Chapter 1
First-Order Logic

However treacherous a ground mathematical logic, strictly interpreted, may be for an amateur, philosophy proper is a subject, on one hand so hopelessly obscure, on the other so astonishingly elementary, that there knowledge hardly counts. If only a question be sufficiently fundamental, the arguments for any answer must be correspondingly crude and simple, and all men may meet to discuss it on more or less equal terms.
G. H. Hardy *Mathematical Proof* [10].

Abstract This book is about a formal approach to mathematical structures. Formal methods are by their very nature formal. When studying mathematical logic, initially one often has to grit ones teeth and absorb certain preliminary definitions on faith. Concepts are given precise definitions, and their meaning is revealed later after one has a chance to see their utility. We will try to follow a different route. Before all formalities are introduced, in this chapter, we will take a detour to see examples of mathematical statements and some elements of the language that is used to express them.

Keywords Arithmetic · Euclid's theorem · Formalization · Vocabulary of first-order-logic · Boolean connectives · Quantifiers · Truth values · Trivial structures

1.1 What We Talk About When We Talk About Numbers

The natural numbers are 0, 1, 2, 3, ...[1] A natural number is *prime* if it is larger than 1 and is not equal to a product of two smaller natural numbers. For example, 11 and 13 are prime, but 15 is not, because $15 = 3 \cdot 5$. Proposition 20 in Book IX of Euclid's *Elements* states: "Prime numbers are more than any assigned multitude of

[1] According to some conventions, zero is not a natural number. For reasons that will be explained later, we will count zero among the natural numbers.

R. Kossak, *Mathematical Logic*, Springer Graduate Texts in Philosophy 3,
https://doi.org/10.1007/978-3-319-97298-5_1

prime numbers." In other words, there are infinitely many prime numbers. This is the celebrated Euclid's theorem. What is this theorem about? In the broadest sense, it is a statement about the world in which some objects are identified as natural numbers, about a particular property of those numbers—primeness, and about inexhaustibility of the numbers with that property. We understand what the theorem says, because we understand its context. We know what natural numbers are, and what it means that there are infinitely many of them. However, none of it is entirely obvious, and we will take a closer look at both issues later. Concerning the infinitude of primes, it occurred to me once when I was about to show the proof of Euclid's theorem in my class, to ask students what they thought about a simpler theorem: "There are infinitely many natural numbers." It was not a fair question, as it immediately takes us away from the solid ground of mathematics into the murky waters of philosophy. The students were bemused, and I was not surprised.

We will formalize Euclid's theorem in a particular way, and to do this we will have to significantly narrow down its context. In a radical approach, the context will be reduced to a bare minimum. We will be talking about certain domains of objects, and in the case of Euclid's theorem the domain is the set of all natural numbers. Once the domain of discourse is specified, we need to decide what features of its elements we want to consider. In school we first learn how to add and how to multiply natural numbers; and we will follow that path. We will express Euclid's theorem as a statement about addition and multiplication in the domain of natural numbers.

We will talk about addition and multiplication using expressions, called *formulas*, in a very restricted vocabulary. We will use *variables*, two operation symbols: $+$ and $\cdot$, and the symbol $=$ for equality. The variables will be lower case letters x, y, z, ... For example, $x + y = z$ is a formula expressing that the result of adding a number x to a number y is some number z. This expression by itself carries no *truth value*. It can be neither true nor false, since we do not assign any specific values to the variables. Later we will see ways in which we can speak about individual elements of a domain, but for now we will only have the option of *quantifying* over the elements of the domain, and that means stating that either something holds for all elements, or that something holds for some. For example:

$$\text{For all } x \text{ and all } y,\ x + y = y + x. \tag{1.1}$$

The sentence above expresses that the result does not depend on the order in which the numbers are added. It is an example of a *universal* statement; it declares that something holds for all elements in the domain.

And here is an example of an *existential* statement, it declares that objects with a certain property exist in the domain:

$$\text{There is an } x \text{ such that } x + x = x. \tag{1.2}$$

This statement is also true. There is an element in the domain of natural numbers that has the required property. In this case there is only one such element, zero. But in general, there can be more elements that witness truth of an existential statement.

For example,

$$\text{There is an } x \text{ such that } x \cdot x = x$$

is a true existential statement about the natural numbers, and there are two witnesses to its veracity, zero, and one.

Interesting statements about numbers often involve comparisons of their sizes. To express such statements, we can enlarge our vocabulary by adding a relation symbol, for example $<$, and interpret expressions of the form $x < y$ as "some number x is less than some number y." Here is an example of a true statement about natural numbers in this extended language.

$$\text{For all } x,\ y,\ \text{and } z,\ \text{if } x < y,\ \text{then } x + z < y + z. \tag{1.3}$$

Notice the grammatical form "if . . . then. . . ."

