building, or everyone in the class, or maybe we mean everyone in the room.

Specifying a universe of discourse eliminates this ambiguity. The $\mathcal{U}$ is the set of things that we are talking about. So if we want to talk about people in Lethbridge, we define $\mathcal{U}$ to be the set of all people in Lethbridge. We write this at the beginning of our symbolization key, like this:

$\mathcal{U}$: the set of all people in Lethbridge

Everything that follows *ranges over* the universe of discourse. Given this $\mathcal{U}$, "everyone" means "everyone in Lethbridge" and "someone" means "someone in Lethbridge."

Each constant names some member of $\mathcal{U}$, so, if $\mathcal{U}$ is the set of people in Lethbridge, then constants Donald, Greg, and Mary can only be used if these three people are all in Lethbridge. If we want to talk about people in places besides Lethbridge, then we need to specify a different universe of discourse.

EXAMPLE 3.2.33. If $\mathcal{U}$ is the set of all Canadian provinces, then

$$\{\, x \mid \text{the English name of } x \text{ has three syllables} \,\} = \{\mathsf{Alberta}, \mathsf{New\ Brunswick}\}.$$

Remark 3.2.34. There is a very close relationship between sets and unary predicates. In general:

- From any unary predicate $P(x)$, we can define the set

$$\{x \mid P(x)\}.$$

- Conversely, from any set A, we can define a unary predicate $P(x)$ to be "x is a member of A."

Because of this, sets are more-or-less interchangeable with unary predicates. For example, the predicate "x is a dog" can be symbolized in two quite different ways:

- Our symbolization key could state that $D(x)$ means "x is a dog."
- Alternatively, our symbolization key could let D be the set of all dogs. Then "x is a dog" would be translated as "$x \in D$."

Mathematicians use sets much more often than unary predicates. We will see in Remark 4.2.1 that this tends to make it easier to translate statements from English into First-Order Logic when quantifiers are involved.

3.3. Operations on sets

There are several important ways that a new set can be made from sets that you already have. Any method of doing this is called a **set operation**.

3.3A. Union and intersection. Two of the most basic operations are *union* and *intersection*. Let us first discuss them in informal terms. Suppose:

- Alice and Bob are going to have a party, and need to decide who should be invited,
- Alice made a list of all the people that she would like to invite, and
- Bob made a list of all the people that he would like to invite.

Here are two of the many possible decisions they could make.

1) One solution would be to invite everyone that is on either of the lists. That is, they could begin their invitation list by writing down all of the names on Alice's list, and then add all of the names from Bob's list (or, more precisely, the names from Bob's list that are not already included in Alice's list). This is the *union* of the lists.
2) A much more conservative solution would be to invite only the people that appear on both of the lists. That is, they could go through Alice's list, and cross off everyone that does not appear on Bob's list. (They would get the same result by going through Bob's list, and crossing off everyone that does not appear on Alice's list.) This is the *intersection* of the lists.

DEFINITION 3.3.1. Suppose A and B are sets.

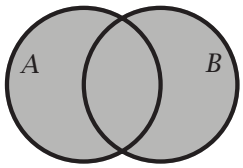
1) The **union** of A and B is the set

$$A \cup B = \{\, x \mid x \in A \textbf{ or } x \in B \,\}.$$

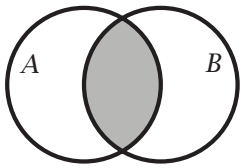
2) The **intersection** of A and B is the set

$$A \cap B = \{\, x \mid x \in A \textbf{ and } x \in B \,\}.$$

Remark 3.3.2. By drawing the sets A and B as overlapping circles, the union and intersection can be represented as follows:



$A \cup B$ is shaded



$A \cap B$ is shaded

Pictures like these are called **Venn diagrams**.

Remark 3.3.3.

1) In ordinary English, the word "intersection" refers to where two things meet. For example, the intersection of two streets is where the two streets come together. We can think of this area as being part of both streets, so this is consistent with the way the term is used in mathematics.
2) In ordinary English, the word "union" refers to joining things together. For example, a marriage is the union of two people — it joins the two people into a single married couple. This is consistent with the way the term is used in mathematics — we could form the union of Alice's list and Bob's list by gluing Bob's list to the end of Alice's list.

EXAMPLE 3.3.4.

1) $\{1,3,5,7,9\} \cup \{1,4,7,10\} = \{1,3,4,5,7,9,10\}$
2) $\{1,3,5,7,9\} \cap \{1,4,7,10\} = \{1,7\}$

EXERCISES 3.3.5. Specify each set by listing its elements. (*You do not need to show your work.*)

1) $\{1,2,3,4\} \cup \{3,4,5,6,7\} =$
2) $\{1,2,3,4\} \cap \{3,4,5,6,7\} =$
3) $\{\mathsf{p},\mathsf{r},\mathsf{o},\mathsf{n},\mathsf{g}\} \cap \{\mathsf{h},\mathsf{o},\mathsf{r},\mathsf{n},\mathsf{s}\} =$
4) $\{\mathsf{p},\mathsf{r},\mathsf{o},\mathsf{n},\mathsf{g}\} \cup \{\mathsf{h},\mathsf{o},\mathsf{r},\mathsf{n},\mathsf{s}\} =$
5) $\big(\{1,3,5\} \cup \{2,3,4\}\big) \cap \{2,4,6\} =$
6) $\big(\{1,3,5\} \cap \{2,3,4\}\big) \cup \{2,4,6\} =$

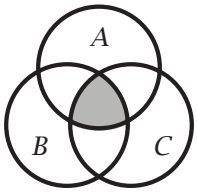
EXERCISES 3.3.6.

1) Assume A and B are sets. Prove that if $c \in A \cap B$, then $c \in A$.
2) Assume X, Y, and Z are sets. Prove that if $r \in (X \cap Y) \cup (X \cap Z)$, then $r \in X$.

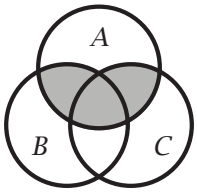
Remark 3.3.7.

1) It is not difficult to see that $\cup$ and $\cap$ are commutative. That is, for all sets A and B, we have
$$A \cup B = B \cup A \quad \text{and} \quad A \cap B = B \cap A.$$
2) It is also not difficult to see that $\cup$ and $\cap$ are associative. That is, for all sets A, B, and C, we have
$$(A \cup B) \cup C = A \cup (B \cup C) \quad \text{and} \quad (A \cap B) \cap C = A \cap (B \cap C).$$
So there is no need for parenthesis when writing $A \cup B \cup C$ or $A \cap B \cap C$. (However, you *do* need parentheses when writing something like $A \cup (B \cap C)$ or $(A \cup B) \cap C$, which uses both $\cup$ and $\cap$.)

EXAMPLE 3.3.8. A Venn diagram can include more than two sets. For example, here are Venn diagrams of $A \cap B \cap C$ and $A \cap (B \cup C)$.



$A \cap B \cap C$ is shaded



$A \cap (B \cup C)$ is shaded

EXERCISE 3.3.9. Draw a Venn diagram of each set. (*You do not need to show your work.*)

1) $A \cup B \cup C$ 2) $A \cup (B \cap C)$ 3) $(A \cup B) \cap C$ 4) $(A \cap C) \cup (B \cap C)$

3.3B. Set difference and complement. The "set difference" is another fundamental operation.

EXAMPLE 3.3.10. If there is a list of people that Alice would like to invite to the party, and also a list of people that Bob refuses to allow to come to the party (the "veto list"), then it would be reasonable to invite the people that are on Alice's list, but not on the veto list. That is, they could start with Alice's list, and remove all of the names that are on Bob's list. This is the **set difference** of Alice's list and Bob's list.

DEFINITION 3.3.11. Suppose A and B are sets.

1) The **set difference** of A and B is the set

$$A \smallsetminus B = \{\, x \in A \mid x \notin B \,\} = \{\, x \mid (x \in A) \,\&\, (x \notin B) \,\}.$$

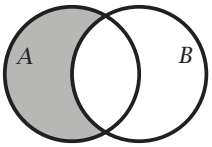
2) The **complement** of B is the set

$$\overline{B} = \mathcal{U} \smallsetminus B = \{\, x \mid x \notin B \,\},$$

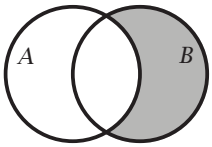
where $\mathcal{U}$ is the universal set, as usual.

OTHER NOTATION. Some authors write $A - B$, instead of $A \smallsetminus B$. Also, some authors write B^c, instead of $\overline{B}$

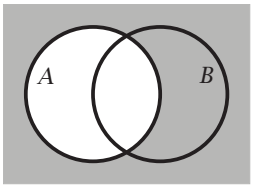
Remark 3.3.12. Here are Venn diagrams.



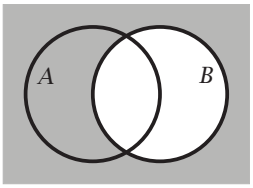
$A \smallsetminus B$ is shaded



$B \smallsetminus A$ is shaded



$\overline{A}$ is shaded



$\overline{B}$ is shaded

EXAMPLE 3.3.13. Suppose $\mathcal{U} = \mathsf{PEOPLE}$ is the set of all people.

1) $\overline{\mathsf{CHILDREN}} = \mathsf{ADULTS}$, because adults are the people who are not children.

2) $\mathsf{FEMALES} \setminus \mathsf{CHILDREN}$ is the set of all adult women.

EXERCISES 3.3.14. Assume $\mathcal{U} = \{1, 2, 3, \ldots, 10\}$. Specify each set by listing its elements. (*You do not need to show your work.*)

1) $\{1, 3, 5, 7, 9\} \setminus \{4, 5, 6, 7\} =$

2) $\{4, 5, 6, 7\} \setminus \{1, 3, 5, 7, 9\} =$

3) $\overline{\{1, 3, 5, 7, 9\}} =$

4) $\overline{\{4, 5, 6, 7\}} =$

EXERCISE 3.3.15. Draw a Venn diagram of each set. (*You do not need to show your work.*)

1) $\overline{A \cup B}$

2) $\overline{A \cap B}$

3) $(A \setminus B) \setminus (A \setminus C)$

4) $A \setminus (B \setminus C)$

5) $(A \cup B) \setminus C$

EXAMPLE 3.3.16. Suppose A and B are sets. Show that if $c \in \overline{A \cup B}$, then $c \in \overline{A} \cap \overline{B}$.

SOLUTION. Assume $c \in \overline{A \cup B}$. By definition of the complement, this means $c \notin A \cup B$. In other words, it is not true that $c \in A \cup B$. From the definition of $A \cup B$, we conclude that

$$\text{it is not true that either } c \in A \text{ or } c \in B.$$

By the rules of negation, this means $c \notin A$ and $c \notin B$. Now:

- since $c \notin A$, we have $c \in \overline{A}$, and
- since $c \notin B$, we have $c \in \overline{B}$.

Therefore, we know $c \in \overline{A}$ and $c \in \overline{B}$, so $c \in \overline{A} \cap \overline{B}$. □

EXERCISE 3.3.17. Suppose A and B are sets. Show that if $c \in \overline{A} \cap \overline{B}$, then $c \in \overline{A \cup B}$.

3.3C. Disjoint sets.

DEFINITION 3.3.18. Two sets A and B are said to be **disjoint** iff their intersection is empty (that is, $A \cap B = \varnothing$). In other words, they have no elements in common:

$$A \text{ and } B \text{ are disjoint} \quad \Leftrightarrow \quad \begin{array}{c}\text{there does } \textit{not} \text{ exist an } x, \\ \text{such that } \big((x \in A) \,\&\, (x \in B)\big).\end{array}$$

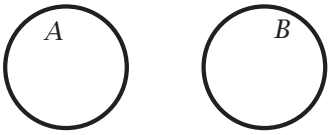
We may also say that A is **disjoint from B**.

EXAMPLE 3.3.19.

1) The sets $\{1, 3, 5\}$ and $\{2, 4, 6\}$ are disjoint, because they have no elements in common.

2) The sets $\{1, 3, 5\}$ and $\{2, 3, 4\}$ are *not* disjoint, because 3 is in their intersection.

3) The following Venn diagram illustrates two disjoint sets A and B (they do not overlap):



A and B are disjoint

Remark 3.3.20. Let us point out some well-known facts that will be formally proved in Chapter 9.

1) If A and B are two disjoint sets, then $\#(A \cup B) = \#A + \#B$.
2) The situation is similar even if there are more than 2 sets: Suppose $A_1, A_2, \ldots, A_n$ are **pairwise-disjoint** sets. (This means that A_i is disjoint from A_j whenever $i \neq j$.) Then
$$\#(A_1 \cup A_2 \cup \cdots \cup A_n) = \#A_1 + \#A_2 + \cdots + \#A_n.$$
3) If A and B are two finite sets that are *not* disjoint, then $\#(A \cup B) < \#A + \#B$.

3.3D. The power set.

EXAMPLE 3.3.21. It is not difficult to list all of the subsets of $\{a, b, c\}$. One way to do this is to consider the possible number of elements in the subset:

0) A subset with 0 elements has no elements at all. It must be the empty set $\varnothing$.
1) Consider a subset with 1 element. That one element must be one of the elements of $\{a, b, c\}$. That is, the element of the set must be a, b, or c. So the 1-element subsets are $\{a\}$, $\{b\}$, and $\{c\}$.
2) Consider the subsets with 2 elements.
   - If a is one of the elements in the subset, then the other element must be either b or c.
   - If a is not in the subset, then the subset must contain both b and c.

   Hence, the 2-element subsets are $\{a, b\}$, $\{a, c\}$, and $\{b, c\}$.
3) A subset with 3 elements must have all of the elements of $\{a, b, c\}$, so the subset must be $\{a, b, c\}$.

(≥ 4) Because $\{a, b, c\}$ has only 3 elements, we know that no subset can have more than 3 elements.

Thus, the subsets of $\{a, b, c\}$ are
$$\varnothing,\ \{a\}, \{b\}, \{c\},\ \{a, b\}, \{a, c\}, \{b, c\},\ \{a, b, c\}.$$
Counting them, we see that there are exactly 8 subsets.

Remark 3.3.22. In general, one can show that any set with n elements has exactly 2^n subsets. In the above example, we have $n = 3$, and the number of subsets is $2^3 = 8$.

EXERCISES 3.3.23. (*You do not need to show your work.*)

1) List the subsets of $\{a\}$.
2) List the subsets of $\{a, b\}$.
3) List the subsets of $\{a, b, c, d\}$.
4) List the subsets of $\varnothing$.

We can make a set by putting set braces at the ends of the above list of subsets of $\{a, b, c\}$:
$$\big\{\varnothing,\ \{a\}, \{b\}, \{c\},\ \{b, c\}, \{a, c\}, \{a, b\},\ \{a, b, c\}\big\}.$$
In general, the set of all subsets of a set is called its *power set*:

DEFINITION 3.3.24. Suppose A is a set. The **power set** of A is the set of all subsets of A. It is denoted $\mathcal{P}(A)$. This means
$$\mathcal{P}(A) = \{\, B \mid B \subset A \,\}.$$

Remark 3.3.25. From Remark 3.3.22, we see that if $\#A = n$, then $\#\mathcal{P}(A) = 2^n$. This formula involving "two-to-the-nth-*power*" can be considered a justification for calling $\mathcal{P}(A)$ the *power* set.

EXERCISES 3.3.26.

1) Describe each of the following sets by listing its elements.
(*You do not need to show your work.*)
(a) $\mathcal{P}(\varnothing)$ (b) $\mathcal{P}\big(\{\mathsf{a}\}\big)$ (c) $\mathcal{P}\big(\{\mathsf{a},\mathsf{b}\}\big)$
(d) $\mathcal{P}\big(\{\mathsf{a},\mathsf{b},\mathsf{c}\}\big)$ (e) $\mathcal{P}\big(\{\mathsf{a},\mathsf{b},\mathsf{c},\mathsf{d}\}\big)$

2) Which of the following are elements of $\mathcal{P}\big(\{\mathsf{a},\mathsf{c},\mathsf{d}\}\big)$?
(*You do not need to show your work.*)
(a) a (b) $\{\mathsf{a}\}$ (c) $\{\mathsf{a},\mathsf{b}\}$ (d) $\{\mathsf{a},\mathsf{c}\}$

3) Suppose A is a set.
(a) Is $\varnothing \in \mathcal{P}(A)$? *Why?* (b) Is $A \in \mathcal{P}(A)$? *Why?*

4) Does there exist a set A, such that $\mathcal{P}(A) = \varnothing$?

5) Let

$$V_0 = \varnothing, \quad V_1 = \mathcal{P}(V_0), \quad V_2 = \mathcal{P}(V_1) = \mathcal{P}\big(\mathcal{P}(V_0)\big), \quad \text{and so forth.}$$

In general, $V_n = \mathcal{P}(V_{n-1})$ whenever $n > 0$.
(a) What are the cardinalities of V_0, V_1, V_2, V_3, V_4, and V_5?
(b) Describe V_0, V_1, V_2, and V_3 by listing their elements.
(c) (harder) Describe V_4 by listing its elements.
(d) Is it reasonable to ask someone to list the elements of V_5? *Why?*

SUMMARY:

- Important definitions:
  - set
  - element
  - subset
  - proper subset
  - predicate
  - union
  - intersection
  - set difference
  - complement
  - disjoint
  - pairwise-disjoint
  - power set
- A set is unordered and without repetition.
- $\varnothing$ and A are subsets of A.
- $A = B$ if and only if we have both $A \subset B$ and $B \subset A$.
- For our purposes, predicates usually have only one or two variables.
- If a predicate has two variables, the order of the variables is important.
- Venn diagrams are a tool for illustrating set operations.
- $\#\mathcal{P}(A) = 2^{\#A}$
- Notation:
  - $\{ \quad \}$
  - $\in$, $\notin$
  - $\varnothing$ (empty set)
  - $\#A$
  - $A \subset B$, $A \not\subset B$, $A \supset B$
  - $P(x)$, $x \; Q \; y$ (predicates)
  - $\{ \, a \in A \mid P(a) \, \}$
  - $\mathcal{U}$ (universe of discourse)
  - $\mathbb{N}$, $\mathbb{N}^+$, $\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{R}$
  - $A \cup B$
  - $A \cap B$
  - $A \setminus B$
  - $\overline{A}$
  - $\mathcal{P}(A)$

Chapter 4

First-Order Logic

Three Logicians walk into a bar, and
the barkeeper asks "Would you all like something to drink?"
The 1st Logician says "I don't know,"
and the 2nd Logician says "I don't know."
Then the 3rd Logician says "yes."

author unknown

4.1. Quantifiers

Earlier, we observed that Propositional Logic cannot fully express ideas involving quantity, such as "some" or "all." In this chapter, we will fill this gap by introducing quantifier symbols. Together with predicates and sets, which have already been discussed, this completes the language of First-Order Logic. We will then use this language to translate assertions from English into mathematical notation.

In our first example we will use this symbolization key:

$\mathcal{U}$: The set of all people.	$x\ R\ y$: x is richer than y
L: The set of all people in Lethbridge.	d: Donald
A: The set of all angry people.	g: Gregor
H: The set of all happy people.	m: Marybeth

Now consider these assertions:

1. Everyone is happy.
2. Everyone in Lethbridge is happy.
3. Everyone in Lethbridge is richer than Donald.
4. Someone in Lethbridge is angry.

It might be tempting to translate Assertion 1 as $(d \in H)\ \&\ (g \in H)\ \&\ (m \in H)$. Yet this would only say that Donald, Gregor, and Marybeth are happy. We want to say that *everyone* is happy, even if we have not listed them in our symbolization key. In order to do this, we introduce the "$\forall$" symbol. This is called the **universal quantifier**.

$\forall x$ means "for all x"

A quantifier must always be followed by a variable, and then a formula that the quantifier applies to. We can translate Assertion 1 as $\forall x, (x \in H)$. Paraphrased in English, this means "For all x, x is happy."

In quantified assertions such as this one, the variable x is serving as a kind of placeholder. The expression $\forall x$ means that you can pick anyone and put them in as x. There is no special

reason to use x rather than some other variable. The assertion "$\forall x, (x \in H)$" means exactly the same thing as "$\forall y, (y \in H)$," "$\forall z, (z \in H)$," or "$\forall x_5, (x_5 \in H)$."

To translate Assertion 2, we use a different version of the universal quantifier:

If X is any set, then $\forall x \in X$ means "for all x in X"

Now we can translate Assertion 2 as $\forall \ell \in L, (\ell \in H)$. (It would also be logically correct to write $\forall x \in L, (x \in H)$, but ℓ is a better name an element of the set L.) Paraphrased in English, our symbolic assertion means "For all ℓ in Lethbridge, ℓ is happy."

Assertion 3 can be paraphrased as, "For all ℓ in Lethbridge, ℓ is richer than Donald." This translates as $\forall \ell \in L, (\ell \; R \; d)$.

To translate Assertion 4, we introduce another new symbol: the **existential quantifier**, $\exists$.

$\exists x$ means "there exists some x, such that"

If X is any set, then $\exists x \in X$ means "there exists some x in X, such that"

We write $\exists \ell \in L, (\ell \in A)$. This means that there exists some ℓ in Lethbridge who is angry. More precisely, it means that there is *at least one* angry person in Lethbridge. Once again, the variable is a kind of placeholder; it would have been logically correct (but poor form) to translate Assertion 4 as $\exists z \in L, (z \in A)$.

EXAMPLE 4.1.1. Consider this symbolization key.

S: The set of all students.

B: The set of all books.

N: The set of all novels.

$x \; L \; y$: x likes to read y.

Then:

1) $\forall n \in N, (n \in B)$ means "every novel is a book," and

2) $\forall s \in S, \big(\exists b \in B, (s \; L \; b)\big)$ means "for every student, there is some book that the student likes to read."

Notice that all of the quantifiers in this example are of the form $\forall x \in X$ or $\exists x \in X$, not $\forall x$ or $\exists x$. That is, all of the variables range over specific sets, rather than being free to range over the entire universe of discourse. Because of this, it is acceptable to omit specifying a universe of discourse. Of course, the universe of discourse (whatever it is) must include at least all students, all books, and all novels.

EXERCISE 4.1.2. Suppose A and B are sets.
Give your answers in the notation of First-Order Logic (not English).

1) What does it mean to say that A is a subset of B?

2) What does it mean to say that A is *not* a subset of B?

4.2. Translating to First-Order Logic

We now have all of the pieces of First-Order Logic. Translating assertions (no matter how complicated) from English to mathematical notation will only be a matter of knowing the right way to combine predicates, constants, quantifiers, connectives, and sets. Consider these assertions:

5. Every coin in my pocket is a dime.

6. Some coin on the table is a dime.
7. Not all the coins in my pocket are dimes.
8. None of the coins on the table are dimes.

In providing a symbolization key, we need to specify $\mathcal{U}$. Since we are not talking about anything besides coins, we may let $\mathcal{U}$ be the set of all coins. (It is not necessary to include all coins in $\mathcal{U}$, but, since we are talking about the coins in my pocket and the coins on the table, $\mathcal{U}$ must at least contain all of those coins.) Since we are not talking about any specific coins, we do not need to define any constants. Since we will be explicitly talking about the coins in my pocket and the coins on the table, it will be helpful to have these defined as sets. The symbolization key also needs to say something about dimes; let's do this with a predicate. So we define this key:

$\mathcal{U}$: The set of all coins.
P: The set of all coins in my pocket.
T: The set of all coins on the table.
$D(x)$: x is a dime.

Assertion 5 is most naturally translated with a universal quantifier. It talks about all of the coins in my pocket (that is, the elements of the set P). It means that, for any coin in my pocket, that coin is a dime. So we can translate it as $\forall p \in P, D(p)$.

Assertion 6 says there is some coin on the table, such that the coin is a dime. So we translate it as $\exists t \in T, D(t)$.

Assertion 7 can be paraphrased as, "It is not the case that every coin in my pocket is a dime." So we can translate it as $\neg\big(\forall p \in P, D(p)\big)$. This is simply the negation of Assertion 5.

Assertion 8 can be paraphrased as, "It is not the case that some coin on the table is a dime." This can be translated as $\neg\big(\exists t \in T, D(t)\big)$. It is the negation of Assertion 6.

Remark 4.2.1. Alternatively, we could have defined a set D, the set of all dimes, instead of the predicate $D(x)$. In this case:

- Assertion 5 would be translated as $\forall p \in P,\ p \in D$.
- Assertion 6 would be translated as $\exists t \in T,\ t \in D$.
- Assertion 7 would be translated as $\neg(\forall p \in P,\ p \in D)$.
- Assertion 8 would be translated as $\neg(\exists t \in T,\ t \in D)$.

Either approach is perfectly legitimate.

However, if we need to quantify over dimes, then it is much better to use a set than a predicate. For example, let's symbolize the assertion "Every dime is in my pocket."

- If D is the set of all dimes, we can write: $\forall d \in D,\ d \in P$.
- It is less straightforward if we use the predicate $D(x)$, because we cannot directly say "for all dimes" (since the set of all dimes is not available). The translation must consider all coins, and then say that the ones that are dimes are in my pocket: $\forall x, \big(D(x) \Rightarrow d \in P\big)$.

For this reason (and others), mathematicians tend to use sets, rather than predicates, and we will do the same.

Remark 4.2.2. If we had defined the predicate $P(x)$ (for "x is in my pocket") instead of the corresponding set P, we would have needed to translate Assertion 5 as $\forall x, \big(P(x) \Rightarrow D(x)\big)$: that is, "for any coin, if it is in my pocket, then it is a dime." Since the assertion is about coins that are both in my pocket *and* that are dimes, it might be tempting to translate it using &. However, the assertion $\forall x, \big(P(x)\ \&\ D(x)\big)$ would mean that everything in $\mathcal{U}$ is both in my

pocket and a dime: All the coins that exist are dimes in my pocket. This is would be a crazy thing to say, and it means something very different than Assertion 5. However, this issue is completely avoided when we use sets. Thus, defining P to be a set, rather than a predicate, is the best approach in this problem (and similarly for T).

EXAMPLE 4.2.3. We can now translate the deduction from page 55, the one that motivated the need for quantifiers:

Merlin is a wizard. All wizards wear funny hats.
$\therefore$ Merlin wears a funny hat.

$\mathcal{U}$: The set of all people.
W: The set of all wizards.
H: The set of all people who wear a funny hat.
m: Merlin

Translating, we get:

Hypotheses:
$m \in W$
$\forall w \in W, (w \in H)$
Conclusion: $m \in H$

This captures the structure that was left out when we translated the deduction into Propositional Logic, and this is a valid deduction in First-Order Logic. We will be able to prove it rigorously after we have discussed the introduction and elimination rules for $\forall$ (and $\exists$) in Section 4.4.

EXERCISES 4.2.4. Using the given symbolization key, translate each English-language assertion into First-Order Logic.

$\mathcal{U}$: The set of all animals.
A: The set of all alligators.
R: The set of all reptiles.
Z: The set of all animals who live at the zoo.
M: The set of all monkeys.
$x \heartsuit y$: x loves y.
a: Amos
b: Bouncer
c: Cleo

1) Amos, Bouncer, and Cleo all live at the zoo.
2) Bouncer is a reptile, but not an alligator.
3) If Cleo loves Bouncer, then Bouncer is a monkey.
4) If both Bouncer and Cleo are alligators, then Amos loves them both.
5) Some reptile lives at the zoo.
6) Every alligator is a reptile.
7) Every animal that lives at the zoo is either a monkey or an alligator.
8) There are reptiles which are not alligators.
9) Cleo loves a reptile.
10) Bouncer loves all the monkeys that live at the zoo.
11) All the monkeys that Amos loves love him back.
12) If any animal is an alligator, then it is a reptile.
13) Every monkey that Cleo loves is also loved by Amos.

14) There is a monkey that loves Bouncer, but Bouncer does not reciprocate this love.

EXERCISES 4.2.5. Using the given symbolization key, translate each English-language assertion into First-Order Logic.

$\mathcal{U}$: The set of all animals.	b: Bertie
D: The set of all dogs.	e: Emerson
S: The set of all animals who like to swim.	f: Fergis
$x\ L\ y$: x is larger than y.	

1) Bertie is a dog who likes to swim.
2) Bertie, Emerson, and Fergis are all dogs.
3) Emerson is larger than Bertie, and Fergis is larger than Emerson.
4) All dogs like to swim.
5) Every animal that likes to swim is a dog.
6) There is a dog that is larger than Emerson.
7) No animal that likes to swim is larger than Emerson.
8) Any animal that does not like to swim is larger than Bertie.
9) There is an animal that is larger than Bertie, but smaller than Emerson.
10) There is no dog that is larger than Bertie, but smaller than Emerson.
11) No dog is larger than itself.

EXERCISES 4.2.6. For each deduction, write a symbolization key and translate the deduction into First-Order Logic.

1) Nothing on my desk escapes my attention. There is a computer on my desk. Therefore, there is a computer that does not escape my attention.
2) All my dreams are black and white. Old TV shows are in black and white. Therefore, some of my dreams are old TV shows.
3) Neither Holmes nor Watson has been to Australia. A person could see a kangaroo only if they had been to Australia or to a zoo. Although Watson has not seen a kangaroo, Holmes has. Therefore, Holmes has been to a zoo.
4) No one expects the Spanish Inquisition. No one knows the troubles I've seen. Therefore, anyone who expects the Spanish Inquisition knows the troubles I've seen.
5) Every antelope is bigger than a bread box. The thing I am thinking of is no bigger than a bread box, and it is either an antelope or a cantaloupe. Therefore, I am thinking of a cantaloupe.

4.2A. Multiple quantifiers.

EXAMPLE 4.2.7. Consider the following symbolization key and the assertions that follow it:

$\mathcal{U}$: The set of all people.	$x\ L\ y$: x likes y.
F: The set of all of Karl's friends.	i: Imre.
N: The set of all of Imre's neighbours.	k: Karl.

9. All of Imre's neighbours like all of Karl's friends.
10. At least one of Karl's friends likes at least one of Imre's neighbours.

11. All of Karl's friends like at least one of Imre's neighbours.
12. There is one of Imre's neighbours, who is a friend of Karl and who likes all of Imre's neighbours.

Beginning to translate Assertion 9, we start with all of Imre's neighbours: $\forall n \in N$. Now we would like to say $n\ L\ f$, where f represents every one of Karl's friends. Before we can do this, we need to introduce the variable f, and give it the desired meaning and the appropriate quantifier: $\forall f \in F$. Thus, Assertion 9 can be translated as $\forall n \in N, \big(\forall f \in F, (n\ L\ f)\big)$.

For Assertion 10, we start with at least one of Karl's friends. Another way to say this is that there is some friend of Karl's: $\exists f \in F$. Similarly, we now need to introduce at least one of Imre's neighbours: $\exists n \in N$. The completed translation is $\exists f \in F, \big(\exists n \in N, (f\ L\ n)\big)\big)$.

For Assertion 11, we start with all of Karl's friends: $\forall f \in F$. Now we need at least one of Imre's neighbours: $\exists n \in N$. The completed translation is $\forall f \in F, \big(\exists n \in N, (f\ L\ n)\big)$.

Finally, for Assertion 12, we start with one of Imre's neighbours: $\exists n \in N$. Now we need this person to be a friend of Karl: $n \in F$. For the next part of the sentence, we need all of Imre's neighbours. It is tempting to write $\forall n \in N$, but we have already used the variable n for a particular one of Imre's neighbours, so we cannot use it again here to mean something else. Let's use n' instead: $\forall n' \in N$. We are now ready to translate Assertion 12:

$$\exists n \in N, \big(n \in F \ \&\ \big[\forall n' \in N, (n\ L\ n')\big]\big).$$

(Alternatively, we could have used n_1 and n_2 instead of n and n'.)

When symbolizing assertions with multiple quantifiers, it is best to proceed by small steps. Figure out who is being discussed in the sentence, and what quantifiers are required to introduce these variables. Paraphrase the English assertion so that the logical structure is readily symbolized in First-Order Logic. Then translate bit by bit, replacing the daunting task of translating a long assertion with the simpler task of translating shorter formulas.

Remark 4.2.8. The equals sign "=" is a part of every symbolization key (even though we do not bother to include it explicitly). It is a binary predicate, and, as you would expect, "$x = y$" means "x is equal to y." This does not mean merely that x and y look very much alike, or that they are indistinguishable, or that they have all of the same properties. Rather, it means that x and y are (different) names for the same object.

WARNING. It is important to put quantifiers in the correct order. For example, consider the following assertions (with $\mathcal{U}$ being the set of all objects):

13. $\forall x, \big(\exists y, (x = y)\big)$
14. $\exists y, \big(\forall x, (x = y)\big)$

They are exactly the same, except for which of $\forall x$ and $\exists y$ is first, and which is second. Assertion 13 is obviously true: "For every thing, there is something that it is equal to." (Namely, every thing is equal to itself.) But Assertion 14 says: "There is some thing (let us call it y), such that everything is equal to y." There is no such y, so this is obviously false.

EXERCISES 4.2.9. Using the symbolization key from Exercise 4.2.5, translate each English-language assertion into First-Order Logic.

1) If there is a dog larger than Fergis, then there is a dog larger than Emerson.
2) Every dog is larger than some dog.
3) There is an animal that is smaller than every dog.
4) If there is an animal that is larger than every dog, then Emerson does not like to swim.
5) For every dog that likes to swim, there is a smaller dog that does not like them.

6) Every dog that likes to swim is larger than every dog that does not like them.

7) Some animal is larger than all of the dogs that like to swim.

8) If there is a dog that does not like to swim, than there is a dog that is larger than every animal that likes to swim.

EXERCISES 4.2.10 (harder). Use the symbolization key to translate each English-language assertion into First-Order Logic.

$\mathcal{U}$: The set of all people.

$x\ C\ y$: x is a child of y.

$x\ S\ y$: x is a sibling of y.

D: The set of all ballet dancers.

e: Elmer

F: The set of all females.

j: Jane

M: The set of all males.

p: Patrick

1) Everyone who dances ballet is the child of someone who dances ballet.

2) Every man who dances ballet is the child of someone who dances ballet.

3) Everyone who dances ballet has a sister who also dances ballet.

4) Jane is an aunt.

5) Patrick's brothers have no children.

4.2B. Uniqueness. Saying "there is a **unique** so-and-so" means not only that there is a so-and-so, but also that there is only one of them—there are not two different so-and-so's. For example, to say that "there is a *unique* person who owes Hikaru money" means

some person owes Hikaru *and* no other person owes Hikaru.

This translates to

$$\exists h \in H, \big(\forall y, (y \neq h \Rightarrow y \notin H)\big);$$

or, equivalently,

$$\exists h \in H, \big(\forall y, (y \in H \Rightarrow y = h)\big).$$

Unfortunately, both of these are quite complicated expressions (and are examples of "multiple quantifiers," because they use both $\exists$ and $\forall$). To simplify the situation, mathematicians introduce a special notation:

"$\exists!\, x$" means "there is a unique x, such that..."

If X is any set, then "$\exists!\, x \in X$" means "there is a unique x in X, such that..."

For example, $\exists!\, h, h \in H$ means exactly the same thing as the complicated expression above.

If we add

R: The set of people who are rich.

to our symbolization key, we can translate "There is a unique rich person who owes Hikaru money." Namely, it translates as:

$$\exists!\, r \in R, (r \in H).$$

EXERCISES 4.2.11. Using the given symbolization key, translate each English-language assertion into First-Order Logic.

$\mathcal{U}$: The set of all creatures.

H: The set of all horses.

W: The set of all creatures with wings.

B: The set of all creatures in Farmer Brown's field.

1) There is a unique winged creature that is in Farmer Brown's field.

2) If some horse has wings, then there is a unique horse that is in Farmer Brown's field.

3) If there is a horse that is in Farmer Brown's field, then there is a unique horse that is in Farmer Brown's and has wings.

4.2C. Bound variables. Recall that an assertion is a statement that is either true or false. For example, consider the following symbolization key:

$\mathcal{U}$: The set of all students. a: Anna

$M(x)$: x is taking a math class.

Then:

- $M(a)$ is an assertion. Either Anna is taking a math class, or she is not.
- $M(x)$ is *not* an assertion. The letter x is a variable, not any particular object. (We call x a **free variable**.) If we plug in a particular value for x (such as a), then we will have an assertion. However, until some value is plugged in for x, we cannot say whether the expression is true or false. So the expression is not an assertion if the variable remains free.
- $\exists x, M(x)$ and $\forall x, M(x)$ are assertions. The letter x is a variable in both of these expressions, but it is no longer free, because it is acted on by the quantifier. (We call x a **bound variable**.)

An important principle of First-Order Logic is that, in an assertion, each variable must be bound by some quantifier:

Assertions cannot have free variables.

EXERCISES 4.2.12. Suppose that p is a constant, but all other lower-case letters represent variables. For each of the following, (a) does it have a free variable? (b) is it an assertion?

1) $\forall x \in X, (x \ L \ y)$

2) $(p \in S) \ \& \ \exists y \in Y, (y \ T \ p)$

3) $\forall v \in V, \big((\exists ! \, y \in Y, v \ R \ y) \Rightarrow (v = z)\big)$

4) $y \in Y \ \& \ \big(\forall x \in X, (x \in T)\big)$

5) $(p \ L \ p) \Rightarrow \exists x, (x \ L \ p)$

6) $\forall x \in X, (x \ L \ x)$

4.3. Negations

Recall part of the symbolization key of Section 4.1:

$\mathcal{U}$: The set of all people.

A: The set of all angry people.

$H(x)$: The set of all happy people.

Consider these further assertions:

15. No one is angry.

16. Not everyone is happy.

Assertion 15 can be paraphrased as, "It is not the case that someone is angry." (In other words, "It is not the case that there exists a person who is angry.") This is the negation of the assertion that there exists an angry person, so it can be translated using "not" and "there exists": $\neg \exists x, (x \in A)$.

It is important to notice that Assertion 15 is equivalent to the assertion that "Everyone is nonangry." This assertion can be translated using "for all" and "not": $\forall x, \neg(x \in A)$, or, in other words, $\forall x, (x \notin A)$. In general:

$\neg\exists x, \mathcal{A}$ is logically equivalent to $\forall x, \neg\mathcal{A}$.

This means that the negation of a "$\exists$" assertion is a "$\forall$" assertion.

Assertion 16 says it is not true that everyone is happy. This is the negation of the assertion that everyone is happy, so it can be translated using "not" and "$\forall$": $\neg\forall x, (x \in H)$.

Moreover, saying that not everyone is happy is the same as saying that someone is not happy. This latter assertion translates to $\exists x, (x \notin H)$. In general:

$\neg\forall x, \mathcal{A}$ is logically equivalent to $\exists x, \neg\mathcal{A}$.

This means that the negation of a "$\forall$" assertion is a "$\exists$" assertion.

Just as for "$\forall x$" and "$\exists x$," the bounded quantifiers "$\forall x \in X$" and "$\exists x \in X$" are interchanged under negation:

$\neg\forall x \in X, \mathcal{A}$ is logically equivalent to "$\exists x \in X, \neg\mathcal{A}$.

$\neg\exists x \in X, \mathcal{A}$ is logically equivalent to "$\forall x \in X, \neg\mathcal{A}$.

There is no fundamental difference between this and the previous examples; we have simply replaced $\mathcal{U}$ with the set X.

In summary: if you need to negate an assertion that starts with a quantifier, switch the quantifier to the other one (from $\exists$ to $\forall$ or vice-versa), and then continue, negating the remainder of the assertion.

To perform the additional negations, you will want to remember the following rules from Exercise 1.7.5:

Rules of Negation

$\neg(A \vee B)$ is logically equivalent to $\neg A \,\&\, \neg B$.

$\neg(A \,\&\, B)$ is logically equivalent to $\neg A \vee \neg B$.

$\neg(A \Rightarrow B)$ is logically equivalent to $A \,\&\, \neg B$.

$\neg\neg A$ is logically equivalent to A.

EXAMPLE 4.3.1. Let us simplify the assertion

$$(*) \qquad \neg\forall s \in S, \Big(\big((s \in A) \vee (s \in B)\big) \,\&\, \big((s \in C) \Rightarrow (s \notin D)\big)\Big).$$

We bring $\neg$ inside the quantifier, switching from $\forall$ to $\exists$:

$$\exists s \in S, \neg\Big(\big((s \in A) \vee (s \in B)\big) \,\&\, \big((s \in C) \Rightarrow (s \notin D)\big)\Big).$$

Now, we switch & to $\vee$, and apply $\neg$ to each of the two terms:

$$\exists s \in S, \Big(\neg\big((s \in A) \vee (s \in B)\big) \vee \neg\big((s \in C) \Rightarrow (s \notin D)\big)\Big).$$

Next, the connective $\vee$ in the left term is changed to & (and $\neg$ is applied to the subterms), and the rule for negating $\Rightarrow$ is implied to the right term:

$$\exists s \in S, \Big(\big(\neg(s \in A) \,\&\, \neg(s \in B)\big) \vee \big((s \in C) \,\&\, \neg(s \notin D)\big)\Big).$$

Finally, we use the abbreviation $\notin$ in the first two terms, and eliminate the double negative in the final term:

$$\exists s \in S, \Big(\big((s \notin A) \,\&\, (s \notin B)\big) \vee \big((s \in C) \,\&\, (s \in D)\big) \Big).$$

This final result is logically equivalent to Assertion $(*)$ above.

The same principles apply to negating assertions in English.

EXAMPLE 4.3.2. Suppose that we want to negate

"Every umbrella either needs a new handle or is not big enough."

We create a symbolization key:

U: The set of all umbrellas.

H: The set of all umbrellas that need a new handle.

B: The set of all umbrellas that are big enough.

