4 The Language of the First-Order Logic

We are now ready to summarize and reflect on what we have accomplished: the development of a formal artificial language capable of representing most kinds of communication that the general mainstream mathematical community is interested in expressing.[1] We call this language the *First-Order Logic*, and it is determined by the syntax and semantics we have specified for

› writing down expressions as well-formed formulae,

› transforming formulae according to the rules of inference,

and the important assumption we have made that the universe of discourse—which we are free to specify depending on the kind of conversation we are interested in having—should never be empty. Our notion of proof is encoded by the $\vdash$ meta-symbol, and the collection of all propositions that are provable within our setup is what *characterizes* our language: any other setup that is capable of proving the same logical theorems we can prove is equivalent to ours.[2] In order to begin practicing *mathematics* the language of the first-order logic, there are certain things we need to introduce and explain so that we know *what* it is we're talking about.

1. A *universe of discourse* Ω that informally describes the objects of our discourse.
2. A list of *variable symbols*, for which we usually use Latin or Greek characters (with or without numeral subsscripts). For our purposes, we will always assume we have an infinite list of variable symbols available to us, and we will not be too careful about what symbols we use if it is clear from context.
3. A list of *constant symbols*, which are special symbols—usually Latin, Greek, or Hebrew—that denote specific distinguished members of Ω.
4. A list of *relational symbols* over Ω that we call *"predicates."* These are things that behave like functions, whose *inputs* are objects from the universe of discourse, and whose *outputs* are truth values.
5. A list of *functional symbols* that denote *"functions"* from Ω to Ω.[3] These are things that behave like functions, whose ***inputs are objects*** from the universe, and whose ***outputs are other objects*** in the universe of discourse.
6. A list of *axioms*, each of which is a *proposition* of the first-order logic, that represent *true* statements we believe without justification.

We have the liberty of specifying, for each of the above types, as many different things as we want.[4] But this liberty is also a burden; if we don't explain what these are *a priori*, we will not be able to speak meaningfully about mathematics using the language of the first-order logic.

Once we have made a choice of universe, introduced our predicate and functional symbols, and written down our axioms, we have specified a *first-order theory* about the kinds of objects in our chosen universe. Since the few theories we will study will all be based on the language of the first-order logic, we will refer to these simply as *theories* for our purposes this semester

[1] The careful reader will note two things here. First, this language is merely *capable* of most mainstream mathematical expression; however, it is very interesting (and, if one is interested in mathematical logic, important) to note that it is not the *only* language we could have constructed, and it actually is not the most "natural" language either—communicating in the first-order logic sometimes requires us to to translateour thoughts in unintuitive ways because of its inherent limitations (for example, it is much more natural to express induction as one sentence rather than as *schema* that lists infinitely many sentences, but arithmetic does not admit a finite first-order axiomatization). Second, not *all* of mathematics can be done in the first-order logic—some important areas of mathematics require moving to a different foundation altogether. In fact, there are fundamental *"self-reflective"* aspects of mathematics that can't be expressed in any formal language.

[2] We have taken the perspective that the first-order logic is the language *"generated by"* the axioms (we have none) and rules of inference (we have many) we established using the $\vdash$ operation. However, a different, simple, and useful perspective we could have taken is that the first-order logic is just *"the collection of strings—built according to our grammar—that we consider to be* provable.*"* This is a common perspective in algorithmic analysis, computability theory, and model theory.

[3] It actually is not *strictly* necessary for us to introduce the concept of *"functional"* symbols—we could technically get away with just having *"relational"* symbols. However, this would require us to contort our language in unnatural ways, so it is convenient to introduce this concept.

[4] For example, we could choose to introduce one constant symbol, zero predicate symbols, and one unary functional symbol; this is what we will do when we study the theory of arithmetic. We could also choose to introduce no constants, twenty different predicate symbols of different arities, and zero functional symbols; that would be a different theory that might be interesting to study for other, unrelated purposes. It's totally up to us to decide what we want to talk about.

4.1 *Predicate Equality*

The first-order logic deliberately does not define any specific predicate symbols up-front in order leave that flexibility to the speaker. There is, however, one exception: the binary predicate for *equality* is thought to be so fundamental that *any* use of the first-order logic *must* end up defining this predicate (or implicitly specifying enough machinery that something equivalent to equality can be implemented).

$=$ The symbol $=$ is universally used in association with this predicate, with the informal definition that $=(x, y)$ means that the symbols x and y *refer to the* ***same*** *object* in the universe of discourse.