The next example is about multiplication. It is an expression without a truth value.

$$1 < x \text{ and for all } y \text{ and } z,\ \text{if } x = y \cdot z,\ \text{then } x = y \text{ or } x = z. \tag{1.4}$$

In statements (1.1), (1.2), and (1.3), all variables were *quantified* by a prefix, either "for all" or "there exists." In (1.4) the variable x is not quantified, it is left *free*; it does not assume any specific value.

Because of the presence of a free variable, (1.4) does not have a truth value, nevertheless it serves a purpose. It *defines* the property of being a prime number in terms of multiplication and the relation $<$. Let me explain how it works.

Think of a prime number, say 7, as a value of x. If I tell you that $7 = y \cdot z$, for some natural numbers y and z, without telling you what these numbers are, then you know that one of them must be 7 and the other is 1, because one cannot break down seven into a product of smaller numbers. It is true "for all y and z," because for all but a couple of them it is not true that $7 = y \cdot x$, and in such cases it does not matter what the rest of the formula says. We only consider the "then" part if indeed $7 = y \cdot z$. If the value of x is not prime, say 6, then $6 = 2 \cdot 3$, so when you think of y as 2 and z as 3, it is true that $6 = y \cdot z$, but neither y nor z is equal to 6, hence the property described in (1.4) does not hold "for all y and z."

If you are familiar with formal logic, I am explaining too much, but if you are not, it is worthwhile to make sure that you see how the formula (1.4) defines primeness. Chose some other candidates for x and see how it works. Also, notice three new additions to the vocabulary: the symbol 1 for the number one; and two connectives "and" and "or."

With the aid of (1.4) we can now write the full statement of Euclid's theorem: For all w, there is an x such that $w < x$, and for all y and z, if $x = y \cdot z$, then $x = y$ or $x = z$.

What is the difference between the statement above and the original "There are infinitely many prime numbers."? First of all, the new formulation includes the definition of primeness in the statement. Secondly, what is more important, the direct reference to infinity is eliminated. Instead, we just say that for every number w there is a prime number greater than it with such and such properties, so it follows that since there are infinitely many natural numbers, there must be infinitely many prime numbers as well. The most important however is that we managed to express an important fact about numbers with modest means, just variables, the symbols $\cdot$ and $<$, the prefixes "for every" and "there is," and the connectives: "and, " "or", and "if ... then...."

We have made the first step towards formalizing mathematics, and we did this informally. The point was to write a statement representing a meaningful mathematical fact in a language that is as unambiguous as possible. We succeeded, by reducing the vocabulary to a few basic elements. This will guide us in our second step, in which we will formally define a certain formal language and its grammar. We will carefully specify the way in which expressions in this language can be formed. Some of those expressions will be statements that can be assigned truth values—true or false—when interpreted in particular structures. The evaluation of those truth values will also be precisely defined. Some other expressions, those that contain free variables, will serve as definitions of properties of elements in structures, and will play an important role. All those properly formed expressions will be referred to as *formulas*. My dictionary explains that a formula is "a mathematical relationship or rule expressed in symbols." The meaning in this book is different. We will talk about relationships, and we will use symbols, but formulas will always represent statements. For example, the expression $b \cdot b - 4a \cdot c$ is a computational rule written in symbols, but it is not a formula in our sense, since it is not a statement about the numbers a, b, and c. In contrast, $d = b \cdot b - 4a \cdot c$ is a formula. It states that if we multiply b by itself and subtract from it the product of four times a times c, the result is d.

1.1.1 How to Choose a Vocabulary?

In the previous section, we formulated an important fact about numbers—Euclid's theorem—using symbols for multiplication a and the ordering ($<$). This is just one example, but how does it work in general? What properties of numbers do we want to talk about? What basic operations or relations can we choose? The answers are very much driven by applications and particular needs and trends in mathematics. In the case of number theory, the discipline that deals with fundamental properties of natural numbers, it turns out that almost any important result can be formulated in a formal language in which one refers only to addition and multiplication.[2] Number

[2]In our example we also used the ordering relation $<$, but in the domain of the natural numbers, the relation $x < y$ can be defined in terms of addition, since for all natural numbers x and y, x is less than y if and only if there is a natural number z such that z is not 0 and $x + z = y$.

theory may be the most difficult and mysterious branch of mathematics. Proofs of many central results are immensely complex, and they often use mathematical machinery that reaches well beyond the natural numbers. Still, a bit surprisingly, a formal system with a few symbols in its vocabulary suffices to express almost all theorems of number theory. It is similar in other branches of mathematics. The mathematical structures, and the facts about them are complex, but the vocabulary and the grammar of the formal system that we will discuss in this book are much simpler.