Now we can translate the assertion as $\forall u \in U, \big((u \in H) \vee (u \notin B)\big)$. Negating this, we have

$$\neg \forall u \in U, \big((u \in H) \vee (u \notin B)\big).$$

We have just learned that this is equivalent to

$$\exists u \in U, \neg\big((u \in H) \vee (u \notin B)\big),$$

which can be simplified to

$$\exists u \in U, \big((u \notin H) \,\&\, \neg(u \notin B)\big),$$

and finally, eliminating the double negative, this is equivalent to

$$\exists u \in U, \big((u \notin H) \,\&\, (u \in B)\big).$$

Now we translate back to English:

"There is some umbrella that does not need a new handle and is big enough."

Applying these rules systematically will enable you to simplify the negation of any assertion (no matter whether it is expressed in English or in First-Order Logic).

English is more open to interpretation and inexactitude than First-Order Logic. Therefore, when we need to negate an English assertion in this chapter, we translate it into First-Order Logic, perform the negation, and translate back. You will also be expected to do this. Later, when you are doing proofs, you might be able to work directly with the English version, although you may find it helpful to keep the First-Order Logic version in mind.

WARNING. To make an assertion, quantifiers must be applied to predicates — they cannot stand by themselves. That is, an assertion must be of the form $\exists x \in X, P(x)$ or $\forall x \in X, P(x)$, not just $\exists x \in X$ or $\forall x \in X$. For example, some students erroneously try to translate the assertion "there exists an umbrella" as "$\exists u \in U$," but that is **not** an assertion. The problem is that it is not a complete sentence: it would translate into English as "there exists an umbrella, such that." (Notice the "such that" left dangling at the end.)

One way to obtain a correct symbolization is to rephrase the original assertion as: "there exists something that is an umbrella." This translates to $\exists x, (x \in U)$, which is a correct symbolization. Its negation simplifies to $\forall x, (x \notin U)$, which means "every thing that exists is not an umbrella."

If $\exists u \in U$ were an assertion, then, by applying the rules for negation, its negation would be "$\forall u \in U, \neg$," which is not a complete sentence: its English translation is "for all u in U, it is not true that." To avoid such mistakes, remember that every quantifier must always be followed by a predicate.

EXERCISES 4.3.3. Negate each of the assertions in Exercise 4.2.4. Express your answer both in the language of First-Order Logic and in English (after simplifying).

EXERCISES 4.3.4. Negate each of the following assertions of First-Order Logic (and simplify, so that $\neg$ is not applied to anything but predicates or assertion variables). *Show your work!*

1) $(L \Rightarrow \neg M)\ \&\ (M \vee N)$

2) $\big((a \in A)\ \&\ (b \in B)\big) \vee (c \in C)$

3) $\forall a \in A, \Big(\big((a \in P) \vee (a \in Q)\big)\ \&\ (a \notin R)\Big)$

4) $\forall a \in A, \Big((a \in T) \Rightarrow \exists c \in C, \big((c \in Q)\ \&\ (c\ R\ a)\big)\Big)$

5) $\forall x, \Big((x \in A)\ \&\ \Big(\exists \ell \in L, \big((x\ B\ \ell) \vee (\ell \in C)\big)\Big)\Big)$

6) $A \Rightarrow \Big(\big(\exists x \in X, (x \in B)\big) \vee \big(\forall e \in E, \exists d \in D, (e\ C\ d)\big)\Big)$

7) $\forall a \in A, \exists b \in B, \exists c \in C, \forall d \in D, \Big((a\ K\ b)\ \&\ \big((a\ Z\ c) \vee (b > d)\big)\Big)$

EXERCISES 4.3.5. Simplify each assertion. *Show your work!*

1) $\neg\forall a \in A, \big((a \in P) \vee (a \in Q)\big)$

2) $\neg\exists a \in A, \big((a \in P)\ \&\ (a \in Q)\big)$

3) $\neg\forall x \in X, \exists y \in Y, \big((x \in A)\ \&\ (x\ C\ y)\big)$

4) $\neg\forall s \in S, \Big((s \in R) \Rightarrow \Big(\exists t \in T, \big((s \neq t)\ \&\ (s\ M\ t)\big)\Big)\Big)$

Remark 4.3.6. Unfortunately, there is no nice, compact way of negating assertions involving uniqueness. For example, if we want to say "It is not the case that there is a unique person who owes Hikaru money," we need to say that "Either no one owes Hikaru money, or more than one person owes Hikaru money." This translates to

$$(H = \varnothing) \vee \big(\exists h_1 \in H, (\exists h_2 \in H, h_2 \neq h_1)\big).$$

In general, if you come across a situation where you want to negate an assertion that involves uniqueness, it is a good idea to rewrite the assertion without using "$\exists!$." You should have no difficulty negating this rephrased assertion.

4.3A. Vacuous truth. Note that if the assertion

$$\exists x \in A, \neg P(x)$$

is true (where A is any set and $P(x)$ is any unary predicate), then there must exist an element a of A, such that $P(a)$ is false. Ignoring the last condition (about $P(a)$), we know that $a \in A$, so $A \neq \varnothing$. That is, we know:

$$\text{If the assertion } \exists x \in A, \neg P(x) \text{ is true, then } A \neq \varnothing.$$

So the contrapositive is also true:

$$\text{If } A = \varnothing, \text{ then the assertion } \exists x \in A, \exists P(x) \text{ is false.}$$

Therefore, the assertion $\exists x \in \varnothing, \neg P(x)$ is false, so its negation is true:

$$\boxed{\text{The assertion } \forall x \in \varnothing, P(x) \text{ is true.}}$$

Since $P(x)$ is an arbitrary predicate, this means that any assertion about *all* of the elements of the empty set is true; we say it is **vacuously true**. The point is that there is nothing in the empty set to contradict whatever assertion you care to make about all of the elements.

EXAMPLE 4.3.7. If you say, "All of the people on Mars have purple skin," and there are not any people on Mars, then you have spoken the truth — otherwise, there would have to be some person on Mars (whose skin is not purple) to provide a counterexample.

In summary:

any assertion about *all* of the elements of the empty set is *vacuously true*.

EXERCISES 4.3.8. Which of the following English assertions are vacuously true (in the real world)?

1) All quintuplets are sickly.
2) All standard playing cards that are numbered fifteen, are green.
3) All prime numbers that are divisible by 12, have 5 digits.
4) All people who have been to the moon are men.
5) All people who do not breathe are dead.

4.4. The introduction and elimination rules for quantifiers

As you know, there are two quantifiers ($\exists$ and $\forall$). Each of these has an introduction rule and an elimination rule, so there are 4 rules to present in this section. Proofs in First-Order Logic can use both of these rules, plus all of the rules of Propositional Logic (such as the rules of negation and the basic theorems, including introduction and elimination rules), and also any other theorems that have been previously proved.

4.4A. $\exists$-introduction. We need to determine how to prove a conclusion of the form $\exists x \in X, \ldots$. For example, in a murder mystery, perhaps Inspector Thinkright gathers the suspects in a room and tells them, "Someone in this room has red hair." That is a $\exists$-statement. (With an appropriate symbolization key, in which P is the set of all of the the people in the room, and $R(x)$ is the predicate "x has red hair," it is the assertion $\exists p \in P, R(p)$.) How would the Inspector convince a skeptic that the claim is true? The easiest way would be to exhibit an explicit example of a person in the room who has red hair. For example, if Jim is in the room, and he has red hair, the Inspector might say,

> "Look, Jim is sitting right there by the door, and now, when I take off his wig, you can see for yourself that he has red hair. So I am right that someone in this room has red hair."

In general, the most straightforward way to prove $\exists p \in P, R(p)$ is true is to find a specific example of a p that makes $R(p)$ true. That is the essence of the $\exists$-intro rule.

Here is a principle to remember:

The proof of an assertion that begins
"there exists $x \in X$, such that..."
will usually be based on the statement "Let $x =$ $\square$,"
where the box is filled with an appropriate element of X.

OTHER TERMINOLOGY. Most mathematicians are not familiar with the terminology of introduction rules and elimination rules. Instead of saying this is the ∃-introduction rule, they would call it "proof by constructing an example," or "giving an explicit example," or other words to the same effect.

Here are some proofs that use ∃-introduction, but we cannot do very much with only one quantifier rule — the examples will be more interesting when we have more rules to work with.

EXAMPLE 4.4.1. Prove there is a natural number n, such that $n^2 = 64$.

PROOF. Let $n = 8 \in \mathbb{N}$. Then $n^2 = 8^2 = 64$. □

EXAMPLE 4.4.2. Prove there is a real number c, such that $5c^2 - 5c + 1 < 0$.

PROOF. Let $c = 1/2 \in \mathbb{R}$. Then

$$5c^2 - 5c + 1 = 5\left(\frac{1}{2}\right)^2 - 5\left(\frac{1}{2}\right) + 1 = \frac{5}{4} - \frac{5}{2} + 1 = -\frac{1}{4} < 0.$$ □

EXAMPLE 4.4.3. Let $N = \{1, 3, 5, 7\}$. Prove there exists $n \in N$, such that $n^3 - 11n^2 + 31n \neq 21$.

PROOF. Let $n = 5 \in N$. Then

$$n^3 - 11n^2 + 31n = 5^3 - 11(5^2) + 31(5) = 125 - 11(25) + 155 = 125 - 275 + 155 = 5 \neq 21.$$ □

EXERCISES 4.4.4.

1) Prove there is a real number r, such that $2r^2 + 9r + 4 = 0$.
2) Let $B = \{1, 3, 5, 7\}$. Prove there exists $b \in B$, such that $3b + 1 = (b - 1)^2$.
3) Prove there exist natural numbers m and n, such that $m^2 = n^3 + 1 > 1$.
   [*Hint:* Try *small* values of m and n.]
4) Show there is an integer n, such that $3n^2 = 5n + 2$.
5) Assume a is a real number. Show there is a real number x, such that $7x - 5 = a$.

4.4B. ∃-elimination. Perhaps Inspector Thinkright knows that one of the men lit a match at midnight, but does not know who it was. The Inspector might say,

> "We know that one of the men lit a match at midnight. Let us call this mysterious gentleman 'Mr. X.' Because right-handed matches are not allowed on the island, we know that Mr. X is left handed. Hence, Mr. X is not a butler, because all of the butlers in this town are right handed. . . ."

and so on, and so on, telling us more and more about Mr. X, based only on the assumption that he lit a match at midnight.

The situation in mathematical proofs is similar. Suppose we know there exists an element of the set A. Then it would be helpful to have a name for this mysterious element, so that we can talk about it. But a mathematician would not call the element "Mr. X": if it is an element of the set A, then he or she would probably call it a (or a_1 if there are going to be other elements of A to talk about). In general, the idea of the ∃-elimination rule is:

> If $\exists x \in X, P(x)$ is known to be true, then we may let x be an element of X, such that $P(x)$ is true.

In the remainder of the proof, we may assume only two things about x: that $x \in X$, and that $P(x)$ is true.

EXAMPLE 4.4.5. Show that if there exists $a \in \mathbb{R}$, such that $a^3 + a + 1 = 0$, then there exists $b \in \mathbb{R}$, such that $b^3 + b - 1 = 0$.

PROOF. Assume there exists $a \in \mathbb{R}$, such that $a^3 + a + 1 = 0$. Let $b = -a$. Then $b \in \mathbb{R}$, and

$$\begin{aligned} b^3 + b - 1 &= (-a)^3 + (-a) - 1 \\ &= -a^3 - a - 1 \\ &= -(a^3 + a + 1) \\ &= -0 && \text{(by the definition of } a\text{)} \\ &= 0, \end{aligned}$$

as desired. □

EXERCISES 4.4.6. Assume A and B are sets of real numbers.

1) Show that if there exists $a \in A$, such that $2a > 5$, then there exists $b \in A$, such that $b > 0$.
2) Show that if there exists $x \in \mathbb{R}$, such that $x - 2 \in A$, then there exists $y \in \mathbb{R}$, such that $3y \in A$.
3) Show that if $A \cap B \neq \varnothing$, then $A \neq \varnothing$.

4.4C. ∀-elimination. Perhaps Inspector Thinkright knows that Jeeves is a butler in the town, and that all of the butlers in the town are right handed. Well, then it is obvious to the Inspector that Jeeves is right handed. This is an example of ∀-elimination: if you know something is true about every element of a set, then it is true about any particular element of the set.

If $\forall x \in X, P(x)$ is true, and $a \in X$, then $P(a)$ is true.

EXAMPLE 4.4.7. Suppose

1) $C \subset \mathbb{R}$, and
2) $\forall x \in \mathbb{R}, \big((x^2 = 9) \Rightarrow (x \in C)\big)$.

Show $\exists c \in \mathbb{R}, c \in C$.

PROOF. Let $c = 3 \in \mathbb{R}$. Then $c^2 = 3^2 = 9$, and letting $x = c$ in Hypothesis (2) tells us that

$$(c^2 = 9) \Rightarrow (c \in C).$$

Therefore $c \in C$. □

EXAMPLE 4.4.8. Assume $A \neq \varnothing$ and $A \subset B$. Prove $\exists b, (b \in B)$.

PROOF. Because A is not the empty set, we know it has at least one element; that is, we have $\exists x, (x \in A)$. Hence, we may let a be some element of A. Now, let $b = a$. Because $A \subset B$, we know that every element of A is an element of B. In particular, since $b = a \in A$, this means that $b \in B$. □

WARNING. When applying ∀-elimination, the variable does not need to be called "x," (it could be y or z or any other variable), and the constant does not need to be called "a" (it could be any element of X). However, if the variable occurs more than once in the formula, it is important to replace *all* of its occurrences with a. For example, if $a \in X$, then, from $\forall x \in X, \big(A(x) \Rightarrow B(x)\big)$, we can conclude $A(a) \Rightarrow B(a)$, but *not* $A(a) \Rightarrow B(x)$ or $A(x) \Rightarrow B(a)$.

EXERCISES 4.4.9.

1) Assume $\forall x \in \mathbb{R}, (x^2 \in Z)$. Show $16 \in Z$.

2) Assume $A \subset B$ and $A \neq \varnothing$. Show $B \neq \varnothing$.

3) Assume

   (a) for every $x \in A$, either $x \in B$, or $x < 0$, and

   (b) there exists $a \in A$, such that $a > 0$.

   Show $B \neq \varnothing$.

4.4D. $\forall$-introduction. If Inspector Thinkright needs to verify that all of the butlers in town have seen the aurora borealis, he would probably get a list of all the butlers, and check them one-by-one. That is a valid approach, but it could be very time-consuming if the list is very long. In mathematics, such one-by-one checking is often not just time-consuming, but impossible. For example, the set $\mathbb{N}$ is infinite, so, if we wish to show $\forall n \in \mathbb{N}, (2n \text{ is even})$, then we would never finish if we tried to go through all of the natural numbers one-by-one. So we need to deal with many numbers at once.

Consider the following simple deduction:

Hypotheses:

Every butler in town got up before 6am today.

Everyone who got up before 6am today, saw the aurora.

Conclusion: Every butler in town saw the aurora.

This is clearly a valid deduction in English. Let us translate it into First-Order Logic to analyze how we were able to reach a conclusion about all of the butlers, without checking each of them individually. Here is a symbolization key:

B: The set of all of the butlers in town.

P: The set of all people.

$U(x)$: x got up before 6am today.

$S(x)$: x saw the aurora.

We can now translate our English deduction, as follows:

Hypotheses:

$\forall b \in B, U(b)$.

$\forall p \in P, \big(U(p) \Rightarrow S(p)\big)$.

Conclusion: $\forall b \in B, S(b)$.

How do we justify the conclusion? Well, suppose for a moment that we start to check every butler in town, and that j represents Jimmy, who is one of the butlers in town. Then our first hypothesis allows us to conclude $U(j)$. Since Jimmy is a person, our second hypothesis allows us to conclude that $U(j) \Rightarrow S(j)$. Then, using $\Rightarrow$-elimination, we conclude $S(j)$. But there was nothing special about our choice of Jimmy. All that we know about him, is that he is a butler in the town. So we could use exactly the same argument to deduce $S(b)$ for any butler b in the town.

This is how we justify a $\forall$-introduction. If we can prove that the desired conclusion is true for an *arbitrary* element of a set, when we assume *nothing* about the element except that it belongs to the set, then the conclusion must be true for every element of the set.

We write the above deduction as follows:

THEOREM 4.4.10. *Assume that every butler in town got up before 6am today. Also assume that everyone who got up before 6am today, saw the aurora. Then every butler in town saw the aurora.*

PROOF. Let b represent an arbitrary butler in town. Then, since all of the butlers got up before 6am, we know that b got up before 6am. By hypothesis, this implies that b saw the aurora. Since b is an arbitrary butler in town, we conclude that every butler in town saw the aurora. □

This reasoning leads to the $\forall$-introduction rule: in order to prove that *every* element of a set X has a certain property, it suffices to show that an *arbitrary* element of X has the desired property. For example, if we wish to prove $\forall b \in B, P(b)$, then our proof should start with the sentence "Let b be an arbitrary element of B." (However, this can be abbreviated to: "Given $b \in B$, ...") After this, our task will be to prove that $P(b)$ is true, without assuming anything about b other than it is an element of B.

> The proof of an assertion that begins "for all $x \in X$,"
> will usually begin with "Let x be an arbitrary element of X"
> (or, for short, "Given $x \in X$,").

WARNING. It is important not to assume anything about x other than that it is an element of X. If you choose x to be a particular element of X that has some special property, then your deduction will not be valid for *all* elements of the set.

EXAMPLE 4.4.11. Suppose we would like to justify the following deduction:

> All of the butlers in town dislike Jimmy, and Jimmy is a butler in town.
> Therefore, all of the butlers in town dislike themselves.

Then it suffices to show, for an arbitrary butler b, that b dislikes b. We might try the following proof:

PROOF ATTEMPT. Let b be Jimmy, who is a butler in town. Then, since all of butlers in town dislike Jimmy, we know that b dislikes Jimmy. Since Jimmy $= b$, this means b dislikes b, as desired. So every butler in town dislikes himself. □

This proof is certainly *not* valid, however. Letting $b =$ Jimmy does not make b an *arbitrary* butler; rather, it makes b a very special butler — the one that everybody dislikes. In this case, conclusions that are true about b are not necessarily true about the other butlers.

Another point that should be emphasized is that an *arbitrary* member of a set is not the same as a *random* member of a set. If we want to prove that all of the butlers have seen the aurora, it is not enough to choose a butler at random, ask if he saw the aurora and draw a conclusion about all of the butlers based on that single answer. It is only if we can determine through logical deduction that *no matter which* butler we choose, that person saw the aurora, that we can conclude that all of the butlers saw the aurora.

EXERCISE 4.4.12. Which of the 4 quantifier rules does each deduction illustrate?

1) Everyone who ate in the cafeteria yesterday is sick today. Oh, no! Susie ate in the cafeteria — she must be sick!
2) Susie ate in the cafeteria yesterday, so I'm sure that *somebody* ate in the cafeteria yesterday.
3) Without knowing which of the athletes it was, I figured out that he or she must have eaten in the cafeteria yesterday. The only way that can be true is if every athlete ate in the cafeteria yesterday.

4) Our food reporter says that some woman knocked over a big box of lima beans while she was eating in the cafeteria yesterday. Whoever she is, let's call her "Ms. Clumsy." So this morning's headline can be "Ms. Clumsy spilled the beans!"

4.5. Some proofs about sets

To find proofs in First-Order Logic, you can use all of the strategies that were helpful in Propositional Logic: working backwards, working forwards, changing what you are looking at, breaking the proof into cases, and proof by contradiction are all important. The introduction and elimination rules for quantifiers just add some new options when you are working forwards or working backwards. In particular:

- If you have $\exists x, \mathcal{A}(x)$, you will probably use $\exists$-elimination: assume $\mathcal{A}(c)$ for some letter c that is not already in use, and then derive a conclusion that does not contain c.
- If the desired conclusion is $\forall x \in X, \mathcal{A}(x)$, then your proof will almost certainly be based on $\forall$-introduction, so the first words of your proof will usually be "Given $x \in X$, ..."
- If you have $\forall x, \mathcal{A}(x)$, and it might be helpful to know $\mathcal{A}(c)$ (for some constant c), then you could use $\forall$-elimination.

For example, now that we have all the rules of First-Order Logic, we can prove the following important fact that was stated on Page 60.

EXAMPLE 4.5.1. Assume A and B are sets. We have $A = B$ if and only if $A \subset B$ and $B \subset A$.

PROOF. ($\Rightarrow$) Assume $A = B$. Every set is a subset of itself (see Remark 3.2.20), so we have

$$A = B \subset B \quad \text{and} \quad B = A \subset A,$$

as desired.

($\Leftarrow$) Assume $A \subset B$ and $B \subset A$. We wish to show $A = B$; in other words, we wish to show

$$\forall x, (x \in A \Leftrightarrow x \in B).$$

Let x be arbitrary.

($\Rightarrow$) Suppose $x \in A$. Since $A \subset B$, this implies $x \in B$.

($\Leftarrow$) Suppose $x \in B$. Since $B \subset A$, this implies $x \in A$.

Therefore, $x \in A \Leftrightarrow x \in B$. Since x is arbitrary, this implies $\forall x, (x \in A \Leftrightarrow x \in B)$, as desired. □

EXAMPLE 4.5.2. Suppose A, B, and C are sets (and $\mathcal{U}$ is the universal set, as usual). Then:

1) $A \cap B \subset A$.
2) $A \subset A \cup B$.
3) $A \cap \mathcal{U} = A$.
4) If $A \subset B$ and $B \subset C$, then $A \subset C$.
5) $A \cap B \subset A \cup B$.
6) If $A \subset B$ and $A \subset C$, then $A \subset B \cap C$.

PROOF. (1) We wish to show that every element of $A \cap B$ is an element of A. Given $x \in A \cap B$, we know, from the definition of $A \cap B$, that $x \in A$ and $x \in B$. In particular, $x \in A$, as desired.

(2) We wish to show that every element of A is an element of $A \cup B$. Given $x \in A$, it is obviously true that either $x \in A$ or $x \in B$ (since, in fact, we know $x \in A$). Therefore $x \in A \cup B$, as desired.

(3) From (1), we know that $A \cap \mathcal{U} \subset A$, so it suffices to show that $A \subset A \cap \mathcal{U}$. Given $a \in A$, we obviously have $a \in A$. Furthermore, since the universal set $\mathcal{U}$ contains every element that is under consideration, we also have $a \in \mathcal{U}$. Hence $a \in A \cap \mathcal{U}$, as desired.

(4) Let a be an arbitrary element of A. Since $A \subset B$, we know $a \in B$. Then, because $B \subset C$, we know $a \in C$. Since a is an arbitrary element of A, this implies $A \subset C$.

(5) Given $x \in A \cap B$, we know $x \in A$ and $x \in B$. In particular, we have $x \in A$, so it is certainly true that either $x \in A$ or $x \in B$. Therefore $x \in A \cup B$. Since x is an arbitrary element of $A \cap B$, this implies $A \cap B \subset A \cup B$, as desired.

Alternate proof of (5). From parts (1) and (2), we know $A \cap B \subset A$ and $A \subset A \cup B$. So (4) implies that $A \cap B \subset A \cup B$, as desired.

(6) Let a be an arbitrary element of A. Since $A \subset B$, we have $a \in B$. Similarly, since $A \subset C$, we also have $a \in C$. Having established that $a \in B$ and $a \in C$, we conclude that $a \in B \cap C$. Since a is an arbitrary element of A, this implies $A \subset B \cap C$. □

EXERCISES 4.5.3. Assume A, B, and C are sets.

1) Show that if $A \subset B$, then $A \cap B = A$.
2) Show if $A \subset B$, then $A \cup B = B$.
3) Show that if $B \subset C$, then $A \cap B \subset A \cap C$.
4) Show that if $A \subset C$ and $B \subset C$, then $A \cup B \subset C$.
5) Show $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$.
6) Show $A \smallsetminus B = A \smallsetminus (A \cap B)$.
7) Let $X = A \cap B$, and show $A \cup B = (A \smallsetminus X) \cup (B \smallsetminus X) \cup X$.
8) Show that if $\mathcal{P}(A \cup B) = \mathcal{P}(A) \cup \mathcal{P}(B)$, then either $A \subset B$ or $B \subset A$. [*Hint:* Prove the contrapositive. Choose $a \in A$ and $b \in B$, such that $a \notin B$ and $b \notin A$. Then $\{a, b\} \notin \mathcal{P}(A) \cup \mathcal{P}(B)$.]

EXERCISES 4.5.4. Suppose A and B are sets.

1) Show $A \smallsetminus B = A \cap \overline{B}$.
2) Show $A = (A \smallsetminus B) \cup (A \cap B)$.
3) Prove De Morgan's Laws:
   (a) $\overline{\overline{A}} = A$. (b) $\overline{A \cap B} = \overline{A} \cup \overline{B}$. (c) $\overline{A \cup B} = \overline{A} \cap \overline{B}$.
4) Show that if $\overline{A} = \overline{B}$, then $A = B$.
   [*Hint:* Follows immediately from one of De Morgan's Laws.]

EXERCISES 4.5.5. Suppose A, B, and C are sets.

1) Show that A is disjoint from B if and only if $A \subset \overline{B}$.
2) Show $A \smallsetminus B$ is disjoint from B.
3) Show that if A is disjoint from B, and C is a subset of B, then A is disjoint from C.
4) Show that $A \smallsetminus B$ is disjoint from $A \cap B$.
5) Show that A is disjoint from $B \cup C$ iff A is disjoint from both B and C.

EXERCISES 4.5.6.

1) Show $A \cup B = (A \smallsetminus B) \cup (B \smallsetminus A) \cup (A \cap B)$.
2) Show the three sets $A \smallsetminus B$, $B \smallsetminus A$, and $A \cap B$ are all disjoint from each other.

Recall from Remark 3.3.20(2) that sets $A_1, A_2, \ldots, A_n$ are **pairwise-disjoint** iff A_i is disjoint from A_j whenever $i \neq j$.

EXERCISE 4.5.7. Suppose the sets $A_1, A_2, \ldots, A_n$ are pairwise-disjoint. Show:

1) The sets $A_1, A_2, \ldots, A_{n-1}$ are pairwise-disjoint, if $n > 1$.
2) A_n is disjoint from $A_1 \cup A_2 \cup \cdots \cup A_{n-1}$, if $n > 1$.

4.6. Counterexamples (reprise)

We have been discussing proofs in this chapter, but you should keep in mind that counterexamples are also an important part of logic:

> To show that a deduction is valid, provide a proof.
>
> To show that a deduction is *not* valid, provide a counterexample.

EXAMPLE 4.6.1. Show that the following deduction is not valid:

$$\exists x, (x \in A), \qquad \therefore \ \forall x, (x \in A).$$

Scratchwork. To get the idea of what is going on, it may be helpful to translate the deduction into English. For example, we could use the symbolization key

$\mathcal{U}$: things on the kitchen table

A: apples on the kitchen table

In this setting, the deduction becomes:

There is an apple on the kitchen table. $\therefore$ Everything on the kitchen table is an apple.

This deduction is obviously not valid: it is easy to imagine a situation in which one thing on the kitchen table is an apple, but something else on the kitchen table is not an apple.

To find the official solution, we will do something analogous, but using the notation of First-Order Logic, instead of talking about apples and tabletops. In order to construct a counterexample, we want the hypothesis of the deduction to be true and the conclusion to be false.

- To make the hypothesis $\exists x, (x \in A)$ true, we need something to be an element of A. For example, we could let $1 \in A$.
- To make the the conclusion $\forall x, (x \in A)$ false, we want its negation to be true: we want $\exists x, (x \notin A)$ to be true. For example, we could arrange that $2 \notin A$.

To satisfy the above two conditions, we let $A = \{1\}$. Since 1 and 2 are the only elements mentioned in the discussion, we can let $\mathcal{U} = \{1, 2\}$. This results in the counterexample we were hoping to find.

SOLUTION. We provide a counterexample. Let

$$\mathcal{U} = \{1, 2\} \text{ and } A = \{1\}.$$

Then:

$1 \in A$ is true, so $\exists x, (x \in A)$ is true, so the hypothesis is true,

but

$2 \notin A$, so $\forall x, (x \in A)$ is false, so the conclusion is false.

Since we have a situation in which the hypothesis is true, but the conclusion is false, the deduction is not valid. □

EXAMPLE 4.6.2. Show that the following deduction is not valid:

Hypotheses:

1. $\forall x, \big((x \in A) \vee (x \in B)\big)$
2. $A \neq \varnothing$
3. $B \neq \varnothing$

Conclusion: $\exists x, \big((x \in A)\,\&\,(x \in B)\big)$.

Scratchwork. In order to construct a counterexample, we want all of the hypotheses of the deduction to be true and the *negation* of the conclusion to be true. The negation of the conclusion is

$$\forall x, \big((x \notin A) \vee (x \notin B)\big),$$

which is logically equivalent to

$$\forall x, \big((x \in A) \Rightarrow (x \notin B)\big). \tag{4.6.3}$$

Now:

- To make Hypothesis 2 true, we may let $1 \in A$.
- To make Hypothesis 3 true, we must put something in the set B. However, it is important to note that (4.6.3) tells us $1 \notin B$, so we must put something else into B. For example, we may let $2 \in B$.
- Now, after A and B have been constructed, we can make Hypothesis 1 true by letting $\mathcal{U} = A \cup B$.

To satisfy all three of these conditions, we may let $A = \{1\}$, $B = \{2\}$, and $\mathcal{U} = A \cup B = \{1, 2\}$.

SOLUTION. We provide a counterexample. Let

$$\mathcal{U} = \{1, 2\}, \quad A = \{1\}, \quad \text{and} \quad B = \{2\}.$$

Then:

1) We have
   - $1 \in A$ is true, so $(1 \in A) \vee (1 \in B)$ is true, and
   - $2 \in B$ is true, so $(2 \in A) \vee (2 \in B)$ is true.

   Since 1 and 2 are the only elements of $\mathcal{U}$, this implies, for every x, that $(x \in A) \vee (x \in B)$ is true. So Hypothesis 1 is true.
2) $1 \in A$, so $A \neq \varnothing$. Hence, Hypothesis 2 is true.
3) $2 \in B$, so $B \neq \varnothing$. Hence, Hypothesis 3 is true.

However:

- $1 \notin B$, so $(1 \in A)\,\&\,(1 \in B)$ is false, and
- $2 \notin A$, so $(2 \in A)\,\&\,(2 \in B)$ is false.

Since 1 and 2 are the only elements of $\mathcal{U}$, this implies there is no x for which the assertion $(x \in A)\,\&\,(x \in B)$ is true. Hence, the assertion $\exists x, \big((x \in A)\,\&\,(x \in B)\big)$ is false; in other words, the conclusion of the deduction is false. Since we have a situation in which the hypothesis is true, but the conclusion is false, the deduction is not valid. □

EXAMPLE 4.6.4. Explain how you know the following deduction is **not** valid:

$$X \cap Y \subset D, \quad \therefore X \subset Y \cup D.$$

SOLUTION. We provide a counterexample. Let $X = \{1\}$, $Y = \{2\}$, and $D = \{3\}$.

Then $X \cap Y = \{1\} \cap \{2\} = \varnothing$, so $X \cap Y \subset D$, because the empty set is a subset of every set. This means the hypothesis is true.

However, $Y \cup D = \{2\} \cup \{3\} = \{2, 3\}$. Therefore

$$1 \in \{1\} = X \text{ and } 1 \notin \{2, 3\} = Y \cup D, \text{ so } X \not\subset Y \cup D.$$

This means the conclusion is false.

Since we have a situation in which the hypothesis is true, but the conclusion is false, the deduction is not valid. □

EXERCISES 4.6.5. Explain how you know that each of the following deductions is not valid.

1) $\exists x, (x \in A),\ \exists x, (x \in B),\ \therefore\ \exists x, \big((x \in A)\ \&\ (x \in B)\big)$

2) $\forall a \in A, \exists b \in B, (a \neq b),\ A \neq \varnothing,\ \therefore\ \forall b \in B, \exists a \in A, (a \neq b).$

3) $A \neq B,\ \therefore\ A \cup B \neq A.$

4) $\forall x \in A, (x \notin B),\ \forall x \in B, (x \notin A),\ \therefore\ A \neq B.$

EXERCISES 4.6.6. Explain how you know that each of these deductions is *not* valid.

1) $A \cup B\ \subset\ E \cup F,\quad \therefore\ A \cap B\ \subset\ E \cap F.$

2) $A \subset B,\ X \subset Y,\qquad \therefore\ A \smallsetminus X\ \subset\ B \smallsetminus Y.$

3) $A \cap B \neq \varnothing,\ B \subset C,\qquad \therefore A \subset C.$

4) $\exists x, \big((x \in P)\ \&\ (x \notin Q)\big),\quad \therefore\ \forall x, \big((x \in P) \Rightarrow (x \notin Q)\big).$

5) $\forall x \in X, \big(x\ R\ x\big),\quad \therefore\ \forall x_1 \in X, \forall x_2 \in X, \big((x_1\ R\ x_2) \Rightarrow (x_2\ R\ x_1)\big)$

EXERCISES 4.6.7. Determine whether each of the following deductions is valid, and justify your answer by giving a proof or a counterexample.

1) $\exists u \in U, (u \notin V),\ \therefore\ \forall u \in U, (u \notin V).$

2) $\forall x, \big((x \in S) \Rightarrow (1 \in T)\big),\quad S \neq \varnothing,\quad \therefore\ 1 \in T.$

3) $\forall a \in A, (a \in B),\quad \forall b \in B, (b \in C),\quad \therefore\ \forall a \in A, (a \in C).$

4) $D \cup E \neq \varnothing,\quad D \subset F,\quad \therefore\ D \cap F \neq \varnothing.$

5) $\forall a_1 \in A, \forall a_2 \in A, \big((a_1\ R\ a_2) \vee (a_2\ R\ a_1)\big),\quad 2 \in A,\quad \therefore\ 2\ R\ 2.$

SUMMARY:

- First-Order Logic includes all of Propositional Logic, plus the quantifiers $\forall$ and $\exists$.
- Translating between English and First-Order Logic.
- The equals sign ($=$) is automatically included in every symbolization key.
- The order of the quantifiers is important, because it can change the meaning of an assertion.
- Uniqueness ($\exists!$)
- Every variable in an assertion must be bound by a quantifier.
- Rules for negating quantifiers:
  - the negation of a "$\forall$" assertion is a "$\exists$" assertion;
  - the negation of a "$\exists$" assertion is a "$\forall$" assertion;
- Any assertion about all elements of $\varnothing$ is "vacuously" true.
- Introduction and elimination rules for quantifiers.
- Just as in Propositional Logic:
  - To show that a deduction is valid, provide a proof.
  - To show that a deduction is *not* valid, provide a counterexample.
- Notation:
  - $\forall x$ (universal quantifier; means "For all x")
  - $\forall x \in X$ (universal quantifier; means "For all x in X")
  - $\exists x$ (existential quantifier; means "There exists some x, such that…")
  - $\exists x \in X$ (existential quantifier; means "There exists some x in X, such that…")
  - $\exists!\, x$ (means "There is a unique x, such that…")
  - $\exists!\, x \in X$ (means "There is a unique x in X, such that…")

Chapter 5

Sample Topics

If people do not believe that mathematics is simple,
it is only because they do not realize how complicated life is.
John von Neumann (1903–1957), Hungarian-American mathematician

This chapter provides exercises from three different mathematical topics (Number Theory, Abstract Algebra, and Real Analysis) that will test your proof-writing skills. To succeed in advanced math classes, you will need to be able to solve problems like these.

TERMINOLOGY 5.0.1. Up to this point, our valid deductions have been called "theorems," but mathematicians usually reserve this name for the ones that are particularly important, and apply some other name to the others. The terminology allows some flexibility, but here are general guidelines:

- Any valid deduction can be referred to as a "result."
- A **theorem** is an important result.
- A **proposition** is a result that is not sufficiently important to be called a theorem.
- A **corollary** is a result that is proved as an easy consequence of some other result.
- A **lemma** is a minor result that is not interesting for its own sake, but will be used as part of the proof of theorem (or other more significant result).

5.1. Number Theory: divisibility and congruence

In this section, we will get some practice with proving properties of integers.

5.1A. Divisibility. Every math student knows that some numbers are even and some numbers are odd; some numbers are divisible by 3, and some are not; etc. Let us introduce a notation that makes it easy to talk about whether or not one number b is divisible by some other number a:

DEFINITION 5.1.1. Suppose $a, b \in \mathbb{Z}$. We say a is a **divisor** of b (and write "$a \mid b$") iff there exists $k \in \mathbb{Z}$, such that $ak = b$. (Since multiplication is commutative and equality is symmetric, this equation can also be written as $b = ka$.)

NOTATION 5.1.2. $a \nmid b$ is an abbreviation for "a does *not* divide b."

Remark 5.1.3. When a is a divisor of b, we may also say:

1) a **divides** b, or
2) a is a **factor** of b, or
3) b is a **multiple** of a, or
4) b is **divisible** by a.

EXAMPLE 5.1.4.

1) We have $5 \mid 30$, because $5 \cdot 6 = 30$, and $6 \in \mathbb{Z}$.

2) We have $5 \nmid 27$, because there is no integer k, such that $5k = 27$.

EXERCISE 5.1.5. Fill each blank with $\mid$ or $\nmid$, as appropriate.

1) 3_____18	2) 4_____18	3) 5_____18
4) 6_____18	5) 18_____6	6) -6_____6

The following definition is perhaps the best known example of divisibility.

DEFINITION 5.1.6. Let $n \in \mathbb{Z}$. We say n is **even** iff $2 \mid n$. We say n is **odd** iff $2 \nmid n$.

Here are some examples of proofs involving divisibility. We will assume the well-known fact that the sum, difference, and product of integers are integers: for all $k_1, k_2 \in \mathbb{Z}$, we know that $k_1 + k_2 \in \mathbb{Z}$, $k_1 - k_2 \in \mathbb{Z}$, and $k_1 k_2 \in \mathbb{Z}$. Also, the negative of any integer is an integer: for all $k \in \mathbb{Z}$, we have $-k \in \mathbb{Z}$.

Our first result is a generalization of the well-known fact that the sum of two even numbers is even.

PROPOSITION 5.1.7. *Suppose $a, b_1, b_2 \in \mathbb{Z}$. If $a \mid b_1$ and $a \mid b_2$, then $a \mid (b_1 + b_2)$.*

Scratchwork. Since $a \mid b_1$ and $a \mid b_2$, we know there is some $k \in \mathbb{Z}$, such that $ak = b_1$, and we know there is some $k \in \mathbb{Z}$, such that $ak = b_2$. However, these are probably two different values of k, so we need to give them different names if we want to talk about both of them at the same time. So let's call the first one k_1 and the second one k_2:

- there exists $k_1 \in \mathbb{Z}$, such that $ak_1 = b_1$, and
- there exists $k_2 \in \mathbb{Z}$, such that $ak_2 = b_2$.

To show $a \mid (b_1 + b_2)$, we want to find some $k \in \mathbb{Z}$, such that

$$ak \stackrel{?}{=} b_1 + b_2.$$

Since $b_1 + b_2 = ak_1 + ak_2 = a(k_1 + k_2)$, this means we want

$$ak \stackrel{?}{=} a(k_1 + k_2).$$

So it is clear that we should let $k = k_1 + k_2$.

PROOF. Since, by assumption, a is a divisor of both b_1 and b_2, there exist $k_1, k_2 \in \mathbb{Z}$, such that $ak_1 = b_1$ and $ak_2 = b_2$. Let $k = k_1 + k_2$. Then $k \in \mathbb{Z}$ and

$$ak = a(k_1 + k_2) = ak_1 + ak_2 = b_1 + b_2,$$

so a is a divisor of $b_1 + b_2$, as desired. □

PROPOSITION 5.1.8. *Suppose $a, b \in \mathbb{Z}$. We have $a \mid b$ iff $a \mid -b$.*

Scratchwork. ($\Rightarrow$) Since $a \mid b$, we know there is some k, such that $ak = b$. To show $a \mid -b$, we want to find some other k — let's call it k' — such that $ak' \stackrel{?}{=} -b$. Since $-b = -(ak) = a(-k)$, this means we want $ak' \stackrel{?}{=} a(-k)$. So we should let $k' = -k$.

($\Leftarrow$) Since $a \mid -b$, we know there is some k, such that $ak = -b$. To show $a \mid b$, we want to find some k', such that $ak' \stackrel{?}{=} b$. Since $-b = ak$, we have $b = -(ak) = a(-k)$, so we can let $k' = -k$. This is the same work as in the proof of ($\Rightarrow$), and the official proof given below avoids the need for repeating this algebra, by appealing to the result that we have already proved.

PROOF. ($\Rightarrow$) By assumption, there is some $k \in \mathbb{Z}$, such that $ak = b$. Then $-k \in \mathbb{Z}$, and we have $a(-k) = -ak = -b$, Therefore, a divides $-b$.

($\Leftarrow$) Assume $a \mid -b$. From the preceding paragraph, we conclude that $a \mid -(-b) = b$, as desired. □

EXERCISES 5.1.9. Assume $a, a', b, b' \in \mathbb{Z}$.

1) Show that if $a \mid b$ and $a \mid b'$, then $a \mid b - b'$.
2) Show that $a \mid b$ iff $-a \mid b$.
3) Show $1 \mid b$.
4) Show $a \mid 0$.
5) Show that if $0 \mid b$, then $b = 0$.
6) Show that if $a \mid b$, then $a \mid bb'$.
7) Show that if $a \mid b$ and $a' \mid b'$, then $aa' \mid bb'$.

PROPOSITION 5.1.10. *Suppose $a, b_1, b_2 \in \mathbb{Z}$. If $a \mid b_1$ and $a \nmid b_2$, then $a \nmid (b_1 + b_2)$.*

PROOF. Assume $a \mid b_1$ and $a \nmid b_2$.