Axiom of Equality 0: Identity of Indiscernibles.
For any *wff* φ, the following universal equivalence holds.

$$\forall x \forall y \Big(x = y \Rightarrow \big(\varphi(x) \leftrightarrow \varphi(y) \big) \Big)$$

An alternative, but equivalent, way to phrase this would be as follows below.

$$\forall x \forall y \Big(\big(x = y \wedge \varphi(x) \big) \rightarrow \varphi(y) \Big)$$ ⌟

This says that, if we know that $x = y$ for two objects x and y, then the idea that *identical* things are *"indiscernible"* is what justifies us saying that any property that x was must also be shared by y; otherwise, we would have a basis for *discerning* between x and y, and they would therefore not be *equal.* The remaining three axioms of equality describe the fundamental fact that equality is a kind of *"equivalence relation."*

Axiom of Equality 1: Reflexivity.
$\forall x (x = x)$. ⌟

Axiom of Equality 2: Symmetry.
$\forall x \forall y ((x = y) \rightarrow (y = x))$. ⌟

Axiom of Equality 3: Transitivity.
$\forall x \forall y \forall z (((x = y) \wedge (y = z)) \rightarrow (x = z))$. ⌟

With these axioms, if we know that $x = y$ for some objects x and y, we can *"substitue"* any occurrence of x with y in any expression in which we encounter it, and *vice versa.*

4.2 *Blurring the Line Between Syntax and Semantics*

Now that we are preparing to study our first mathematical theory, there is one final note about syntax that we should be aware of. The honest reality is that most mathematicians have not deeply studied logic. Even if they have, a lot of the technical considerations logicians concern themselves with are not really problems for mainstream mathematics because of the fact that most *"everyday"* mathematics can be expressed using the tools of the classical first-order[1] logic. Now, this doesn't mean we learned all of this for nothing, and it doesn't permit us to be careless. There are serious dangers to relaxing when engaged in these kinds of intellectual arts, as we will find out later. Yet, our choice of logical foundation here comes with some almost unbelievable benefits, and we should take a minute to describe and understand them. We can think of the first-order logic as being the *strongest*

[1] Or using a suitably conservative extension of classical first-order logic, such as classical second-order logic.

"effective" logical system with certain properties.[1] Not all of these properties are uniquely characteristic of the first-order logic, and honestly we are not yet ready to understand all of these properties, but there are two important ones that we *can* wrap our heads around right now.[2] The first theorem is table-stakes for any logical system. We will first say this informally, but we also give a more formal statement of the result afterwards for those interested.

[1] We are paraphrasing a pair of results here attributed to Per Lindström, commonly called *"Lindström's theorems."*

[2] While the second of these two results is due to Kurt Gödel, there are many famous mathematicians whose names *should* be mentioned in this discussion, but which we will omit for the sake of not turning this into an encyclopedia.

Idea 4.1: The First-Order Logic is Consistent.
The first-order logic does not prove any contradictions; in particular, there is no way to write a proof of $\bot$ (or anything equivalent to $\bot$) in the first-order logic. ⌟

Theorem 4.1: The First-Order Logic is Consistent.
Fix an arbitrary theory with a (nonempty) universe of discourse Ω. If φ is a *wff* and the first-order logic *with equality* is capable of proving φ in this universe Ω with the axioms this theory establishes, then $\varphi \not\equiv \bot$, which is equivalent to saying $\varphi \equiv \top$. ⌟

This means it is impossible to prove $\bot$ in the first-order logic, and this follows from the fact that our rules of inference are *truth-preserving* transformations of sentences: if you fix a theory (by specifying all of the things mentioned in the first section of this article) and take a finite collection of sentences that are all *true* and apply a rule of inference, then the resulting expression must also be *true.* The presence of $\bot$ as either a theorem or an axiom would destroy our logical system because of the fact that, for any proposition p, we know $\bot \vdash p$, so *every* statement would become a theorem, making the concepts of *"proof"* and *"truth"* meaningless.

While the previous theorem says that every proof is about something *true,* our second theorem tells us the converse: every universally *true* statement is *provable.*

Idea 4.2: The First-Order Logic is Complete.
Every universally *true* proposition has a proof. ⌟

Theorem 4.2: The First-Order Logic is Complete.
If φ is a *wff* such that, for any (nonempty) universe of discourse Ω and any theory over Ω, we know that $\varphi \equiv \top$ in that theory, then φ has a proof within that theory. ⌟

The *completeness theorem* acts like the converse of the *soundness theorem,* and together they help us understand the equivalence between *"truth"* and *"proof"* in the classical first-order logic. There are many other (sometimes *"higher-order"*) logics that are stronger, or more expressive, than the first-order logic, but most of them lose one of these two important properties.[3]

[3] The other ones lose some of the more esoteric—but still important—properties that we will not discuss here.

As a result of these metalogical properties, and thanks to the fact that we know that

$$(\varphi \vdash \psi) \vdash (\varphi \to \psi)$$
$$(\varphi \to \psi) \vdash (\varphi \vdash \psi)$$

both hold for any two *wffs* φ and ψ, we can essentially ignore the formal distinction between the $\to$ and $\vdash$ symbols and replace them with just one symbol: the $\Rightarrow$ double-stroked single-headed arrow. Similarly, the distinction between $\leftrightarrow$ and $\dashv\vdash$ (and $\equiv$) can be ignored, and we can replace those with just one symbol: the $\Leftrightarrow$ double-stroked double-headed arrow.