The *real numbers* will be defined precisely later. For the moment, you can think of them as all numbers representing geometric distances and their negative opposites. The following statement is written in a rigorous, but informal language of mathematics. It involves the concept of one-to-one correspondence. A one-to-one correspondence between two sets A and B is a matching that to every element of A assigns exactly on element of B in such a way that every element of B has a match.

> Let A be an infinite set of real numbers. Then either there is a one-to-one correspondence between A and the set of all natural numbers, or there is a one-to-one correspondence A and the set of all real numbers.

This is a variant of what is known as the *Continuum Hypothesis*. The Continuum Hypothesis can also be stated in terms of sizes of infinite sets. In the 1870s, Georg Cantor found a way to measure sizes of infinite sets by assigning to them certain infinite objects, which he called *cardinal numbers*. The smallest infinite cardinal number is $\aleph_0$ and it is the size of the set of all natural numbers. It was Cantor's great discovery that the size of the set of all real numbers, denoted by $\mathfrak{c}$, is larger that $\aleph_0$. Another way to state the Continuum Hypothesis is: if A is an infinite set of real numbers, then the cardinality of A is either $\aleph_0$ or $\mathfrak{c}$. The hypothesis was proposed by Georg Cantor in the 1870s, and David Hilbert put it prominently at the top of his list of open problems in mathematics presented to the International Congress of Mathematicians in Paris in 1900. The Continuum Hypothesis is about numbers, but it is not about arithmetic. It is about infinite sets, and about one-to-one correspondences between them. What are those objects, and how can we know anything about them? What is an appropriate language in which facts about infinite objects can be expressed? What principles can be used in proofs? Precisely such questions led David Hilbert to the idea of formalizing and axiomatizing mathematics. There is a short historical note about Hilbert's program for foundations of mathematics in Appendix B.

The Continuum Hypothesis is a statement about sets or real numbers and their correspondences. To express it formally one needs to consider a large domain in which all real numbers, their sets, and matchings between them are elements. Remarkably, it turned out that the vocabulary of a formal system in which one can talk about all those different elements, and much more, can be reduced to logical symbols of the kind we used for the domain of the natural numbers, and just one symbol for the set membership relation $\in$. All that will be discussed in detail in Chap. 6.

So for now we just have two examples of vocabularies, one with symbols + and · for arithmetic, and the other with just one symbol ∈ for set theory. We will see more examples later, and our focus will be on the number structures. In general, for every mathematical structure, and for every collection of mathematical structures of a particular kind, there is a choice of symbols that is sometimes natural and obvious, and sometimes arrived at with a great effort.

A digression: Once formalized, mathematical proofs become strings of symbols that are manipulated according to well-defined syntactic rules. In this form, they themselves become subjects of mathematical inquiry. One can ask whether such and such formal statement can be derived formally from a given set of premises. The whole discipline known as proof theory deals with such questions with remarkable successes. In 1940, Kurt Gödel proved that the Continuum Hypothesis cannot be disproved on the basis of our, suitably formalized and commonly accepted axioms of set theory, and in 1963, Paul Cohen proved that it cannot be proved from those axioms either. This is all remarkable, and was a result of a great effort in foundational studies.

1.2 Symbolic Logic

Mathematical logic is sometimes called *symbolic logic*, since in logical formulas ordinary expressions are replaced with formal symbols. We will introduce those symbols in the next section. Henri Poincaré, the great French mathematician, who was strongly opposed to formal methods in mathematics, wrote in 1908: "It is difficult to admit that the word *if* acquires, when written $\supset$, a virtue it did not possess when written *if*."[3] Poincaré was right. Nothing is gained conceptually by just replacing words with symbols, but the introduction of symbols is just a first step. The more important feature is a precise definition of the grammar of the formalized language. We are going to pay close attention to the shape of logical formulas, and the logical symbols will help. It is very much as in algebra: $x + y$, simply means x *plus* y, and $x \cdot y$ means x *times* y, but if you thought that our formalized expression for Euclid's theorem was complicated, think how complicated it would have been if we did not use $+$ and $\cdot$.

There are many advantages of the symbolic notation. It is precise and concise. One not only saves space by using symbols; sometimes symbolic notation allows one to express complex statements that would be hard to understand in the natural language. The most common symbolic system of mathematics is called first-order logic. It will be defined in this section and it will be extensively used in the rest of the book.