Suppose $a \mid (b_1 + b_2)$. (This will lead to a contradiction.) Then a is a divisor of both $b_1 + b_2$ and (by assumption) b_1. So Exercise 5.1.9(1) tells us

$$a \mid \big((b_1 + b_2) - b_1\big) = b_2.$$

This contradicts the assumption that $a \nmid b_2$.

Because it leads to a contradiction, our hypothesis that $a \mid (b_1 + b_2)$ must be false. This means $a \nmid (b_1 + b_2)$. □

It is well known that 1 and -1 are the only integers whose reciprocal is also an integer. In the language of divisibility, this can be restated as the following useful fact:

For any integer n, we have $n \mid 1$ iff $n = \pm 1$.

EXERCISE 5.1.11. Prove:

1) For all $a \in \mathbb{Z}$, we have $a \mid a$.
2) For all $a, b, c \in \mathbb{Z}$, if $a \mid b$ and $b \mid c$, then $a \mid c$.
3) It is *not* true that, for all $a, b \in \mathbb{Z}$, we have $a \mid b$ iff $b \mid a$.

Remark 5.1.12. The three parts of the preceding exercise show that divisibility is "reflexive" and "transitive," but not "symmetric." These terms will be defined in Definition 7.1.9.

EXERCISES 5.1.13. Assume $a, b, x, y \in \mathbb{Z}$. Prove:

1) If $a \mid x$ and $b \mid y$, then $ab \mid 2xy$.
2) If $a, b \in \mathbb{N}^+$, and $a \mid b$, then $a \leq b$.

EXERCISES 5.1.14. Prove or disprove each assertion.

1) For all $a, b_1, b_2 \in \mathbb{Z}$, if $a \nmid b_1$ and $a \nmid b_2$, then $a \nmid (b_1 + b_2)$.
2) For all $a, b_1, b_2 \in \mathbb{Z}$, if $a \nmid b_1$ and $a \nmid b_2$, then $a \nmid b_1 b_2$.
3) For all $a, b, c \in \mathbb{Z}$, if $a \nmid b$ and $b \nmid c$, then $a \nmid c$.
4) For all $a, b \in \mathbb{Z}$, if $a \nmid b$, then $a \nmid -b$.

5.1B. Congruence modulo n.

DEFINITION 5.1.15. Suppose $a, b, n \in \mathbb{Z}$. We say a is **congruent to b modulo n** iff $a - b$ is divisible by n. The notation for this is: $a \equiv b \pmod{n}$.

EXAMPLE 5.1.16.

1) We have $22 \equiv 0 \pmod{2}$, because $22 - 0 = 22 = 11 \times 2$ is a multiple of 2. (More generally, for $a \in \mathbb{Z}$, one can show that $a \equiv 0 \pmod{2}$ iff a is even.)
2) We have $15 \equiv 1 \pmod{2}$, because $15 - 1 = 14 = 7 \times 2$ is a multiple of 2. (More generally, for $a \in \mathbb{Z}$, one can show that $a \equiv 1 \pmod{2}$ iff a is odd.)
3) We have $28 \equiv 13 \pmod{5}$, because $28 - 13 = 15 = 3 \times 5$ is a multiple of 5.
4) For any $a, n \in \mathbb{Z}$, it is not difficult to see that $a \equiv 0 \pmod{n}$ iff a is a multiple of n.

EXERCISES 5.1.17. Fill each blank with $\equiv$ or $\not\equiv$, as appropriate.

1) 14_____5 (mod 2) 2) 14_____5 (mod 3) 3) 14_____5 (mod 4)
4) 14_____32 (mod 2) 5) 14_____32 (mod 3) 6) 14_____32 (mod 4)

EXERCISES 5.1.18. ("Congruence modulo n is reflexive, symmetric, and transitive.")
Let $n \in \mathbb{Z}$. Prove:

1) For all $a \in \mathbb{Z}$, we have $a \equiv a \pmod{n}$.
2) For all $a, b \in \mathbb{Z}$, we have $a \equiv b \pmod{n}$ iff $b \equiv a \pmod{n}$.
3) For all $a, b, c \in \mathbb{Z}$, if $a \equiv b \pmod{n}$ and $b \equiv c \pmod{n}$, then $a \equiv c \pmod{n}$.

EXERCISES 5.1.19. Assume $a_1 \equiv a_2 \pmod{n}$ and $b_1 \equiv b_2 \pmod{n}$. Show:

1) $a_1 + b_1 \equiv a_2 + b_2 \pmod{n}$.
2) $a_1 - b_1 \equiv a_2 - b_2 \pmod{n}$.
3) $a_1 b_1 \equiv a_2 b_2 \pmod{n}$. [*Hint:* $a_1 b_1 - a_2 b_2 = a_1(b_1 - b_2) + (a_1 - a_2)b_2$.]

Children are taught that if a number a is divided by a number n, then there may be a remainder, but the remainder is always smaller than n. That idea is said more precisely in the following theorem:

THEOREM 5.1.20 (Division Algorithm). *Suppose $a, n \in \mathbb{Z}$, and $n \neq 0$. Then there exist unique integers q and r in $\mathbb{Z}$, such that:*

1) $a = qn + r$, *and*
2) $0 \leq r < |n|$.

DEFINITION 5.1.21. In the situation of Theorem 5.1.20, the number r is called the **remainder** when a is divided by n.

The following exercise reveals the close relationship between congruence and remainders.

EXERCISE 5.1.22. Suppose $a, b, n \in \mathbb{Z}$ (and $n \neq 0$).

1) Let r be the remainder when a is divided by n. Show $a \equiv r \pmod{n}$.
2) Show that $a \equiv b \pmod{n}$ iff a and b have the same remainder when divided by n.

Remark 5.1.23. From the second half of parts (1) and (2) of Example 5.1.16, we see that every integer is congruent to either 0 or 1 modulo 2 (and not both).

$$\boxed{\begin{array}{c} n \text{ is even iff } n \equiv 0 \pmod 2. \\ n \text{ is odd iff } n \equiv 1 \pmod 2. \end{array}}$$

Exercise 5.1.22(1) generalizes this to congruence modulo numbers other than 2: if $n \in \mathbb{N}^+$, then every integer is congruent (modulo n) to some number in $\{0, 1, 2, \ldots, n-1\}$.

EXAMPLE 5.1.24. Let us show that if n is odd, then $9n+6$ is also odd. To see this, note that:

- $9 \equiv 1 \pmod 2$, because $9 = 4(2) + 1$,
- $n \equiv 1 \pmod 2$, because n is odd, and
- $6 \equiv 0 \pmod 2$, because $6 = 3(2) + 0$.

Therefore, using Exercise 5.1.19, we have

$$9n + 6 \equiv (1)(1) + 0 \equiv 1 \pmod 2.$$

The same method can be applied in the following exercises:

EXERCISES 5.1.25. Let $n \in \mathbb{Z}$.

1) Show $6n + 3$ is odd.
2) Show that if n is even, then $5n + 3$ is odd.
3) Show that if n is odd, then $5n + 3$ is even.

PROPOSITION 5.1.26. *Let $n \in \mathbb{Z}$. Then $n^2 + n$ is even.*

PROOF. From Remark 5.1.23, we know that n is congruent to either 0 or 1 modulo 2. We consider these two possibilities as separate cases.

Case 1. Assume $n \equiv 0 \pmod 2$. By the assumption of this case, we have $n = 2q$, for some $q \in \mathbb{Z}$. Therefore

$$n^2 + n = (2q)^2 + 2q = 4q^2 + 2q = 2(2q^2 + q)$$

is divisible by 2.

Case 2. Assume $n \equiv 1 \pmod 2$. By the assumption of this case, we have $n = 2q + 1$, for some $q \in \mathbb{Z}$. Therefore

$$n^2 + n = (2q+1)^2 + (2q+1) = (4q^2 + 4q + 1) + (2q + 1) = 4q^2 + 6q + 2 = 2(2q^2 + 3q + 1)$$

is divisible by 2. □

EXERCISES 5.1.27. Let $n \in \mathbb{Z}$.

1) Show that if n is even, then $n^2 \equiv 0 \pmod 4$. [*Hint:* We have $n = 2q$, for some $q \in \mathbb{Z}$.]
2) Show that if n is odd, then $n^2 \equiv 1 \pmod 8$. [*Hint:* We have $n = 2q + 1$, for some $q \in \mathbb{Z}$.]
3) Show that if n^2 is even, then n is even.

5.1C. Examples of irrational numbers. Every rational number is a real number (in other words, $\mathbb{Q} \subset \mathbb{R}$), but it is perhaps not so obvious that some real numbers are not rational (in other words, $\mathbb{R} \not\subset \mathbb{Q}$). Such numbers are said to be **irrational**. We now give one of the simplest examples of an irrational number.

PROPOSITION 5.1.28. *$\sqrt{2}$ is irrational.*

PROOF BY CONTRADICTION. Suppose $\sqrt{2}$ is rational. (This will lead to a contradiction.) By definition, this means $\sqrt{2} = a/b$ for some $a, b \in \mathbb{Z}$, with $b \neq 0$. By reducing to lowest terms, we may assume that a and b have no common factors. In particular,

$$\text{it is not the case that both } a \text{ and } b \text{ are even.}$$

We have

$$\frac{a^2}{b^2} = \left(\frac{a}{b}\right)^2 = \sqrt{2}^2 = 2,$$

so $a^2 = 2b^2$ is even. Then Exercise 5.1.27(3) tells us that

$$a \text{ is even,}$$

so we have $a = 2k$, for some $k \in \mathbb{Z}$. Then

$$2b^2 = a^2 = (2k)^2 = 4k^2,$$

so $b^2 = 2k^2$ is even. Then Exercise 5.1.27(3) tells us that

$$b \text{ is even.}$$

We have now shown that a and b are even, but this contradicts the fact, mentioned above, that it is not the case that both a and b are even. □

EXERCISES 5.1.29.

1) For $n \in \mathbb{Z}$, show that if $3 \nmid n$, then $n^2 \equiv 1 \pmod 3$.

2) Show that $\sqrt{3}$ is irrational.

3) Is $\sqrt{4}$ irrational?

Remark 5.1.30. The famous numbers $\pi = 3.14159\ldots$ and $e = 2.71828\ldots$ are also irrational, but these examples are much harder to prove than Proposition 5.1.28.

5.2. Abstract Algebra: commutative groups

Every schoolchild learns about addition ($+$), subtraction ($-$), and multiplication ($\times$). Each of these is a "binary operation" on the set of real numbers, which means that it takes two numbers, and gives back some other number. In this section, we discuss binary operations on an arbitrary set; that is, we consider various ways of taking two elements of the set and giving back some other element of the set. (The official definition of the term "binary operation" is in Example 6.3.6, but it suffices to have an informal understanding for the present purposes.)

DEFINITION 5.2.1 (unofficial). Let A be a set. We say that $+$ is a **binary operation** on A if, for every $a, b \in A$, we have a corresponding element $a + b$ of A.

(The element $a + b$ must exist for *all* $a, b \in A$. Furthermore, the sum $a + b$ must depend only on the values of a and b, not on any other information.)

EXAMPLE 5.2.2.

1) Addition ($+$), subtraction ($-$), and multiplication ($\times$) are examples of binary operations on $\mathbb{R}$. They also provide binary operations on $\mathbb{Q}$ and $\mathbb{Z}$. However, subtraction ($-$) does *not* provide a binary operation on $\mathbb{N}$, because $x - y$ is not in $\mathbb{N}$ when $x < y$ (whereas the values of a binary operation on a set must all belong to the given set).

2) Division ($\div$) is *not* a binary operation on $\mathbb{R}$. This is because $x \div y$ does not exist when $y = 0$ (whereas a binary operation on a set needs to be defined for *all* pairs of elements of the set). (On the other hand, division is a binary operation on the set $\mathbb{R} \smallsetminus \{0\}$ of all nonzero real numbers.)

3) Union ($\cup$), intersection ($\cap$), and set difference ($\smallsetminus$) are binary operations on the collection of all sets.

4) If a set does not have too many elements, then a binary operation on it can be specified by providing an "addition table." For example, the following table defines a binary operation $+$ on $\{\mathsf{a}, \mathsf{b}, \mathsf{c}, \mathsf{d}, \mathsf{e}, \mathsf{f}\}$:

$+$	a	b	c	d	e	f
a	a	b	c	d	e	f
b	b	c	a	e	f	d
c	c	a	b	f	d	e
d	d	e	f	a	b	c
e	e	f	d	b	c	a
f	f	d	e	c	a	b

To calculate $x + y$, find the row that has x at its left end, and find the column that has y at the top. The value $x + y$ is the entry of the table that is in that row and that column. For example, f is at the left of the bottom row and e is at the top of the second-to-last column, so $\mathsf{f} + \mathsf{e} = \mathsf{a}$, because a is the second-to-last entry of the bottom row.

DEFINITION 5.2.3. Let $+$ be a binary operation on a set G.

1) $+$ is **commutative** iff $g + h = h + g$ for all $g, h \in G$.

2) $+$ is **associative** iff $g + (h + k) = (g + h) + k$ for all $g, h, k \in G$.

3) An element 0 of G is an **identity element** iff $g + 0 = g$, for all $g \in G$.

4) For $g \in G$, a **negative** of g is an element $-g$ of G, such that $g + (-g) = 0$, where 0 is an identity element of G.

5) We say $(G, +)$ is a **commutative group** iff all three of the following conditions (or "axioms") are satisfied:

- $+$ is commutative and associative,
- there is an identity element, and
- every element of G has a negative.

OTHER TERMINOLOGY. For historical reasons, most mathematicians use the term "abelian group," rather than "commutative group," but they would agree that "commutative group" is also acceptable.

EXAMPLES 5.2.4.

1) $(\mathbb{R}, +)$ is a commutative group: we all learned in elementary school that addition is commutative and associative, that $g + 0 = g$, and that $g + (-g) = 0$. (The same is true for $(\mathbb{Q}, +)$ and $(\mathbb{Z}, +)$.)

2) $(\mathbb{N}, +)$ is *not* a commutative group (even though addition is commutative and associative, and 0 is an identity element), because no nonzero element has a negative in $\mathbb{N}$.

3) $(\mathbb{R}, -)$ is *not* a commutative group, because subtraction is not commutative: $g - h$ is usually not equal to $h - g$. (Another reason $(\mathbb{R}, -)$ is not a commutative group is that subtraction is not associative: $(g - h) - k$ is usually not equal to $g - (h - k)$.)

4) $(\mathbb{R}, \times)$ is *not* a commutative group (even though addition is commutative and associative, and 1 is an identity element for multiplication), because 0 does not have a "negative" or "multiplicative inverse": there does not exist $h \in \mathbb{R}$, such that $0 \times h = 1$.

OTHER TERMINOLOGY. Our condition to be an identity element is what is usually called a "right identity element." For 0 to be an identity element, it should be true not only that $g + 0 = g$, but also that $0 + g = g$. However, the only binary operations of interest to us are commutative groups, where the second condition is also satisfied (see Exercise 5.2.11(1) below), so there is no need for us to make this distinction.

EXERCISE 5.2.5. For the binary operation on $\{\mathsf{a}, \mathsf{b}, \mathsf{c}, \mathsf{d}, \mathsf{e}, \mathsf{f}\}$ in Example 5.2.2(4), verify that:

1) a is an identity element, and
2) every element has a negative, namely: $-\mathsf{a} = \mathsf{a}$, $-\mathsf{b} = \mathsf{c}$, $-\mathsf{c} = \mathsf{b}$, $-\mathsf{d} = \mathsf{d}$, $-\mathsf{e} = \mathsf{f}$, $-\mathsf{f} = \mathsf{e}$.

The following result tells us that, instead of *an* identity element of a commutative group, we may speak of *the* identity element.

PROPOSITION 5.2.6. *The identity element of any commutative group is unique.*

PROOF. Suppose 0 and θ are any identity elements of a commutative group $(G, +)$. Then

$$\begin{aligned} 0 &= 0 + \theta && (\theta \text{ is an identity element}) \\ &= \theta + 0 && (+ \text{ is commutative}) \\ &= \theta && (0 \text{ is an identity element}). \end{aligned}$$

Since 0 and θ are arbitrary identity elements of $(G, +)$, this implies that all identity elements are equal to each other, so the identity element is unique (there is only one of them). □

NOTATION 5.2.7. In this section, the symbol 0 will always represent the identity element of whatever commutative group is under consideration.

Warning: This means that 0 will usually *not* represent the number zero. For example, in Example 5.2.2(4), we have $0 = \mathsf{a}$.

Similarly, instead of *a* negative, we may speak of *the* negative of an element of G:

PROPOSITION 5.2.8. *Let $(G, +)$ be a commutative group. For each $g \in G$, the negative of g is unique.*

PROOF. Suppose $-g$ and h are any negatives of g. Then

$$\begin{aligned} -g &= -g + 0 && (0 \text{ is the identity element}) \\ &= -g + (g + h) && (h \text{ is a negative of } g) \\ &= (-g + g) + h && (+ \text{ is associative}) \\ &= h + \big(g + (-g)\big) && (+ \text{ is commutative}) \\ &= h + 0 && (-g \text{ is a negative of } g) \\ &= h && (0 \text{ is the identity element}). \end{aligned}$$

Therefore, all negatives of g are equal, so the negative is unique. □

PROPOSITION 5.2.9. *For every commutative group $(G, +)$, we have $-0 = 0$.*

PROOF. Let $g = 0$. Then g is the identity element, so $0 + g = 0$. By definition of the negative, this means that $g = -0$. Since $g = 0$, we conclude that $0 = -0$. □

NOTATION 5.2.10. Assume $(G, +)$ is a commutative group. For $g, h \in G$, we use $g - h$ as an abbreviation for $g + (-h)$.

EXERCISES 5.2.11. Assume $(G,+)$ is a commutative group, and $g,h \in G$.
Carefully justify each step of your proofs, using only the axioms stated in Definition 5.2.3. Do not assume any other properties of addition that you were taught in school.

1) Show $0+g=g$.
2) Show $(-g)+g=0$.
3) Show $g-g=0$.
4) Show $-(-g)=g$.
5) Show $(-g)+h=h-g$.
6) Show $(g-h)+h=g$.

The associative law tells us that $(g+h)+k=g+(h+k)$. Therefore, we learn in elementary school to simply write $g+h+k$, because it does not matter where the parentheses go. Officially, the associative law In fact (as we also learn in elementary school), there is no need to include parentheses in any sum (even if it has more than three terms). Also, the commutative law tells us that $g+h=h+g$. We learn in elementary school that this allows us to rearrange the terms in a sum of any length, and the same is true for commutative groups. For example:

$$g_1+g_2+g_3+g_4+g_5=g_4+g_3+g_1+g_5+g_2.$$

Here is an official statement of these observations:

PROPOSITION 5.2.12. *If $(G,+)$ is a commutative group, $n \in \mathbb{N}^+$, and $g_1,g_2,\ldots,g_n \in G$, then:*

1) ($+$ *is* **associative**) *The expression $g_1+g_2+\cdots+g_n$ represents a well-defined element of G, which does not depend on how the expression is parenthesized.*
2) ($+$ *is* **commutative**) *If $h_1,h_2,\ldots,h_n$ is a list of the same elements of G, but perhaps in a different order, then*

$$h_1+h_2+\cdots+h_n=g_1+g_2+\cdots+g_n.$$

EXAMPLE 5.2.13. Assume $(G,+)$ is a commutative group, and $g,h \in G$. Show

$$-(g+h)=(-g)+(-h).$$

PROOF. We have

$$\begin{aligned}
(g+h)+\big((-g)+(-h)\big) &= g+h+(-g)+(-h) && (+ \text{ is associative})\\
&= g+(-g)+h+(-h) && (+ \text{ is commutative})\\
&= \big(g+(-g)\big)+\big(h+(-h)\big) && (+ \text{ is associative})\\
&= 0+0 && (\text{definition of } -g \text{ and } -h)\\
&= 0 && (0 \text{ is the identity element}).
\end{aligned}$$

So $(-g)+(-h)$ is the negative of $g+h$. In other words, $(-g)+(-h)=-(g+h)$. □

EXERCISES 5.2.14. Assume $(G,+)$ is a commutative group, and $g,h,a \in G$.

1) Show $-(g-h)=h-g$.
2) Show that if $g+a=h+a$, then $g=h$.

DEFINITION 5.2.15. Let $(G,+)$ be a commutative group. A subset H of G is a **subgroup** of $(G,+)$ iff

1) $H \neq \varnothing$,
2) (**closed under negatives**) $-h \in H$, for all $h \in H$, and
3) (**closed under addition**) $h_1+h_2 \in H$, for all $h_1,h_2 \in H$.

EXAMPLE 5.2.16. $\mathbb{Z}$ and $\mathbb{Q}$ are subgroups of $(\mathbb{R}, +)$, but $\mathbb{N}$ is **not** a subgroup (because it is not closed under negatives).

PROPOSITION 5.2.17. *If H is a subgroup of a commutative group $(G, +)$, then $0 \in H$ (where, as usual, 0 is the identity element of $(G, +)$).*

PROOF. We know that $H \neq \varnothing$ (from the definition of subgroup), so there is some $h \in H$. Since H is closed under negatives (because it is a subgroup) this implies $-h \in H$. Then, since H is closed under addition (because it is a subgroup), we have $h + (-h) \in H$. Since $h + (-h) = 0$ (by the definition of $-h$), this means $0 \in H$. □

EXERCISES 5.2.18. Assume H is a subgroup of a commutative group $(G, +)$, and $h, k \in H$.

1) Show $h - k \in H$.

2) For all $a \in G$, show that if $a \notin H$, then $a + h \notin H$.

EXERCISES 5.2.19. Assume H is a subgroup of a commutative group $(G, +)$, and $a, b \in G$. Let $a + H = \{ a + h \mid h \in H \}$ and $b + H = \{ b + h \mid h \in H \}$.

1) Show that if $a + H$ is a subgroup of G, then $a \in H$. [*Hint:* Every subgroup contains 0.]

2) Show that if $a \in H$, then $a + H = H$.

3) Show that $a + H = b + H$ iff $a - b \in H$.

4) Show that if $(a + H) \cap (b + H) \neq \varnothing$, then $a + H = b + H$.

PROPOSITION 5.2.20. *Assume $(G, +)$ is a commutative group, and let $T = \{t \in G \mid t+t = 0\}$. Then T is a subgroup of G.*

PROOF. It suffices to show that T is: nonempty, closed under negatives, and closed under addition.

(nonempty) We have $0 \in T$ (because it is immediate that $0 + 0 = 0$), so $T \neq \varnothing$.

(closed under negatives) Given $t \in T$, we have

$$\begin{aligned} (-t) + (-t) &= -(t + t) && \text{(Example 5.2.13)} \\ &= -0 && (t \in T) \\ &= 0 && \text{(Proposition 5.2.9).} \end{aligned}$$

This means $-t \in T$, so T is closed under negatives.

(closed under addition) Given $s, t \in T$, we have $s + s = 0$ and $t + t = 0$. Therefore

$$\begin{aligned} (s + t) + (s + t) &= (s + t) + (t + s) && (+ \text{ is commutative}) \\ &= s + \big(t + (t + s)\big) && (+ \text{ is associative}) \\ &= s + \big((t + t) + s\big) && (+ \text{ is associative}) \\ &= s + (0 + s) && (t \in T) \\ &= s + s && (0 \text{ is the identity element}) \\ &= 0 && (s \in T), \end{aligned}$$

so $s + t \in T$. Therefore T is closed under addition. □

EXERCISES 5.2.21. Assume H and K are subgroups of a commutative group $(G, +)$.

1) Show that $H \cap K$ is a subgroup of $(G, +)$.

2) Let $H + K = \{ h + k \mid h \in H, k \in K \}$. Show that $H + K$ is a subgroup of $(G, +)$.

5.3. Real Analysis: convergent sequences

NOTATION 5.3.1. For $x \in \mathbb{R}$, $|x|$ denotes the **absolute value** of x:

$$|x| = \begin{cases} x & \text{if } x \geq 0, \\ -x & \text{if } x < 0. \end{cases}$$

You may assume the following basic properties of absolute value (without proof):

LEMMA 5.3.2. *For $x, y, z \in \mathbb{R}$, we have:*

1) $|x| \geq 0$ *(and $|x| = 0 \Leftrightarrow x = 0$).*
2) $|x| = |-x|$.
3) $|x+y| \leq |x|+|y|$. ("**triangle inequality**")
4) $|xy| = |x| \cdot |y|$.
5) $-|x| \leq x \leq |x|$.
6) $\exists N \in \mathbb{N}$, $N > |x|$.
7) *If $|x| < |y|$ and $z \neq 0$, then $|xz| < |yz|$.*
8) *If $|x| > |y| \neq 0$, then $1/|x| < 1/|y|$.*

DEFINITION 5.3.3. Assume $a_1, a_2, a_3, \ldots$ is an infinite sequence of real numbers, and $L \in \mathbb{R}$. We say that the sequence **converges** to L (and write $a_n \to L$) iff

$$\forall \epsilon > 0,\ \exists N \in \mathbb{N},\ \forall n > N,\ |a_n - L| < \epsilon.$$

OTHER TERMINOLOGY. When $a_n \to L$, we can also say that the **limit** of the sequence is L.

EXAMPLE 5.3.4. Let $t \in \mathbb{R}$. If $a_n = t$ for all n, then $a_n \to t$.

PROOF. Given $\epsilon > 0$, let $N = 0$. Given $n > N$, we have $|a_n - t| = |t - t| = |0| = 0 < \epsilon$. □

EXAMPLE 5.3.5. If $a_n = 1/n$ for all n, then $a_n \to 0$.

Scratchwork. To prove $a_n \to 0$, we want:

$$|a_n - 0| \stackrel{?}{<} \epsilon \qquad 1/n \stackrel{?}{<} \epsilon \qquad 1/\epsilon \stackrel{?}{<} n$$

Since $n > N$, it suffices to choose $N > 1/\epsilon$.

PROOF. Given $\epsilon > 0$, Lemma 5.3.2(6) tells us there exists $N \in \mathbb{N}$, such that $N > 1/\epsilon$. Given $n > N$, we have

$$\begin{aligned} |a_n - 0| &= 1/n && (a_n = 1/n > 0) \\ &< 1/N && (n > N \text{ and Lemma 5.3.2(8)}) \\ &< \epsilon && (N > 1/\epsilon \text{ and Lemma 5.3.2(8)}). \end{aligned}$$ □

EXERCISE 5.3.6. Show that if $a_n = n/(n+1)$ for all n, then $a_n \to 1$.

PROPOSITION 5.3.7. *If $a_n \to L$ and $b_n \to M$, then $a_n + b_n \to L + M$.*

Scratchwork. To prove $a_n + b_n \to L + M$,

$$\text{we want to make } |(a_n + b_n) - (L + M)| \text{ small (less than } \epsilon).$$

What we know is that we can make $|a_n - L|$ and $|b_n - M|$ as small as we like. By the triangle inequality, we have

$$\big|(a_n - L) + (b_n - M)\big| < |a_n - L| + |b_n - M|.$$

By simple algebra, the left-hand side is equal to $|(a_n + b_n) - (L + M)|$, so we just need to make the right-hand side less than ϵ. This will be true if $|a_n - L|$ and $|b_n - M|$ are both less than $\epsilon/2$.

Since $a_n \to L$, there is some large N_a, such that $|a_n - L| < \epsilon/2$ for all $n > N_a$. Similarly, since $b_n \to M$, there is some large N_b, such that $|b_n - M| < \epsilon/2$ for all $n > N_b$. Now, we just need know that n will be larger than both N_a and N_b whenever $n > N$. So we should choose N to be whichever of N_a and N_b is larger. That is, we let N be the **maximum** of N_a and N_b, which is denoted $\max(N_a, N_b)$.

PROOF. Given $\epsilon > 0$, we know that $\epsilon/2 > 0$. Hence:

- Since $a_n \to L$, we know $\exists N_a \in \mathbb{N},\ \forall n > N_a,\ |a_n - L| < \epsilon/2$.
- Since $b_n \to M$, we know $\exists N_b \in \mathbb{N},\ \forall n > N_b,\ |b_n - M| < \epsilon/2$.

Let $N = \max(N_a, N_b) \in \mathbb{N}$, so $N \geq N_a$ and $N \geq N_b$.

Given $n > N$:

$(*)$ We have $n > N \geq N_a$, so $|a_n - L| < \epsilon/2$.

$(**)$ We have $n > N \geq N_b$, so $|b_n - M| < \epsilon/2$.

Therefore

$$\begin{aligned}
\big|(a_n + b_n) - (L + M)\big| &= \big|(a_n - L) + (b_n - M)\big| && \text{(high-school algebra)} \\
&\leq |a_n - L| + |b_n - M| && \text{(triangle inequality)} \\
&< \epsilon/2 \ + \ \epsilon/2 && ((*) \text{ and } (**)) \\
&= \epsilon.
\end{aligned}$$

□

EXERCISES 5.3.8. Assume $a_n \to L$, and $c \in \mathbb{R}$.
Do these proofs directly from the definition of convergence.

1) Show $-a_n \to -L$.
2) Show $a_n + c \to L + c$.
3) Show $2a_n \to 2L$.
4) Show $ca_n \to cL$ if $c \neq 0$.
5) Show that if $L > 0$, then $\exists N \in \mathbb{N}$, such that $a_n > 0$ for all $n > N$.
6) (*harder*) Show that if $L = 1$, then $1/a_n \to 1$.

EXERCISES 5.3.9. Assume $a_n \to L$ and $b_n \to M$.

1) Show that if $M = 0$, and $|a_n| \leq 2$ for all n, then $a_n b_n \to 0$.
2) (*harder*) Show $a_n b_n \to LM$.

SUMMARY:

- A valid deduction (or "result") is usually called a theorem, proposition, corollary, or lemma.
- **Divisibility and congruence**
  - Important definitions:
    - divisor, multiple
    - congruent modulo n: $a \equiv b \pmod{n}$
    - remainder
    - irrational number
  - Congruence (mod n) is reflexive, symmetric, and transitive
  - $a \equiv b \pmod{n}$ iff a and b have the same remainder when divided by n
  - $\sqrt{2}$ is irrational
  - Notation:
    - $a \mid b$, $a \nmid b$
    - $a \equiv b \pmod{n}$
- **Commutative groups**
  - Important definitions:
    - commutative group (commutative, associative, identity element, negatives)
    - subgroup (closed under negatives and addition)
  - The identity element of a group is unique.
  - The negative of each element of a group is unique.
  - Notation:
    - 0 (identity element)
    - $-g$ (negative)
- **Convergent sequences**
  - Important definitions:
    - absolute value
    - converges
  - triangle inequality
  - Notation:
    - $|x|$
    - $a_n \to L$

Part III

Other Fundamental Concepts

Chapter 6

Functions

It is the pervading law of all things ... that form ever follows function. This is the law.

Louis Sullivan (1856–1924), American architect
The tall office building artistically considered

6.1. Cartesian product

We discussed unions and intersections in Section 3.3. The Cartesian product is another important set operation. Before introducing it, let us recall the notation for an ordered pair.

NOTATION 6.1.1. For any objects x and y, mathematicians use (x, y) to denote the **ordered pair** whose first coordinate is x and whose second coordinate is y. It is important to know that the order matters: (x, y) is usually not the same as (y, x). (That is why these are called *ordered* pairs. Notice that sets are not like this: sets are unordered, so $\{x, y\}$ *is* always the same as $\{y, x\}$.) It is important to realize that:

$$(x_1, y_1) = (x_2, y_2) \qquad \Leftrightarrow \qquad x_1 = x_2 \text{ and } y_1 = y_2$$

EXAMPLE 6.1.2. A special case of the Cartesian product is familiar to all algebra students: recall that

$$\mathbb{R}^2 = \{\, (x, y) \mid x \in \mathbb{R},\, y \in \mathbb{R} \,\} \tag{6.1.3}$$

is the set of all ordered pairs of real numbers. This is the "coordinate plane" (or "xy-plane") that is used to draw the graphs of functions.

The formula $y = f(x)$ often appears in elementary algebra, and, in that subject, the variables x and y represent real numbers. However, advanced math courses allow x and y to be elements of any sets A and B, not just from $\mathbb{R}$. Therefore, it is important to generalize the above example by replacing the two appearances of $\mathbb{R}$ in the right-hand side of Equation 6.1.3 with arbitrary sets A and B:

DEFINITION 6.1.4. For any sets A and B, we let

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

This notation means, for all x, that

$$x \in A \times B \text{ iff } \exists a \in A, \exists b \in B,\ x = (a, b).$$

The set $A \times B$ is called the **Cartesian product** of A and B.

EXAMPLE 6.1.5.

1) $\mathbb{R} \times \mathbb{R} = \mathbb{R}^2$.

2) $\{1, 2, 3\} \times \{\mathsf{a}, \mathsf{b}\} = \big\{(1, \mathsf{a}), (1, \mathsf{b}), (2, \mathsf{a}), (2, \mathsf{b}), (3, \mathsf{a}), (3, \mathsf{b})\big\}$.

3) $\{\mathsf{a}, \mathsf{b}\} \times \{1, 2, 3\} = \big\{(\mathsf{a}, 1), (\mathsf{a}, 2), (\mathsf{a}, 3), (\mathsf{b}, 1), (\mathsf{b}, 2), (\mathsf{b}, 3)\big\}$.

By comparing (2) and (3), we see that $\times$ is *not* commutative: $A \times B$ is usually *not* equal to $B \times A$.

EXERCISES 6.1.6. Specify each set by listing its elements.

1) $\{\mathsf{a}, \mathsf{i}\} \times \{\mathsf{n}, \mathsf{t}\} =$

2) $\{\mathsf{Q}, \mathsf{K}\} \times \{\clubsuit, \diamondsuit, \heartsuit, \spadesuit\} =$

3) $\{1, 2, 3\} \times \{3, 4, 5\} =$

Remark 6.1.7. We will prove in Theorem 9.1.18 that

$$\#(A \times B) = \#A \cdot \#B.$$

In other words:

> The cardinality of a Cartesian product is the product of the cardinalities.

For now, let us just give an informal justification:

Suppose $\#A = m$ and $\#B = n$. Then, by listing the elements of these sets, we may write

$$A = \{a_1, a_2, a_3, \ldots, a_m\} \quad \text{and} \quad B = \{b_1, b_2, b_3, \ldots, b_n\}.$$

The elements of $A \times B$ are:

$$\begin{array}{ccccc}
(a_1, b_1), & (a_1, b_2), & (a_1, b_3), & \cdots & (a_1, b_n), \\
(a_2, b_1), & (a_2, b_2), & (a_2, b_3), & \cdots & (a_2, b_n), \\
(a_3, b_1), & (a_3, b_2), & (a_3, b_3), & \cdots & (a_3, b_n), \\
\vdots & \vdots & \vdots & \ddots & \vdots \\
(a_m, b_1), & (a_m, b_2), & (a_m, b_3), & \cdots & (a_m, b_n).
\end{array}$$

In this array,

- each row has exactly n elements, and
- there are m rows,

so the number of elements is the product $mn = \#A \cdot \#B$.

Here are some examples of proofs involving Cartesian products.

EXAMPLE 6.1.8. If A and B are nonempty sets, and $A \times B = B \times A$, then $A = B$.

PROOF. Assume A and B are nonempty sets, such that $A \times B = B \times A$. It suffices to show $A \subset B$ and $B \subset A$. By symmetry, we need only show $A \subset B$.

Let a_0 be an arbitrary element of A. Since B is nonempty, there exists some $b_0 \in B$. Then

$$(a_0, b_0) \in A \times B = B \times A = \{\, (b, a) \mid b \in B, a \in A \,\},$$

so there exist $b \in B$ and $a \in A$, such that $(a_0, b_0) = (b, a)$. Therefore, $a_0 = b$ (and $b_0 = a$, but we do not need that fact). Hence $a_0 = b \in B$. □

EXAMPLE 6.1.9. If B is disjoint from C, then $A \times B$ is disjoint from $A \times C$.

PROOF. We prove the contrapositive: Assume $A \times B$ is *not* disjoint from $A \times C$, and we will show B is *not* disjoint from C.

By assumption, the intersection of $A \times B$ and $A \times C$ is not empty, so we may choose some

$$x \in (A \times B) \cap (A \times C).$$

Then:

- Since $x \in A \times B$, there exist $a_1 \in A$ and $b \in B$, such that $x = (a_1, b)$.
- Since $x \in A \times C$, there exist $a_2 \in A$ and $c \in C$, such that $x = (a_2, c)$.

Hence $(a_1, b) = x = (a_2, c)$, so $b = c$. Now $b \in B$ and $b = c \in C$, so $b \in B \cap C$. Therefore $B \cap C \neq \varnothing$, so, as desired, B and C are *not* disjoint. □

Remark 6.1.10. When reading the proof above, you may have noticed that the variable x (a single letter) was used to represent an ordered pair (a, b). There is nothing wrong with this, because the ordered pair is a single object, and a variable can represent any mathematical object at all, whether it is an element of a set, or an entire set, or a function, or an ordered pair, or something else.

EXAMPLE 6.1.11. Assume A, B, and C are sets. Prove $A \times (B \cup C) = (A \times B) \cup (A \times C)$.

PROOF. ($\supset$) Given $x \in (A \times B) \cup (A \times C)$, we have $x \in A \times B$ or $x \in A \times C$. By symmetry, we may assume $x \in A \times B$, so $x = (a, b)$ for some $a \in A$ and $b \in B$. Note that $b \in B \cup C$, so we have $a \in A$ and $b \in B \cup C$. Therefore

$$x = (a, b) \in A \times (B \cup C).$$

Since x is an arbitrary element of $(A \times B) \cup (A \times C)$, this implies $(A \times B) \cup (A \times C) \subset A \times (B \cup C)$.

($\subset$) Given $(a, x) \in A \times (B \cup C)$, we have $a \in A$, and either $x \in B$ or $x \in C$. By symmetry, we may assume $x \in B$. Then $(a, x) \in A \times B \subset (A \times B) \cup (A \times C)$, so $(a, x) \in (A \times B) \cup (A \times C)$. Since (a, x) is an arbitrary element of $A \times (B \cup C)$, this implies $A \times (B \cup C) \subset (A \times B) \cup (A \times C)$. □

EXERCISES 6.1.12.

1) Suppose A, B, and C are sets.
   (a) Show that if $B \subset C$, then $A \times B \subset A \times C$.
   (b) Show that if $A \times B = A \times C$, and $A \neq \varnothing$, then $B = C$.

2) Suppose A is a set.
   (a) Show $A \times \varnothing = \varnothing$.
   (b) Show $A \times A = \varnothing$ if and only if $A = \varnothing$.

3) To say that $\times$ is distributive over $\cup$ means that, for all sets A, B, and C, we have

$$A \times (B \cup C) = (A \times B) \cup (A \times C) \quad \text{and} \quad (B \cup C) \times A = (B \times A) \cup (C \times A).$$

   The first equation was established in Example 6.1.11. Complete the proof that $\times$ is distributive over $\cup$ by proving the second equation.

4) Show that $\times$ is distributive over $\cap$. That is, for all sets A, B, and C, we have
   (a) $A \times (B \cap C) = (A \times B) \cap (A \times C)$, and
   (b) $(B \cap C) \times A = (B \times A) \cap (C \times A)$.

6.2. Informal introduction to functions

You have seen many examples of functions in your previous math classes. Most of these were probably given by formulas (such as $f(x) = x^3$), but functions can also be given in other ways. The key property of a function is that it accepts inputs, and provides a corresponding output value for each possible input.

EXAMPLE 6.2.1. For the function $f(x) = x^3$, the input x can be any real number. Plugging a value for x into the formula yields an output value, which is also a real number. For example, using $x = 2$ as the input yields the output value $f(2) = 2^3 = 8$.

DEFINITION 6.2.2 (unofficial). Suppose f is any function.

1) The set of allowable inputs of f is called the **domain** of f.
2) If A is the domain of f, and B is any set that contains all of the possible outputs of f, then we say that f is a **function from A to B**. In the case of the function $f(x) = x^3$, we may take A and B to both be the set of real numbers; thus, f is a function from $\mathbb{R}$ to $\mathbb{R}$.

EXAMPLE 6.2.3. $g(x) = 1/x$ is *not* a function from $\mathbb{R}$ to $\mathbb{R}$. This is because 0 is an element of $\mathbb{R}$, but the formula does not define a value for $g(0)$. Thus, 0 cannot be in the domain of g. To correct this problem, one could say that g is a function from the set $\{\, x \in \mathbb{R} \mid x \neq 0 \,\}$ of *nonzero* real numbers, to $\mathbb{R}$.

Intuitively, a function from A to B can be thought of being any process that accepts inputs from the set A, and assigns an element of the set B to each of these inputs. The process need not be given by a formula. Indeed, most of the functions that arise in science or in everyday life are not given by any formula.

EXAMPLE 6.2.4.

1) Each point on the surface of the earth has a particular temperature right now, and the temperature (in degrees centigrade) is a real number. Thus, temperature defines a function temp from the surface of the earth to $\mathbb{R}$: $\mathsf{temp}(x)$ is the temperature at the point x.
2) The items in a grocery store each have a particular price, which is a certain number of cents, so price can be thought of as a function from the set of items for sale to the set $\mathbb{N}$ of all natural numbers: $\mathsf{price}(x)$ is the price of item x (in cents).
3) If we let People be the set of all people (alive or dead), then mother is a function from People to People. For example,

$$\mathsf{mother}(\mathsf{Prince\ Charles}) = \mathsf{Queen\ Elizabeth}.$$

(To avoid ambiguity, perhaps we should clarify that, by "mother," we mean "biological mother.")

4) In contrast, $\mathsf{grandmother}$ is *not* a function from People to People. This is because people have not just one grandmother, but two (a maternal grandmother and a paternal grandmother). For example, if we say that Prince Charles wrote a poem for his grandmother, we do not know whether he wrote the poem for the mother of Queen Elizabeth, or for his other grandmother. A function is not ever allowed to have such an ambiguity. (In technical terms, $\mathsf{grandmother}$ is a "relation," not a function. This will be explained in Section 7.1.)

Functions are often given by a *table* of values.

EXAMPLE 6.2.5. The list of prices in a store is an example of this:

item	price (in cents)
apple	65
banana	83
cherry	7
donut	99
egg	155

In this example:

- The domain of price is $\{\mathsf{apple}, \mathsf{banana}, \mathsf{cherry}, \mathsf{donut}, \mathsf{egg}\}$.
- $\mathsf{price}(\mathsf{banana}) = 83$.
- $\mathsf{price}(\mathsf{guava})$ does not exist, because guava is not in the domain of the function.

Instead of making a table, mathematicians prefer to represent each row of the table by an ordered pair. For example, the first row of the table is apple | 65. This has apple on the left and 65 on the right, so we represent it by the ordered pair $(\mathsf{apple}, 65)$, which has apple on the left and 65 on the right. The second row is represented by $(\mathsf{banana}, 83)$. Continuing in this way yields a total of 5 ordered pairs (one for each row). To keep them gathered together, a mathematician puts them into a set. Thus, instead of writing a table, a mathematician would represent this function as:

$$\big\{\, (\mathsf{apple}, 65), (\mathsf{banana}, 83), (\mathsf{cherry}, 7), (\mathsf{donut}, 99), (\mathsf{egg}, 155) \,\big\}.$$

The set of ordered pairs contains exactly the same information as a table of values, but the set is a more convenient form for mathematical manipulations.

EXERCISE 6.2.6. At right is a function f given by a table of values.
(*You do not need to show your work on any parts of this problem.*)

x	$f(x)$
1	7
2	3
3	2
4	4
5	9

1) What is the domain of f?

2) What is $f(3)$?

3) Represent f as a set of ordered pairs.

4) Find a formula to represent f.
[*Hint:* There is a formula of the form $f(x) = ax^2 + bx + c$.]

EXAMPLE 6.2.7. Not every table of values represents a function. For example, suppose we have the following price list, which is a slight change from Example 6.2.5:

item	price (in cents)
apple	65
banana	83
cherry	7
donut	99
banana	155

There is a problem here, because there are two possible prices for a banana, depending on which line of the table is looked at. (So you might pick up a banana, expecting to pay 83 cents, and end up having the cashier charge you $1.55.) This is not allowed in a function: each input must have exactly one output, not a number of different possible outputs. Thus, if a table represents a function, and an item appears in the left side of more than one row, then all of those rows must have the same output listed on the right side.

Remark 6.2.8. A 2-column table represents a function from A to B if and only if:

1) every value that appears in the left column of the table is an element of A,

2) every value that appears in the right column of the table is an element of B,

3) every element of A appears in the left side of the table, and

4) no two rows of the table have the same left side, but different right sides.

EXAMPLE 6.2.9. Which of the following are functions from $\{1,2,3\}$ to $\{\mathsf{w},\mathsf{h},\mathsf{o}\}$? (If it is not such a function, then explain why not.)

1) $\{(1,\mathsf{w}),(1,\mathsf{h}),(1,\mathsf{o})\}$
2) $\{(1,\mathsf{h}),(2,\mathsf{h}),(3,\mathsf{h})\}$
3) $\{(1,\mathsf{h}),(2,\mathsf{o}),(3,\mathsf{w})\}$
4) $\{(\mathsf{w},1),(\mathsf{h},2),(\mathsf{o},3)\}$

SOLUTION.

(1) This is not a function. Since $(1,\mathsf{w})$, $(1,\mathsf{h})$, and $(1,\mathsf{o})$ are all in the set, there are three different elements b (not a *unique* b), such that $(1,b)$ is in the set.

(2) This is such a function.

(3) This is such a function.

(4) This is not such a function, because, for the element $(\mathsf{w},1)$ of the set, there do not exist elements a of $\{1,2,3\}$ and b of $\{\mathsf{w},\mathsf{h},\mathsf{o}\}$, such that $(\mathsf{w},1)=(a,b)$. (Instead, we would need to take a in $\{\mathsf{w},\mathsf{h},\mathsf{o}\}$ and b in $\{1,2,3\}$, which is backwards from what is required. In fact, f is a function from $\{\mathsf{w},\mathsf{h},\mathsf{o}\}$ to $\{1,2,3\}$, not from $\{1,2,3\}$ to $\{\mathsf{w},\mathsf{h},\mathsf{o}\}$.) □

EXERCISE 6.2.10. Let

- $A=\{\mathsf{a},\mathsf{b},\mathsf{c},\mathsf{d},\mathsf{e}\}$, and
- $B=\{1,3,5,7,9,11\}$.

Which of the following sets of ordered pairs are functions from A to B? (For those that are not, explain why.)

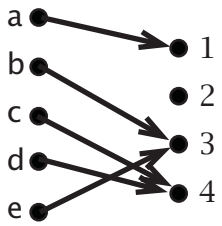
1) $\{(\mathsf{a},1),(\mathsf{b},3),(\mathsf{c},5),(\mathsf{d},7),(\mathsf{e},9)\}$
2) $\{(\mathsf{a},1),(\mathsf{b},2),(\mathsf{c},3),(\mathsf{d},4),(\mathsf{e},5)\}$
3) $\{(\mathsf{a},1),(\mathsf{b},3),(\mathsf{c},5),(\mathsf{d},3),(\mathsf{e},1)\}$
4) $\{(\mathsf{a},1),(\mathsf{b},3),(\mathsf{c},5),(\mathsf{d},7),(\mathsf{e},9),(\mathsf{a},11)\}$
5) $\{(\mathsf{a},1),(\mathsf{b},3),(\mathsf{c},5),(\mathsf{e},7)\}$
6) $\{(\mathsf{a},1),(\mathsf{b},1),(\mathsf{c},1),(\mathsf{d},1),(\mathsf{e},1)\}$
7) $\{(\mathsf{a},\mathsf{a}),(\mathsf{b},\mathsf{a}),(\mathsf{c},\mathsf{a}),(\mathsf{d},\mathsf{a}),(\mathsf{e},\mathsf{a})\}$
8) $\{(\mathsf{a},1),(\mathsf{b},3),(\mathsf{c},5),(\mathsf{d},5),(\mathsf{e},3),(\mathsf{a},1)\}$
9) $\{(1,\mathsf{a}),(3,\mathsf{a}),(5,\mathsf{a}),(7,\mathsf{a}),(9,\mathsf{a}),(11,\mathsf{a})\}$
10) $\{(\mathsf{c},1),(\mathsf{b},3),(\mathsf{e},5),(\mathsf{a},7),(\mathsf{d},9)\}$

Remark 6.2.11. It is sometimes helpful to represent a function $f\colon A\to B$ by drawing an **arrow diagram**:

- a dot is drawn for each element of A and each element of B, and
- an arrow is drawn from a to $f(a)$, for each $a\in A$.

For example, suppose

- $A=\{\mathsf{a},\mathsf{b},\mathsf{c},\mathsf{d},\mathsf{e}\}$,
- $B=\{1,2,3,4\}$, and
- $f=\{(\mathsf{a},1),(\mathsf{b},3),(\mathsf{c},4),(\mathsf{d},4),(\mathsf{e},3)\}$.



Then the picture at right is an arrow diagram of f.

Notice that:

1) There is exactly one arrow coming out of each element of A. This is true for the arrow diagram of any function.
2) There can be any number of arrows coming into each element of B (perhaps none, perhaps one, or perhaps more than one).

6.3. Official definition

The preceding section provided some intuition about how and why functions are represented as sets of ordered pairs, but it is not at all authoritative. Here are the official definitions.

DEFINITION 6.3.1. Suppose A and B are sets.

1) A set f is a **function from A to B** iff
   (a) each element of f is an ordered pair (a, b), such that $a \in A$ and $b \in B$, and
   (b) for each $a \in A$, there is a *unique* $b \in B$, such that $(a, b) \in f$.
2) If f is a function from A to B, then
   - A is called the **domain** of f, and
   - B is a **codomain** of f.
3) We write "$f\colon A \to B$" to denote that f is a function from A to B.

EXERCISE 6.3.2. We can express the definition of a function in First-Order Logic:

1) Translate the assertion of Definition 6.3.1(1a) into First-Order Logic.
2) Translate the assertion of Definition 6.3.1(1b) into First-Order Logic.

NOTATION 6.3.3. Suppose $f\colon A \to B$.

1) For $a \in A$, it is convenient to have a name for the element b of B, such that $(a, b) \in f$. The name we use is $f(a)$:

$$f(a) = b \text{ if and only if } (a, b) \in f.$$

2) Each element a of A provides us with an element $f(a)$ of B. The **range** of f is the set that collects together all of these elements $f(a)$. That is,

$$b \text{ is in the range of } f \text{ iff there is some } a \in A, \text{ such that } b = f(a).$$

The range can be denoted $\{\, f(a) \mid a \in A \,\}$.

EXAMPLE 6.3.4. Suppose the function f is defined by $f(x) = x^2$, on the domain $\{0, 1, 2, 4\}$. Then:

1) To represent f as a set of ordered pairs, each element of the domain must appear exactly once as a first coordinate, with the corresponding output given in the second coordinate. Since there are four elements in the domain, there will be four ordered pairs: $f = \{(0, 0), (1, 1), (2, 4), (4, 16)\}$.
2) To give a table for f, we include one row for every element of the domain. The table will be:

n	$f(n)$
0	0
1	1
2	4
4	16

3) If we are asked what is $f(3)$, the answer is that $f(3)$ *does not exist*, because 3 is not in the domain of f. Even though we know that $3^2 = 9$, the formula we gave for f only applies to elements that are in the domain of f! It is not true that $f(3) = 9$.

4) The range of f is the set of possible outputs: in this case, the range of f is $\{0, 1, 4, 16\}$.

5) If we are asked what is $f(2)$, the answer is $f(2) = 4$.

6) Is f a function from $\{n \in \mathbb{N} \mid n \leq 4\}$ to $\{0, 1, 4, 16\}$? The answer is no, because the first set is $\{0, 1, 2, 3, 4\}$, which includes the value 3, but 3 is not in the domain of f.

7) Is f a function from $\{0, 1, 2, 4\}$ to $\{n \in \mathbb{N} \mid n \leq 16\}$? The answer is yes; even though the second set has many values that are not in the range, it is a possible codomain for f. A codomain can be any set that contains all of the elements of the range, so every function has many different codomains (but only one domain and only one range).

EXERCISES 6.3.5.

1) The table at right describes a certain function g.

n	$g(n)$
2	7
4	9
6	11
8	13
10	15

(a) What is the domain of g?

(b) What is the range of g?

(c) What is $g(6)$?

(d) What is $g(7)$?

(e) Represent g as a set of ordered pairs.

(f) Draw an arrow diagram to represent g.

(g) Write down a formula that describes g.
(Express $g(n)$ in terms of n, by using simple arithmetic operations.)

2) Suppose

- f is a function whose domain is $\{0, 2, 4, 6\}$, and
- $f(x) = 4x - 5$, for every x in the domain.

Describe the function in each of the following ways:

(a) Make a table. (b) Draw an arrow diagram. (c) Use ordered pairs.

3) For the given sets A and B:

(i) Write each function from A to B as a set of ordered pairs.

(ii) Write down the range of each function.

(a) $A = \{\mathsf{a}, \mathsf{b}, \mathsf{c}\}$, $B = \{\mathsf{d}\}$ (b) $A = \{\mathsf{a}, \mathsf{b}\}$, $B = \{\mathsf{c}, \mathsf{d}\}$

(c) $A = \{\mathsf{a}\}$, $B = \{\mathsf{b}, \mathsf{c}, \mathsf{d}\}$ (d) $A = \{\mathsf{a}, \mathsf{b}\}$, $B = \{\mathsf{c}, \mathsf{d}, \mathsf{e}\}$

[*Hint:* For (i), you may assume, without proof, that if A has exactly m elements, and B has exactly n elements, then the number of functions from A to B is n^m. (Do you see why this is the correct number?)]

4) Which of the following sets of ordered pairs are functions from $\{\mathsf{x}, \mathsf{y}, \mathsf{z}\}$ to $\{\mathsf{a}, \mathsf{b}, \mathsf{c}, \mathsf{d}, \mathsf{e}\}$?

- If it is such a function, then what is its range?

• If it is not such a function, then explain why not.

(a) $\{(\mathsf{y},\mathsf{a}),(\mathsf{x},\mathsf{b}),(\mathsf{y},\mathsf{c})\}$ (b) $\{(\mathsf{y},\mathsf{a}),(\mathsf{x},\mathsf{b}),(\mathsf{z},\mathsf{c})\}$ (c) $\{(\mathsf{y},\mathsf{a}),(\mathsf{x},\mathsf{c}),(\mathsf{z},\mathsf{a})\}$

EXAMPLE 6.3.6. If A is a set, then any function from $A \times A$ to A is called a **binary operation** on A. However, when $+$ is a binary operation and $a, b \in A$, we write $a + b$ for the value of the function at (a, b), instead of writing $+\big((a, b)\big)$. See Section 5.2 for some examples and many exercises about binary operations.

6.4. One-to-one functions

We begin this section with an example.

EXAMPLE 6.4.1.

1) Suppose Inspector Thinkright knows two facts:
   (a) Alice is the thief's wife, and
   (b) Alice is Bob's wife.

   Then the Inspector can arrest Bob for theft, because a woman cannot be the wife of more than one man.
2) On the other hand, suppose the Inspector knows:
   (a) Alice is the forger's mother, and
   (b) Alice is Charlie's mother.

   Then the Inspector does not know enough to be sure who the forger is, because it could be some other child of Alice, instead of being Charlie.

This example illustrates a fundamental difference between the wife function and the mother function: two different people can have the same mother, but only one person can have any particular person as their wife. (For example, if Bud and Charlie have the same wife, then "Bud" must be a nickname for Charlie.) In mathematical terms, this important property of the wife function is expressed by saying that the wife function is "one-to-one."

The notion is formalized in the following definition:

DEFINITION 6.4.2. Suppose $f\colon A \to B$. We say f is **one-to-one** iff, for all $a_1, a_2 \in A$, such that $f(a_1) = f(a_2)$, we have $a_1 = a_2$.

EXERCISES 6.4.3. Assume $f\colon A \to B$ and $g\colon X \to Y$. Translate each assertion into First-Order Logic.

1) f is one-to-one.
2) g is one-to-one.
3) f is *not* one-to-one.
4) g is *not* one-to-one.

(Simplify your answers in (3) and (4) so that $\neg$ is applied only to predicates.)

Remark 6.4.4. If you have an arrow diagram of a function, then it is easy to tell whether or not the function is one-to-one. For example:

1) The function f of Figure 6A(a) is *not* one-to-one. This is because the arrow from b and the arrow from d go to the same place, so $f(\mathsf{b}) = f(\mathsf{d})$. In general, if arrows from two different elements of the domain go to the same element of the range, then the function is not one-to-one.

2) The function g of Figure 6A(b) is one-to-one. This is because the arrows from two different elements of the domain never go to the same element of the range. In short, there is only *one* element of the domain that goes *to* any *one* element of the range. (This is the reason for the terminology "one-to-one." A function is "two-to-one" if there are two elements of the domain mapping to each element of the range, as is true of the function h in Figure 6A(c), but we do not need this terminology.)

3) *Warning.* Although the arrow diagram of a one-to-one function never has more than one arrow pointing to the same element of the codomain, this does not mean that every element of the codomain has *exactly* one arrow into it. For example, the function g of Figure 6A(b) is one-to-one (because there is never more than one arrow into any point), but there is a point in the codomain that does not have any arrows to it.

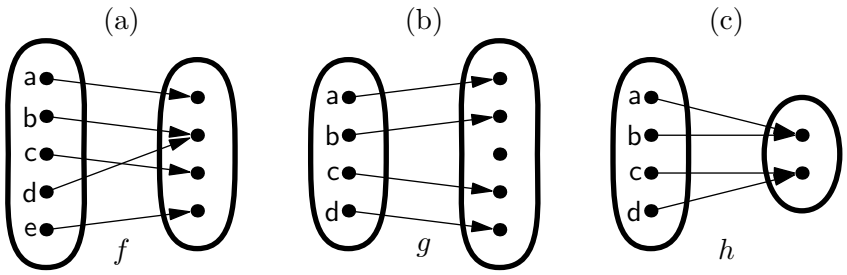


Figure 6A. Arrow diagrams of three functions f, g, and h.

EXAMPLE 6.4.5. Without giving official proofs, let us determine which of the following functions are one-to-one.

1) $f\colon \mathbb{R} \to \mathbb{R}$, defined by $f(x) = x + 1$.

This is one-to-one. For any real numbers x and y, $f(x) = f(y)$ means that $x + 1 = y + 1$. Subtracting 1 from both sides of the equation, we conclude that $x = y$ whenever $f(x) = f(y)$.

2) $g\colon \mathbb{R} \to \mathbb{R}$, defined by $g(x) = |x|$.

This is not one-to-one. We demonstrate this by finding two distinct real numbers whose image is the same:

$$g(1) = |1| = 1 = |-1| = g(-1),$$

but $1 \neq -1$. This shows that g is *not* one-to-one.

3) $f\colon \{1, 2, 3\} \to \{\mathsf{a}, \mathsf{b}, \mathsf{c}\}$ defined by $f = \{(1, b), (2, a), (3, a)\}$.

This is not one-to-one. We demonstrate this by finding two distinct values in $\{1, 2, 3\}$ whose image is the same:

$$f(2) = a = f(3),$$

but $2 \neq 3$. This shows that f is *not* one-to-one.

4) $h\colon \mathbb{N} \to \mathbb{N}$, defined by $h(x) = |x|$.

This is one-to-one. Since all natural numbers are nonnegative, we have $|x| = x$ for every natural number x. So if $h(x) = h(y)$, then

$$x = |x| = h(x) = h(y) = |y| = y.$$

Remark 6.4.6. These examples demonstrate the general pattern of how to prove a function is (or is not) one-to-one:

- To prove that a function $f\colon A \to B$ *is* one-to-one, we need to demonstrate that for *every* $a_1, a_2 \in A$, if $f(a_1) = f(a_2)$, then $a_1 = a_2$.
- To prove that a function $f\colon A \to B$ is *not* one-to-one, we need only find a single pair of values $a_1, a_2 \in A$, for which $f(a_1) = f(a_2)$, but $a_1 \neq a_2$.

EXERCISES 6.4.7. *Explain why your answers are correct* (*but you do not need to give formal proofs*).

1) Each formula defines a function from $\mathbb{R}$ to $\mathbb{R}$. Which of the functions are one-to-one?

   (a) $f(x) = 1$. (b) $g(x) = x$. (c) $h(x) = x^2$.
   (d) $i(x) = 3x + 2$. (e) $j(x) = 1/\big(|x| + 1\big)$.

2) Each of the following sets of ordered pairs is a function from $\{1, 2, 3, 4\}$ to $\{\mathsf{a}, \mathsf{b}, \mathsf{c}, \mathsf{d}, \mathsf{e}\}$. Which are one-to-one?

   (a) $f = \{(1, \mathsf{a}), (2, \mathsf{b}), (3, \mathsf{d}), (4, \mathsf{e})\}$ (b) $g = \{(1, \mathsf{c}), (2, \mathsf{d}), (3, \mathsf{d}), (4, \mathsf{e})\}$
   (c) $h = \{(1, \mathsf{e}), (2, \mathsf{d}), (3, \mathsf{c}), (4, \mathsf{b})\}$ (d) $i = \{(1, \mathsf{e}), (2, \mathsf{e}), (3, \mathsf{e}), (4, \mathsf{e})\}$
   (e) $j = \{(1, \mathsf{a}), (2, \mathsf{c}), (3, \mathsf{e}), (4, \mathsf{c})\}$ (f) $k = \{(1, \mathsf{a}), (2, \mathsf{c}), (3, \mathsf{e}), (4, \mathsf{d})\}$

Here is an example of a formal proof that a function is one-to-one.

EXAMPLE 6.4.8. Let $f\colon \mathbb{R} \to \mathbb{R}$ be defined by $f(x) = 2x + 1$. Then f is one-to-one.

Scratchwork. By definition, we wish to show $\forall x_1, x_2 \in \mathbb{R}, \big(f(x_1) = f(x_2) \Rightarrow x_1 = x_2\big)$. Thus, the proof will use $\forall$-introduction: the first words in the proof will be "Given $x_1, x_2 \in \mathbb{R}$" (or other words to that effect). Then, because we wish to show $f(x_1) = f(x_2) \Rightarrow x_1 = x_2$, we will assume $f(x_1) = f(x_2)$, and the proof will be complete as soon as we can prove $x_1 = x_2$.

By the definition of f, the assumption $f(x_1) = f(x_2)$ means that

$$2x_1 + 1 = 2x_2 + 1.$$

Subtracting 1 from both sides, we see that

$$2x_1 = 2x_2.$$

Dividing both sides by 2, we conclude that $x_1 = x_2$, as desired.

Since readers of this textbook are expected to have a good command of logic and high-school algebra, our official proof can omit the comments that are unnecessary for such a well-educated reader. For example, the reader can be expected to be able to easily verify that the equation $2x + 1 = 2x_2 + 1$ can be simplified to the equation $2x_1 = 2x_2$, without being told that they should subtract 1 from both sides.

PROOF. Given $x_1, x_2 \in \mathbb{R}$, such that $f(x_1) = f(x_2)$, we have

$$2x_1 + 1 = 2x_2 + 1,$$

so

$$2x_1 = 2x_2,$$

so $x_1 = x_2$. □

A very similar argument applies to other linear functions.

EXAMPLE 6.4.9. Define $p\colon \mathbb{R} \to \mathbb{R}$ by $p(z) = 6z - 100$. Show that p is one-to-one.

PROOF. Given $z_1, z_2 \in \mathbb{R}$, such that $p(z_1) = p(z_2)$, we have

$$6z_1 - 100 = 6z_2 - 100,$$

so

$$6z_1 = 6z_2,$$

so $z_1 = z_2$. □

EXERCISES 6.4.10. Prove that each function is one-to-one.

1) $f\colon \mathbb{R} \to \mathbb{R}$ defined by $f(x) = 3x + 5$.
2) $f\colon \mathbb{R} \to \mathbb{R}$ defined by $f(x) = 7x - 2$.
3) $g\colon \mathbb{R} \to \mathbb{R}$ defined by $g(t) = 4t + 9$.
4) $h\colon \mathbb{R} \to \mathbb{R}$ defined by $h(s) = 7 - 8s$.
5) $i\colon \mathbb{R} \to \mathbb{R}$ defined by $i(r) = (5r - 2)/7$.
6) $j\colon \mathbb{Z} \to \mathbb{Z}$ defined by $j(n) = 2n + 11$.

The fact that the wife function is one-to-one can be restated as the fact that two different people cannot have the same wife. In general, a function is one-to-one iff two different elements of the domain always map to two different elements of the range:

(6.4.11) A function $f\colon A \to B$ is one-to-one if and only if $\forall a_1, a_2 \in A, \big(a_1 \neq a_2 \Rightarrow f(a_1) \neq f(a_2)\big)$.

(The notation "$\forall a_1, a_2 \in A$" is short for "$\forall a_1 \in A, \forall a_2 \in A$.")

We should justify the assertion in this box with a proof. The implication $\Rightarrow$ is proved in the following theorem; the other direction is an exercise.

THEOREM 6.4.12. *If a function $f\colon A \to B$ is one-to-one, then*

$$\forall a_1, a_2 \in A, \big(a_1 \neq a_2 \Rightarrow f(a_1) \neq f(a_2)\big).$$

PROOF. Let $f\colon A \to B$ be one-to-one. Given $a_1, a_2 \in A$, we know, from the definition of one-to-one, that

$$f(a_1) = f(a_2) \Rightarrow a_1 = a_2.$$

So the contrapositive of this implication is also true. That is,

$$a_1 \neq a_2 \Rightarrow f(a_1) \neq f(a_2).$$ □

EXERCISES 6.4.13.

1) Prove the converse of Theorem 6.4.12. More precisely, assume $f\colon A \to B$, and show that if

$$\forall a_1, a_2 \in A, \big(a_1 \neq a_2 \Rightarrow f(a_1) \neq f(a_2)\big),$$

then f is one-to-one.

2) Assume:

(a) $f\colon A \to B$, (b) f is one-to-one, (c) $a_1, a_2 \in A$,
(d) $g\colon B \to C$, (e) g is one-to-one, (f) $b_1, b_2 \in B$,
(g) $f(a_1) = b_1$, (h) $f(a_2) = b_2$, and (i) $g(b_1) = g(b_2)$.

Show $a_1 = a_2$. [*Hint:* First use the fact that g is one-to-one, then the fact that f is one-to-one.]

Remark 6.4.14 (alternative terminology). Many mathematicians use the word "*injective*," rather than "one-to-one." (This comes from French.) Also, a function that is one-to-one can be called an *injection*.

6.5. Onto functions

In an arrow diagram of a function $f\colon A \to B$, the definition of a function requires that there is exactly one arrow out of each element of A, but it says nothing about the number of arrows into each element of B. There may be elements of B with lots of arrows into them (unless the function is one-to-one), and there may be other elements of B that have no arrows into them. The function is called "onto" if all of the elements of B are hit by arrows; none are missed.

EXAMPLE 6.5.1. Figure 6B shows arrow diagrams of various functions, some onto and some not.

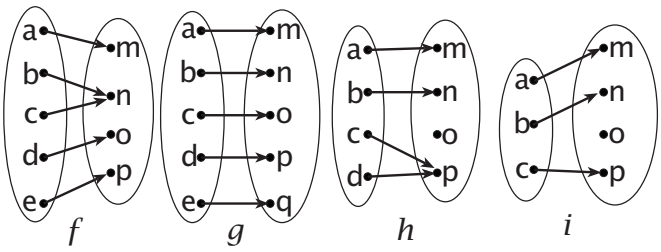


Figure 6B.
f is onto, but not one-to-one.
g is both one-to-one and onto.
h is neither one-to-one nor onto.
i is one-to-one, but not onto.

EXAMPLE 6.5.2. Not every woman is a mother. This means that if you draw an arrow from each person to his or her mother, there will be some women who have no arrows into them. So the function

$$\mathsf{mother}\colon \mathsf{People} \to \mathsf{Women}$$

is *not* onto.

The following official definition of "onto" formalizes the ideas described above.

DEFINITION 6.5.3. Suppose $f\colon A \to B$. We say f is **onto** iff, for all $b \in B$, there is some $a \in A$, such that $f(a) = b$.

EXERCISES 6.5.4. Suppose $f\colon A \to B$ and $g\colon X \to Y$. Translate each of the following assertions into First-Order Logic:

1) f is onto. 2) f is *not* onto. 3) g is onto. 4) g is *not* onto.

(Simplify your answers in (3) and (4) so that $\neg$ is applied only to predicates.)

EXAMPLE 6.5.5. Without giving formal proofs, let us demonstrate that each of the following functions is not onto.

1) $f\colon \mathbb{R} \to \mathbb{R}$, defined by $f(x) = |x|$.

Recall that the absolute value of a real number can never be negative. In particular, we can never have $|x| = -1$ for any real number x. Thus, there does not exist $x \in \mathbb{R}$, such that $f(x) = -1$. This shows that f is *not* onto.

2) $g\colon \{1,2,3\} \to \{\mathsf{a},\mathsf{b},\mathsf{c}\}$ defined by $g = \{(1,b),(2,a),(3,a)\}$.

Notice that c never appears as the second coordinate of an ordered pair in this function. This means there does not exist any x, such that $g(x) = \mathsf{c}$. This means that g is *not* onto.

EXERCISE 6.5.6. Each of the following sets of ordered pairs is a function from $\{1,2,3,4,5\}$ to $\{\clubsuit, \diamondsuit, \heartsuit, \spadesuit\}$. Which of the functions are onto? *Briefly justify your answers.*

1) $a = \{(1,\clubsuit),(2,\diamondsuit),(3,\heartsuit),(4,\spadesuit),(5,\clubsuit)\}$
2) $b = \{(1,\clubsuit),(2,\heartsuit),(3,\clubsuit),(4,\heartsuit),(5,\clubsuit)\}$
3) $c = \{(1,\heartsuit),(2,\heartsuit),(3,\heartsuit),(4,\heartsuit),(5,\heartsuit)\}$
4) $d = \{(1,\diamondsuit),(2,\spadesuit),(3,\heartsuit),(4,\spadesuit),(5,\clubsuit)\}$
5) $e = \{(1,\clubsuit),(2,\spadesuit),(3,\heartsuit),(4,\spadesuit),(5,\clubsuit)\}$

Let us see how to prove that a function $f\colon A \to B$ is onto. By definition, we wish to show:

$$\text{for all } b \in B, \text{ there is some } a \in A, \text{ such that } f(a) = b.$$

In other words: "$\forall b \in B, \exists a \in A, \big(f(a) = b\big)$."

The first quantifier is $\forall$; we are required to prove something about every element of B. Hence, we use $\forall$-introduction, so our proof should start with the sentence "Let b be an arbitrary element of B." (However, this can be abbreviated to: "Given $b \in B$, ...") After this, our task will be to prove "$\exists a \in A, \big(f(a) = b\big)$."

At this point, the quantifier that concerns us is $\exists$; we are required to prove that some element of A has a certain property. The tool to use for this is $\exists$-introduction: we find (or "construct") an appropriate element of A, and then verify that it does what it is supposed to. Thus, the next step in the proof is "Let a = ???" (where ??? needs to be replaced with an appropriate expression). Then all that remains is to verify that the value we assigned to a does the job it is required to do: calculate that $f(a)$ is indeed equal to b.

So here what a typical "onto" proof looks like:

$$\text{Given } b \in B, \text{ let } a = \boxed{}. \text{ Then } f(a) = \cdots = b.$$

An appropriate value for a needs to be put in the box (perhaps a formula that depends on b), and the dots need to be filled in with a calculation that shows the value of $f(a)$ is b. (Also, of course, some of the letters will need to be changed if the name of the function is not f, or if the sets are not called A and B.)

EXAMPLE 6.5.7. Define $g\colon \mathbb{R} \to \mathbb{R}$ by $g(x) = 5x - 2$. Show g is onto.

Scratchwork. We wish to show $\forall y \in \mathbb{R}, \exists x \in \mathbb{R}, \big(g(x) = y\big)$. By $\forall$-introduction, the first words of the proof are easy: "Given $y \in \mathbb{R}$." Then we need to find a value of x that makes $g(x) = y$. The appropriate value of x is probably not obvious, so we will do some scratchwork. We postulate the desired equation $g(x) = y$ and use algebra to solve it:

$$\begin{aligned} g(x) &= y \\ 5x - 2 &= y \\ 5x &= y + 2 \\ x &= \frac{y+2}{5}. \end{aligned}$$

Now that we know the correct value of x, it is easy to write the rest of the proof.

PROOF. Given $y \in \mathbb{R}$, let $x = (y+2)/5 \in \mathbb{R}$. Then

$$g(x) = 5x - 2 = 5\left(\frac{y+2}{5}\right) - 2 = (y+2) - 2 = y. \qquad \square$$

EXERCISES 6.5.8. Each formula defines a function from $\mathbb{R}$ to $\mathbb{R}$. Show the function is onto.

1) $f(x) = 2x + 1$. 2) $g(x) = 7x - 3$. 3) $h(t) = 4t + 9$.
4) $i(z) = 6 - 11z$. 5) $j(r) = (3r - 4)/5$.

Remark 6.5.9. Some "onto" proofs are more complicated than what is described above, because it may not be possible to go directly from "given $b \in B$" to "let $a = \square$." The problem is that it is sometimes necessary to insert calculations (or other explanations) between "given b" and "let a." Some examples of this will be seen in Exercise 6.8.14.

To complete the discussion, let us also see how to prove that a function $f\colon A \to B$ is *not* onto. By negating the definition of "onto," we see that we wish to prove "$\exists b \in B, \forall a \in A, \big(f(a) \neq b\big)$."

The first quantifier is $\exists$; we are required to prove that some element of B has a certain property. The tool to use for this is $\exists$-introduction: we find an appropriate element of B, and then we will need to verify that it does what it is supposed to. Thus, the first step in the proof is "Let $b = ???$" (where ??? needs to be replaced with an appropriate expression). After this, our task will be to prove "$\forall a \in A, \big(f(a) \neq b\big)$."

At this point, the quantifier that concerns us is $\forall$; we are required to prove something about every element of A. Hence, we use $\forall$-introduction, so the next step in our proof is the sentence "Let a be an arbitrary element of A" (or, for short, "Given $a \in A$, ..."). Then all that remains is to verify that $f(a) \neq b$.

So here what a typical "not onto" proof looks like:

$$\text{Let } b = \square \in B. \text{ Given } a \in A, \text{ we have } \ldots, \text{ so } f(a) \neq b.$$

An appropriate value for b needs to be put in the box, and the dots need to be filled in with an explanation that leads to the conclusion $f(a) \neq b$. (Also, as usual, some of the letters will need to be changed if the name of the function is not f, or if the sets are not called A and B.)

EXAMPLE 6.5.10. Define $e\colon \mathbb{R} \to \mathbb{R}$ by $e(r) = 1/\big(|r| + 1\big)$. Show e is not onto.

Scratchwork. We wish to prove $\exists y \in \mathbb{R}, \forall r \in \mathbb{R}, \big(e(r) \neq y\big)$, so we need to come up with an appropriate value of y. To do this, let's attempt to prove e is onto. Hopefully, we will run into trouble, and this difficulty will point us to a good choice for y. Namely, if e were onto, we would be able to solve the equation

$$e(r) = y$$

Let's put in the definition of $e(r)$, and use algebra to try to solve this equation:

$$\begin{aligned} \frac{1}{|r| + 1} &= y \\ |r| + 1 &= \frac{1}{y} \\ |r| &= \frac{1}{y} - 1. \end{aligned}$$

Now the absolute value $|r|$ is never negative, but the right-hand side of the equation could be negative. For example, if $y = -1$, then

$$\frac{1}{-1} - 1 = -1 - 1 = -2 < 0.$$

This suggests that we should let $y = -1$. With this in mind, we can write the proof.

PROOF. Let $y = -1$. Given $r \in \mathbb{R}$, we have $|r| \geq 0$, so $|r| + 1 \geq 0 + 1 = 1 > 0$. Therefore

$$e(r) = \frac{1}{|r| + 1} \geq 0 > -1 = y,$$

so $e(r) \neq y$. Since r is an arbitrary element of the domain $\mathbb{R}$, this implies that e is not onto. □

And sometimes you will need to decide whether a function is onto or not.

EXAMPLE 6.5.11. Define $m\colon \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ by $m(x, y) = x + y$. Is m onto?

Scratchwork. Let's try to prove m is onto. (If we fail, this is evidence that m is not onto, and we will try to prove that.) Given $z \in \mathbb{R}$, we try to solve the equation

$$m(x, y) = z.$$

In other words:

$$x + y = z.$$

It is easy to find values of x and y that satisfy the equation: perhaps the easiest solution is to let $y = 0$ and $x = z$. We can use these values to prove that m is onto.

SOLUTION. m is onto.

PROOF. Given $z \in \mathbb{R}$, let $(x, y) = (z, 0) \in \mathbb{R} \times \mathbb{R}$. Then $m(x, y) = x + y = z + 0 = z$. □

EXERCISES 6.5.12. Each formula defines a function from $\mathbb{R}$ to $\mathbb{R}$. Which of the functions are onto? *Prove that your answers are correct.*

1) $f(x) = 1$. 2) $a(x) = x$. 3) $b(t) = t^2$.

4) $c(s) = 3s + 2$. 5) $d(r) = \sqrt[3]{r + 5} - 5$.

EXERCISE 6.5.13. Suppose $f\colon A \to B$. Show that f is onto if and only if the range of f is B.

Remark 6.5.14 (alternative terminology). Some mathematicians say "*surjective*," rather than "onto." (Like "injective" in place of "one-to-one," this comes from French.) Also, a function that is onto can be called a *surjection*.

6.6. Bijections

The best functions are both one-to-one *and* onto. These are called "bijections."

DEFINITION 6.6.1. A function is a **bijection** iff it is both one-to-one and onto.

EXAMPLE 6.6.2. Consider a hypothetical country Married, in which

- everyone is married (to only one person — there is no polygamy!), and
- every marriage is between a man and a woman (there are no same-sex marriages).

Let

- Men be the set of men in the country, and
- Women be the set of women in the country.

Then wife: Men → Women is a bijection:

- Two different men cannot have the same wife, so we know that wife is one-to-one.
- Every woman is the wife of some man (because everyone is married), so wife is also onto.

Similarly, the function $\mathsf{husband}\colon \mathsf{Women} \to \mathsf{Men}$ is also a bijection.

Remark 6.6.3. In the country Married described above, it is clear that the number of men is exactly equal to the number of women. (If there were more men than women, then not every man could have a wife; if there were more women than men, then not every women could have a husband.) This is an example of the following important principle that will be discussed in the later chapter on "cardinality":

> If there is a **bijection** from A to B, then
> the two sets A and B must have exactly the same number of elements.

Finding a bijection is the most common way to show two sets have the same number of elements.

Remarks 6.6.4.

1) You may recall that a one-to-one function can be called an "injection," and an onto function can be called a "surjection." The term "bijection" comes from having both of these two properties.
2) Some textbooks use the term "one-to-one correspondence" for a bijection, but we will avoid that terminology, because it is too easy to confuse with "one-to-one function," which does not mean the same thing.

Remark 6.6.5. Showing that a function is a bijection requires two things: showing that the function is one-to-one, and showing that the function is onto. So a proof that a function is a bijection will (usually) have two parts:

1) Show that the function is one-to-one.
2) Show that the function is onto.

The two parts can come in either order: it is perfectly acceptable to first prove that the function is onto, and then prove that it is one-to-one.

EXAMPLE 6.6.6. *Define $f\colon \mathbb{R} \to \mathbb{R}$ by $f(x) = 5x - 7$. Then f is a bijection.*

PROOF. It suffices to show that f is both one-to-one and onto.

(one-to-one) Given $x_1, x_2 \in \mathbb{R}$, such that $f(x_1) = f(x_2)$, we have

$$5x_1 + 7 = 5x_2 + 7,$$

so

$$5x_1 = 5x_2,$$

so

$$x_1 = x_2.$$

Therefore f is one-to-one.

(onto) Given $y \in \mathbb{R}$, let $x = (y + 7)/5$. Then

$$f(x) = 5x - 7 = 5\left(\frac{y+7}{5}\right) - 7 = (y + 7) - 7 = y.$$

Therefore f is onto.

Since f is both one-to-one and onto, it is a bijection. □

EXERCISE 6.6.7. Each formula defines a function from $\mathbb{R}$ to $\mathbb{R}$. Show that the function is a bijection.

1) $a(x) = 5x + 2$ 2) $b(x) = 2x - 5$ 3) $c(x) = 12x - 15$
4) $d(x) = -15x - 12$ 5) $e(x) = x^3$ 6) $f(x) = \sqrt[3]{x - 4}$

NOTATION 6.6.8. For any set A, define the **identity map** $I_A \colon A \to A$ by $I_A(a) = a$ for every $a \in A$.

EXERCISE 6.6.9. Let A be a set. Show that the identity map I_A is a bijection from A to A.

EXERCISE 6.6.10. Each formula defines a function from $\mathbb{R}$ to $\mathbb{R}$. Which of the functions are bijections? *Show that your answers are correct.*

1) $a(x) = 7$. 2) $b(x) = 4x - 7$. 3) $c(x) = x^2$.
4) $d(r) = 3r + 2$. 5) $e(s) = 3|s| + 2$. 6) $f(t) = \sqrt{t^2 + 1}$
7) $g(u) = \sqrt[3]{u} - 5$.

EXAMPLE 6.6.11. Define $f \colon \mathbb{N} \times \mathbb{N} \to \mathbb{N}$ by $f(m, n) = m^2 + n$.

1) Show that f is onto. 2) Show that f is *not* one-to-one.

SOLUTION. (1) Given $k \in \mathbb{N}$, let $x = (0, k) \in \mathbb{N} \times \mathbb{N}$. Then

$$f(x) = f(0, k) = 0^2 + k = k.$$

Since k is an arbitrary element of $\mathbb{N}$, we conclude that f is onto.

(2) Let $x_1 = (1, 0)$ and $x_2 = (0, 1)$. Then $x_1 \neq x_2$, but

$$f(x_1) = f(1, 0) = 1^2 + 0 = 1 = 0^2 + 1 = f(0, 1) = f(x_2),$$

so f is not one-to-one. □

EXERCISES 6.6.12. Define $g \colon \mathbb{Z} \times \mathbb{Z} \to \mathbb{Z} \times \mathbb{Z}$ by $g(m, n) = (m + n, m - n)$.

1) Show that g is *not* onto. [*Hint:* $(m + n) + (m - n) = 2m$. Can this be odd?]

2) Show that g is one-to-one.

PROPOSITION 6.6.13. *Suppose $f \colon A \to B$. Show f is a bijection iff, for each $b \in B$, there is a* unique *$a \in A$, such that $f(a) = b$. In other words, f is a bijection if and only if*

$$\forall b \in B, \exists! \, a \in A, \big(f(a) = b\big).$$

PROOF. ($\Rightarrow$) Let b be an arbitrary element of B. Since f is a bijection, it is onto, so there exists $a \in A$, such that $f(a) = b$. All that remains is to show that a is unique. To this end, let $a' \in A$, such that $f(a') = b$. Then

$$f(a') = b = f(a).$$

Since f is a bijection, it is one-to-one, so we conclude that $a' = a$. Thus, a is unique.