The mathematical notation with all its symbols and abbreviations is the language of modern mathematics that has to be learned as any other language, and learning a

[3] In Poincaré's time, $\supset$ was used to denote implication.

language takes time. I will try to limit notation to a minimum, but you have to bear with me. The language that you will learn serves communication among mathematicians well, but the fact that so much of mathematics requires it creates problems in writing about mathematics for non-mathematicians. Moreover, mathematicians have created their language in a rather chaotic historical process without particular regard to the needs of beginners. Alexandre Borovik wrote [6]: "Why are we so sure that the *alphabet* of mathematics, as we teach it—all that corpus of terminology, notation, symbolism—is natural?... We have nothing to compare our mathematical language with. How do we know that it is optimal?"

In the previous section we saw how mathematical statements are written using logical connectives and quantifiers. Now we will be writing them using *logical symbols*. Those are $\wedge$, $\vee$, and $\neg$, representing *and*, *or*, and *not*, respectively. The connectives will be used to group together statements about relations, and those statements will be composed of variables and *relation symbols*. The prefixes of the form "for all x ..." and "there is an x such that ..." are the quantifiers, the first is called *universal* and it will be written $\forall x \ldots$; the second is called *existential*, and it will be written $\exists x \ldots$ (think of ∀ll, and ∃xist).

There are two ways of introducing relation symbols. One could first define an infinite collection of symbols, and then for each structure choose only particular symbols specific to the structure. This would give us a "one language—all structures" model. Alternatively, one can first make a choice of symbols for a particular structure, or a class of structures, and use only those. The latter "one kind of structures—one language" model does not need some of the small technicalities that the former requires, so we will adopt it. Since at first we want to discuss number structures, we choose the following three relation symbols: A for addition, M for multiplication, and L for the "less than" relation. By the standard convention, regardless of the choice of other relation symbols, the equality relation symbol $=$ is also always included in the vocabulary.

For an important technical reason, we will need infinitely many variables. We will index them by natural numbers: x_0, x_1, x_2, x_3, and so on. Each formula will only use finitely many variables, but there is no limit on the number of variables that can be used. This is an important feature of first-order logic so we have to keep all those infinitely many variables in mind, and from time to time there will be a need to refer to all of them. To simplify notation, we will often drop the subscripts, and we will use other letters as well.

We are used to thinking of addition and multiplication as functions, or operations on numbers. Now I will ask you to think of them as relations.

In full generality, the language of first-order logic includes relation symbols and function symbols, but to avoid some technicalities we will not use function symbols. The word "technicalities" is one of those treacherous expressions that often hides some important issues that the author is trying to sweep under the rug, so let me offer an explanation. In mathematics one studies both functions and relations. We use mathematical functions to model processes and operations. Metaphorically speaking, a function "takes" an object as input and "produces" another object as an output. Addition is a two argument function, the input is a pair of numbers, say 1 and 3, and the output is their sum 4. Functions are useful when change is involved; when, for example, some quantity changes as a function of time. Relations are more like databases—they record relationships. Both concepts have their formalizations, and in mathematical practice the distinction between them is not sharp. A relation can evolve in time; a function can be considered as a relation, relating inputs to outputs. The technicalities hinted at above are the rules that must be obeyed when we compose functions, i.e. when we apply one function after another in a specified order. Those rules are not complicated, but at this level of exposition they would require a more careful treatment. We will not do that, and, since every function can be represented as a relation, the price that will be paid will not be great.

Let us see how addition and multiplication can be represented as relations. As we noted earlier, addition of natural numbers is a function. To each pair of natural numbers, the function assigns a value that is their sum. The inputs are pairs of numbers m, n, and the outputs are the sums $m + n$. Let us name this function f. Using function notation, we can write $2+2 = 4$ as $f(2, 2) = 4$, and $100+0 = 100$ as $f(100, 0) = 100$. We are not concerned here with any actual process of adding numbers, we think of f as a device that instantly provides a correct answer in each case. But addition also determines a relationship between ordered triples of numbers as follows. We can say that the numbers k, m, and n are related if and only if[4] $f(k, m) = n$, or, in other words, $k + m = n$. In this sense, the numbers 2, 3, and 5 are related, and so are 100, 0, and 100, but 0, 0, and 1 are not. Notice that we must be careful about the order in which we list the numbers. For example, 2, 2, and 4 are related, but 2, 4, and 2 are not. Addition as a relation carries exactly the same information as the function f does.

To go further, we must now define the rules that generate all formulas of first-order logic. A formula is a formal expression that can be generated (constructed) in a process that starts with basic formulas, according to precise rules. The definition itself is an example of a formal mathematical definition. It is an *inductive definition*. In an inductive definition, one first defines a basic collection of objects, and then describes the rules by which new objects can be constructed from those objects we already have constructed. The definition also declares that only objects obtained this way qualify.