($\Leftarrow$) It suffices to show that f is both onto and one-to-one.

(onto) Given $b \in B$, we are assuming that there is a (unique) element a of A, such that $f(a) = b$. Therefore f is onto.

(one-to-one) Given $a_1, a_2 \in A$, such that $f(a_1) = f(a_2)$, let $b = f(a_1)$. Then $f(a_1) = b$ and $f(a_2) = b$. From the uniqueness of the element a of A, such that $f(a) = b$, we conclude that $a_1 = a_2$. Since a_1 and a_2 are arbitrary elements of A, such that $f(a_1) = f(a_2)$, this implies that f is one-to-one. □

Remark 6.6.14. Officially, $\times$ is *not* associative, because

$$(A \times B) \times C = \{\, \big((a,b),c\big) \mid a \in A, b \in B, c \in C \,\}$$

and

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

are (usually) not the same sets: an element of $(A \times B) \times C$ must have an ordered pair (a,b) as its first coordinate, whereas an element of $A \times (B \times C)$ can have any element of A as its first coordinate.

EXERCISE 6.6.15. Suppose A, B, and C are sets. Define

$$f\colon (A \times B) \times C \to A \times (B \times C) \quad \text{by} \quad f\big((a,b),c\big) = \big(a,(b,c)\big).$$

Show that f is a bijection.

Remark 6.6.16. One can define the Cartesian product of more than two sets. For example,

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

Although $A \times B \times C$ is not the same as $(A \times B) \times C$ or $A \times (B \times C)$, the difference between them can often be ignored in practice.

6.7. Inverse functions

Backwards poets write inverse.

Author unknown

All students of mathematics have experience with solving an equation for x. Inverse functions are a special case of this.

EXAMPLE 6.7.1. In Example 6.6.6, it was shown that $f(x) = 5x - 7$ is a bijection. A look at the proof reveals that the formula $(y+7)/5$ plays a key role. The reason this formula is so important is that (solving for x) we have

$$y = 5x - 7 \qquad \Leftrightarrow \qquad x = \frac{y+7}{5}.$$

In order to see this as an "inverse function," we translate into the language of functions, by letting $g\colon \mathbb{R} \to \mathbb{R}$ be defined by $g(y) = (y+7)/5$. Then the above assertion can be restated as:

$$y = f(x) \qquad \Leftrightarrow \qquad x = g(y). \tag{6.7.2}$$

This tells us that g does exactly the opposite of what f does: if f takes x to y, then g takes y to x. We will say that g is the "inverse" of f.

The following exercise provides a restatement of (6.7.2) that will be used in the official definition of inverse functions. However, we usually use A for the domain of a generic function (and B for the codomain), so it replaces the variables x and y with a and b.

EXERCISE 6.7.3. Suppose $f\colon A \to B$ and $g\colon B \to A$. Show that if

$$\forall a \in A, \forall b \in B, \big(b = f(a) \Leftrightarrow a = g(b)\big),$$

then

a) $g\big(f(a)\big) = a$ for all $a \in A$, and

b) $f\big(g(b)\big) = b$ for all $b \in B$.

DEFINITION 6.7.4. Suppose

- $f\colon A \to B$, and
- $g\colon B \to A$.

We say that g is the **inverse** of f iff:

a) $g\big(f(a)\big) = a$ for all $a \in A$, and

b) $f\big(g(b)\big) = b$ for all $b \in B$.

EXAMPLE 6.7.5. Suppose $z\colon S \to T$ and $k\colon T \to S$. What does it mean to say that k is the inverse of z?

SOLUTION. It means that two things are true:

a) $k\big(z(s)\big) = s$ for all $s \in S$, and

b) $z\big(k(t)\big) = t$ for all $t \in T$. □

EXERCISE 6.7.6. Suppose $c\colon U \to V$ and $d\colon V \to U$. What does it mean to say that d is the inverse of c?

NOTATION 6.7.7. The inverse of f is denoted f^{-1}.

EXAMPLE 6.7.8. Note that:

- the husband of the wife of any married man is the man himself, i.e.,

$$\mathsf{husband}\big(\mathsf{wife}(m)\big) = m,$$

and

- the wife of the husband of any married woman is the woman herself, i.e.,

$$\mathsf{wife}\big(\mathsf{husband}(w)\big) = w.$$

This means the husband function is the inverse of the wife function. That is, $\mathsf{wife}^{-1} = \mathsf{husband}$.

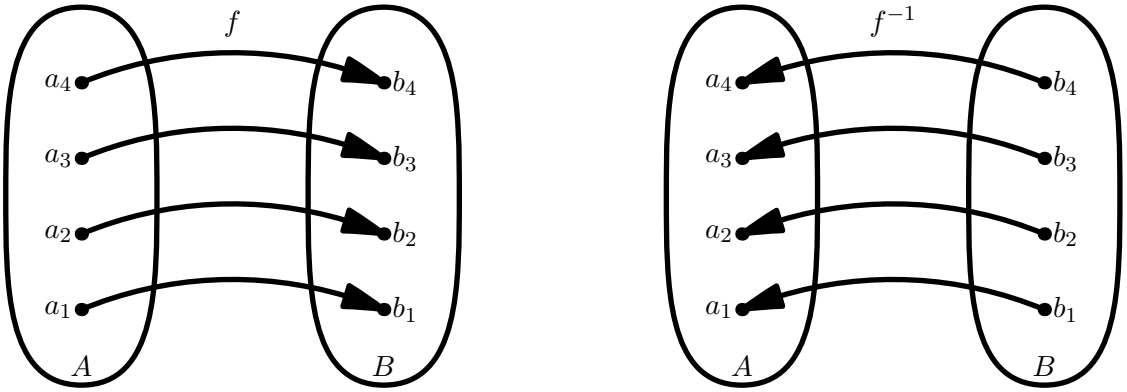
Remark 6.7.9. The inverse is easy to describe in terms of arrow diagrams. Namely, from the fact that

$$b = f(a) \quad \Leftrightarrow \quad a = f^{-1}(b),$$

we see that

$$f \text{ has an arrow from } a \text{ to } b \quad \Leftrightarrow \quad f^{-1} \text{ has an arrow from } b \text{ to } a.$$

Therefore, the arrow diagram of f^{-1} is obtained by just reversing all the arrows in the arrow diagram of f:



EXAMPLE 6.7.10. Define $f\colon \mathbb{R} \to \mathbb{R}$ and $g\colon \mathbb{R} \to \mathbb{R}$ by $f(x) = 7x - 4$ and $g(x) = (x + 4)/7$. Verify that g is the inverse of f.

PROOF. It suffices to show:

a) $g\big(f(x)\big) = x$ for all $x \in \mathbb{R}$, and

b) $f\big(g(y)\big) = y$ for all $y \in \mathbb{R}$.

(a) Given $x \in \mathbb{R}$, we have

$$g\big(f(x)\big) = \frac{f(x) + 4}{7} = \frac{\big(7x - 4\big) + 4}{7} = \frac{7x}{7} = x.$$

(b) Given $y \in \mathbb{R}$, we have

$$f\big(g(y)\big) = 7\,g(y) - 4 = 7\left(\frac{y + 4}{7}\right) - 4 = \big(y + 4\big) - 4 = y. \qquad \square$$

EXERCISES 6.7.11. In each case, verify that g is the inverse of f.

1) $f\colon \mathbb{R} \to \mathbb{R}$ is defined by $f(x) = 9x - 6$ and
$g\colon \mathbb{R} \to \mathbb{R}$ is defined by $g(x) = (x + 6)/9$.

2) $f\colon \mathbb{R}^+ \to \mathbb{R}^+$ is defined by $f(x) = x^2$ and
$g\colon \mathbb{R}^+ \to \mathbb{R}^+$ is defined by $g(x) = \sqrt{x}$.

3) $f\colon \mathbb{R}^+ \to \mathbb{R}^+$ is defined by $f(x) = 1/x$ and
$g\colon \mathbb{R}^+ \to \mathbb{R}^+$ is defined by $g(x) = 1/x$.

4) $f\colon \mathbb{R}^+ \to \mathbb{R}^+$ is defined by $f(x) = \sqrt{x + 1} - 1$ and
$g\colon \mathbb{R}^+ \to \mathbb{R}^+$ is defined by $g(x) = x^2 + 2x$.

Most functions do *not* have an inverse. In fact, only bijections have an inverse:

THEOREM 6.7.12. *Suppose $f\colon A \to B$. If f has an inverse $f^{-1}\colon B \to A$, then f is a bijection.*

PROOF. Assume there is a function $f^{-1}\colon B \to A$ that is an inverse of f. Then

a) $f^{-1}\big(f(a)\big) = a$ for all $a \in A$, and

b) $f\big(f^{-1}(b)\big) = b$ for all $b \in B$.

We wish to show that f is a bijection. This is left as an exercise for the reader. [*Hint:* This is very similar to many of the previous proofs that functions are bijections, but with the equation $a = f^{-1}(b)$ in place of an explicit formula for a. For example, if $f(a_1) = f(a_2)$, then $f^{-1}\big(f(a_1)\big) = f^{-1}\big(f(a_2)\big)$. What is each side of this equation equal to?] $\square$

EXERCISES 6.7.13.

1) Prove that the inverse of a bijection is a bijection.

2) Prove the converse of Exercise 6.7.3.

3) Show that the inverse of a function is *unique*: if g_1 and g_2 are inverses of f, then $g_1 = g_2$. (This is why we speak of *the* inverse of f, rather than *an* inverse of f.)

Remark 6.7.14. If f is a function that has an inverse, then it is easy to find f^{-1} as a set of ordered pairs. Namely,

$$f^{-1} = \{\, (b, a) \mid (a, b) \in f \,\}.$$

This is simply a restatement of the fact that

$$b = f(a) \Leftrightarrow a = f^{-1}(b)$$

(or the fact that the arrow diagram of f^{-1} is obtained by reversing the arrows in the arrow diagram of f).

EXERCISE 6.7.15. Prove the converse of Theorem 6.7.12. [*Hint:* Find f^{-1} as a set of ordered pairs.]

EXERCISE 6.7.16. Suppose $f\colon A \to B$ is a bijection. Show that the inverse of f^{-1} is f. That is, $(f^{-1})^{-1} = f$.

6.8. Composition of functions

Nothing goes by luck in composition. It allows of no tricks.
Henry David Thoreau (1817–1862), American author

The term "composition" is a name that mathematicians use for an idea that comes up fairly often in everyday life.

EXAMPLE 6.8.1.

1) The father of the mother of a person is the grandfather of the person. (To be precise, it is the *maternal* grandfather of the person — and his or her other grandfather is *paternal*.) To express the relationship in a mathematical formula, we can write:

$$\forall x, \Big(\mathsf{grandfather}(x) = \mathsf{father}\big(\mathsf{mother}(x)\big)\Big).$$

A mathematician abbreviates this formula by writing

$$\mathsf{grandfather} = \mathsf{father} \circ \mathsf{mother}$$

and says that the (maternal) grandfather function is the *composition* of father and mother.

2) The brother of the mother of a person is an uncle of the person, so uncle is the composition of brother and mother:

$$\forall x, \Big(\mathsf{uncle}(x) = \mathsf{brother}\big(\mathsf{mother}(x)\big)\Big),$$

or, more briefly,

$$\mathsf{uncle} = \mathsf{brother} \circ \mathsf{mother}.$$

(For the sake of this example, let us ignore the issue that uncle and brother are not functions, because some people have no uncle or no brother, or have more than one.)

3) The daughter of a child is a granddaughter, so granddaughter is a composition of daughter and child:

$$\mathsf{granddaughter} = \mathsf{daughter} \circ \mathsf{child}.$$

(We ignore the fact that granddaughter, daughter, and child are not functions.)

EXERCISES 6.8.2. State the usual name for each composition. (Ignore the fact that sister, daughter, and many of the other relations are not functions.)

1) husband ∘ sister =
2) husband ∘ mother =
3) husband ∘ wife =
4) husband ∘ daughter =
5) mother ∘ sister =
6) daughter ∘ sister =
7) parent ∘ parent =
8) child ∘ child =
9) parent ∘ parent ∘ parent =
10) child ∘ brother ∘ parent =

DEFINITION 6.8.3. Suppose $f\colon A \to B$ and $g\colon B \to C$. The **composition** $g \circ f$ of g and f is the function from A to C defined by

$$(g \circ f)(a) = g\big(f(a)\big) \text{ for all } a \in A.$$

EXAMPLE 6.8.4. Define $f\colon \mathbb{R} \to \mathbb{R}$ and $g\colon \mathbb{R} \to \mathbb{R}$ by $f(x) = 3x$ and $g(x) = x^2$. Then $g \circ f$ and $f \circ g$ are functions from $\mathbb{R}$ to $\mathbb{R}$. For all $x \in \mathbb{R}$, we have

$$(g \circ f)(x) = g\big(f(x)\big) = g(3x) = (3x)^2 = 9x^2$$

and

$$(f \circ g)(x) = f\big(g(x)\big) = f(x^2) = 3(x^2) = 3x^2.$$

Notice that (in this example) $f \circ g \neq g \circ f$, so *composition is* ***not*** *commutative.*

EXERCISE 6.8.5. The formulas define functions f and g from $\mathbb{R}$ to $\mathbb{R}$. Find formulas for $(f \circ g)(x)$ and $(g \circ f)(x)$.

1) $f(x) = 3x + 1$ and $g(x) = x^2 + 2$.
2) $f(x) = 3x + 1$ and $g(x) = (x - 1)/3$.
3) $f(x) = ax + b$ and $g(x) = cx + d$ (where $a, b, c, d \in \mathbb{R}$).
4) $f(x) = |x|$ and $g(x) = x^2$.
5) $f(x) = |x|$ and $g(x) = -x$.

WARNING. To calculate the value of the function $g \circ f$ at the point a, do *not* begin by calculating $g(a)$. Instead, you need to calculate $f(a)$. Then plug that value into the function g.

EXAMPLE 6.8.6. Figure 6C provides an arrow diagram to illustrate the composition $g \circ f$.

- Starting from any point of A, follow the arrow (for the function f) that starts there to arrive at some point of B.
- Then follow the arrow (for the function g) that starts there to arrive at a point of C.

For example, the f-arrow from a leads to m and the g-arrow from m leads to u. So $(g \circ f)(a) = u$.

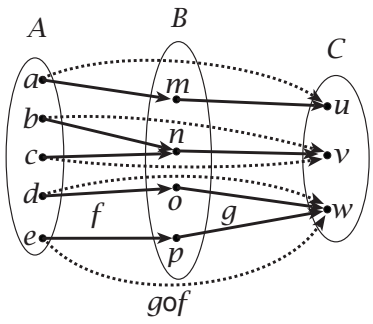


Figure 6C. Arrows for the composition $g \circ f$ are dotted.

Notice that even though g appears on the left in the expression $g \circ f$, the arrow diagram for g appears on the right in the figure. This is an unfortunate consequence of the way that we calculate $g\big(f(x)\big)$ — see the warning above.

EXERCISES 6.8.7. Let $A = \{1, 2, 3, 4\}$, $B = \{\mathsf{a}, \mathsf{b}, \mathsf{c}, \mathsf{d}\}$, and $C = \{\clubsuit, \diamondsuit, \heartsuit, \spadesuit\}$. The sets of ordered pairs in each part are functions $f\colon A \to B$ and $g\colon B \to C$. Represent $g \circ f$ as a set of ordered pairs.

1) $f = \{(1, \mathsf{a}), (2, \mathsf{b}), (3, \mathsf{c}), (4, \mathsf{d})\}$,
$g = \{(\mathsf{a}, \clubsuit), (\mathsf{b}, \diamondsuit), (\mathsf{c}, \heartsuit), (\mathsf{d}, \spadesuit)\}$

2) $f = \{(1, \mathsf{a}), (2, \mathsf{b}), (3, \mathsf{c}), (4, \mathsf{d})\}$,
$g = \{(\mathsf{a}, \clubsuit), (\mathsf{b}, \clubsuit), (\mathsf{c}, \clubsuit), (\mathsf{d}, \clubsuit)\}$

3) $f = \{(1, \mathsf{b}), (2, \mathsf{c}), (3, \mathsf{d}), (4, \mathsf{a})\}$,
$g = \{(\mathsf{a}, \clubsuit), (\mathsf{b}, \spadesuit), (\mathsf{c}, \heartsuit), (\mathsf{d}, \diamondsuit)\}$

4) $f = \{(1, \mathsf{a}), (2, \mathsf{b}), (3, \mathsf{c}), (4, \mathsf{d})\}$,
$g = \{(\mathsf{a}, \clubsuit), (\mathsf{b}, \clubsuit), (\mathsf{c}, \heartsuit), (\mathsf{d}, \spadesuit)\}$

5) $f = \{(1, \mathsf{a}), (2, \mathsf{b}), (3, \mathsf{a}), (4, \mathsf{b})\}$,
$g = \{(\mathsf{a}, \clubsuit), (\mathsf{b}, \clubsuit), (\mathsf{c}, \heartsuit), (\mathsf{d}, \spadesuit)\}$

EXERCISE 6.8.8. The definition of $g \circ f$ requires that the domain of g is equal to the codomain of f. (They are both called B in the definition, so they are required to be equal.) *Why?*

Here are some examples of proofs that combine composition with other important ideas that we have seen.

EXAMPLE 6.8.9. Suppose $f\colon A \to B$ and $g\colon B \to C$. Show that if $g \circ f$ is one-to-one, then f is one-to-one.

PROOF. Given $a_1, a_2 \in A$, such that $f(a_1) = f(a_2)$, we have

$$g\big(f(a_1)\big) = g\big(f(a_2)\big),$$

so

$$(g \circ f)(a_1) = (g \circ f)(a_2).$$

Since $g \circ f$ is one-to-one, this implies $a_1 = a_2$. So f is one-to-one. □

EXAMPLE 6.8.10. Suppose $f\colon A \to B$ and $g\colon B \to C$. Show that if f and g are onto, then $g \circ f$ is onto.

PROOF. Let c be an arbitrary element of C. Since g is onto, there is some $b \in B$, such that $g(b) = c$. Then, since f is onto, there is some $a \in A$, such that $f(a) = b$. Therefore

$$(g \circ f)(a) = g\big(f(a)\big) = g(b) = c.$$

Since c is an arbitrary element of C, this implies that $g \circ f$ is onto. □

EXAMPLE 6.8.11. Suppose $f\colon A \to B$ and $g\colon B \to C$. Show that if f and $g \circ f$ are bijections, then g is a bijection.

PROOF. It suffices to show that g is both one-to-one and onto.

(one-to-one) Let b_1 and b_2 be arbitrary elements of B, such that $g(b_1) = g(b_2)$. Since f is a bijection, it is onto, so there exist $a_1, a_2 \in A$, such that $f(a_1) = b_1$ and $f(a_2) = b_2$. Then

$$(g \circ f)(a_1) = g\big(f(a_1)\big) = g(b_1) = g(b_2) = g\big(f(a_2)\big) = (g \circ f)(a_2).$$

Since $g \circ f$ is a bijection, it is one-to-one, so we conclude that $a_1 = a_2$. Therefore

$$b_1 = f(a_1) = f(a_2) = b_2.$$

Since b_1 and b_2 are arbitrary elements of B, such that $g(b_1) = g(b_2)$, this implies that g is one-to-one.

(onto) Let c be an arbitrary element of C. Since $g \circ f$ is a bijection, it is onto, so there exists $a \in A$, such that $(g \circ f)(a) = c$. Let $b = f(a)$. Then

$$g(b) = g\big(f(a)\big) = (g \circ f)(a) = c.$$

Since c is an arbitrary element of C, we conclude that g is onto. □

EXERCISES 6.8.12. Suppose $f\colon A \to B$ and $g\colon B \to C$.

1) Show that if f and g are bijections, then $g \circ f$ is a bijection.
2) Show that if g and $g \circ f$ are bijections, then f is a bijection.
3) Show that if f and g are bijections, then $(g \circ f)^{-1} = f^{-1} \circ g^{-1}$.

EXERCISES 6.8.13. Assume $f\colon A \to B$ and $g\colon B \to A$ (and see Notation 6.6.8 for the definition of the identity maps I_A and I_B).

1) Show that g is the inverse of f if and only if $f \circ g = I_B$ and $g \circ f = I_A$.
2) What are $f \circ I_A$ and $I_B \circ f$?

EXERCISES 6.8.14.

1) Give an example of functions $f\colon A \to B$ and $g\colon B \to C$, such that $g \circ f$ is onto, but f is *not* onto. [*Hint:* Let $A = B = \mathbb{R}$, $C = [0, \infty)$, $f(x) = x^2$, and $g(x) = x^2$.]
2) Define $f\colon [0, \infty) \to \mathbb{R}$ by $f(x) = x$ and $g\colon \mathbb{R} \to \mathbb{R}$ by $g(x) = |x|$. Show that $g \circ f$ is one-to-one, but g is *not* one-to-one.
3) (*harder*) Suppose $f\colon A \to B$ and $g\colon B \to C$. Write a definition of $g \circ f$ purely in terms of sets of ordered pairs. That is, find a predicate $P(x, y)$, such that

$$g \circ f = \{\, (a, c) \in A \times C \mid P(a, c) \,\}.$$

The predicate cannot use the notation $f(x)$ or $g(x)$. Instead, it should refer to the ordered pairs that are elements of f and g.

6.9. Image and pre-image

It is sometimes necessary to collect together many values of a function. Let us start with a real-world example.

EXAMPLE 6.9.1. Suppose the astronomy club at an elementary school decides to have a Father's Day party. Then they should make a list of all of their fathers, so that invitations can be sent. Mathematically speaking, they want to make a set that contains precisely the people who are the father of someone in the club. That is, if A_1 is the set of people in the club, then they are interested in

$$\left\{\, x \;\middle|\; \exists a \in A_1, \big(x = \mathsf{father}(a)\big) \,\right\}.$$

Another way of thinking of this is that they should apply the father function to every element of the set A_1, and gather all of the resulting values into a set. The mathematical notation for this set that gathers together the values is

$$\{\, \mathsf{father}(a) \mid a \in A_1 \,\}.$$

In English, we could call this set "the fathers of the elements of A_1," but mathematicians abbreviate this to "$\mathsf{father}(A_1)$." In summary, the same set has three names:

$$\mathsf{father}(A_1) = \{\, \mathsf{father}(a) \mid a \in A_1 \,\} = \left\{\, x \;\middle|\; \exists a \in A_1, \big(x = \mathsf{father}(a)\big) \,\right\}.$$

A similar idea can be applied to any function $f\colon A \to B$. Namely, if $A_1 \subset A$, then we can apply f to every element of the set A_1, and gather all of the resulting values into a set. We call this set $f(A_1)$.

DEFINITION 6.9.2. Suppose $f\colon A \to B$, and $A_1 \subset A$. The **image** of A_1 under f is

$$f(A_1) = \{\, f(a) \mid a \in A_1 \,\}.$$

It is a subset of B. The notation means that, for all x, we have

$$x \in f(A_1) \quad \Leftrightarrow \quad \exists a \in A_1, \big(x = f(a)\big).$$

Remark 6.9.3. We can take the image of any subset of the domain of f, and the result will be some subset of the range of f. In the special case where we take the entire domain of f as our set A_1, we obtain the entire range of f as the image.

You are expected to be able to combine the definition of "image" with the proof techniques that you already know.

EXAMPLE 6.9.4. Assume $f\colon A \to B$. Show that if A_1 and A_2 are subsets of A, and f is one-to-one, then $f(A_1) \cap f(A_2) \subset f(A_1 \cap A_2)$.

PROOF. Given $b \in f(A_1) \cap f(A_2)$, we know $b \in f(A_1)$ and $b \in f(A_2)$. Therefore, since $b \in f(A_1)$, we know there is some $a_1 \in A_1$, such that $b = f(a_1)$. Also, since $b \in f(A_2)$, we know there is some $a_2 \in A_2$, such that $b = f(a_2)$. Then

$$f(a_1) = b = f(a_2).$$

Since f is one-to-one, this implies $a_1 = a_2 \in A_2$. Since we also know that $a_1 \in A_1$, this implies $a_1 \in A_1 \cap A_2$. So $f(a_1) \in f(A_1 \cap A_2)$. Since $b = f(a_1)$, this means $b \in f(A_1 \cap A_2)$. Since b is an arbitrary element of $f(A_1) \cap f(A_2)$, we conclude that $f(A_1) \cap f(A_2) \subset f(A_1 \cap A_2)$. □

EXERCISES 6.9.5. Assume $f\colon A \to B$.

1) Show that if A_1 and A_2 are subsets of A, such that $A_2 \subset A_1$, then $f(A_2) \subset f(A_1)$.

2) Assume f is one-to-one, and $a \in A$. Show that if $f(a) \in f(A_1)$, then $a \in A_1$.

Taking the image of a subset of the domain yields a subset of the codomain. Sometimes we need to go the other direction.

EXAMPLE 6.9.6. Perhaps we would like to make a list of all the people whose father is a friend of the pop singer Bono. If B_1 is the set of Bono's friends, then the mathematical notation for the set of these people is

$$\{\, x \in \mathsf{PEOPLE} \mid \mathsf{father}(x) \in B_1 \,\}.$$

Notice that if the father function had an inverse, then the same set could be obtained by applying father^{-1} to the elements of B_1. That is, the set would be $\mathsf{father}^{-1}(B_1)$. Mathematicians use this notation for the set even if there is no inverse function.

DEFINITION 6.9.7. Suppose $f\colon A \to B$, and $B_1 \subset B$. The **pre-image** (or **inverse image**) of B_1 under f is

$$f^{-1}(B_1) = \{\, a \in A \mid f(a) \in B_1 \,\}.$$

It is a subset of A. When $B_1 = \{b\}$ has only one element, we usually write $f^{-1}(b)$, instead of $f^{-1}\big(\{b\}\big)$.

WARNING. The fact that we write $f^{-1}(B_1)$ does not imply that f has an inverse, or that f^{-1} is a function. This is simply a notation that refers to the set we have defined.

EXAMPLE 6.9.8.

1) For the function $\mathsf{mother}\colon \mathsf{PEOPLE} \to \mathsf{WOMEN}$, $\mathsf{mother}^{-1}(m)$ is the set of all children of m.

2) For the function $f\colon \mathbb{R} \to \mathbb{R}$ defined by $f(x) = x^2$:

   (a) We have $f^{-1}(4) = \{2, -2\}$, because 2 and -2 are all of the square roots of 4.

   (b) We have $f^{-1}\big([0,4]\big) = [-2,2]$, because $0 \le x^2 \le 4$ iff $-2 \le x \le 2$.

Here are examples of proofs involving inverse images:

EXAMPLE 6.9.9. Suppose $f\colon A \to B$ and $B_1 \subset B$.

1) We have $f\big(f^{-1}(B_1)\big) \subset B_1$.

2) If f is onto, then $f\big(f^{-1}(B_1)\big) = B_1$.

PROOF. (1) Let $b \in f\big(f^{-1}(B_1)\big)$. By definition, we have

$$f\big(f^{-1}(B_1)\big) = \big\{\, f(a) \;\big|\; a \in f^{-1}(B_1) \,\big\},$$

so we must have $b = f(a_1)$, for some $a_1 \in f^{-1}(B_1)$. From the definition of $f^{-1}(B_1)$, we know that $f(a_1) \in B_1$. Therefore $b = f(a_1) \in B_1$. Since b is an arbitrary element of $f\big(f^{-1}(B_1)\big)$, this implies that $f\big(f^{-1}(B_1)\big) \subset B_1$, as desired.

(2) Assume f is onto. We know, from (1), that $f\big(f^{-1}(B_1)\big) \subset B_1$, so it suffices to show that $B_1 \subset f\big(f^{-1}(B_1)\big)$.

Let $b \in B_1$ be arbitrary. Because f is onto, we know there exists $a_1 \in A$, such that $f(a_1) = b$. Then $f(a_1) = b \in B_1$, so $a_1 \in f^{-1}(B_1)$. Therefore

$$f(a_1) \in \big\{\, f(a) \;\big|\; a \in f^{-1}(B_1) \,\big\} = f\big(f^{-1}(B_1)\big).$$

Since $f(a_1) = b$, we conclude that $b \in f\big(f^{-1}(B_1)\big)$. Since b is an arbitrary element of B_1, this implies that $B_1 \subset f\big(f^{-1}(B_1)\big)$, as desired. □

EXERCISES 6.9.10. Suppose that $f\colon A \to B$, that $A_1 \subset A$, and that $B_1 \subset B$.

1) Show that if $B_2 \subset B_1$, then $f^{-1}(B_2) \subset f^{-1}(B_1)$.

2) Show $A_1 \subset f^{-1}\big(f(A_1)\big)$.

EXERCISE 6.9.11. Assume $f\colon X \to Y$, $A \subset Y$, and $B \subset Y$. Show

$$f^{-1}(A) \cap f^{-1}(B) = f^{-1}(A \cap B).$$

EXERCISE 6.9.12. Assume $f\colon X \to Y$, $g\colon Y \to Z$, $X_1 \subset X$, $Z_1 \subset Z$, and $(g \circ f)(X_1) \subset Z_1$. Show $f(X_1) \subset g^{-1}(Z_1)$.

SUMMARY:

- Important definitions:
  - Cartesian product
  - function
  - domain
  - codomain, range
  - one-to-one
  - onto
  - bijection
  - identity map
  - inverse function
  - composition
  - image
  - pre-image
- Notation:
  - $A \times B$
  - $f\colon A \to B$
  - $f(a)$
  - $\forall a_1, a_2 \in A$
  - I_A
  - f^{-1}
  - $g \circ f$
  - $\{\, f(a) \mid a \in A \,\}$
  - $f(A_1)$
  - $f^{-1}(B_1)$
- A function $f\colon A \to B$ has an inverse $f^{-1}\colon B \to A$ iff f is a bijection.
- The inverse of a bijection is a bijection.

Chapter 7

Equivalence Relations

An idea which can be used once is a trick.
If it can be used more than once it becomes a method.
George Pólya (1887–1985) and Gábor Szegő (1895–1985), Hungarian mathematicians
Problems and Theorems in Analysis

7.1. Binary relations

Recall that, by definition, any function $f\colon A \to B$ is a set of ordered pairs. More precisely, each element of f is an ordered pair (a, b), such that $a \in A$ and $b \in B$. Therefore, every element of f is an element of $A \times B$, so f is a subset of $A \times B$.

Every function from A to B is a subset of $A \times B$.

EXAMPLE 7.1.1. The function mother: PEOPLE $\to$ PEOPLE is represented by the set

$$\{\, (p, m) \in \text{PEOPLE} \times \text{PEOPLE} \mid m \text{ is the mother of } p \,\}.$$

Many other relationships can also be represented by subsets of PEOPLE $\times$ PEOPLE, even though they are not functions. For example, son is not a function, because some people have more than one son (or because some people have no sons at all). However, we can represent this relation by the set

$$\{\, (p, s) \in \text{PEOPLE} \times \text{PEOPLE} \mid s \text{ is a son of } p \,\}.$$

In fact, any relationship that you can define between two people (or, to say this in the official language of logic, any binary predicate on the set PEOPLE) can be represented by a subset of PEOPLE $\times$ PEOPLE. A few examples of possible relationships are:

- x is a sister of y
- x knew y in high school
- x is taller than y
- x and y are in the same math class
- etc.

In recognition of this, mathematicians simply *define* a relation to be a set of ordered pairs; that is, a relation is any subset of $A \times B$. Unlike the case of functions, there are no restrictions — every subset is a relation.

DEFINITION 7.1.2. Suppose A and B are sets.

1) Any subset of $A \times B$ is called a **relation from A to B**.

2) For the special case where $A = B$, any subset of $A \times A$ is called a **binary relation** on A.

We will mostly be concerned with binary relations, not relations from some set A to some other set B.

EXAMPLE 7.1.3. Some examples of binary relations on PEOPLE are: brother, sister, aunt, uncle, mother, father, grandfather, cousin, etc.

DEFINITION 7.1.4. We can draw a picture to represent any given binary relation on any given set A:

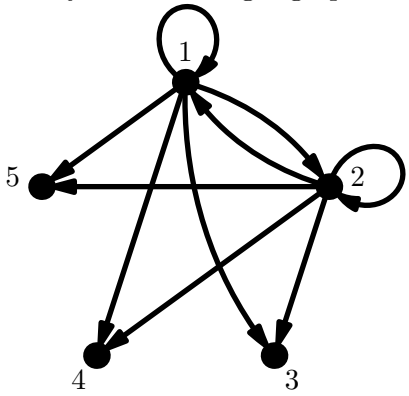
- Draw a dot for each element of A.
- For $a, b \in A$, draw an arrow from a to b if and only if (a, b) is an element of the relation.

The resulting picture is called a **digraph**. (The word is pronounced "DIE-graff" — it is short for "directed graph.")

EXAMPLE 7.1.5. Let $A = \{1, 2, 3, 4, 5\}$. We can define a binary relation R on A by letting

$$R = \{\, (x, y) \mid x^2 + y < 10 \,\}.$$

This binary relation is represented by the following digraph:



For example, note that $(x, 4) \in R$ iff $x \in \{1, 2\}$, and the digraph has arrows from 1 to 4 and from 2 to 4.

EXERCISE 7.1.6. Let B be the set consisting of you, your siblings, your parents, and your grandparents. Draw a digraph that represents each of the following binary relations on B.

1) The relation "has ever had the same last name as."

2) The relation "is a child of."

3) The relation "has ever been married to."

EXAMPLE 7.1.7. This book (like other mathematics textbooks) deals mainly with relations on sets of mathematical objects. Here are a few well-known examples:

1) The less-than relation "$<$" is a binary relation on $\mathbb{R}$.
That is, for any real numbers x and y, the assertion $x < y$ is either true or false.

2) The equality relation "$=$" is a binary relation on the entire universe of discourse $\mathcal{U}$.

3) The subset relation "$\subset$" is a binary relation on the collection of all sets in $\mathcal{U}$.

4) The relation "x is disjoint from y" is also a binary relation on the collection of all sets in $\mathcal{U}$.

NOTATION 7.1.8. Suppose R is a binary relation on a set A. For $a_1, a_2 \in A$:

1) To signify that $(a_1, a_2) \in R$, we may write $a_1 \; R \; a_2$.

2) To signify that $(a_1, a_2) \notin R$, we may write $a_1 \not R \; a_2$.

There are three basic properties that any given binary relation may or may not have:

DEFINITION 7.1.9. Suppose R is a binary relation on a set A.

1) We say that R is **reflexive** iff $\quad \forall a \in A, (a \; R \; a)$.

2) We say that R is **symmetric** iff $\quad \forall a, b \in A, \big((a \; R \; b) \Rightarrow (b \; R \; a)\big)$.

3) We say that R is **transitive** iff $\quad \forall a, b, c \in A, \Big(\big((a \; R \; b) \; \& \; (b \; R \; c)\big) \Rightarrow (a \; R \; c)\Big)$.

EXAMPLE 7.1.10.

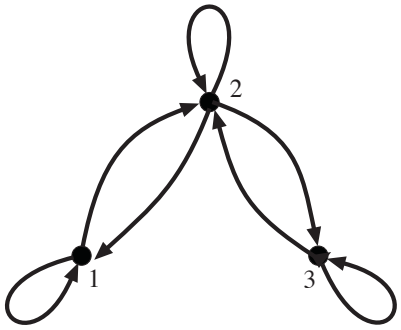
1) "=" is reflexive, symmetric, and transitive.

2) "$<$" is transitive, but neither reflexive nor symmetric.

3) "$\subset$" is transitive and reflexive, but not symmetric.

EXAMPLE 7.1.11. Consider the following binary relation R on $\{1, 2, 3\}$:

$$R = \big\{(1,1), (2,2), (3,3), (1,2), (2,1), (2,3), (3,2)\big\}$$



1) R is reflexive, because $1 \; R \; 1$, $2 \; R \; 2$, and $3 \; R \; 3$.

2) R is symmetric, because, for each $(a, b) \in R$, the reversal (b, a) is also in R.

3) R is *not* transitive, because $1 \; R \; 2$ and $2 \; R \; 3$, but $1 \not R \; 3$.

EXERCISE 7.1.12. Find binary relations on $\{1, 2, 3\}$ that are:

1) symmetric, but neither reflexive nor transitive.

2) reflexive, but neither symmetric nor transitive.

3) transitive and symmetric, but not reflexive.

4) neither reflexive, nor symmetric, nor transitive.

(*Express each relation as a set of ordered pairs, draw the corresponding digraph, and briefly justify your answers.*)

7.2. Definition and basic properties of equivalence relations

People often need to sort through a collection of objects, putting similar objects together in a group.

EXAMPLE 7.2.1. When making an inventory of the animals in a zoo, we may wish to count the number of antelopes, the number of baboons, the number of cheetahs, and so forth. In this case, all of the animals of the same species might be grouped together. Mathematically speaking, we would define a binary relation S on the set of animals in the zoo by

$$x \; S \; y \quad \text{iff} \quad x \text{ and } y \text{ are in the same species.}$$

When x and y are placed in the same group (that is, when $x \; S \; y$ in the above example), we may say that x is "equivalent" to y. This means that x and y are the same in all respects that are of interest to us. (In the above example, we are interested only in the species of an animal, not its weight, or its age, or anything else.) We call the corresponding binary relation an "equivalence relation." Thus, the binary relation S in the above example is an equivalence relation.

EXAMPLE 7.2.2. Here are additional examples:

1) If we are interested only in first names, we could define an equivalence relation N on the set of all people by

$$x \; N \; y \text{ iff } x \text{ has the same first name as } y.$$

2) Suppose a candy store has bins for different kinds of candy that they sell. To an employee restocking the bins, all pieces of candy that go into the same bin are the same, so he or she is dealing with an equivalence relation on the set of candy.

3) In geometry, one is often interested only in the shape of a triangle, not its location (or its colour, or anything else). Therefore, mathematicians define an equivalence relation $\cong$ on the set of all triangles by

$$T_1 \cong T_2 \text{ iff } T_1 \text{ is congruent to } T_2.$$

Remark 7.2.3. Suppose $\sim$ is an equivalence relation. (That is, we have $x \sim y$ iff x and y are the same in all respects that are of interest to us.) Then we would expect:

1) $\sim$ is reflexive (x is the same as x),

2) $\sim$ is symmetric (if x is the same as y, then y is the same as x), and

3) $\sim$ is transitive (if x is the same as y, and y is the same as z, then x is the same as z).

This motivates the following definition:

DEFINITION 7.2.4. An **equivalence relation** on a set A is a binary relation on A that is reflexive, symmetric, and transitive.

Remark 7.2.5. Instead of representing an equivalence relation by a letter, it is traditional to use the symbol $\sim$ (or sometimes $\equiv$ or $\cong$).

EXAMPLE 7.2.6. For any $n \in \mathbb{Z}$, we know that congruence modulo n is reflexive, symmetric, and transitive (see Exercise 5.1.18). Therefore, congruence modulo n is an equivalence relation.

EXAMPLE 7.2.7. Define a binary relation $\sim$ on $\mathbb{R}$ by $x \sim y$ iff $x^2 = y^2$. Then $\sim$ is an equivalence relation.

PROOF. We wish to show that $\sim$ is reflexive, symmetric, and transitive.

(reflexive) Given $x \in \mathbb{R}$, we have $x^2 = x^2$, so $x \sim x$.

(symmetric) Given $x, y \in \mathbb{R}$, such that $x \sim y$, we have $x^2 = y^2$. Since equality is symmetric, this implies $y^2 = x^2$, so $y \sim x$.

(transitive) Given $x, y, z \in \mathbb{R}$, such that $x \sim y$ and $y \sim z$, we have $x^2 = y^2$ and $y^2 = z^2$. Therefore $x^2 = y^2 = z^2$, so $x^2 = z^2$. Hence $x \sim z$. □

EXAMPLE 7.2.8. Define a binary relation $\sim$ on $\mathbb{N} \times \mathbb{N}$ by $(a_1, b_1) \sim (a_2, b_2)$ iff $a_1 + b_2 = a_2 + b_1$. Then $\sim$ is an equivalence relation.

PROOF. We wish to show that $\sim$ is reflexive, symmetric, and transitive.

(reflexive) Given $(a, b) \in \mathbb{N} \times \mathbb{N}$, we have $a + b = a + b$, so $(a, b) \sim (a, b)$.

(symmetric) Given $(a_1, b_1), (a_2, b_2) \in \mathbb{N} \times \mathbb{N}$, such that $(a_1, b_1) \sim (a_2, b_2)$, the definition of $\sim$ tells us that $a_1 + b_2 = a_2 + b_1$. Since equality is symmetric, this implies $a_2 + b_1 = a_1 + b_2$, so $(a_2, b_2) \sim (a_1, b_1)$.

(transitive) Given $(a_1, b_1), (a_2, b_2), (a_3, b_3) \in \mathbb{N} \times \mathbb{N}$, such that

$$(a_1, b_1) \sim (a_2, b_2) \quad \text{and} \quad (a_2, b_2) \sim (a_3, b_3),$$

we have

$$a_1 + b_2 = a_2 + b_1 \text{ and } a_2 + b_3 = a_3 + b_2. \tag{7.2.9}$$

Therefore

$$\begin{aligned} (a_1 + b_3) + (a_2 + b_2) &= (a_1 + b_2) + (a_2 + b_3) && \text{(rearrange terms)} \\ &= (a_2 + b_1) + (a_3 + b_2) && \text{(7.2.9)} \\ &= (a_3 + b_1) + (a_2 + b_2) && \text{(rearrange terms).} \end{aligned}$$

Subtracting $a_2 + b_2$ from both sides of the equation, we conclude that $a_1 + b_3 = a_3 + b_1$, so $(a_1, b_1) \sim (a_3, b_3)$. □

EXERCISES 7.2.10. Show that each of these binary relations is an equivalence relation.