Here is an example of a simple inductive definition. Everyone knows what a finite sequence of 0's and 1's is. It is enough to see an example or two. Here is one: 100011101. Here is another: 1111111. We recognize such sequences when we see them, but notice that this rests on an intuitive understanding of the concept of finite sequence. The inductive definition will not make any explicit references to finiteness, instead the finite character of the concept will be built into the definition. This aspect is not just a philosophical nicety, it has practical consequences. We use inductive definition in a specific way prove results about the defined concepts. The advantage of inductive definitions is that they give an insight into the internal

[4]The phrase "if and only if" is commonly used in mathematics to connect two equivalent statements.

structure of the objects they define. They show us how they are made in a step-by-step process.

Let us now define sequences of 0's and 1's inductively. We begin by saying that 0 and 1 are finite sequences.[5] These are our basic objects. Then comes the inductive rule: if s is a finite sequence then so are $s0$ and $s1$. Finally, we declare that the finite sequences of 0's and 1's are only those objects that are obtained from the basic sequences 0 and 1 by applying the inductive rule (over and over again).

Now we go back to formulas. Recall that we chose A, M, and L for relation symbols. A and M are *ternary*—they bind three variables, and L is *binary*—it binds two variables. "Binding" is a technical term, and you should not put more meaning to it beyond what is written in the following definition of basic formulas, which we will call *atomic*. Atomic formulas are all expressions of the form $A(x_i, x_j, x_k)$, $M(x_i, x_j, x_k)$, $L(x_i, x_j)$, and $x_i = x_j$, where i, j, and k are arbitrary natural numbers. For example, $A(x_0, x_1, x_2)$, $M(x_5, x_3, x_1)$, $M(x_1, x_0, x_0)$, $L(x_1, x_2)$ are atomic formulas. In other words, an atomic formula is a relation symbol followed by a list of three arbitrary variables, in the case of A and M, or two arbitrary variables, in the case of L. Notice that since there are infinitely many variables, there are also infinitely many atomic formulas.

When we discuss particular examples of formulas, for greater readability we will usually drop the subscripts and use other letters for variables, so expressions such as $A(x, y, z)$ or $L(y, x)$ (not a typo) will also be considered atomic formulas, although in the strict sense, according to the definition they are not.

Along with formal expressions, defined according to strict rules, we will also use other common mathematical expressions and those will also often use symbols. The two different kinds of symbols should not be confused. We will need names for many different objects, including formulas and sentences of first-order logic. We will see this in the definition below. The Greek characters φ (phi), ψ (psi), and other, are used as names for formulas of the language we define. They are not a part of the formalism.

Definition 1.1

1. Every atomic formula is a formula.
2. If φ and ψ are formulas, then so are $(\varphi) \wedge (\psi)$, $(\varphi) \vee (\psi)$, and $\neg(\varphi)$.
3. If φ is a formula, then, for each n, $\exists x_n(\varphi)$ and $\forall x_n(\varphi)$, are formulas.
4. There are no other formulas.

Notice the use of parentheses. They play an important role. They guarantee that every first-order formula can be read in only one way.[6]

[5] We could actually start one level lower. We could say that the empty sequence, with no symbols at all, is a finite sequence of 0's and 1's.

[6] This unique readability of first-order formulas is not an obvious fact and requires a proof, which is not difficult, but we will not present it here.

Let us see how Definition 1.1 works. Per clause (1.1), the expressions $x = y$ and $A(x, z, y)$ (no typos here) are formulas, because they are atomic formulas.[7] Per rule (1.2), $\neg(x = y)$ and $\exists z(A(x, y, z))$ are formulas. By applying (1.2), we see that

$$(\neg(x = y)) \wedge (\exists z(A(x, y, z)))$$

is a formula as well. Let us call this formula $\varphi(x, y)$.[8] The only displayed variables are x and y, because they are free in $\varphi(x, y)$. The third variable z is bound by the existential quantifier $\exists z$. Think of A as the addition relation of the natural numbers. For what values of x and y does $\varphi(x, y)$ become a true statements. First of all they must be different, as declared by the first component of $\varphi(x, y)$, but also x must be less than y, because only then there is a natural number z such that $x + z = y$.

An important caveat. According to the rules, we are free to choose any variables we like to form atomic formulas, so for example $A(x, z, z)$, and $L(x, x)$ are well-formed formulas. If the same free variable is used in different places in a formula, when we interpret the formula in a structure, that variable will always be evaluated by the same element, but this does not mean that if the variables are different that they represent different objects. We are free to evaluate any free variable by any object, in particular we can use the same object for different variables.