1) A binary relation $\sim$ on $\mathbb{R}$ is defined by $x \sim y$ iff $x^2 - 3x = y^2 - 3y$.
2) A binary relation $\sim$ on $\mathbb{R}$ is defined by $x \sim y$ iff $x - y \in \mathbb{Z}$.
   [*Hint:* You may assume (without proof) that the negative of any integer is an integer, and that the sum of any two integers is an integer. For transitivity, notice that $x - z = (x - y) + (y - z)$.]
3) A binary relation $\sim$ on $\mathbb{N}^+ \times \mathbb{N}^+$ is defined by $(a_1, b_1) \sim (a_2, b_2)$ iff $a_1 b_2 = a_2 b_1$.
   [*Hint:* Similar to the proof in Example 7.2.8, but with multiplication in place of addition.]

Any time we have a function, we can use it to make an equivalence relation on the domain of the function:

EXAMPLE 7.2.11.

1) Every animal has only one species, so Species is a function that is defined on the set of all animals. The equivalence relation S of Example 7.2.1 can be characterized by
$$x \; S \; y \quad \text{iff} \quad \mathsf{Species}(x) = \mathsf{Species}(y).$$
2) If we assume that every person has a first name, then FirstName is a function on the set of all people. The equivalence relation N of Example 7.2.2(1) can be characterized by
$$x \; N \; y \quad \text{iff} \quad \mathsf{FirstName}(x) = \mathsf{FirstName}(y).$$

The following result generalizes this idea to all functions.

EXERCISE 7.2.12. Suppose $f\colon A \to B$. If we define a binary relation $\sim$ on A by

$$a_1 \sim a_2 \quad \text{iff} \quad f(a_1) = f(a_2),$$

then $\sim$ is an equivalence relation.

EXERCISE 7.2.13. Show that if $\sim$ is an equivalence relation on $\mathbb{R}$, then there exist $a, b \in \mathbb{R}$, such that $a \sim b$ and $a + b = 6$.

7.3. Equivalence classes

If we are interested in first names (as in Example 7.2.2(1)), then we may also be interested in the set of all people who have the same first name as you. This is called your "equivalence class."

DEFINITION 7.3.1. Suppose $\sim$ is an equivalence relation on a set A. For each $a \in A$, the **equivalence class** of a is the following subset of A:

$$[a] = \{\, a' \in A \mid a' \sim a \,\}.$$

EXAMPLE 7.3.2. For the equivalence relation N described in Example 7.2.2(1), we have

$$[\mathsf{Justin\ Timberlake}] = \{\, x \in \mathsf{People} \mid \mathsf{FirstName}(x) = \mathsf{FirstName}(\mathsf{Justin\ Timberlake}) \,\}.$$

In other words, [Justin Timberlake] is the set of all people whose first name is Justin.

WARNING. The notation $[a]$ does not tell us which equivalence relation is being used. This can be confusing if more than one equivalence relation is under consideration.

EXAMPLE 7.3.3. Suppose $A = \{1, 2, 3, 4, 5\}$ and

$$R = \left\{ \begin{matrix} (1,1), (1,3), (1,4), (2,2), (2,5), (3,1), (3,3), \\ (3,4), (4,1), (4,3), (4,4), (5,2), (5,5) \end{matrix} \right\}.$$

One can verify that R is an equivalence relation on A. The equivalence classes are:

$$[1] = \{1, 3, 4\}, \qquad [2] = \{2, 5\}, \qquad [3] = \{1, 3, 4\}$$

$$[4] = \{1, 3, 4\}, \qquad [5] = \{2, 5\}.$$

EXERCISES 7.3.4. *You do not need to show your work.*

1) Let $B = \{1, 2, 3, 4, 5\}$ and

$$S = \left\{ \begin{matrix} (1,1), (1,4), (2,2), (2,3), (3,2), \\ (3,3), (4,1), (4,4), (5,5) \end{matrix} \right\}.$$

Find the equivalence class of each element of B. (You may assume without proof that S is an equivalence relation on B.)

2) Let $C = \{1, 2, 3, 4, 5\}$ and define T by

$$x \ T \ y \text{ iff } x + y \text{ is even.}$$

Find the equivalence class of each element of C. (You may assume without proof that T is an equivalence relation on C.)

The following theorem presents some very important properties of equivalence classes:

THEOREM 7.3.5. *Suppose $\sim$ is an equivalence relation on a set A. Then:*

1) *For all $a \in A$, we have $a \in [a]$.*
2) *For all $a \in A$, we have $[a] \neq \varnothing$.*
3) *The union of the equivalence classes is all of A. That is, we have $A = \bigcup_{a \in A} [a]$, where*

$$\bigcup_{a \in A} [a] = \{\, x \mid \exists a \in A, (x \in [a]) \,\}.$$

4) *For any $a_1, a_2 \in A$, such that $a_1 \sim a_2$, we have $[a_1] = [a_2]$.*
5) *For any $a_1, a_2 \in A$, such that $a_1 \not\sim a_2$, we have $[a_1] \cap [a_2] = \varnothing$.*

EXERCISE 7.3.6. Prove Theorem 7.3.5.
[*Hint:* Use the fact that $\sim$ is reflexive, symmetric and transitive.]

Remark 7.3.7. Suppose $\sim$ is an equivalence relation on a set A. The above theorem implies that any two equivalence classes are either equal or disjoint; that is, either they have exactly the same elements, or they have no elements in common.

PROOF. Given two equivalence classes $[a_1]$ and $[a_2]$ that are not disjoint, we wish to show $[a_1] = [a_2]$. Since the equivalence classes are not disjoint, their intersection is nonempty, thus, there is some $a \in [a_1] \cap [a_2]$. Hence, $a \in [a_1]$ and $a \in [a_2]$. By definition of the equivalence classes, this means $a \sim a_1$ and $a \sim a_2$. Hence, Theorem 7.3.5(4) tells us that $[a] = [a_1]$ and $[a] = [a_2]$. Therefore $[a_1] = [a] = [a_2]$, as desired. □

7.4. Modular arithmetic

Suppose, as usual, that $\sim$ is an equivalence relation on a set A. Writing $a \sim b$ means that a is "equivalent" to b. In this case, we may want to think of a as being *equal* to b. But that would not be right, because a and b are (probably) two different things. However, we have the following fundamental property of equivalence classes:

$$a \sim b \quad \text{iff} \quad [a] = [b].$$

Thus, by putting square brackets around a and b, we can turn mere equivalence into true equality. That is what makes equivalence classes so important. A good example is provided by congruence modulo n.

7.4A. The integers modulo 3. For any $n \in \mathbb{Z}$, we know that congruence modulo n is an equivalence relation (see Exercise 5.1.18). As an example, let us consider the case where $n = 3$. To emphasize the fact that $n = 3$, we will include a subscript $_3$ in the notation for an equivalence class: we write $[k]_3$, instead of $[k]$.

We all know that when an integer is divided by 3, the remainder must be either 0, 1, or 2, so Exercise 5.1.22(1) tells us that every integer is congruent (modulo 3) to either 0, 1, or 2. Thus,

- for every $k \in \mathbb{Z}$, the equivalence class $[k]_3$ must be either $[0]_3$, $[1]_3$, or $[2]_3$.

On the other hand, it is easy to check that no two of 0, 1, and 2 are congruent (modulo 3), so

- $[0]_3$, $[1]_3$, and $[2]_3$ are three distinct equivalence classes.

Thus, we see that there are exactly three equivalence classes, namely, $[0]_3$, $[1]_3$, and $[2]_3$. The set of these equivalence classes is called the **integers modulo 3**. It is denoted $\mathbb{Z}_3$.

NOTATION 7.4.1. The notation $[k]_3$ (or even just $[k]$) is rather cumbersome. For convenience, we may write $\overline{k}$ for the equivalence class of k. Thus,

$$\mathbb{Z}_3 = \{\overline{0}, \overline{1}, \overline{2}\}.$$

DEFINITION 7.4.2. We can do arithmetic (add, subtract, and multiply) on these equivalence classes, just as we do for ordinary integers. This is called **arithmetic modulo 3**. The rules are:

- $[a]_3 + [b]_3 = [a+b]_3$ (or $\overline{a} + \overline{b} = \overline{a+b}$),
- $[a]_3 - [b]_3 = [a-b]_3$ (or $\overline{a} - \overline{b} = \overline{a-b}$), and
- $[a]_3 \times [b]_3 = [ab]_3$ (or $\overline{a} \times \overline{b} = \overline{ab}$).

(Actually, we should write $+_3$, $-_3$, and $\times_3$, to indicate that the arithmetic is being done modulo 3, but we will usually not bother.)

EXAMPLE 7.4.3. We have $[1]_3 + [2]_3 = [1+2]_3 = [3]_3$. However, since $3 \equiv 0 \pmod 3$, we have $[3]_3 = [0]_3$, so the above equation can also be written as $[1]_3 + [2]_3 = [0]_3$. Equivalently, $\overline{1} + \overline{2} = \overline{0}$.

This is an example of the following general principle:

If r is the remainder when $a+b$ is divided by 3, then $\overline{a} +_3 \overline{b} = \overline{r}$.

EXAMPLE 7.4.4. Here is a table that shows the results of addition modulo 3:

$+_3$	$\overline{0}$	$\overline{1}$	$\overline{2}$
$\overline{0}$	$\overline{0}$	$\overline{1}$	$\overline{2}$
$\overline{1}$	$\overline{1}$	$\overline{2}$	$\overline{0}$
$\overline{2}$	$\overline{2}$	$\overline{0}$	$\overline{1}$

EXERCISES 7.4.5. Make a table that shows the results of:

1) subtraction modulo 3 2) multiplication modulo 3

(*Write each of the entries of your table as either $\overline{0}$, $\overline{1}$, or $\overline{2}$.*)

7.4B. The integers modulo n. The preceding discussion can be generalized to apply with any integer n in place of 3. This results in **modular arithmetic**.

DEFINITION 7.4.6. Fix some nonzero natural number $n \in \mathbb{N}^+$.

1) For any integer k, we use $[k]_n$ to denote the equivalence class of k under congruence modulo n. When n is clear from the context, we may write $\overline{k}$, instead of $[k]_n$.
2) The set of these equivalence classes is called the **integers modulo n**. It is denoted $\mathbb{Z}_n$.
3) Addition, subtraction, and multiplication modulo n are defined by:
   - $\overline{a} +_n \overline{b} = \overline{a+b}$,
   - $\overline{a} -_n \overline{b} = \overline{a-b}$, and
   - $\overline{a} \times_n \overline{b} = \overline{ab}$.

   (When n is clear from the context, we usually write $+$, $-$, and $\times$, rather than $+_n$, $-_n$, and $\times_n$.)

Note that $\#\mathbb{Z}_n = n$. More precisely:

PROPOSITION 7.4.7. *For any $n \in \mathbb{N}^+$, we have*

$$\mathbb{Z}_n = \{\overline{0}, \overline{1}, \overline{2}, \ldots, \overline{n-1}\}$$

and $\overline{0}, \overline{1}, \overline{2}, \ldots, \overline{n-1}$ are all distinct.

EXAMPLE 7.4.8. Simplify $(\overline{17} - \overline{5}) \times (\overline{21} + \overline{11})$ in $\mathbb{Z}_7$.

SOLUTION. We have

$$\begin{aligned}(\overline{17} - \overline{5}) \times (\overline{21} + \overline{11}) &= (\overline{3} - \overline{5}) \times (\overline{0} + \overline{4}) = (\overline{3-5}) \times (\overline{0+4}) \\ &= \overline{-2} \times \overline{4} = \overline{5} \times \overline{4} = \overline{5 \times 4} = \overline{20} = \overline{6}.\end{aligned}$$

□

EXERCISES 7.4.9.

1) Simplify $(\overline{17} - \overline{5}) \times (\overline{21} + \overline{11})$ in $\mathbb{Z}_5$.

2) Simplify $\overline{32} + (\overline{23} \times \overline{16})$ in $\mathbb{Z}_9$.

3) Simplify $(\overline{25} \times \overline{35}) + (\overline{18} - \overline{12})$ in $\mathbb{Z}_{12}$.

4) Make tables that show the results of:
   (a) addition modulo 4.
   (b) subtraction modulo 5.
   (c) multiplication modulo 6.

5) Find $x, y \in \mathbb{Z}_{12}$, such that $x \neq \overline{0}$ and $y \neq \overline{0}$, but $xy = \overline{0}$.

7.5. Functions need to be well-defined

The discussion of modular arithmetic ignored a very important point: the operations of addition, subtraction, and multiplication need to be **well-defined**. That is, if $\overline{a_1} = \overline{a_2}$ and $\overline{b_1} = \overline{b_2}$, then we need to know that

1) $\overline{a_1} +_n \overline{b_1} = \overline{a_2} +_n \overline{b_2}$,

2) $\overline{a_1} -_n \overline{b_1} = \overline{a_2} -_n \overline{b_2}$, and

3) $\overline{a_1} \times_n \overline{b_1} = \overline{a_2} \times_n \overline{b_2}$.

Fortunately, these statements are all true. Indeed, they follow easily from Exercise 5.1.19:

1) Since $\overline{a_1} = \overline{a_2}$ and $\overline{b_1} = \overline{b_2}$, we have $a_1 \equiv a_2 \pmod{n}$ and $b_1 \equiv b_2 \pmod{n}$, so Exercise 5.1.19(1) tells us that $a_1 + b_1 \equiv a_2 + b_2 \pmod{n}$. Therefore $\overline{a_1 + b_1} = \overline{a_2 + b_2}$, as desired.

The proofs for $-_n$ and $\times_n$ are similar.

EXAMPLE 7.5.1. One might try to define an exponentiation operation by:

$$\overline{a} \wedge_n \overline{b} = \overline{a^b} \quad \text{for } \overline{a}, \overline{b} \in \mathbb{Z}_n.$$

Unfortunately, this does not work, because $\wedge_n$ is not well-defined:

EXERCISE 7.5.2. Find $a_1, a_2, b_1, b_2 \in \mathbb{Z}$, such that $[a_1]_3 = [a_2]_3$ and $[b_1]_3 = [b_2]_3$, but $\left[a_1^{b_1}\right]_3 \neq \left[a_2^{b_2}\right]_3$.

EXERCISES 7.5.3. Assume $m, n \in \mathbb{N}^+$.

1) Show that if $n > 2$, then absolute value does *not* provide a well-defined function from $\mathbb{Z}_n$ to $\mathbb{Z}_n$. That is, show there exist $a, b \in \mathbb{Z}$, such that $[a]_n = [b]_n$, but $\big[|a|\big]_n \neq \big[|b|\big]_n$.

2) Show that if $m \mid n$, then there is a well-defined function

$$f\colon \mathbb{Z}_n \to \mathbb{Z}_m, \text{ given by } f\big([a]_n\big) = [a]_m.$$

3) Show that if we try to define a function $g\colon \mathbb{Z}_3 \to \mathbb{Z}_2$ by $g\big([a]_3\big) = [a]_2$, then the result is *not* well-defined.

7.6. Partitions

It often happens that someone divides up a set into several disjoint subsets. This is called a "partition" of the set.

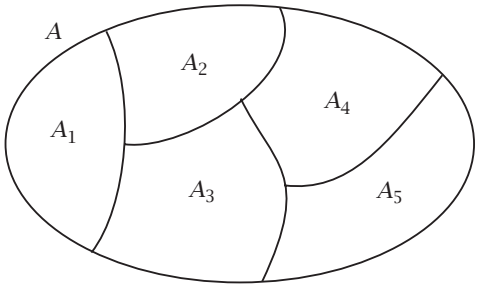


Figure 7A. A partition of A into subsets $A_1, \ldots, A_5$. (Each element of A is in one and only one of the subsets.)

EXAMPLE 7.6.1. Mary is leaving for university, and does not want her childhood toys any more, so she will divide them up among her younger siblings: Alice, Bob, and Cindy. Let

- T be the set of all of Mary's toys, and
- A, B, and C be the set of toys that she will give to Alice, to Bob, and to Cindy, respectively.

Then A, B, and C are subsets of T, and they should be chosen so that:

1) the union of A, B and C is T (that is, $A \cup B \cup C = T$), so all of the toys are given away, and

2) the sets A, B, and C are pairwise disjoint (that is, $A \cap B = \varnothing$, $A \cap C = \varnothing$, and $B \cap C = \varnothing$), so there will not be any confusion about who is the new owner of each toy.

Thus, we see that Mary should partition T into three disjoint subsets.

DEFINITION 7.6.2. A **partition** of a set A is a collection of nonempty subsets of A, such that each element of A is in exactly one of the subsets. In other words:

1) the union of the subsets in the collection is all of A, and

2) the subsets in the collection are pairwise disjoint.

EXAMPLE 7.6.3. In Example 7.6.1, the collection $\{A, B, C\}$ is a partition of T.*

EXAMPLE 7.6.4. In Example 7.3.3, the equivalence classes are $\{1, 3, 4\}$ and $\{2, 5\}$. Since $1, 2, 3, 4, 5$ each belong to exactly one of these sets, we see that the set

$$\big\{\{1, 3, 4\}, \{2, 5\}\big\}$$

of equivalence classes is a partition of $\{1, 2, 3, 4, 5\}$.

The following result is an immediate consequence of Theorem 7.3.5. It says that equivalence classes always provide a partition.

COROLLARY 7.6.5. *Suppose $\sim$ is an equivalence relation on a set A. Then*

$$\{\, [a] \mid a \in A \,\}$$

is a partition of A.

PROOF. From parts (2), (3), and (5) of Theorem 7.3.5, we know that the equivalence classes are nonempty, that their union is A, and that they are pairwise disjoint. □

Remark 7.6.6. Corollary 7.6.5 tells us that every equivalence relation gives us a partition. Conversely, the following result shows that any partition comes from an equivalence relation. Thus, equivalence relations and partitions are just two different ways of looking at the same thing.

EXERCISE 7.6.7. Suppose $\mathcal{P}$ is a partition of a set A. Define a binary relation $\sim$ on A by

$$a \sim b \quad \text{iff} \quad \exists C \in \mathcal{P}, (a \in C \text{ and } b \in C).$$

Show that:

1) $\sim$ is an equivalence relation on A, and

2) the set of equivalence classes is the partition $\mathcal{P}$.

Recall that $\mathbb{Z}_n$ replaces integers a and b that are congruent modulo n with objects $\overline{a}$ and $\overline{b}$ that are exactly equal to each other. This was achieved by letting $\mathbb{Z}_n$ be the set of all equivalence classes. The set $\mathbb{Z}_n$ applies only to congruence modulo n, but the same thing can be done for any equivalence relation:

DEFINITION 7.6.8. Suppose $\sim$ is an equivalence relation on a set A. The set of all equivalence classes is called ***A* modulo $\sim$**. It is denoted $A/\!\sim$.

EXAMPLE 7.6.9. Suppose we define an equivalence relation $\sim$ on $\mathbb{Z}$ by $a \sim b$ iff $a \equiv b \pmod{n}$. Then $\mathbb{Z}/\!\sim$ is simply another name for $\mathbb{Z}_n$.

*Actually, this may not be correct, because, for a partition, we require the sets A, B, and C to be nonempty, but it is possible that one (or more) of Mary's siblings will not be given any toys.

SUMMARY:

- Important definitions:
  - relation, binary relation
  - reflexive, symmetric, transitive
  - equivalence relation
  - equivalence class
  - modular arithmetic
  - integers modulo n
  - well-defined
  - partition
- Modular arithmetic is an important example of the use of equivalence classes.
- Functions must be well-defined.
- Every binary relation can be drawn as a digraph.
- Every partition gives rise to an equivalence relation, and vice versa.
- Notation:
  - $\sim$, $\cong$, or $\equiv$ are used for equivalence relations
  - $[a]$, or $\overline{a}$
  - $\mathbb{Z}_n$

Chapter 8

Proof by Induction

Mathematicians aren't satisfied because they know there are no solutions up to four million or four billion, they really want to know that there are no solutions up to infinity.

attributed to Andrew Wiles (1953–), British mathematician

You are familiar with many of the properties of natural numbers, such as:

- the commutative laws: $x + y = y + x$ and $xy = yx$,
- the associative laws: $(x + y) + z = x + (y + z)$ and $(xy)z = x(yz)$, and
- the distributive laws: $x(y + z) = xy + xz$ and $(y + z)x = yx + zx$.

These properties are also true for integers, for rational numbers, and for real numbers.

In this chapter, we discuss a very useful property of $\mathbb{N}$ that is **not** true of $\mathbb{Z}$ or $\mathbb{Q}$ or $\mathbb{R}$. It is often very useful for proving assertions about natural numbers, and requires an understanding of sets and predicates (which were introduced in Chapter 3), but not the full theory of First-Order Logic.

8.1. The Principle of Mathematical Induction

REMINDER 8.1.1. To say that $P(n)$ is a **predicate** of natural numbers, means that, for each natural number n, we have an assertion $P(n)$ that is either true or false. Some examples of predicates are:

- Let $P_{\text{odd}}(n)$ be the assertion "n is odd."
- Let $P_{\text{big}}(n)$ be the assertion "$n > 1000$."
- Let $P_{\text{square}}(n)$ be the assertion "$\exists k \in \mathbb{N}, (n = k^2)$."
- Let $P_{\text{prime}}(n)$ be the assertion "n is a prime number."

Mathematicians accept the truth of the following assertion as a basic fact about the natural numbers:

AXIOM 8.1.2 (Principle of Mathematical Induction). *Suppose $P(n)$ is a predicate of natural numbers. If*

(i) *$P(1)$ is true, and*

(ii) *for every $k \geq 2$, $\big(P(k-1) \Rightarrow P(k)\big)$,*

then $P(n)$ is true for all $n \in \mathbb{N}^+$.

Although we cannot *prove* Axiom 8.1.2, it can be given an informal justification that may convince you to accept it as a valid property of $\mathbb{N}$:

INFORMAL JUSTIFICATION OF AXIOM 8.1.2. Let n be an arbitrary element of $\mathbb{N}^+$.

- From (i), we know $P(1)$ is true.
- From (ii), we know $P(2-1) \Rightarrow P(2)$ is true.
  Since $P(2-1) = P(1)$ is true, we conclude, by $\Rightarrow$-elimination, that $P(2)$ is true.
- From (ii), we know $P(3-1) \Rightarrow P(3)$ is true.
  Since $P(3-1) = P(2)$ is true, we conclude, by $\Rightarrow$-elimination, that $P(3)$ is true.

  $\vdots$
- From (ii), we know $P(n-1) \Rightarrow P(n)$ is true.
  Since $P(n-1)$ is true, we conclude, by $\Rightarrow$-elimination, that $P(n)$ is true.

Since n is an arbitrary element of $\mathbb{N}^+$, we conclude that $P(n)$ is true for all $n \in \mathbb{N}^+$. □

TERMINOLOGY 8.1.3.

- In a proof using Mathematical Induction, establishing (i) is called the **base case**, and establishing (ii) is the **induction step**.
- In the induction step, we are proving $P(k-1) \Rightarrow P(k)$, so we assume that $P(k-1)$ is true (and establish $P(k)$). This assumption $P(k-1)$ is called the **induction hypothesis**.

Here is an example of how mathematical induction can be used.

PROPOSITION 8.1.4. *For every $n \in \mathbb{N}^+$, we have* $1+2+3+\cdots+n = \dfrac{n(n+1)}{2}$.

PROOF BY INDUCTION. Define $P(n)$ to be the assertion

$$1+2+3+\cdots+n = \frac{n(n+1)}{2}.$$

(i) *Base case.* For $n = 1$, we have

$$1+2+3+\cdots+n = 1 \quad \text{and} \quad \frac{n(n+1)}{2} = \frac{1(1+1)}{2} = 1.$$

Since these are equal, $P(1)$ is true.

(ii) *Induction step.* Assume $P(k-1)$ is true (and $k \geq 2$). This means that

$$1+2+3+\cdots+(k-1) = \frac{(k-1)\big((k-1)+1\big)}{2}.$$

Hence

$$\begin{aligned}
&1+2+3+\cdots+k \\
&\quad = \big(1+2+3+\cdots+(k-1)\big)+k \\
&\quad = \frac{(k-1)\big((k-1)+1\big)}{2}+k && \text{(Induction Hypothesis)} \\
&\quad = \frac{(k-1)k}{2}+k \\
&\quad = k\left(\frac{k-1}{2}+1\right) \\
&\quad = k\left(\frac{k+1}{2}\right) \\
&\quad = \frac{k(k+1)}{2},
\end{aligned}$$

so $P(k)$ is true.

Therefore, by the Principle of Mathematical Induction, we know $P(n)$ is true for all n. This means

$$1 + 2 + 3 + \cdots + n = \frac{n(n+1)}{2}$$

for every $n \in \mathbb{N}^+$. □

Remark 8.1.5. A proof by induction is often used to show that two functions $f(n)$ and $g(n)$ are equal. (Proposition 8.1.4 is an example of this, with $f(n) = 1 + 2 + 3 + \cdots + n$ and $g(n) = n(n+1)/2$.) The base case is usually easy: calculate $f(1)$ and $g(1)$, then notice that they are equal. On the other hand, it is usually not immediately obvious how to do the induction step, so it is a good idea to start by doing some scratch work:

- Write down the desired equality $f(k) \stackrel{?}{=} g(k)$.
- Then use algebraic simplifications to arrive at a true statement. At some point in these manipulations, you will use the induction hypothesis to replace $f(k-1)$ with $g(k-1)$, or vice-versa.

A proof can then be obtained by rewriting these algebraic steps in a logical order (preferably, as a "one-line proof" — a string of equalities that starts with $f(k)$ and ends with $g(k)$).

For example, the scratch work that led to the above proof of Proposition 8.1.4 can be found in Figure 8A below. In this case, the algebraic manipulations are fairly simple, but some problems are considerably more difficult.

$$1 + 2 + 3 + \cdots + k \stackrel{?}{=} \frac{k(k+1)}{2}$$

$$\big(1 + 2 + 3 + \cdots + (k-1)\big) + k \stackrel{?}{=} \frac{k(k+1)}{2}$$

$$\frac{(k-1)\big((k-1)+1\big)}{2} + k \stackrel{?}{=} \frac{k(k+1)}{2} \qquad \begin{pmatrix}\text{Induction}\\ \text{Hypothesis}\end{pmatrix}$$

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

$$\left(\frac{(k-1)}{2} + 1\right) k \stackrel{?}{=} \frac{k(k+1)}{2}$$

$$\left(\frac{(k+1)}{2}\right) k \stackrel{?}{=} \frac{k(k+1)}{2} \qquad \checkmark$$

Figure 8A. Scratch work for the proof of Proposition 8.1.4.

Here is another example that is fairly straightforward:

PROPOSITION 8.1.6. *For every $n \in \mathbb{N}^+$, we have*

$$3 + 7 + 11 + \cdots + (4n-1) = 2n^2 + n.$$

PROOF BY INDUCTION. Define $P(n)$ to be the assertion

$$3 + 7 + 11 + \cdots + (4n - 1) = 2n^2 + n.$$

(i) *Base case.* For $n = 1$, we have

$$3 + 7 + 11 + \cdots + (4n - 1) = 3 \quad \text{and} \quad 2n^2 + n = 2(1^2) + 1 = 3.$$

Since these are equal, $P(1)$ is true.

(ii) *Induction step.* Assume $P(k - 1)$ is true (and $k \geq 2$). This means that

$$3 + 7 + 11 + \cdots + \big(4(k - 1) - 1\big) = 2(k - 1)^2 + (k - 1).$$

Hence

$$\begin{aligned}
&3 + 7 + 11 + \cdots + (4k - 1) \\
&\quad = \Big(3 + 7 + 11 + \cdots + \big(4(k - 1) - 1\big)\Big) + (4k - 1) \\
&\quad = \Big(2(k - 1)^2 + (k - 1)\Big) + (4k - 1) && \text{(Induction Hypothesis)} \\
&\quad = \Big(2(k^2 - 2k + 1) + (k - 1)\Big) + (4k - 1) \\
&\quad = (2k^2 - 4k + 2) + (k - 1) + (4k - 1) \\
&\quad = 2k^2 + k,
\end{aligned}$$

so $P(k)$ is true.

Therefore, by the Principle of Mathematical Induction, we know $P(n)$ is true for all n. This means

$$3 + 7 + 11 + \cdots + (4n - 1) = 2n^2 + n$$

for every $n \in \mathbb{N}^+$. □

EXERCISES 8.1.7. Prove each formula by Mathematical Induction.

1) $2 + 4 + 6 + 8 + \cdots + 2n = n(n + 1)$.

2) $1 + 3 + 5 + 7 + \cdots + (2n - 1) = n^2$.

3) $2 + 7 + 12 + 17 + \cdots + (5n - 3) = \dfrac{n(5n - 1)}{2}$.

NOTATION 8.1.8. It is often necessary to add up a long list of numbers (as in the above exercises), so it is convenient to have a good notation for this: if $a_1, a_2, \ldots, a_n$ is any sequence of numbers, then the sum $a_1 + a_2 + \cdots + a_n$ can be denoted by

$$\sum_{k=1}^{n} a_k.$$

(The symbol $\sum$ is a capital sigma, the Greek version of the letter S — it stands for "sum.")

EXAMPLE 8.1.9. Let $a_1, a_2, \ldots, a_n$ be a sequence of numbers. Then:

1) $\displaystyle\sum_{k=1}^{1} a_k = a_1, \qquad \sum_{k=1}^{2} a_k = a_1 + a_2, \qquad \text{and} \qquad \sum_{k=1}^{3} a_k = a_1 + a_2 + a_3.$

2) $\displaystyle\sum_{k=1}^{1} k = 1, \qquad \sum_{k=1}^{2} k = 1 + 2 = 3, \qquad \sum_{k=1}^{3} k = 1 + 2 + 3 = 6.$

3) $\displaystyle\sum_{k=1}^{n} k = 1 + 2 + 3 + \cdots + n.$

4) For any $n \in \mathbb{N}^+$, we have $\displaystyle\sum_{k=1}^{n} a_k = \left(\sum_{k=1}^{n-1} a_k\right) + a_n.$

EXERCISES 8.1.10.

1) In Exercise 8.1.7, the left-hand side of each formula is a sum. Write each of these sums in $\sum$-notation.

2) Show that (1) and (4) of Example 8.1.9 imply $\displaystyle\sum_{k=1}^{0} a_k = 0.$

Remark 8.1.11. In the Induction Step of a proof by induction, we wish to prove

$$\forall k \geq 2, \big(P(k-1) \Rightarrow P(k)\big).$$

Since k is a bound ("dummy") variable in this assertion, there is no harm in replacing it with a different letter: for example, if you prefer, it is perfectly acceptable to prove, say,

$$\forall i \geq 2, \big(P(i-1) \Rightarrow P(i)\big), \quad \text{or} \quad \forall n \geq 2, \big(P(n-1) \Rightarrow P(n)\big).$$

This is important to keep in mind when the variable k is already being used for something else.

EXAMPLE 8.1.12. Show that $\displaystyle\sum_{k=1}^{n}(2k-5) = n^2 - 4n.$

PROOF BY INDUCTION. Define $P(n)$ to be the assertion

$$\sum_{k=1}^{n}(2k-5) = n^2 - 4n.$$

(i) *Base case.* For $n = 1$, we have

$$\sum_{k=1}^{n}(2k-5) = \sum_{k=1}^{1}(2k-5) = 2(1) - 5 = -3 = 1^2 - 4(1) = n^2 - 4n.$$

So $P(1)$ is true.

(ii) *Induction step.* Assume $n \geq 2$ and $P(n-1)$ is true. This means that

$$\sum_{k=1}^{n-1}(2k-5) = (n-1)^2 - 4(n-1).$$

Hence

$$\begin{aligned}
\sum_{k=1}^{n}(2k-5) &= \left(\sum_{k=1}^{n-1}(2k-5)\right) + (2n-5) \\
&= \big((n-1)^2 - 4(n-1)\big) + (2n-5) && \text{(Induction Hypothesis)} \\
&= \big((n^2 - 2n + 1) - 4n + 4\big) + (2n-5) \\
&= n^2 - 4n,
\end{aligned}$$

so $P(n)$ is true.

Therefore, by the Principle of Mathematical Induction, we conclude that $P(n)$ is true for every $n \in \mathbb{N}^+$. □

EXERCISES 8.1.13. Prove each formula by induction.

1) $\sum_{k=1}^{n}(6k+7) = 3n^2 + 10n.$

2) $\sum_{k=1}^{n}(4k-5) = 2n^2 - 3n.$

3) $\sum_{k=1}^{n}(12k-19) = 6n^2 - 13n.$

4) $\sum_{k=1}^{n}(3k+11) = \dfrac{3n^2+25n}{2}.$

5) $\sum_{k=1}^{n}3^k = \dfrac{3^{n+1}-3}{2}.$

6) $\sum_{k=0}^{n}ar^k = a\dfrac{r^{n+1}-1}{r-1}.$

7) $\sum_{k=1}^{n}k^2 = \dfrac{n(n+1)(2n+1)}{6}.$

8) (harder) $\sum_{k=1}^{n}k^3 = \left(\dfrac{n(n+1)}{2}\right)^2.$

Remark 8.1.14. If you wish to prove that $P(k)$ is true for all k, then the Principle of Induction can be applied with k in the role of n. This is called "inducting on k." Similarly, any other letter can be used in place of n.

EXAMPLE 8.1.15. Show, for all $k \in \mathbb{N}^+$, that

$$\sum_{i=1}^{k}(3i^2 - 3i + 1) = k^3.$$

PROOF BY INDUCTION. Define $P(k)$ to be the assertion

$$\sum_{i=1}^{k}(3i^2 - 3i + 1) = k^3.$$

(i) *Base case.* For $k = 1$, we have

$$\sum_{i=1}^{k}(3i^2 - 3i + 1) = \sum_{i=1}^{1}(3i^2 - 3i + 1) = 3(1)^2 - 3(1) + 1 = 1 = 1^3 = k^3.$$

So $P(1)$ is true.

(ii) *Induction step.* Assume $k \geq 2$ and $P(k-1)$ is true. This means that

$$\sum_{i=1}^{k-1}(3i^2 - 3i + 1) = (k-1)^3.$$

Hence

$$\begin{aligned}
\sum_{i=1}^{k}(3i^2 - 3i + 1) &= \left(\sum_{i=1}^{k-1}(3i^2 - 3i + 1)\right) + (3k^2 - 3k + 1) \\
&= (k-1)^3 + (3k^2 - 3k + 1) && \text{(Induction Hypothesis)} \\
&= (k^3 - 3k^2 + 3k - 1) + (3k^2 - 3k + 1) \\
&= k^3,
\end{aligned}$$

so $P(k)$ is true.

By the Principle of Mathematical Induction, we conclude that $P(k)$ is true for all $k \in \mathbb{N}^+$. □

8.2. Other proofs by induction

Not all proofs by induction are about sums:

EXAMPLE 8.2.1. Suppose $a, b, n \in \mathbb{Z}$, with $a \equiv b \pmod{n}$. Show $a^k \equiv b^k \pmod{n}$, for all $k \in \mathbb{N}^+$.

PROOF BY INDUCTION. We induct on k. Define $P(k)$ to be the assertion

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

(i) *Base case.* Since $a^1 = a$ and $b^1 = b$, the hypothesis $a \equiv b \pmod{n}$ tells us that

$$a^1 \equiv b^1 \pmod{n},$$

so $P(1)$ is true.

(ii) *Induction step.* Assume $P(k-1)$ is true. This means that

$$a^{k-1} \equiv b^{k-1} \pmod{n}.$$

By assumption, we also have

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

Exercise 5.1.19(3) tells us that the product of congruent quantities is congruent, so we can multiply the above congruences, to conclude that

$$(a^{k-1})(a) \equiv (b^{k-1})(b) \pmod{n}.$$

In other words,

$$a^k \equiv b^k \pmod{n},$$

so $P(k)$ is true.

Therefore, by the Principle of Mathematical Induction, $P(k)$ is true for every $k \in \mathbb{N}^+$. □

EXERCISES 8.2.2. Prove each of the following assertions by induction.

1) $5^k \equiv 5 \pmod{4}$, for every $k \in \mathbb{N}^+$.

2) $n^3 \equiv n \pmod{3}$ for every $n \in \mathbb{N}^+$.

The Principle of Mathematical Induction is an important tool for proving things about sequences of numbers in which each term is defined from preceding terms. (Such sequences are said to be defined "**recursively**" or "inductively.") Fibonacci numbers are one famous example of this. In this case, each term is the sum of the two preceding terms:

DEFINITION 8.2.3. The **Fibonacci numbers** $F_1, F_2, F_3, \ldots$ are defined by:

- $F_1 = 1$,
- $F_2 = 1$, and
- $F_n = F_{n-1} + F_{n-2}$ for $n \geq 3$.

(For example, $F_3 = F_{3-1} + F_{3-2} = F_2 + F_1 = 1 + 1 = 2$.) In general, each Fibonacci number (after F_2) is the sum of the two preceding Fibonacci numbers, so the first few Fibonacci numbers are:

n	1	2	3	4	5	6	7	$\cdots$
F_n	1	1	2	3	5	8	13	$\cdots$

EXAMPLE 8.2.4. Prove $\displaystyle\sum_{k=1}^{n} F_k = F_{n+2} - 1$ for all $n \in \mathbb{N}^+$.

PROOF BY INDUCTION. Define $P(n)$ to be the assertion

$$\sum_{k=1}^{n} F_k = F_{n+2} - 1.$$

(i) *Base case.* For $n = 1$, we have

$$\sum_{k=1}^{n} F_k = \sum_{k=1}^{1} F_k = F_1 = 1 = 2 - 1 = F_3 - 1 = F_{1+2} - 1 = F_{n+2} - 1,$$

so $P(1)$ is true.

(ii) *Induction step.* Assume $P(n-1)$ is true (and $n \geq 2$). Then

$$\begin{aligned}
\sum_{k=1}^{n} F_k &= \left(\sum_{k=1}^{n-1} F_k\right) + F_n \\
&= \Big(F_{(n-1)+2} - 1\Big) + F_n && \text{(Induction Hypothesis)} \\
&= (F_{n+1} - 1) + F_n \\
&= (F_{n+1} + F_n) - 1 \\
&= F_{n+2} - 1 && \left(\begin{array}{c}\text{definition of}\\ \text{Fibonacci number}\end{array}\right).
\end{aligned}$$

Therefore, by the Principle of Mathematical Induction, $P(n)$ is true for every n. This means $\sum_{k=1}^{n} F_k = F_{n+2} - 1$ for all $n \in \mathbb{N}^+$. □

EXERCISES 8.2.5. Prove each assertion by induction.

1) $\sum_{k=1}^{n} F_k^2 = F_n F_{n+1}$.

2) F_{3k} is even, for all $k \in \mathbb{N}^+$.

3) F_{4k} is divisible by 3, for all $k \in \mathbb{N}^+$.

Induction can also be applied to other sequences that are defined recursively:

EXAMPLE 8.2.6. Define a sequence $\{a_n\}$ by:

- $a_1 = 1$, and
- $a_n = 2a_{n-1} + 1$ for $n \geq 2$.

Show $a_n = 2^n - 1$ for all $n \in \mathbb{N}^+$.

PROOF BY INDUCTION. Define $P(n)$ to be the assertion

$$a_n = 2^n - 1.$$

(i) *Base case.* For $n = 1$, we have

$$a_n = a_1 = 1 \quad \text{and} \quad 2^n - 1 = 2^1 - 1 = 2 - 1 = 1.$$

Since these are equal, we know that $P(1)$ is true.

(ii) *Induction step.* Assume $P(k-1)$ is true (and $k \geq 2$). This means that

$$a_{k-1} = 2^{k-1} - 1.$$

Then

$$\begin{aligned} a_k &= 2a_{k-1} + 1 && \text{(definition of } a_k\text{)} \\ &= 2\Big(2^{k-1} - 1\Big) + 1 && \text{(Induction Hypothesis)} \\ &= \Big(2^k - 2\Big) + 1 \\ &= 2^k - 1, \end{aligned}$$

so $P(k)$ is true.

Therefore, by the Principle of Mathematical Induction, $P(n)$ is true for every n. This means $a_n = 2^n - 1$ for all $n \in \mathbb{N}^+$. □

Historical remark 8.2.7. The Italian mathematician Fibonacci discovered the Fibonacci numbers in 1202, as the answer to a problem about the population growth of rabbits. Namely, assume that:

- You start with 1 pair of newborn rabbits (a male and a female).
- Each female gives birth to another pair of rabbits each month (a male and a female), starting when she is two months old. (And rabbits never die.)

This means that the pairs of rabbits you have in the nth month consist of the pairs you already had last month, plus one new pair for each pair that you had two months ago. Therefore, if F_n is the number of pairs in the nth month, then $F_n = F_{n-1} + F_{n-2}$. You can read more about Fibonacci numbers on *Wikipedia*.

EXERCISES 8.2.8.

1) Define a sequence $\{b_n\}$ by:
   - $b_1 = 4$, and
   - $b_n = 3b_{n-1} - 2$ for $n \geq 2$.

   Show $b_n = 3^n + 1$ for all $n \in \mathbb{N}^+$.

2) Define a sequence $\{c_n\}$ by:
   - $c_1 = 25$, and
   - $c_n = 4c_{n-1} + 5^n$ for $n \geq 2$.

   Show $c_n = 5^{n+1}$ for all $n \in \mathbb{N}^+$.

3) Define a sequence $\{d_n\}$ by:
   - $d_1 = 3$, and
   - $d_n = 2d_{n-1} - n + 2$ for $n \geq 2$.

   Show $d_n = 2^n + n$ for all $n \in \mathbb{N}^+$.

4) Define a sequence $\{e_n\}$ by:
   - $e_1 = 2$, and
   - $e_n = 2e_{n-1} - n + 1$ for $n \geq 2$.