Let us recapitulate. The list of symbols of the first-order logic is: $\wedge$, $\vee$, $\neg$, $\exists$, $\forall$, and $=$, and then for each particular structure, in addition, it includes a collection of relation symbols. In our case, we chose A, M, and L. Each relation symbol has a prescribed *arity* which is given in the definition of the atomic formulas. The symbols A and M are of arity three, and L is of arity two. This means that, for example, $A(x_0, x_1, x_2)$ is a well-formed atomic formula, but $A(x_0, x_1)$ and $L(x_0, x_1, x_2)$ are not, because the number of variables does not match the arity of the symbol.

The attentive reader will ask: But what about all those parentheses and commas? Yes, they are also formal symbols of our language, and their use is entirely determined by Definition 1.1. The rules for commas are hidden in clause (1.1). One could do without them. For example $Ax_0x_1x_2$ also represents a uniquely recognizable string of symbols, and this is all we want, but for greater readability, and to avoid additional conventions that we would have to introduce in the absence of parentheses, we use parentheses and commas. Often we will also use "[," and "]," and sometimes "{" and "}." They all have the same status as "(," and ")."

[7] This is an example of mathematical pedantry. Of course, you would say, they are formulas. They are even atomic formulas! But when we defined atomic formulas, we defined a special kind of expression, and called expressions of this kind "atomic formulas." When we did that, the formal concept of formula had not been defined yet. To know what a formula of first-order logic is one has to wait for a formal definition of the kind we gave here. To avoid this whole discussion we could have called atomic formulas atoms. If we did that, then clause (1.1) of the definition above would say "Every atom is a formula," but since the term "atomic formula" is commonly used, we did not have that choice.

[8] This is another example of an informal abbreviation.

Each application of rules (1.2) and (1.3) introduces a new layer of parentheses. If we continue this way, formulas quickly become unreadable, but this formalism is not designed for the human eye. We sacrifice easy reading, but we gain much in return. One bonus is that it is now easy to check and correct grammar. The only grammatical rules are those in Definition 1.1. In particular, in every formula the number of left parentheses must be equal to the number of right parentheses. If it is not, the sequence of symbols is not properly formed and it is not a formula.

The most important aspect of Definition 1.1 is that it shows how all formulas are generated in a step-by-step process in which more and more complex formulas are generated. This is a crucial feature, that opens the door to investigations of formal languages by mathematical means. Another essential feature is that the set of all formulas is generated without any regard to what those formulas may express. In fact, most formulas do not express anything interesting at all. For example

$$((x = x) \vee (x = x) \wedge (x = x))$$

is a proper, but uninteresting formula, and so is $\exists x L(y, z)$.

What is the point of allowing meaningless formulas? What we are after are formulas and sentences that express salient properties of structures and their elements, but we would be at a loss trying to give a mathematical definition of a meaningful formula. It is much easier to accept them all, whatever they may be expressing. There is something profound in treating all formulas this way. Meaningfulness is a vague concept. A sentence of no interest today, may turn out to be most important tomorrow, so it would make no sense to eliminate any of them in advance, but this is not the main point. Most mechanically formed formulas are not only uninteresting, they actually make no sense at all. Still we want to keep them in, because it is the price to pay for the clarity of the definition. Moreover, there are also some unexpected technical applications. If φ is a formula, then, according to rule (1.2), so are $(\varphi) \wedge (\varphi)$ and $(\varphi) \wedge ((\varphi) \wedge (\varphi))$, and $(\varphi) \wedge ((\varphi) \wedge ((\varphi) \wedge (\varphi)))$, and so on. Nothing new is expressed, but there are some important results in mathematical logic that depend in an essential way on existence of such statements.

The definition of the syntax of first-order logic is completed. Now it is time to define the *semantics*, i.e. the procedure that gives meaning and truth values to formulas when interpreted in a structure. We already did that informally, when we talked about Euclid's theorem and interpretations of formulas in the natural numbers. Full definition of semantics for first-order logic is based on Alfred Tarski's famous *definition of truth* from 1933 [34]. It is formulated in a set-theoretic setting that we will discuss later. For now, we will show how it all works using examples. For a full formal definition consult any textbook on mathematical logic. A good source online is the Stanford Encyclopedia of Philosophy [13].

We will interpret formulas in the domain of the natural numbers. To begin with, for any three numbers m, n, and k, we need to assign truth values (true or false) to all atomic formulas $A(x, y, z)$ and $M(x, y, z)$, $L(x, y)$, and $x = y$, when x, y, and z are interpreted as m, n, and k respectively. We declare $A(x, y, z)$ to be true if and only if $m + n = k$, $M(x, y, z)$ to be true, if and only if $m \cdot n = k$, $L(x, y)$ to be true if and only if m is less than n, and finally $x = y$ to be true if and only if m equals n. We are exceedingly pedantic here, and for a good reason. We just described the definition of truth for the atomic formulas.

What makes notation complicated in the explanations above is the reference to evaluation of the variables. To simplify matters, one is tempted to assign truth values directly to expressions such as $A(m, n, k)$. There is a problem with that. The expression $A(x, y, z)$ is a formula. It is just a string of symbols of the language of first-order logic. The expression $A(m, n, k)$ is not a formula. The letters m, n, and k, as used here are informal names for numbers. No rule in Definition 1.1 allows inserting names of objects into formulas. In the expression $A(m, n, k)$ two worlds are mixed. The relation symbol A, the parentheses and commas, come from the world of syntax; m, n, and k are not symbols of the formal language, they are informal names of elements of the domain of the structure.

What is the difference between the statement "$A(x, y, z)$ is true, when x, y and z are interpreted as m, n, and k" and the statement "$m + n = k$"? The former states that a certain *truth value* is assigned to a certain formula under certain conditions. The latter is a statement about the state of affairs in a certain structure. While the definition is telling us under what conditions certain statements are true, it has nothing to do with whether we can actually check if those conditions are satisfied. In the case of checking whether m plus n equals k, think of numbers so incredibly large that there is not enough space to write them down. We are not talking of any practical aspects of computation here. Still, it makes sense to define the truth values of interpretations of formulas this way. The definition is precise, and it is exactly this definition that makes a bridge between the syntax and the world of mathematical objects in which it is interpreted.

Once the definition of truth values for atomic formulas is established, truth values for more complex formulas are determined in a way parallel to the rules for generating formulas in Definition 1.1. For example, $\neg(\varphi)$ is true if and only if φ is false; $(\varphi) \wedge (\psi)$ is true, if and only if both φ and ψ are true; and $\exists x(\varphi)$ is true if and only if there is an evaluation of the variable x under which φ becomes true. Here we take advantage of the inductive form of Definition 1.1. In the same way in which the more complex formulas are inductively built from simpler ones, the truth values of more complex formulas are inductively determined by the truth values assigned to their simpler components, with the atomic case as the base. As was mentioned earlier, the full formal definition of this process is somewhat technical, and we will omit it.

In our discussion of Euclid's theorem, we included "if ... then ..." among the logic connectives. Conditional statements of the form "if φ then ψ" abbreviated by $(\varphi) \implies (\psi)$, are essential in mathematics, but Definition 1.1 has no provision for them. One could add another clause there explaining how $(\varphi) \implies (\psi)$ is to be interpreted, but this is not necessary. In classical logic, the formula $(\varphi) \implies (\psi)$ is defined as an abbreviation of $\neg(\varphi) \vee (\psi)$, hence its interpretation is already covered by the Definition 1.1. Let us see it on an example we already discussed.

The formula (1.4) defining prime numbers in the previous section included the following conditional statement:

$$\text{For all } y \text{ and all } z, \text{ if } x = y \cdot z, \text{ then } x = y \text{ or } x = z.$$

Its symbolic version is

$$\forall y \forall z[(M(y, z, x)) \Longrightarrow [(y = x) \vee (z = x)]], \tag{1.4$'$}$$

which in turn is equivalent to

$$\forall y \forall z[\neg(M(y, z, x)) \vee [(y = x) \vee (z = x)]]. \tag{1.4$''$}$$

Convince yourself (1.4$''$) is true only if x is interpreted as either 1 or a prime number.

Another common logical connective is "if and only if." The symbol of it is $\Longleftrightarrow$, and $(\varphi) \Longleftrightarrow (\psi)$ is defined as an abbreviation for $((\varphi) \Longrightarrow (\psi)) \wedge ((\psi) \Longrightarrow (\varphi))$. For an example, see Exercise 1.5.

A first-order property is a property that can be expressed in first-order logic, which means it can be defined by a formula in the formalism we just described. This whole book is about mathematical structures and their first-order properties. Not all properties are first-order. For example, here is a property of natural numbers that is not defined in a first-order way

$$\text{Every set of natural numbers has a least element.} \tag{1.5}$$

In (1.5) we quantify over sets of numbers, and that makes this statement *second-order*. In first-order logic we can only quantify over individual elements of domains, but we cannot quantify over sets of elements. Quantification over sets is allowed in second-order logic with its special syntax and semantics. There is a third-order logic that allows quantification over sets of sets of elements. There are higher-order logics, each with stronger expressive powers. There is more about this in Chap. 14.