   Show $e_n = n + 1$ for all $n \in \mathbb{N}^+$.

Induction is not only for proving that things are equal. For example, it can also be used to prove inequalities:

EXAMPLE 8.2.9. Prove $2^n \geq n$ for all $n \in \mathbb{N}^+$.

Scratchwork.

$$\begin{aligned} 2^n &\overset{?}{>} n \\ 2 \cdot 2^{n-1} &\overset{?}{>} n \\ 2(n-1) &\overset{?}{\geq} n && \text{(Induction Hypothesis)} \\ n &\overset{?}{\geq} 2 && \checkmark \end{aligned}$$

PROOF BY INDUCTION. Define $P(n)$ to be the assertion

$$2^n > n.$$

(i) *Base case.* For $n = 1$, we have

$$2^n = 2^1 = 2 > 1 = n,$$

so $P(1)$ is true.

(ii) *Induction step.* Assume $P(n-1)$ is true (and $n \geq 2$). This means that

$$2^{n-1} > n - 1.$$

Then

$$\begin{aligned} 2^n &= 2 \cdot 2^{n-1} \\ &> 2(n-1) && \text{(Induction Hypothesis)} \\ &= n + (n-2) \\ &\geq n + 0 && (n \geq 2, \text{ so } n - 2 \geq 0) \\ &= n, \end{aligned}$$

so $P(n)$ is true.

Therefore, by the Principle of Mathematical Induction, $P(n)$ is true for every natural number n. □

EXERCISES 8.2.10.

1) Prove $3^n \geq 3n$ for all $n \in \mathbb{N}^+$. [*Hint:* Note that $3^n - 3^{n-1} > 3n - 3(n-1)$ if $n \geq 2$.]

2) Prove $(1+x)^n \geq 1 + nx$ for all $x \in \mathbb{R}^+$ and all $n \in \mathbb{N}^+$.

Here is a standard piece of advice:

SUGGESTION 8.2.11. *Whenever you need to prove a statement with an n in it, you should consider using induction.*

EXERCISE 8.2.12 (*assumes familiarity with polynomials***).** Prove by induction on n that the polynomial $x^n - y^n$ is divisible by $x - y$, for all $n \in \mathbb{N}^+$. [*Hint:* What is $(x-y)x^n + y(x^n - y^n)$?]

EXERCISE 8.2.13 (*assumes familiarity with commutative groups***).** Suppose $(G, +)$ is a commutative group. For $g \in G$ and $n \in \mathbb{N}^+$, we define ng recursively, by:

$$1g = g \quad \text{and} \quad (n+1)g = ng + g.$$

Prove by induction on n that $(m+n)g = mg + ng$ for all $m, n \in \mathbb{N}^+$.

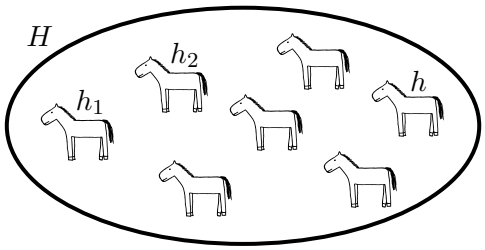
EXERCISE 8.2.14. Explain what is wrong with the following famous "proof" that all horses have the same colour.

ATTEMPT AT A PROOF BY INDUCTION. Define $P(n)$ to be the assertion

"In every set of n horses, all of the horses have the same colour."

(i) *Base case.* For $n = 1$, let H be any set of n horses. Since $n = 1$, there is only one horse in H, so it is obvious that all of the horses in H have the same colour.

(ii) *Induction step.* Assume, for every set of $n - 1$ horses, that all of the horses have the same colour (and $n \geq 2$). Let H be any set of n horses.



Remove one horse h_1 from H to form a set H_1 of $n-1$ horses. By the induction hypothesis, we know that

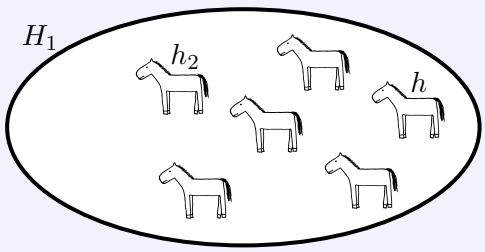
(8.2.15) all of the horses in H_1 have the same colour.

Now, remove some other horse h_2 from H to form a different set H_2 of $n-1$ horses. By applying the induction hypothesis again, we know that

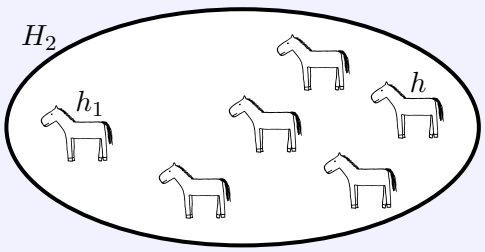
(8.2.16) all of the horses in H_2 have the same colour.

Now, choose h to be some other horse (neither h_1 nor h_2). Since $h \neq h_1$, we know $h \in H_1$, so, from (8.2.15), we know that all of the horses in H_1 have the same colour as h. Similarly, since $h \neq h_2$, we know $h \in H_2$, so, from (8.2.16), we know that all of the horses in H_2 also have the same colour as h. This means all of the horses in $H_1 \cup H_2$ have the same colour (namely, the colour of horse h). Since it is clear that $H = H_1 \cup H_2$ (because H_1 contains every horse except h_1, which is in H_2), we conclude that all of the horses in H have the same colour.

By the Principle of Mathematical Induction, we conclude that, in every (finite) set of horses, all of the horses have the same colour. □



All of the horses in H_1 have the same colour, so they all have the same colour as horse h.



All of the horses in H_2 have the same colour, so they all have the same colour as horse h.

Since every horse is in either H_1 or H_2, this implies that all of the horses have the same colour as h. So all of the horses have the same colour.

WARNING. Mathematical induction is a method that is used extensively by mathematicians and computer scientists. However, other scientists (and also philosophers) use the word "induction" to refer to a quite different method of reasoning: *scientific induction* (or "inductive reasoning") is the process of deriving a general rule from specific examples. (It is the opposite of deductive reasoning, where specific conclusions are derived from general rules.) For example, a scientist might measure the length and the width of very many rectangles, and compare with the areas of the rectangles. He or she would find that the area always came out to be the product of the length with the width. The scientist would then conclude (by inductive reasoning) that the area of every rectangle is the product of its length and its width. However, this does *not* constitute a *mathematical proof* of the formula for the area of a rectangle.

8.3. Other versions of induction

It is sometimes difficult to apply the Principle of Mathematical Induction in the form we have stated in Axiom 8.1.2. The following proposition provides some alternative versions that are more useful in some of those situations. All of them follow quite easily from the approach using "well-ordering" that will be discussed in the following section.

PROPOSITION 8.3.1. *Suppose $P(n)$ is a predicate of natural numbers, and $m \in \mathbb{N}^+$.*

1) (Strong induction) *If*
   (i) $P(1)$ *is true, and*
   (ii) *for every* $n \geq 2$,
   $$\Big(\big(\textit{for every } k \in \{1, 2, \ldots, n-1\}, P(k)\big) \Rightarrow P(n)\Big),$$
   then $P(n)$ is true for all $n \in \mathbb{N}^+$.

2) (Generalized induction) *If*
   (i) $P(m)$ *is true, and*
   (ii) *for every* $n > m$, $\big(P(n-1) \Rightarrow P(n)\big)$,
   then $P(n)$ is true for all $n \geq m$.

3) (Strong induction with multiple base cases) *If*
   (i) $P(k)$ *is true for all* $k \in \{1, 2, \ldots, m\}$, *and*
   (ii) *for every* $n > m$,
   $$\Big(\big(\textit{for every } k \in \{1, 2, \ldots, n-1\}, P(k)\big) \Rightarrow P(n)\Big),$$
   then $P(n)$ is true for all $n \in \mathbb{N}^+$.

4) *If*
   (i) $P(1)$ *is true, and*
   (ii) *for every* $k \in \mathbb{N}^+$, $P(k) \Rightarrow P(k+1)$,
   then $P(n)$ is true for all $n \in \mathbb{N}^+$.

5) *Suppose $S \subset \mathbb{N}^+$. If*
   (i) $1 \in S$, *and*
   (ii) *for every* $n \in S$, $\big(n+1 \in S\big)$,
   then $S = \mathbb{N}^+$.

Remark 8.3.2. There are many other versions of induction. For example, if you wish to prove that $P(n)$ is true for all $n \in \mathbb{N}$ (rather than only for all $n \in \mathbb{N}^+$), then

i) your base case would be to prove $P(0)$, and

ii) your induction step would be to prove $P(n-1) \Rightarrow P(n)$, for all $n \geq 1$.

EXAMPLE 8.3.3. Prove $F_n < 2^n$, for every $n \in \mathbb{N}^+$.

PROOF BY INDUCTION. Define $P(n)$ to be the assertion

$$F_n < 2^n.$$

We use strong induction with 2 base cases.

(i) *Base cases.* We have

$$F_1 = 1 < 2 = 2^1,$$

and

$$F_2 = 1 < 4 = 2^2,$$

so $P(1)$ and $P(2)$ are true.

(ii) *Induction step.* Assume $n \geq 3$, and that $P(n-1)$ and $P(n-2)$ are true. We have

$$\begin{aligned} F_n &= F_{n-1} + F_{n-2} \\ &< 2^{n-1} + 2^{n-2} && \text{(Induction Hypotheses)} \\ &< 2^{n-1} + 2^{n-1} \\ &= 2^n, \end{aligned}$$

so $P(n)$ is true.

By the Principle of Mathematical Induction (in the form of strong induction with multiple base cases), we conclude that $P(n)$ is true for all $n \in \mathbb{N}^+$. □

8.4. The natural numbers are well-ordered

Induction is a very powerful tool, but it is sometimes hard to apply (and there are many different versions to keep track of, as we saw in the preceding section). In this section, we demonstrate a closely related technique that is often less cumbersome to use.

DEFINITION 8.4.1. Let $S \subset \mathbb{R}$ and $a \in \mathbb{R}$. We say a is the **smallest** element of S iff:

- $a \in S$, and
- $\forall s \in S$, $a \leq s$.

EXAMPLE 8.4.2.

- The smallest element of $\{2, 4, 6, 8\}$ is 2.
- The smallest element of $\{12, 9, 18, 5, 13\}$ is 5.

It is important to realize that not every set of numbers has a smallest element:

EXERCISE 8.4.3. Show that the given set does not have a smallest element.

1) $\mathbb{Z}$ [*Hint:* If $n \in \mathbb{Z}$, then $n - 1 < n$.]
2) $\mathbb{R}^+ = \{ x \in \mathbb{R} \mid x > 0 \}$ [*Hint:* If $x \in \mathbb{R}^+$, then $x/2 \in \mathbb{R}^+$.]
3) $\varnothing$ [*Hint:* A set with no elements cannot have a *smallest* element.]

This problem does not arise for subsets of $\mathbb{N}$:

THEOREM 8.4.4 ($\mathbb{N}$ is well-ordered). *Every nonempty subset of $\mathbb{N}$ has a smallest element.*

This rather obvious observation is as powerful as all of the many variations of the Principle of Mathematical Induction. Namely, if $P(n)$ can be proved for all $n \in \mathbb{N}^+$ by using any one of the many forms of Mathematical Induction, then it can also be proved by applying Theorem 8.4.4 to the set

$$S = \{ n \in \mathbb{N}^+ \mid \neg P(n) \}.$$

More precisely, suppose $P(n)$ is not true for all $n \in \mathbb{N}^+$. Then the fact that $\mathbb{N}$ is well-ordered tells us that there is a *smallest* n, such that $P(n)$ is not true. This means that:

i) $P(n)$ is not true, but

ii) $P(k)$ is true for all $k < n$ (such that $k \in \mathbb{N}^+$).

Obtaining a contradiction from these two assumptions will complete the proof that $P(n)$ is true for all $n \in \mathbb{N}^+$.

EXAMPLE 8.4.5 (Alternate proof of Example 8.3.3). Suppose it is not true that $F_n < 2^n$ for all $n \in \mathbb{N}^+$. (This will lead to a contradiction.) Then, since $\mathbb{N}$ is well-ordered, there is a smallest n, such that $F_n \geq 2^n$. This means that

i) $F_n \geq 2^n$, but

ii) $F_k < 2^k$ for all $k < n$ (such that $k \in \mathbb{N}^+$).

Note that, since

$$F_1 = 1 < 2 = 2^1 \quad \text{and} \quad F_2 = 1 < 4 = 2^2,$$

we see from (i) that $n \notin \{1, 2\}$, so $n \geq 3$. Therefore $n - 1 \in \mathbb{N}^+$ and $n - 2 \in \mathbb{N}^+$. Then, since $n - 1$ and $n - 2$ are less than n, we see from (ii) that

$$(*) \qquad F_{n-1} < 2^{n-1} \text{ and } F_{n-2} < 2^{n-2}.$$

Now, we have

$$\begin{aligned} F_n &= F_{n-1} + F_{n-2} && \text{(definition of Fibonacci sequence)} \\ &< 2^{n-1} + 2^{n-2} && ((*)\text{: the minimality of } n) \\ &< 2^{n-1} + 2^{n-1} && (n-2 < n-1) \\ &= 2^n. \end{aligned}$$

This contradicts (i). □

EXERCISES 8.4.6.

1) Prove $F_{n+4} + F_n = 3F_{n+2}$ for all $n \in \mathbb{N}^+$.

2) Prove $2^n > n^2$ for every $n \geq 5$.

3) Prove $3^n > 2^n + 2n$, for every $n \geq 2$.

EXERCISE 8.4.7.

1) Show that the Principle of Mathematical Induction (8.1.2) follows from the fact that $\mathbb{N}$ is well-ordered.

2) Use induction to prove Theorem 8.4.4. [*Hint:* Define $P(n)$ to be the assertion "If $S \subset \mathbb{N}$ and S contains an element $\leq n$, then S has a smallest element."]

8.5. Applications in Number Theory

Induction (or the fact that $\mathbb{N}$ is well-ordered) can be used to prove many important properties of natural numbers. Here are just three examples.

DEFINITION 8.5.1. An element p of $\mathbb{N}^+$ is **prime** iff $p > 1$ and p is not divisible by any element of $\mathbb{N}^+$ other than 1 and p.

PROPOSITION 8.5.2. *If $n \in \mathbb{N}$ and $n > 1$, then n is divisible by a prime number.*

PROOF. Suppose there is some natural number $n > 1$, such that n is not divisible by a prime number. (This will lead to a contradiction.) Since $\mathbb{N}$ is well-ordered, we may assume that n is the smallest such number, so:

If $1 < k < n$ (and $k \in \mathbb{N}$), then k is divisible by a prime number.

Since $n \mid n$, but (by assumption) n is not divisible by any prime number, we know that n is not prime. By definition, this means there exists $k \in \mathbb{N}$, such that $k \mid n$ and $1 < k < n$. From the minimality of n, we know that k is divisible by some prime number p. Then $p \mid k$ and $k \mid n$, so $p \mid n$. This contradicts the fact that n is not divisible by a prime number. □

THEOREM 8.5.3 (Fundamental Theorem of Arithmetic). *Every natural number (other than 0 and 1) is a product of prime numbers (or is itself a prime).*

PROOF BY CONTRADICTION. Suppose there is some natural number $n > 1$, such that n is not a product of prime numbers (and is not a prime). Since $\mathbb{N}$ is well-ordered, we may assume that n is the smallest such number, so:

$$\text{If } 1 < k < n \text{ (and } k \in \mathbb{N}\text{), then } k \text{ is a product of prime numbers.}$$

Since n is not prime, it is divisible by some natural number k, with $1 < k < n$. This means we may write $n = km$, for some $m \in \mathbb{N}^+$. Since $m = n/k$ and $1 < k < n$, we see that $1 < m < n$. Therefore, the minimality of n implies that k and m are products of prime numbers: say $k = p_1 p_2 \cdots p_r$ and $m = q_1 q_2 \cdots q_s$. Then

$$n = km = (p_1 p_2 \cdots p_r)(q_1 q_2 \cdots q_s)$$

is a product of prime numbers. This is a contradiction. □

Remark 8.5.4. In fact, every natural number can be written in only one way as a product of prime numbers (up to rearranging the order of the factors), but we will not prove this fact.

DEFINITION 8.5.5. Let $a, b \in \mathbb{N}^+$. We say a and b are **relatively prime** iff they have no divisors in common, other than 1. (I.e., if $k \in \mathbb{N}^+$, and k is a divisor of both a and b, then $k = 1$. In other words, the "greatest common divisor" of a and b is 1.)

THEOREM 8.5.6. *Let $a, b \in \mathbb{N}^+$. If a and b are relatively prime, then there exist $m, n \in \mathbb{Z}$, such that $ma + nb = 1$.*

PROOF. Let

$$S = \{\, ma + nb \mid m, n \in \mathbb{Z}\ k = \} \cap \mathbb{N}^+.$$

It is easy to see that $a \in S$ (by letting $m = 1$ and $n = 0$), so $S \neq \varnothing$. Therefore, since $\mathbb{N}$ is well-ordered, we may let d be the smallest element of S. Then $d \in S$, so we have $d = m_0 a + n_0 b$ for some $m_0, n_0 \in \mathbb{Z}$.

By the Division Algorithm (5.1.20), we may write

$$a = qd + r \text{ with } 0 \leq r < d.$$

So

$$r = a - qd = a - q(m_0 a + n_0 b) = (1 - qm_0)a + qn_0 b = ma + nb,$$

where $m = 1 - qm \in \mathbb{Z}$ and $n = qn \in \mathbb{Z}$. On the other hand, since $r < d$, and d is the smallest element of S, we know $r \notin S$. From the definition of S, we conclude that $r = 0$. So $d \mid a$.

By repeating the same argument with a and b interchanged (and m_0 and n_0 also interchanged) we see that $d \mid b$.

Therefore, d is a divisor of both a and b. Since a and b are relatively prime, we conclude that $d = 1$. Since $d \in S$, this means $1 \in S$, which establishes the desired conclusion. □

EXERCISE 8.5.7. Prove the converse of Theorem 8.5.6.

Theorem 8.5.6 is of fundamental importance in **Number Theory**, the mathematical study of properties of $\mathbb{N}$ and $\mathbb{Z}$. Here are a few of its many consequences:

EXERCISES 8.5.8. Assume $a, b \in \mathbb{N}^+$.

1) Show a and b are relatively prime iff there exists $x \in \mathbb{Z}$, such that $xa \equiv 1 \pmod{b}$.
2) Show a and b are relatively prime iff for all $y \in \mathbb{Z}$, there exists $x \in \mathbb{Z}$, such that $xa \equiv y \pmod{b}$.
3) (**Chinese Remainder Theorem**) Suppose a and b are relatively prime. For all $y_1, y_2 \in \mathbb{Z}$, show there exists $x \in \mathbb{Z}$, such that $x \equiv y_1 \pmod{a}$ and $x \equiv y_2 \pmod{b}$.

Proposition 8.5.2 also has important consequences. For example:

COROLLARY 8.5.9. *There are infinitely many prime numbers.*

PROOF BY CONTRADICTION. Suppose there are only finitely many prime numbers. Then we can make a list of all of them:

$$\text{The set of all prime numbers is } \{p_1, p_2, \ldots, p_n\}.$$

Let

$$N = p_1 p_2 \cdots p_n.$$

From Proposition 8.5.2, we know there is some prime p, such that $p \mid (N+1)$.

Since $p_1, p_2, \ldots, p_n$ is a list of all the prime numbers, we know $p = p_i$, for some i. Therefore $p = p_i$ is one of the factors in the product that defines N, so $p \mid N$. Therefore, p divides both N and $N+1$, so (from 5.1.9(1)) we have

$$p \mid \big((N+1) - N\big) = 1.$$

This implies $p = \pm 1$ (see page 97), which contradicts the fact that p, being a prime number, must be > 1. □

SUMMARY:

- Important definitions:
  - proof by induction
  - base case, induction step
  - induction hypothesis
  - relatively prime
- Whenever you need to prove a statement with an n in it, you should consider using induction.
- Sequences of numbers are sometimes defined "recursively," which means that the value of a term may depend on previous terms.
- There are several alternate forms of induction, including strong induction, generalized induction, and strong induction with multiple base cases.
- $\mathbb{N}$ is well-ordered.
- If a and b are relatively prime, then $ma + nb = 1$, for some $m, n \in \mathbb{Z}$.
- Notation:
  - $\sum_{k=1}^{n} a_k = a_1 + a_2 + \cdots + a_n$.

Chapter 9

Cardinality

Our knowledge can only be finite,
while our ignorance must necessarily be infinite.
Karl Popper (1902–1994), Austrian-born British philosopher
Conjectures and Refutations: The Growth of Scientific Knowledge

9.1. Definition and basic properties

Informally, the *cardinality* of a set is the number of elements that it contains. (This was mentioned in Notation 3.2.14.)

EXERCISES 9.1.1. What is the cardinality of each of these sets?
(*You do* not *need to show your work or justify your answers.*)

1) $\#\{1, 2, 3, 4\} =$
2) $\#\{\mathsf{a}, \mathsf{e}, \mathsf{i}, \mathsf{o}, \mathsf{u}\} =$
3) $\#\{\mathsf{a}, \mathsf{l}, \mathsf{b}, \mathsf{e}, \mathsf{r}, \mathsf{t}, \mathsf{a}\} =$
4) $\#\varnothing =$
5) $\#\{\varnothing\} =$
6) $\#\big\{k \in \{1, 2, \ldots, 10\} \;\big|\; k \neq 7\big\} =$

An informal understanding of cardinality is not sufficient in advanced mathematics courses, so we need to study this idea more thoroughly.

EXAMPLE 9.1.2. The cardinality of $\{\mathsf{a}, \mathsf{b}, \mathsf{c}\}$ is 3. Children learn to verify this by counting:

$$1 \text{ (for } \mathsf{a}), \quad 2 \text{ (for } \mathsf{b}), \quad 3 \text{ (for } \mathsf{c}).$$

They are taught to assign exactly one number to each element of the set (and not skip any numbers as they count up). Mathematicians express this idea in a more sophisticated way: if we define a function $f\colon \{\mathsf{a}, \mathsf{b}, \mathsf{c}\} \to \{1, 2, 3\}$ by

$$f(\mathsf{a}) = 1, \; f(\mathsf{b}) = 2, \; f(\mathsf{c}) = 3,$$

then f is a bijection.

In general, if a set A has n elements, then counting the elements one-by-one defines a bijection from A to $\{1, 2, 3, \ldots, n\}$. This observation leads to the following official definition.

DEFINITION 9.1.3.

1) Let A be a set and n be a natural number. We say that the **cardinality** of A is n (and write $\#A = n$) iff there is a bijection from A to $\{1, 2, 3, \ldots, n\}$.
2) A set A is **finite** iff there is some $n \in \mathbb{N}$, such that $\#A = n$.
3) A set A is **infinite** iff it is *not* finite.

Remarks 9.1.4.

1) Not all sets are finite. For example, $\mathbb{N}$ is infinite (see Exercise 9.2.8(5)).

2) We will see in Theorem 9.1.20 that every subset of a finite set is finite.

Since the definition of $\#A$ is based on bijections, every proof about cardinality will rely on facts about bijections.

EXAMPLE 9.1.5. Show $\#\{1, 2, 3, \ldots, n\} = n$, for every $n \in \mathbb{N}$.

Scratchwork. The definition of cardinality tells us that if A is any set, and we need to show $\#A = n$, then we need to find a bijection from A to $\{1, 2, 3, \ldots, n\}$. In this problem, we have $A = \{1, 2, 3, \ldots, n\}$. Therefore, we need to find a bijection from $\{1, 2, 3, \ldots, n\}$ to $\{1, 2, 3, \ldots, n\}$. Thus, we need to find a bijection from a set to itself. Exercise 6.6.9 tells us that the identity map is such a function.

PROOF. From Exercise 6.6.9, we know that the identity map $I_{\{1,2,\ldots,n\}}$ is a bijection from $\{1, 2, \ldots, n\}$ to $\{1, 2, \ldots, n\}$, so $\#\{1, 2, \ldots, n\} = n$. □

ALTERNATE PROOF. Define $f\colon \{1, 2, 3, \ldots, n\} \to \{1, 2, 3, \ldots, n\}$ by $f(k) = k$.

We claim that f is a bijection. In other words, we claim that f is one-to-one and onto.

(one-to-one) Given $i_1, i_2 \in \{1, 2, 3, \ldots, n\}$, such that $f(i_1) = f(i_2)$, we have $i_1 = i_2$. So f is one-to-one.

(onto) Given $j \in \{1, 2, 3, \ldots, n\}$, let $i = j$. Then $f(i) = i = j$. So f is onto.

Since f is both one-to-one and onto, it is a bijection. This completes the proof of the claim.

Therefore, there is a bijection (namely, f) from $\{1, 2, 3, \ldots, n\}$ to $\{1, 2, 3, \ldots, n\}$. Hence, $\#\{1, 2, 3, \ldots, n\} = n$. □

Remark 9.1.6. Since the empty set has no elements, its cardinality should be 0. Although it may not be obvious, Definition 9.1.3 does agree with this observation. To verify this, it is important to realize that, for any $n \in \mathbb{N}$, the notation $\{1, 2, \ldots, n\}$ is just another name for the set $\{\, i \in \mathbb{N} \mid 1 \leq i \leq n \,\}$. In particular, if $n = 0$, then

$$\{1, 2, \ldots, n\} = \{\, i \in \mathbb{N} \mid 1 \leq i \leq n \,\} = \{\, i \in \mathbb{N} \mid 1 \leq i \leq 0 \,\} = \varnothing.$$

Therefore, letting $n = 0$ in Example 9.1.5 tells us that $\#\varnothing = 0$, as expected.

The definition of $\#A$ specifies that there is a bijection from A to $\{1, 2, 3, \ldots, n\}$. The following exercise shows there is no harm if you choose to have the bijection go the other way.

EXERCISE 9.1.7. Let A be a set. For $n \in \mathbb{N}$, show $\#A = n$ iff there is a bijection from $\{1, 2, 3, \ldots, n\}$ to A. [*Hint:* Use Exercise 6.7.13(1).]

Bijections can be used to show that two sets have the same cardinality, without knowing how many elements they have.

EXAMPLE 9.1.8. In the society Married of Example 6.6.2 (where every man is married to a woman, and vice-versa), it is clear that there must be exactly the same number of men and women. (If there were more women than men, then either some woman would be unmarried, or more than one woman would have to be married to the same man. Similarly, if there were more men than women.) This is true, even though we have no idea how many women or men there are in the society. All we know is that however many women there are is exactly the same as the number of men.

This observation is formalized in the following proposition.

PROPOSITION 9.1.9. *Suppose A and B are finite sets. Then $\#A = \#B$ if and only if there is a bijection from A to B.*

PROOF. ($\Rightarrow$) Let n be the cardinality of A. By definition, this means

$$\text{there is a bijection } f\colon A \to \{1, 2, \ldots, n\}.$$

By assumption, n is also the cardinality of B, so

$$\text{there is also a bijection } g\colon B \to \{1, 2, \ldots, n\}.$$

The inverse of a bijection is a bijection (see Exercise 6.7.13(1)), and the composition of bijections is a bijection (see Exercise 6.8.12(1)), so $g^{-1} \circ f$ is a bijection from A to B.

($\Leftarrow$) We leave this as an exercise. □

EXERCISE 9.1.10. Prove Proposition 9.1.9($\Leftarrow$). [*Hint:* Use Exercise 6.8.12(1).]

EXERCISES 9.1.11.

1) Show that if $\#A_1 = \#A_2$, then $\#(A_1 \times B) = \#(A_2 \times B)$.
   [*Hint:* If $f\colon A_1 \to A_2$ is a bijection, define $g\colon A_1 \times B \to A_2 \times B$ by $g(a_1, b) = \big(f(a_1), b\big)$.]
2) Show that if $\{a_0\}$ is any set with only one element, then $\#\big(\{a_0\} \times B\big) = \#B$.
   [*Hint:* Define $f\colon B \to \{a_0\} \times B$ by $f(b) = (a_0, b)$.]
3) Suppose $f\colon A \to B$ is one-to-one, and $X \subset A$. Show $\#f(X) = \#X$.

EXAMPLE 9.1.12. In elementary school, we learn that if Alice has m apples and Bob has n apples, then the sum $m+n$ is the total number of apples that the two of them have. However, this simple example assumes that Alice and Bob are not sharing any of the apples; the set of Alice's apples must be *disjoint* from the set of Bob's apples.

The following result generalizes this example.

PROPOSITION 9.1.13. *If A and B are disjoint finite sets, then*

$$\#(A \cup B) = \#A + \#B.$$

PROOF. Let $m = \#A$ and $n = \#B$. Then there exist bijections

$$f\colon \{1, 2, \ldots, m\} \to A \quad \text{and} \quad g\colon \{1, 2, \ldots, n\} \to B.$$

Define a function $h\colon \{1, 2, \ldots, m+n\} \to (A \cup B)$ by

$$h(k) = \begin{cases} f(k) & \text{if } k \leq m \\ g(k-m) & \text{if } k > m \end{cases}$$

(Notice that if $k \in \{1, 2, \ldots, m+n\}$, and $k > m$, then $m+1 \leq k \leq m+n$, so $1 \leq k-m \leq n$; therefore, $k-m$ is in the domain of g, so the expression $g(k-m)$ makes sense.)

To complete the proof, it suffices to show that h is a bijection; thus, we we need only show that h is one-to-one and onto.

(onto) Given $y \in A \cup B$, we have either $y \in A$ or $y \in B$, and we consider these two possibilities as separate cases.

1) Suppose $y \in A$. Since f is onto, there is some $k \in \{1, 2, \ldots, m\}$ with $f(k) = y$. Then, because $k \leq m$, we have
   $$h(k) = f(k) = y.$$
2) Suppose $y \in B$. Since g is onto, there is some $k \in \{1, 2, \ldots, n\}$ with $g(k) = y$. Then $k + m \in \{1, 2, \ldots, m+n\}$ and $k + m > m$, so
   $$h(k+m) = g\big((k+m) - m\big) = g(k) = y.$$

Since y is an arbitrary element of $A \cup B$, we conclude that h is onto.

(one-to-one) We leave this as an exercise. □

EXERCISE 9.1.14. Show that the function h defined in the proof of Proposition 9.1.13 is one-to-one.

EXERCISES 9.1.15. Assume B is a finite set.

1) For all $b \in B$, show that $\#\big(B \smallsetminus \{b\}\big) = \#B - 1$.
2) Show that if $A \subset B$, then $\#A \leq \#B$. [*Hint:* $\#B = \#A + \#(B \smallsetminus A) \geq \#A$.]
3) Show that if A_1 and A_2 are disjoint subsets of B, then $\#A_1 + \#A_2 \leq \#B$.
4) (harder) Assume $B \neq \varnothing$. Show that if $S \subset \mathcal{P}(B)$, and no two elements of S are disjoint, then $\#S \leq \frac{1}{2}\#\mathcal{P}(B)$.
   [*Hint:* You may assume (without proof) that $\mathcal{P}(B)$ is finite. If $X \in S$, then $\overline{X} \notin S$.]
5) Show $\#B = 0$ if and only if $B = \varnothing$.

The following generalization of Proposition 9.1.13 applies to the union of any number of sets, not just two.

EXERCISE 9.1.16. If $A_1, A_2, \ldots, A_n$ are pairwise-disjoint finite sets, then

$$\#(A_1 \cup A_2 \cup \cdots \cup A_n) = \#A_1 + \#A_2 + \cdots + \#A_n.$$

[*Hint:* Induct on n, using Proposition 9.1.13 and Exercise 4.5.7.]

It was pointed out in Remark 6.1.7 that if A and B are finite sets, then $\#(A \times B) = \#A \cdot \#B$. We can now prove this, after using Exercise 9.1.16 to solve the following exercise:

EXERCISE 9.1.17. Suppose $A_1, A_2, \ldots, A_m$ are pairwise-disjoint finite sets. Show that if $A = A_1 \cup \cdots \cup A_m$, then

$$\#(A \times B) = \#(A_1 \times B) + \#(A_2 \times B) + \cdots + \#(A_m \times B).$$

[*Hint:* The sets $A_i \times B$ are pairwise-disjoint, and their union is $A \times B$.]

THEOREM 9.1.18. *For any finite sets A and B, we have*

$$\#(A \times B) = \#A \cdot \#B.$$

PROOF. Let $m = \#A$. Then there is no harm in assuming $A = \{1, 2, \ldots, m\}$ (see Exercise 9.1.11(1)). Therefore

$$A = \{1\} \cup \{2\} \cup \cdots \cup \{m\},$$

and the sets $\{1\}, \{2\}, \cdots, \{m\}$ are pairwise-disjoint, so

$$\begin{aligned}
\#(A \times B) &= \#\big(\{1\} \times B\big) + \#\big(\{2\} \times B\big) + \cdots + \#\big(\{m\} \times B\big) && \text{(Exercise 9.1.17)} \\
&= \#B + \#B + \cdots + \#B \qquad (m \text{ summands}) && \text{(Exercise 9.1.11(2))} \\
&= m \cdot \#B \\
&= \#A \cdot \#B.
\end{aligned}$$

□

EXERCISES 9.1.19. Suppose A and B are finite sets, and $m, n \in \mathbb{N}$. Prove:

1) If $m \leq n$, then there exists a one-to-one function $f\colon \{1, 2, \ldots, m\} \to \{1, 2, \ldots, n\}$.
2) If $\#A \leq \#B$, then there exists a one-to-one function $f\colon A \to B$.
   [*Hint:* Use the preceding exercise.]
3) If $m \geq n > 0$, then there exists an onto function $f\colon \{1, 2, \ldots, m\} \to \{1, 2, \ldots, n\}$.

4) If A and B are nonempty, and $\#A \geq \#B$, then there exists an onto function $f\colon A \to B$. [*Hint:* Use the preceding exercise.]

The converses of the exercises in (9.1.19) are true, and important. They will be discussed in Section 9.2.

We end the section with a basic fact that may seem to be obvious (and was mentioned in Remark 9.1.4(2)), but that would be difficult or impossible to prove without using induction.

THEOREM 9.1.20. *Every subset of any finite set is finite.*

PROOF. Define $P(n)$ to be the assertion

If A is any set with $\#A = n$, then every subset of A is finite.

(i) *Base case.* Assume $n = 0$, and let B be any subset of A. Now $\#A = n = 0$, so $A = \varnothing$. Since the only subset of the empty set is the empty set, we have $B = \varnothing$. Hence $\#B = 0$, so B is finite.

(ii) *Induction step.* Assume $n \geq 1$, and that $P(n-1)$ is true, and let B be any subset of A. We may assume B is a proper subset of A. (Otherwise, we have $B = A$, so $\#B = \#A = n$, which means that B is finite.) Thus, there exists $a \in A$, such that $a \notin B$. Let $A' = A \smallsetminus \{a\}$. Then $\#A' = n - 1$, so the induction hypothesis tells us that every subset of A' is finite. Since $B \subset A'$, we conclude that B is finite. □

9.2. The Pigeonhole Principle

If a mail carrier has m letters to distribute among n mailboxes (or "pigeonholes"), and $m > n$, then it is clear that at least one of the mailboxes will have to get more than one letter. This important observation is known as the "Pigeonhole Principle." (See Exercise 9.3.6 for the proof.) In the language of sets, it can be stated as follows.

PROPOSITION 9.2.1 (Pigeonhole Principle). *Let B and $A_1, A_2, \ldots, A_n$ be finite sets. If*

$$B \subset A_1 \cup A_2 \cup \ldots \cup A_n,$$

and $\#B > n$, then $\#A_i \geq 2$, for some i.

Here are a few of the many applications of the Pigeonhole Principle. In these real-world examples, our explanations will be a bit informal.

EXAMPLE 9.2.2. Bob's sock drawer has many, many socks in it, and they come in 4 colours. Unfortunately, the light in his room has burned out, so he cannot see anything. How many socks should he grab from the drawer, so that he can be sure at least two of them are of the same colour?

SOLUTION. Bob should grab 5 (or more) socks.

To see this, note, first, that taking 4 socks may not be enough: If Bob grabs only 4 socks, it is possible that he has one sock of each of the 4 different colours. Then he would not have two socks of the same colour.

Now suppose Bob grabs (at least) 5 socks. He can sort them into 4 piles, by colour. Since $5 > 4$, one of the piles must have more than one sock. So there are (at least) 2 socks of the same colour. □

EXAMPLE 9.2.3. If you pick 50 numbers from 1 to 98, then it is guaranteed that two of them will add up to exactly 99.

SOLUTION. The numbers from 1 to 98 can be divided into 49 pigeonholes:

$$\{1, 98\},\ \{2, 97\},\ \{3, 96\},\ \ldots,\ \{49, 50\}.$$

(So two different numbers x and y are in the same pigeonhole iff $x+y = 99$.) Since $50 > 49$, two of the numbers we chose must be in the same pigeonhole. Then the sum of these two numbers is exactly 99. □

EXAMPLE 9.2.4. If you pick 5 points on the surface of a (spherical) orange, then there is always a way to cut the orange exactly in half, such that at least 4 of your points are in the same half. (We assume any point that is exactly on the cut is considered to belong to both halves.)

SOLUTION. Any two of the points will lie on a great circle of the sphere, so we can cut the orange so that 2 of the points are exactly on the cut. The other 3 points are distributed in some way among the two halves of the orange. By the Pigeonhole Principle, at least two of those three points are in the same half. Then that half contains the 2 points on the cut, plus these additional 2 points, for a total of (at least) 4 of the points you picked. □

EXERCISES 9.2.5.

1) There are ten people in a skating rink, playing hockey. Explain how you know that two of them were born on the same day of the week.
2) If there are 400 students in an elementary school, explain how you know that (at least) two of them have the same birthday.
3) If there are 700 employees at a company, explain how you know that there are two of them with the same initials. (That is, their first names start with the same letter, and their last names start with the same letter.)
4) It is known that:
   - No one has more than 300,000 hairs on their head.
   - More than a million people live in Calgary.

   Show that there are two people in Calgary who have exactly the same number of hairs on their heads.

In addition to the above real-world examples, the Pigeonhole Principle has important applications in theoretical mathematics.

COROLLARY 9.2.6. *Suppose A and B are finite sets.*

1) *If there exists a one-to-one function $f\colon A \to B$, then $\#A \leq \#B$.*
2) *If there exists an onto function $f\colon A \to B$, then $\#A \geq \#B$.*

PROOF. Let $m = \#A$ and $n = \#B$.

(1) Suppose $f\colon A \to B$ is one-to-one, and $m > n$. Assume without loss of generality that $B = \{1, 2, \ldots, n\}$, so we may let

$$A_i = f^{-1}(i) \quad \text{for } i = 1, 2, \ldots, n.$$

For any $a \in A$, we have $a \in f^{-1}\big(f(a)\big) = A_{f(a)}$, so $a \in A_1 \cup A_2 \cup \cdots \cup A_n$. Since a is an arbitrary element of A, this implies $A \subset A_1 \cup A_2 \cup \ldots \cup A_n$. Because $\#A = m > n$, we conclude that $\#A_i \geq 2$ for some i. This means $\#f^{-1}(i) > 1$, which contradicts the fact that f is one-to-one.

(2) Suppose $f\colon A \to B$ is onto, and $m < n$. There is no harm in assuming $A = \{1, 2, \ldots, m\}$, and then we may let

$$B_i = \{f(i)\}$$

for $i = 1, 2, \ldots, m$. Since f is onto, we know, for any $b \in B$, there is some $i \in A$, such that $f(i) = b$. This means $b \in B_i$; hence, $b \in B_1 \cup B_2 \cup \cdots \cup B_m$. Since b is an arbitrary element of B, this implies $B \subset B_1 \cup B_2 \cup \ldots \cup B_m$. Because $\#B = n > m$, we conclude that $\#B_i \geq 2$ for some i. This contradicts the fact that $\#B_i = 1$ (because $B_i = \{f(i)\}$ has only one element). □

Remark 9.2.7. Instead of Proposition 9.2.1, many mathematicians consider the contrapositive of Corollary 9.2.6(1) to be the Pigeonhole Principle:

If $\#A > \#B$, then there does not exist a one-to-one function from A to B.

EXERCISES 9.2.8.

1) Suppose A is a set of 10 natural numbers between 1 and 100 (inclusive). Show that two different subsets of A have the same sum. For example, if
$$A = \{2, 6, 13, 30, 45, 59, 65, 82, 88, 97\},$$
then the subsets $\{6, 13, 45, 65\}$ and $\{2, 30, 97\}$ both add up to 129. [*Hint:* Compare the answers to two questions: How many subsets of A are there? Since there are only 10 elements of A, and all of them are ≤ 100, how many different possible sums are there?]

2) Show that if you put 5 points into an equilateral triangle of side length 2 cm, then there are two of the points that are no more than 1 cm apart.
[*Hint:* Divide the triangle into 4 equilateral triangle of side length 1 cm.]

3) The numbers 1, 11, 111, 1111, etc. are called **repunits**. (Their decimal representation consists entirely of 1's.) Show that some repunit is divisible by 2017.
[*Hint:* If $n \times 10^k$ is divisible by 2017, for some $k \in \mathbb{N}$, then n is divisible by 2017. *Why?*]

4) Show that $\#A$ is well-defined. That is, if $\#A = m$ and $\#A = n$, for some $m, n \in \mathbb{N}$, then $m = n$. [*Hint:* Apply Corollary 9.2.6 with $B = A$.]

5) Show $\mathbb{N}$ is infinite. [*Hint:* Proof by contradiction. Apply Corollary 9.2.6(1).]

Remark 9.2.9. Here are two generalizations of the Pigeonhole Principle that are often useful.