We will stick to first-order logic for two reasons. One is that even with its restrictions, first-order logic is a strong enough formal framework for a substantive analysis of mathematical structures in general, but there is also another appealing reason. First-order logic is based on rudimentary principles. It only uses simple connectives "and," "or," and "not," and the quantification only allows us to ask whether some property holds for all elements in a domain ($\forall$), or if there is an element in a domain with a given property ($\exists$). In other words, it is a formalization of the most basic elements of logic, and one could argue that it also captures some basic features of perception. Let's think of collections of elements and some of their properties that can be visually recognized. If I see a set of elements having a property $\varphi(x)$, I also see its complement consisting of the elements that do not have that property, which is the same as having the property $\neg\varphi(x)$. If some elements have a property $\varphi(x)$, and some have another property $\psi(x)$, then I can see the collection of elements with both properties, i.e. the set of defined by property $\varphi(x) \wedge \psi(x)$.

Similarly I can see the elements having one property or the other: $\varphi(x) \vee \psi(x)$. If all elements have a property $\varphi(x)$, I see that $\forall x \varphi(x)$. To see that it is not the case, it is enough to notice one element that does not have the property, so it is enough to see that $\exists x \neg\varphi(x)$. The first-order approach provides a basic framework for what I will call *logical visibility*. Equipped with this framework, we will try to find out what can, and what cannot be seen in structures through the eyes of logic.

Here is a rough outline of what we will do next. To define a structure, we start with a collection of individual objects sharing certain features. In each structure, the objects in the collection are related to one another in various ways. We will give those relations names, and then we will try to see what properties of the structure and its individual elements are first-order. An analysis of the complexity of formulas and sentences of first-order logic will allow us to apply geometric intuitions, and to see geometric patterns in the structure. In this sense, the formalism will allow us to go back to more natural, unformalized ways of thinking about the structure, and to "logically see" some of its features, that otherwise might have stayed invisible. This works well in mathematics and we will examine some examples.

1.2.1 Trivial Structures

The simplest structures are domains with no relations on them. Think of a domain with five objects. If those objects are not related to one another in any way, this is an example of a *trivial* structure. What can be said about it? Not much more than what we have said already, but it is good to keep trivial structures in mind for further discussion. They are a good source of examples and counterexamples. Due to our convention, the equality relation is always among the relation symbols for any structure. Hence, even though a trivial structure has no relations of its own, it still has the equality relation, and it allows us to express specific facts about it in the first-order way. Consider the sentence:

$$\exists x \, \exists y \, [\neg(x = y) \wedge \forall z \, (z = x) \vee (z = y)].$$

It says that there are two distinct elements and any element in the structure must be one of them. In other words, it expresses that the structure has exactly two elements. In a similar way, for any number n, one can write a first-order sentence expressing that the structure has exactly n elements.

It is an interesting fact, that follows from the compactness theorem for first-order logic, that while for each number n, having exactly n elements is a first-order property, having a finite number of elements is not. The compactness theorem and an argument showing why finiteness is not a first-order property are presented in Chap. 11.

Exercises

Exercises marked by the asterisk are more advanced.

Exercise 1.1 *Write first-order sentences expressing the following:*

1. *There are at least three elements.*
2. *There are at most five elements.*
3. *There are either three, four, or five elements.*

Exercise 1.2 *Write the Euclid's theorem as a first-order sentence using the ternary relation symbols A and M and the binary symbol L.*

Exercise 1.3 *Twin primes are prime numbers that differ by* 2*. For example,* 3 *and* 5 *are twin, and so are* 11 *and* 13*. The Twin Primes conjecture says that there are infinitely twin prime numbers. We do not know if the conjecture is true, although there has been recent progress in number theory suggesting that it may be. Express the Twin Primes conjecture by a first-order sentence using the ternary relation symbols A and M and the binary symbol L.*

Exercise 1.4 *The Goldbach conjecture says that every even number greater than two is a sum of two prime numbers. For example:* $4 = 2 + 2$, $6 = 3 + 3$, $8 = 3 + 5$, $100 = 3 + 97$. *Express the Goldbach conjecture by a first-order sentence using the ternary relation symbols A and M, and the binary symbol L.*

Exercise 1.5 ** Later in the book, instead of fully formal expressions, we will use more readable notation. For example, instead of $A(x, y, z)$ we will simply write $x + y = z$. We will also use abbreviations. In the example below, $P(x)$ stands for a first-order sentence expressing that x is prime. Fermat's theorem says that a prime number p is a sum of two squares if and only if $p = 4m + 1$, for some natural number m. To express the theorem in a more formal way, one can write*

$$\forall x[P(x) \Longrightarrow (\exists y \exists z\ (x = y^2 + z^2) \Longleftrightarrow \exists t\ (x = 4t + 1)].$$

Try to write this sentence using the relation symbols A, M, and L, as in the previous exercises. It will be long.