1) If a mail carrier has m letters to distribute among n mailboxes, and $m > kn$, then at least one of the mailboxes has to get more than k letters.

2) Suppose a mail carrier has m letters to distribute among n mailboxes. If
$$k_1, k_2, \ldots, k_n \in \mathbb{N} \text{ and } m > k_1 + k_2 + \cdots + k_n,$$
then there must be some i, such that the ith mailbox gets more than k_i letters.

EXERCISE 9.2.10. State (1) and (2) of Remark 9.2.9 analogously to:

a) Proposition 9.2.1 (that is, in terms of a set B contained in $A_1 \cup A_2 \cup \cdots \cup A_n$), and

b) Corollary 9.2.6(1) (that is, in terms of a function $f\colon A \to B$).

EXERCISE 9.2.11.

1) Strengthen the conclusion of Exercise 9.2.5(4): show there is a collection of 4 people in Calgary who all have exactly the same number of hairs on their head.

2) As in Example 9.2.2, Bob's sock drawer has many, many socks in it, and they come in 4 colours. How many socks should he grab from the drawer, so that he can be sure at least 12 of them are of the same colour?

3) Betty's sock drawer has 6 blue socks, 10 red socks, 14 white socks, and 18 black socks. How many socks should she grab from the drawer, so that she can be sure at least 12 of them are of the same colour?

9.3. Cardinality of a union

We know that if A and B are disjoint, then the cardinality of $A \cup B$ is $\#A + \#B$. Here is a formula that works even when the sets are not disjoint:

PROPOSITION 9.3.1. *For any finite sets A and B, we have*

$$\#(A \cup B) = \#A + \#B - \#(A \cap B).$$

PROOF. From Exercise 4.5.6, we know that $A \smallsetminus B$, $B \smallsetminus A$, and $A \cap B$ are pairwise-disjoint, and that their union is $A \cup B$, so

$$\#(A \smallsetminus B) + \#(B \smallsetminus A) + \#(A \cap B) = \#\big((A \smallsetminus B) \cup (B \smallsetminus A) \cup (A \cap B)\big) = \#(A \cup B).$$

Also, we have

$$\begin{aligned} \#A &= \#\big((A \smallsetminus B) \cup (A \cap B)\big) && \text{(Exercise 4.5.4(2))} \\ &= \#(A \smallsetminus B) + \#(A \cap B) && \text{(Exercise 4.5.5(4)).} \end{aligned}$$

Similarly, we have

$$\#B = \#(B \smallsetminus A) + \#(A \cap B).$$

Therefore

$$\begin{aligned} \#A + \#B &= \big(\#(A \smallsetminus B) + \#(A \cap B)\big) + \big(\#(B \smallsetminus A) + \#(A \cap B)\big) \\ &= \#(A \smallsetminus B) + \#(B \smallsetminus A) + 2\#(A \cap B) \\ &= \#(A \cup B) + \#(A \cap B). \end{aligned}$$

The desired conclusion is obtained by subtracting $\#(A \cap B)$ from both sides. □

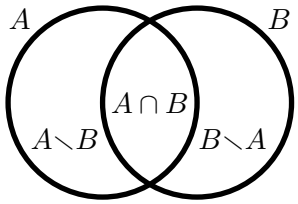


Figure 9A. The sum $\#A + \#B$ includes all the elements of $A \cup B$, but counts the elements of $A \cap B$ twice, so $\#A + \#B = \#(A \cup B) + \#(A \cap B)$. Therefore $\#(A \cup B) = \#A + \#B - \#(A \cap B)$.

EXAMPLE 9.3.2. Let $A = \{\mathsf{p}, \mathsf{r}, \mathsf{o}, \mathsf{n}, \mathsf{g}\}$ and $B = \{\mathsf{h}, \mathsf{o}, \mathsf{r}, \mathsf{n}, \mathsf{s}\}$. Then

$$\#A = 5,\ \#B = 5, \text{ and } \#(A \cap B) = \#\{\mathsf{r}, \mathsf{o}, \mathsf{n}\} = 3,$$

so Proposition 9.3.1 tells us that

$$\#(A \cup B) = \#A + \#B - \#(A \cap B) = 5 + 5 - 3 = 7.$$

This is correct, since

$$\#(A \cup B) = \#\{\mathsf{p}, \mathsf{r}, \mathsf{o}, \mathsf{n}, \mathsf{g}, \mathsf{h}, \mathsf{s}\} = 7.$$

EXAMPLE 9.3.3. Every one of the 4000 students at Modern U owns either a cell phone or an iPod (or both). Surveys show that:

- 3500 students own a cell phone, and
- 1000 students own an iPod.

How many students own *both* a cell phone and an iPod?

SOLUTION. Let

- S be the set of all students at Modern U,
- C be the set of students who own a cell phone, and
- I be the set of students who own an iPod.

Then, by assumption,

$$\#S = 4000, \qquad \#C = 3500, \qquad \#I = 1000.$$

Since every student owns either a cell phone or an iPod, we have $S = C \cup I$. Therefore, Proposition 9.3.1 tells us that

$$\#S = \#(C \cup I) = \#C + \#I - \#(C \cap I),$$

so

$$\#(C \cap I) = \#C + \#I - \#S = 3500 + 1000 - 4000 = 500.$$

Hence, there are exactly 500 students who own both a cell phone and an iPod. □

EXERCISES 9.3.4.

1) Assume $\#U = 15$, $\#V = 12$, and $\#(U \cap V) = 4$. Find $\#(U \cup V)$.
2) Assume $\#R = 13$, $\#S = 17$, and $\#(R \cup S) = 25$. Find $\#(R \cap S)$.
3) Assume $\#J = 300$, $\#(J \cup L) = 500$, and $\#(J \cap L) = 150$. Find $\#L$.
4) At a small university, there are 90 students that are taking either Calculus *or* Linear Algebra (or both). If the Calculus class has 70 students, and the Linear Algebra class has 35 students, then how many students are taking both Calculus *and* Linear Algebra?
5) (*harder*) Suppose A, B, and C are finite sets. Show

$$\begin{aligned}\#(A \cup B \cup C) = \#A + \#B + \#C \\ - \#(A \cap B) - \#(A \cap C) - \#(B \cap C) \\ + \#(A \cap B \cap C).\end{aligned}$$

[*Hint:* We have formulas for $\#\big((A \cup B) \cup C\big)$ and $\#(A \cup B)$. Another useful formula comes from the equality $(A \cup B) \cap C = (A \cap C) \cup (B \cap C)$.]

The following exercises are examples of another type of application of the formula for the cardinality of a union.

EXERCISES 9.3.5.

1) Suppose A and B are subsets of a finite set C. Show that if $\#A + \#B > \#C$, then $A \cap B \neq \varnothing$. [*Hint:* Use Proposition 9.3.1 to show that $\#(A \cap B) \neq 0$.]
2) Show that if A is a set of at least 600 natural numbers that are less than 1000, then two of the numbers in A differ by exactly 100. [*Hint:* Let $B = \{\, a + 100 \mid a \in A \,\}$, and use the preceding exercise to show that $A \cap B \neq \varnothing$.]

EXERCISES 9.3.6.

1) Suppose $A_1, A_2, \ldots, A_n$ are finite sets. Show

$$\#(A_1 \cup A_2 \cup \cdots \cup A_n) \leq \#A_1 + \#A_2 + \cdots + \#A_n.$$

[*Hint:* The induction step uses Proposition 9.3.1.]

2) Prove the Pigeonhole Principle (9.2.1).
[*Hint:* The contrapositive states that if $\#A_i \leq 1$ for every i, then $\#(A_1 \cup A_2 \cup \cdots \cup A_n) \leq n$.]

Remark 9.3.7. Generalizing Exercise 9.3.4(5), the **Inclusion-Exclusion** formula calculates the cardinality of the union of any number of sets:

$$|A_1 \cup \cdots \cup A_n| = \sum_{i=1}^{n} |A_i| - \sum_{1 \le i < j \le n} |A_i \cap A_j| + \sum_{1 \le i < j < k \le n} |A_i \cap A_j \cap A_k| - \cdots + (-1)^{n+1} |A_1 \cap \cdots \cap A_n|.$$

(The sign $+/-$ depends on the number of sets being intersected: odd intersections are added, and even intersections are subtracted, which agrees with both Proposition 9.3.1 and Exercise 9.3.4(5).) We do not need this formula, but you can read about it in a Combinatorics textbook (or on *Wikipedia*).

9.4. Hotel Infinity and the cardinality of infinite sets

The above discussion of cardinality included the following important fact that appeared in Proposition 9.1.9:

> Two finite sets A and B have the same cardinality if and only if there is a bijection from A to B.

Extending this property to all sets (not just the finite ones) is the *definition* of cardinality for infinite sets:

DEFINITION 9.4.1. Two sets A and B have the **same cardinality** iff there exists a bijection from A to B.

Two main ideas will be developed in the remainder of this chapter:

1) Many sets have the same cardinality as $\mathbb{N}^+$. These are the smallest infinite sets, and they are said to be **countably** infinite.

2) Not all sets are countably infinite: some sets are more infinite than others! These sets are said to be **uncountable**.

Let us begin with an informal discussion of some of the ideas that are involved in countability, rather than looking at the official definition right away. First, a simple example involving finite sets.

EXAMPLE 9.4.2. Suppose a hotel has n rooms, numbered $1, 2, 3, \ldots, n$.

1) If A is a tour group of n people $a_1, a_2, \ldots, a_n$, then the hotel clerk will obviously have no trouble assigning each of the people a room: a_i can be put in room i. There will be no empty rooms left.

room #	1	2	3	$\cdots$	n
occupant	a_1	a_2	a_3	$\cdots$	a_n

2) Now, if another person b arrives who wants a room, then the situation is hopeless. There is no way to give each of these $n+1$ people a room, without making two of them share a room. In general:

> If there are more guests than hotel rooms, then not everyone can have a room.

This is a restatement of the Pigeonhole Principle (9.2.1).

EXAMPLE 9.4.3 (Hotel Infinity). Now consider a hotel with infinitely many rooms, numbered $1, 2, 3, \ldots$. (There is one room for each $i \in \mathbb{N}^+$.)

1) If A is a tour group of n people $a_1, a_2, \ldots, a_n$, then the hotel clerk will obviously have no trouble giving each of the people a room: a_i can be put in room i. There will be lots of empty rooms left over.

1	2	3	$\cdots$	n	$n+1$	$n+2$	$\cdots$
a_1	a_2	a_3	$\cdots$	a_n			$\cdots$

2) Even if A is infinite, rather than finite, with people $a_1, a_2, \ldots$, the hotel clerk can accommodate all of them, by putting a_i in room i. There will be no empty rooms left.

1	2	3	4	$\cdots$
a_1	a_2	a_3	a_4	$\cdots$

3) Now suppose that, in addition to this tour group, there is another person b who also wants a room. Even though the hotel is already full, the hotel clerk can handle this situation quite easily, by shifting everyone to the next room, to make space for b. That is, the clerk can put

$$b \text{ into room } 1 \quad \text{and} \quad a_i \text{ in room } i+1.$$

Everyone will have his or her own room.

1	2	3	4	$\cdots$
b	a_1	a_2	a_3	$\cdots$

4) The same idea works, even if, instead of just one person, there is a whole group B of n people $b_1, b_2, \ldots, b_n$ that want rooms. The clerk can put

$$b_j \text{ in room } j, \quad \text{and} \quad a_i \text{ in room } i+n.$$

1	2	3	$\cdots$	n	$n+1$	$n+2$	$n+3$	$\cdots$
b_1	b_2	b_3	$\cdots$	b_n	a_1	a_2	a_3	$\cdots$

5) It may seem that there would be a problem if the second group B consists of infinitely many people $b_1, b_2, b_3, \ldots$, but a clever hotel clerk can accommodate even this situation. Note that there are infinitely many odd-numbered rooms, so all of A can be put in those rooms, and there are also infinitely many even-numbered rooms, so all of B can be put in there. More precisely, the clerk can put

$$a_i \text{ in room } 2i-1, \quad \text{and} \quad b_j \text{ in room } 2j.$$

1	2	3	4	$\cdots$	$2n-1$	$2n$	$2n+1$	$2n+2$	$\cdots$
a_1	b_1	a_2	b_2	$\cdots$	a_n	b_n	a_{n+1}	b_{n+1}	$\cdots$

6) Even if there are several of these countably infinite tour groups, not just 2 of them, they can all be accommodated. Namely, suppose there are n tour groups $A_1, A_2, A_3, \ldots, A_n$, and make a table (or matrix) with n infinitely long rows that lists the elements of A_i in the ith row. That is, if $a_{i,1}, a_{i,2}, a_{i,3}, \ldots$ is a list of the people in A_i, then the table looks like:

$$\begin{array}{lccccccc}
A_1: & a_{1,1} & a_{1,2} & a_{1,3} & a_{1,4} & a_{1,5} & \cdots \\
A_2: & a_{2,1} & a_{2,2} & a_{2,3} & a_{2,4} & a_{2,5} & \cdots \\
A_3: & a_{3,1} & a_{3,2} & a_{3,3} & a_{3,4} & a_{3,5} & \cdots \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\
A_n: & a_{n,1} & a_{n,2} & a_{n,3} & a_{n,4} & a_{n,5} & \cdots
\end{array}$$

We can assign rooms $1, 2, 3, \ldots$ to the entries of this table by counting down the 1st column (from 1 to n), then the 2nd column (starting at $n+1$), then the 3rd column (continuing from where we left off after the 2nd column), etc., as indicated here:

$$\begin{array}{ccccccc}
\begin{array}{c}1\\ a_{1,1}\end{array} & \begin{array}{c}n+1\\ a_{1,2}\end{array} & \begin{array}{c}2n+1\\ a_{1,3}\end{array} & \begin{array}{c}3n+1\\ a_{1,4}\end{array} & \begin{array}{c}4n+1\\ a_{1,5}\end{array} & \cdots \\
\downarrow & \downarrow & \downarrow & \downarrow & \downarrow & \cdots \\
\begin{array}{c}2\\ a_{2,1}\end{array} & \begin{array}{c}n+2\\ a_{2,2}\end{array} & \begin{array}{c}2n+2\\ a_{2,3}\end{array} & \begin{array}{c}3n+2\\ a_{2,4}\end{array} & \begin{array}{c}4n+2\\ a_{2,5}\end{array} & \cdots \\
\downarrow & \downarrow & \downarrow & \downarrow & \downarrow & \cdots \\
\begin{array}{c}3\\ a_{3,1}\end{array} & \begin{array}{c}n+3\\ a_{3,2}\end{array} & \begin{array}{c}2n+3\\ a_{3,3}\end{array} & \begin{array}{c}3n+3\\ a_{3,4}\end{array} & \begin{array}{c}4n+3\\ a_{3,5}\end{array} & \cdots \\
\downarrow & \downarrow & \downarrow & \downarrow & \downarrow & \cdots \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\
\downarrow & \downarrow & \downarrow & \downarrow & \downarrow & \cdots \\
\begin{array}{c}n\\ a_{n,1}\end{array} & \begin{array}{c}2n\\ a_{n,2}\end{array} & \begin{array}{c}3n\\ a_{n,3}\end{array} & \begin{array}{c}4n\\ a_{n,4}\end{array} & \begin{array}{c}5n\\ a_{n,5}\end{array} & \cdots
\end{array}$$

This assigns a different room number to each of the people, so everyone has their own room.

Remarks.

(a) Another way for the hotel clerk to find this solution is to note that there are infinitely many numbers that are congruent to i modulo n, so all of A_i can be put in those rooms.

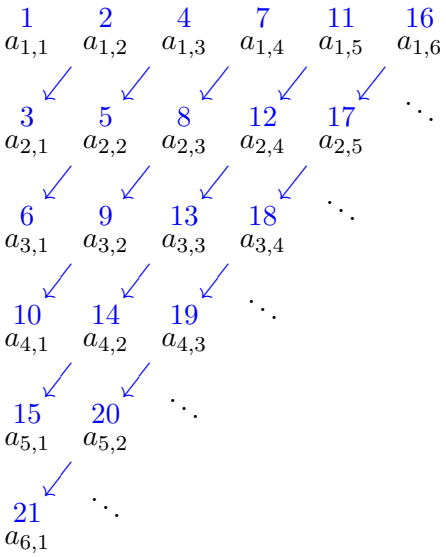
(b) It can be seen that guest $a_{i,j}$ is assigned to room $i+(j-1)n$, but we have no need for this formula.

7) Going further, even if there were infinitely many infinite tour groups $A_1, A_2, \ldots$, they could all be accommodated. (We assume that each tour group is countably infinite, and that the number of groups is countably infinite.) To see this, start by considering an infinitely large table (or matrix) that lists the elements of A_i in the ith row:

$$\begin{array}{lccccccc}
A_1: & a_{1,1} & a_{1,2} & a_{1,3} & a_{1,4} & a_{1,5} & \cdots \\
A_2: & a_{2,1} & a_{2,2} & a_{2,3} & a_{2,4} & a_{2,5} & \cdots \\
A_3: & a_{3,1} & a_{3,2} & a_{3,3} & a_{3,4} & a_{3,5} & \cdots \\
A_4: & a_{4,1} & a_{4,2} & a_{4,3} & a_{4,4} & a_{4,5} & \cdots \\
A_5: & a_{5,1} & a_{5,2} & a_{5,3} & a_{5,4} & a_{5,5} & \cdots \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots
\end{array}$$

We can assign rooms $1, 2, 3, \ldots$ to the entries of this table as follows:

- Begin with 1 in the top left corner.
- Then place 2 at the top of the second column and move diagonally (down and to the left) to place 3.
- Then place 4 at the top of the third column, and move diagonally (down and to the left) to place 5 and 6.
- Then place the next number (namely, 7) at the first open spot in the top row (namely, at the top of the fourth column), and move diagonally (down and to the left) to place the following numbers (namely, 8, 9, and 10), until a number (namely, 10) is placed in the first column.
- Continue by moving to the first open spot in the top row, and repeating infinitely.



No entries of the table are omitted from the numbering, and no room numbers are repeated, so each guest has his or her own room.

Remark. It can be shown that guest $a_{i,j}$ is put into room $\binom{i+j-1}{2} + i$, but we have no need for this formula.

It might seem that Hotel Infinity could accommodate every set of tourists, but that is not the case. For example, we will see in Section 9.6 that if all of the real numbers want rooms at the hotel, then some of them will have to share. In other words, *the set* $\mathbb{R}$ *of real numbers is uncountable.*

OTHER TERMINOLOGY. Hotel Infinity is often called "Hilbert's Hotel," in honour of the German mathematician David Hilbert (1862–1943), who was apparently the first to talk about such a hotel (in lectures in Germany in 1924).

9.5. Countable sets

For use in proving theorems, the ideas encountered in the discussion of Hotel Infinity need to be stated more formally. Let us begin with the definitions that form the foundation of the subject.

DEFINITION 9.5.1. Suppose A and B are sets.

1) A and B **have the same cardinality** iff there is a bijection from A to B.
2) A is **countably infinite** iff it has the same cardinality as $\mathbb{N}^+$.
3) A is **countable** iff either A is finite or A is countably infinite.
4) A is **uncountable** iff A is *not* countable.

Remarks 9.5.2.

1) In the terminology of the preceding section, a set is countable if and only if all of its elements can be given rooms in Hotel Infinity (see Exercise 9.5.6(3)).
2) If you are told to show that A is countably infinite, directly from the definition, then you should find a bijection from A to $\mathbb{N}^+$. However, because it is so well known that the inverse of a bijection is a bijection, it is acceptable to find a bijection from $\mathbb{N}^+$ to A, instead.

Remarks 9.5.3. One can (and should) think of countable sets as the sets whose elements can be listed in a sequence. The sequence may have only finitely many terms, or may continue forever:

1) A set A is finite iff its elements can be listed in a sequence $a_1, a_2, \ldots, a_n$, for some n.
2) If the elements of A can be listed in an infinite sequence $a_1, a_2, a_3, \ldots$, then we may define a bijection $f\colon \mathbb{N}^+ \to A$ by $f(i) = a_i$. Therefore, A is countably infinite.
3) Conversely, if A is countably infinite, then there is a bijection $f\colon \mathbb{N}^+ \to A$, so letting $a_i = f(i)$ yields an infinite sequence $a_1, a_2, a_3, \ldots$ that lists the elements of A.

EXERCISE 9.5.4. Define a binary relation $\approx$ on the collection of all sets by

$$A \approx B \quad \text{iff} \quad A \text{ and } B \text{ have the same cardinality.}$$

1) Show that $\approx$ is an equivalence relation. 2) What is the equivalence class of $\mathbb{N}^+$?

The following fundamental result shows that the smallest infinite sets are the countable ones:

THEOREM 9.5.5.

1) *Every infinite set contains a countably infinite subset.*
2) *Every subset of a countable set is countable.*

PROOF. (1) Given an infinite set A, it suffices to construct an infinite sequence $a_1, a_2, a_2, \ldots$ of distinct elements of A, for then $\{a_1, a_2, a_3, \ldots\}$ is a countably infinite subset of A.

1) Since A is infinite, it is certainly not empty, so it has some elements. Let a_1 be any of these elements of A.
2) Since A is infinite, we know that a_1 is not its only element. Let a_2 be any element of A other than a_1. Then a_1 and a_2 are distinct elements of A.
3) Since A is infinite, we know that a_1 and a_2 are not its only elements. Let a_3 be any element of A other than a_1 and a_2. Then a_1, a_2, and a_3 are distinct elements of A.

$\vdots$

(i) Since A is infinite, we know that $a_1, a_2, a_3, \ldots, a_{i-1}$ are not its only elements. Let a_i be any element of A other than $a_1, a_2, a_3, \ldots, a_{i-1}$. Then the elements $a_1, a_2, a_3, \ldots, a_i$ are distinct.

$\vdots$

Continuing this inductive process yields the desired infinite sequence $a_1, a_2, a_3, \ldots$ of distinct elements of A.

(2) Given a subset M of a countable set A, we wish to show that M is countable. We may assume M is infinite, for otherwise it is obviously countable. So we wish to show there is a sequence $m_1, m_2, m_3, \ldots$ that lists all the elements of M.

To simplify the notation, let us first consider the case where $A = \mathbb{N}^+$.

Case 1. Assume $A = \mathbb{N}^+$. Let

1) m_1 be the smallest element of M,
2) m_2 be the smallest element of $M \smallsetminus \{m_1\}$,
3) m_3 be the smallest element of $M \smallsetminus \{m_1, m_2\}$,
4) m_4 be the smallest element of $M \smallsetminus \{m_1, m_2, m_3\}$,

$\vdots$

(i) m_i be the smallest element of $M \setminus \{m_1, m_2, \ldots, m_{i-1}\}$.
(In other words, m_i is the smallest element of M that is not in $\{m_1, m_2, \ldots, m_{i-1}\}$.)
$\vdots$

For each $m \in M$, notice that if we let k be the number of elements of M that are less than m, then $m_{k+1} = m$. Therefore, every element m of M appears in the sequence $m_1, m_2, m_3, \ldots$.

Case 2. The general case. This is left as an exercise. □

EXERCISES 9.5.6.

1) Complete Case 2 in the proof of Theorem 9.5.5.
[*Hint:* If M is infinite, then A must be infinite (why?). List the elements of A in a sequence $a_1, a_2, a_3, \ldots$, and apply the argument of Case 1.]

2) Suppose that $f\colon A \to B$, that f is one-to-one, and that B is countable. Show that A is countable.
[*Hint:* If f is not onto, you will want to use the fact that every subset of a countable set is countable.]

3) Show that a set A is countable iff there exists a one-to-one function $f\colon A \to \mathbb{N}^+$.

Remark 9.5.7. Mathematicians think of countable sets as being small — even though they may be infinite, they are almost like finite sets. Consider the following basic properties of finite sets:

1) Any subset of a finite set is finite.
2) The union of two finite sets is finite.
3) More generally, the union of finitely many finite sets is finite.
4) If you have two finite sets, then you can make only finitely many ordered pairs from them. (That is, the Cartesian product of two finite sets is finite.)
5) If you have only finitely many darts to throw, then you can hit only finitely many things with them. That is, if $f\colon A \to B$, and A is finite, then the image $f(A)$ is finite.

All of the above assertions remain true when the word "finite" is replaced with "countable." The first assertion was established in Thm. 9.5.5(2); the others are contained in the following important theorem:

THEOREM 9.5.8.

1) *A countable union of countable sets is countable.*
2) *The cartesian product of two countable sets is countable.*
3) *The image of a countable set is countable.*

Remark 9.5.9. Here are more precise statements of the assertions of Theorem 9.5.8:

1) If $A_1, A_2, A_3, \ldots$ is any sequence of countable sets, then the union

$$\bigcup_{i=1}^{\infty} A_i = A_1 \cup A_2 \cup A_3 \cup \cdots$$

is countable. Also, if $A_1, A_2, A_3, \ldots, A_n$ is any finite sequence of countable sets, then the union

$$\bigcup_{i=1}^{n} A_i = A_1 \cup A_2 \cup A_3 \cup \cdots \cup A_n$$

is countable.

2) If A and B are any countable sets, then $A \times B$ is countable.

3) If $f\colon A \to B$, and A is countable, then $f(A)$ is countable.

PROOF OF THEOREM 9.5.8. (1) Given either an infinite sequence $A_1, A_2, A_3, \ldots$ of countable sets, or a finite sequence $A_1, A_2, A_3, \ldots, A_n$ of countable sets, we wish to show that the the union of the sets is countable. Subsets of a countable set are countable, so there is no harm in assuming:

- the sequence is infinite (because adding additional terms to the sequence will make the union larger), and
- each of the sets is infinite (because replacing A_i with an infinite superset will make the union larger).

Now, the numbering method of Eg. 9.4.3(7) shows there is an onto function $g\colon \mathbb{N}^+ \to \bigcup_{i=1}^{\infty} A_i$. So, from (3), we conclude that $\bigcup_{i=1}^{\infty} A_i$ is countable.

(2) Given countable sets A and B, we wish to show that $A \times B$ is countable. Subsets of a countable set are countable, so there is no harm in assuming that A and B are infinite (because replacing A and B with infinite supersets will make the cartesian product larger). Let

- $a_1, a_2, a_3, \ldots$ be a list of the elements of A, and
- $b_1, b_2, b_3, \ldots$ be a list of the elements of B,

Then the elements of $A \times B$ are listed in the following table (or matrix):

$$\begin{array}{cccccc}
(a_1,b_1) & (a_1,b_2) & (a_1,b_3) & (a_1,b_4) & (a_1,b_5) & \cdots \\
(a_2,b_1) & (a_2,b_2) & (a_2,b_3) & (a_2,b_4) & (a_2,b_5) & \cdots \\
(a_3,b_1) & (a_3,b_2) & (a_3,b_3) & (a_3,b_4) & (a_3,b_5) & \cdots \\
(a_4,b_1) & (a_4,b_2) & (a_4,b_3) & (a_4,b_4) & (a_4,b_5) & \cdots \\
(a_5,b_1) & (a_5,b_2) & (a_5,b_3) & (a_5,b_4) & (a_5,b_5) & \cdots \\
\vdots & \vdots & \vdots & \vdots & \vdots & \ddots
\end{array}$$

The numbering method of Eg. 9.4.3(7) defines a bijection from $A \times B$ to $\mathbb{N}^+$. So $A \times B$ is countable.

(3) Suppose $f\colon A \to B$, and A is countable. By replacing B with $f(A)$, we may assume f is onto; then we wish to show that B is countable.

By Exercise 9.5.6(2), it suffices to define a one-to-one function $g\colon B \to A$. The function f is onto, so, for each $b \in B$, there is some $a \in A$, such that $f(a) = b$; thus, for each $b \in B$, we may choose $g(b)$ to be an element of A such that

$$f\big(g(b)\big) = b.$$

Then $g\colon B \to A$, and all that remains is to show that g is one-to-one. Given $b_1, b_2 \in B$, such that $g(b_1) = g(b_2)$, we have $f\big(g(b_1)\big) = b_1$ and $f\big(g(b_2)\big) = b_2$. Therefore

$$b_1 = f\big(g(b_1)\big) = f\big(g(b_2)\big) = b_2.$$

So g is one-to-one, as desired. □

The theorems of this section make it easy to show that many sets are countable. Here are a few important examples:

EXERCISE 9.5.10. Show that each of the following sets is countable.

1) $\mathbb{N}^+$.

2) $\mathbb{N}$. [*Hint:* $\mathbb{N} = \mathbb{N}^+ \cup \{0\}$.]

3) $\mathbb{Z}$. [*Hint:* Let $\mathbb{Z}^- = \{\, n \in \mathbb{Z} \mid n < 0 \,\}$, so $\mathbb{Z} = \mathbb{N} \cup \mathbb{Z}^-$. The set $\mathbb{Z}^-$ is the image of $\mathbb{N}^+$ under a function.]

4) $\mathbb{Q}$. [*Hint:* Let $\mathbb{Z}^+$ be the set of positive integers, so $\mathbb{Q}$ is the image of $\mathbb{Z} \times \mathbb{Z}^\times$ under the function $f(a, b) = a/b$.]

It is very important to remember that $\mathbb{Q}$ is countable.
Since $\mathbb{N}$ and $\mathbb{Z}$ are subsets of $\mathbb{Q}$, this implies that $\mathbb{N}$ and $\mathbb{Z}$ are also countable.

EXERCISES 9.5.11.

1) Suppose A is countably infinite, and $b \notin A$. Show, directly from the definition, that $A \cup \{b\}$ is countably infinite.

2) Suppose A is countably infinite, and $a \in A$. Show, directly from the definition, that $A \smallsetminus \{a\}$ is countably infinite.

3) Suppose A and B are countably infinite and disjoint. Show, directly from the definition, that $A \cup B$ is countably infinite.

4) Suppose

$$\begin{array}{lll} A_1 \text{ is disjoint from } B_1, & & A_1 \text{ and } A_2 \text{ have the same cardinality,} \\ A_2 \text{ is disjoint from } B_2, & \text{and} & B_1 \text{ and } B_2 \text{ have the same cardinality.} \end{array}$$

Show that $(A_1 \cup B_1)$ has the same cardinality as $(A_2 \cup B_2)$.

5) Suppose A is infinite. Show there is a *proper* subset B of A, such that B has the same cardinality as A. [*Hint:* Combine Theorem 9.5.5(1) with Exercises (2) and (4).]

OTHER TERMINOLOGY. Some mathematicians do not consider finite sets to be "countable," so the terms "countable" and "countably infinite" are synonymous to them. Then, a set that is either finite or countably infinite is said to be "at most countable." Other mathematicians say that a countably infinite set is "denumerable."

9.6. Uncountable sets

The preceding section showed that many sets (including $\mathbb{N}$, $\mathbb{Z}$, and $\mathbb{Q}$) are countable. We will now see that some sets are *not* countable.

9.6A. The reals are uncountable. Here is perhaps the most important example of an uncountable set:

THEOREM 9.6.1. *The set $\mathbb{R}$ of real numbers is uncountable.*

If $\mathbb{R}$ were countable, then all of its subsets would be countable. Thus, in order to establish Theorem 9.6.1, it suffices to find an uncountable subset of $\mathbb{R}$. Here are some well-known examples of subsets:

NOTATION 9.6.2. For $a, b \in \mathbb{R}$ with $a < b$:

- **open interval**: $(a, b) = \{\, x \in \mathbb{R} \mid a < x < b \,\}$.
- **closed interval**: $[a, b] = \{\, x \in \mathbb{R} \mid a \le x \le b \,\}$.
- **half-open interval**: $[a, b) = \{\, x \in \mathbb{R} \mid a \le x < b \,\}$ or $(a, b] = \{\, x \in \mathbb{R} \mid a < x \le b \,\}$.

We will see that all of these intervals are uncountable. Here is one example:

THEOREM 9.6.3. *The interval $[0, 1)$ is uncountable.*

PROOF BY CONTRADICTION. Suppose $[0,1)$ is countable. (This will lead to a contradiction.) This means there is a list $x_1, x_2, x_3, \ldots$ of all the numbers in $[0,1)$. To obtain a contradiction, we will use a method called the *Cantor Diagonalization Argument*. It was discovered by the mathematician Georg Cantor in the 19th century.

Each number in $[0,1)$ can be written as a decimal of the form $0.d_1d_2d_3\ldots$, where each d_k is a digit (0, 1, 2, 3, 4, 5, 6, 7, 8, or 9). In particular, we can write each x_i in this form:

$$x_i = 0.x_{i,1}x_{i,2}x_{i,3}x_{i,4}x_{i,5}\ldots$$

Then we can make a list of all of these decimals (omitting the leading 0 in each one):

$$\begin{aligned}
x_1 &= .x_{1,1}x_{1,2}x_{1,3}x_{1,4}x_{1,5}\ldots \\
x_2 &= .x_{2,1}x_{2,2}x_{2,3}x_{2,4}x_{2,5}\ldots \\
x_3 &= .x_{3,1}x_{3,2}x_{3,3}x_{3,4}x_{3,5}\ldots \\
x_4 &= .x_{4,1}x_{4,2}x_{4,3}x_{4,4}x_{4,5}\ldots \\
x_5 &= .x_{5,1}x_{5,2}x_{5,3}x_{5,4}x_{5,5}\ldots \\
&\vdots \qquad\qquad\qquad \vdots
\end{aligned}$$

The right-hand side can be thought of as an array of digits, and we now focus on the diagonal entries $x_{i,i}$ of this array, which are circled in the following picture:

$$\begin{aligned}
x_1 &= .x_{1,1}x_{1,2}x_{1,3}x_{1,4}x_{1,5}\ldots \\
x_2 &= .x_{2,1}x_{2,2}x_{2,3}x_{2,4}x_{2,5}\ldots \\
x_3 &= .x_{3,1}x_{3,2}x_{3,3}x_{3,4}x_{3,5}\ldots \\
x_4 &= .x_{4,1}x_{4,2}x_{4,3}x_{4,4}x_{4,5}\ldots \\
x_5 &= .x_{5,1}x_{5,2}x_{5,3}x_{5,4}x_{5,5}\ldots \\
&\vdots \qquad\qquad\qquad \vdots
\end{aligned}$$

They form a sequence $x_{1,1}, x_{2,2}, x_{3,3}, \ldots$.

The key to the proof is to make a new sequence $d_1, d_2, d_3, \ldots$ of digits, such that

$$d_1 \neq x_{1,1}, \quad d_2 \neq x_{2,2}, \quad d_3 \neq x_{3,3}, \quad \text{etc.}$$

This means that every term of the new sequence is different from the corresponding term of the diagonal sequence. (This idea of choosing a sequence that is completely different from the diagonal is called **Cantor diagonalization**, because it was invented by the mathematician Georg Cantor.) Also, to avoid problems coming from the fact that $.999\cdots = 1.000\cdots$, you should not use the digits 0 and 9. The sequence $\{d_i\}$ can be constructed in many ways: just be sure to choose each d_i to be a digit that is not $x_{i,i}$ (and is not 0 or 9). For example, we could let

$$d_i = \begin{cases} 1 & \text{if } x_{i,i} \neq 1 \\ 5 & \text{if } x_{i,i} = 1. \end{cases}$$

Now, let

$$d = 0.d_1d_2d_3\ldots \in [0,1).$$

For each i, we made sure that $d_i \neq x_{i,i}$, which means that the ith digit of d is different from the ith digit of x_i. Therefore, for each i, we have $d \neq x_i$.* So d is an element of $[0,1)$ that is not in the list $x_1, x_2, x_3, \ldots$. This contradicts the fact that $x_1, x_2, x_3, \ldots$ is a list of *all* the numbers in $[0,1)$. □

*The digits of d are only 1's and 5's, so it is not a problem that numbers ending $000\ldots$ can also be expressed as a different decimal that ends $999\ldots$.

EXERCISES 9.6.4.

1) Show that the interval $(0, 1)$ is uncountable. [*Hint:* $[0, 1) = (0, 1) \cup \{0\}$ is uncountable.]
2) Suppose $a, b \in \mathbb{R}$. Show that if $a < b$, then the interval (a, b) has the same cardinality as $(0, 1)$. [*Hint:* Define $f\colon (0, 1) \to (a, b)$ by $f(x) = a + (b - a)x$.]
3) Suppose $a \in \mathbb{R}$. Show that the interval (a, ∞) has the same cardinality as $(0, 1)$. [*Hint:* Define $f\colon (0, 1) \to (a, \infty)$ by $f(x) = (1/x) + a - 1$.]

COROLLARY 9.6.5. *If $a, b \in \mathbb{R}$ and $a < b$, then the intervals (a, b), $[a, b]$, $[a, b)$, and $(a, b]$ are uncountable.*

EXERCISES 9.6.6. Decide which of the following sets are countable, and which are uncountable. (*You do not need to justify your answers.*)

1) The closed interval $[3, 3.1]$
2) $\{1, 2, 3, \ldots, 1000\}$
3) $\mathbb{Z} \times \mathbb{Z}$
4) $\mathbb{Z} \times \mathbb{Q}$
5) $\mathbb{R} \times \mathbb{N}$
6) $\mathbb{R} \smallsetminus \mathbb{Q}$
7) $\mathbb{Q} \smallsetminus \mathbb{R}$
8) $(\mathbb{R} \smallsetminus \mathbb{Q}) \cap (\mathbb{Q} \smallsetminus \mathbb{R})$
9) $(\mathbb{R} \smallsetminus \mathbb{Q}) \cup (\mathbb{Q} \smallsetminus \mathbb{R})$
10) $(\mathbb{R} \smallsetminus \mathbb{Q}) \times (\mathbb{Q} \smallsetminus \mathbb{R})$
11) $\{\, x \in \mathbb{R} \mid 2 < x < 3 \,\}$
12) $\{\, x \in \mathbb{R} \mid x^2 = 3 \,\}$
13) $\{\, x \in \mathbb{R} \mid x^2 + 12 < 3 \,\}$
14) $\{\, x \in \mathbb{R} \mid x^2 + 3 < 12 \,\}$
15) $\{\, x \in \mathbb{Q} \mid x^2 + 3 < 12 \,\}$
16) $\{\, (x, y) \in \mathbb{R} \times \mathbb{R} \mid x - y = 1 \,\}$
17) $\{\, (x, y) \in \mathbb{R} \times \mathbb{R} \mid x - y \in \mathbb{Z} \,\}$
18) $\{\, (x, y) \in \mathbb{Z} \times \mathbb{R} \mid x - y \in \mathbb{Q} \,\}$
19) $\{\, (x, y) \in \mathbb{R}^2 \mid x^2 + y^2 = 1 \,\}$
20) $\{\, (x, y) \in \mathbb{R}^2 \mid x^2 + y^2 = -1 \,\}$

9.6B. The cardinality of power sets. If A is a finite set, then the set $\mathcal{P}(A)$ of all subsets of A is also finite. (Indeed, $\#\mathcal{P}(A) = 2^{\#A}$.) However, this assertion does *not* remain true when the word "finite" is replaced with "countable":

EXERCISE 9.6.7. Show that $\mathcal{P}(\mathbb{N}^+)$ is uncountable.
[*Hint:* For any $f\colon \mathbb{N}^+ \to \mathcal{P}(\mathbb{N}^+)$, the set $\{\, i \in \mathbb{N}^+ \mid i \notin f(i) \,\}$ is not in the image of f.]

For every set A, not just the countable ones, the same argument shows that the cardinality of $\mathcal{P}(A)$ is greater than the cardinality of A. Thus, there is no "largest" infinite set. For every set, there is always some set that has *much* larger cardinality.

EXERCISE 9.6.8. For every set A, show there does not exist an onto function $f\colon A \to \mathcal{P}(A)$.
[*Hint:* The set $\{\, a \in A \mid a \notin f(a) \,\}$ is not in the image of f.]

The preceding exercises are very closely related to a classical paradox:

EXAMPLE 9.6.9 ("Barber Paradox"). Suppose there is a town with only one barber, and that barber is a man. Furthermore, assume the barber shaves precisely those men in the town who do not shave themselves. More precisely, if B is the Barber, and M is any man in the town, then

$$B \text{ shaves } M \qquad \text{iff} \qquad M \text{ does not shave } M.$$

Now, we ask:

Does the barber shave himself?

This question is a paradox:

- If the answer is yes, then the barber shaves himself. But the barber does *not* shave men who shave themselves, so this means that the barber does not shave himself. But we already said that the barber does shave himself, so this is nonsense.

- If the answer is no, then the barber does not shave himself. But the barber *does* shave any man who does not shave himself, so this means that the barber does shave himself. But we already said that the barber does not shave himself, so this is nonsense.

The upshot of this discussion is that the hypothesized situation leads to a contradiction, so it is impossible.

The same reasoning shows there is no set that contains precisely the sets that do not contain themselves. (Otherwise, the question "*Does the set contain itself?*" would be a paradox. Do you see why?) The solutions of Exercises 9.6.7 and 9.6.8 are based on the same idea.

SUMMARY:

- Important definitions:
  - cardinality
  - finite, infinite
  - countable, countably infinite
  - uncountable
- A and B have the same cardinality iff there is a bijection from A to B.
- Pigeonhole Principle
- For finite sets A and B, we have $\#(A \times B) = \#A \cdot \#B$.
- Inclusion-Exclusion: $A \cup B = \#A + \#B - \#(A \cap B)$.
- Properties of countable sets, including:
  - a countable union of countable sets is countable; and
  - the cartesian product of two countable sets is countable.
- $\mathbb{N}$, $\mathbb{Z}$, and $\mathbb{Q}$ are countable, but $\mathbb{R}$ is uncountable.
- The power set $\mathcal{P}(A)$ has larger cardinality than A, for any set A.
- Notation:
  - $\#A$
  - intervals (a, b), $[a, b]$, $[a, b)$, $(a, b]$
  - power set $\mathcal{P}(A)$

Index of Definitions

List of Notation