First-Order Logic: an Introductory Guide

April 17, 2024

Section 0: Overview

Logic effectively forms the basis for the entire field of mathematics. Whenever you see a mathematical result, there must always be a proof behind that result, and behind that proof there must always be a logic which connects observation to conclusion, line to line. Therefore, it is only natural that mathematicians would want a way to formalize and see clearly the underlying logic behind all mathematical proof.

Of course, there are quite a lot of basic logical notions in mathematics, so today we will be restricting ourselves to first-order logic, which excludes certain concepts (such as quantifying over subsets) for the sake of being able to make certain very convenient claims (such as the compactness theorem). In order to study logic, we need some terminology to describe what we're seeing, and it turns out that there are quite a lot of symbols employed in order to write things concisely. We will discuss this in the Section 1, before moving on to what we can actually do with logic in future sections.

Section 1: Definitions and Symbols

In this section, we will introduce the symbols used in first-order logic and their definitions. This will be somewhat fast-paced, so be ready!

Proposition/Sentence (usually p or q by default, going down the alphabet if needed)**:** A proposition (or a sentence) is a statement that is either true or false. For instance, "2=1+1" is true[citation needed], so if p were the proposition "2=1+1," then we would say "p" is a true statement.

Not (¬)**:** This symbol indicates the negation of a statement. For instance, if p were the proposition "1517 is prime," we would say that "¬p is true." For technical reasons, it is probably better to define the word "false" in terms of the negation of a statement being true. I'll go into a little more detail about this when we get to contradictions. We are assuming the Law of the Excluded Middle here, so for any proposition p, p and ¬(¬p) are equivalent.

And (∧)**:** When we say that p∧q is true, this is equivalent to saying that p is true and q is true. This definition is a little circular, but that's because at a fundamental level questioning the definition of "and" is best left to philosophers. You can also think of it as the "and" between p and q in the statement "p and q" or "p∧q" being a more "mathematical" sense of the word, while the "and" in the definition is just the plain english that is required to form any definition.

Or (∨)**:** When we say that p∨q is true, this is equivalent to saying that *at least one* of p and q is true. This is an inclusive or; there is no single common mathematical symbol (at least to my knowledge) for exclusive or (although we'll see a way to represent the exclusive or when we get to "if and only if." Note that p∨q and ¬(¬p∧¬q) are logically equivalent; in this sense, the symbol is a little redundant (but saving space like this is very important).

If/Implies (→)**:** It is perhaps best to *define* p→q as ¬p∨q, or perhaps ¬(p∧¬q). This latter formulation makes it most clear what is going on here; effectively it means that p→q is true when there is not a counterexample to the claim "if p, then q" in a vernacular sense. There are a few things to note about this.

1) A case where q is true and p is false (i.e. ¬p is true) does not contradict p→q. For instance, the existence of the number 2 does not contradict that if an integer is divisible by 4, it is divisible by 2.
2) As a consequence of our definition, if p is always false or q is always true, then p→q is always true. For instance, if the sun is neon green, then the Riemann Zeta Hypothesis is true. Conversely, if the sun is not neon green, then 23 is prime.
3) As the above examples show, "if" does not necessarily have anything to do with causality. Therefore, the term "implies" is sometimes favored.
4) An interesting thing about "if" is that, in order to be meaningful in the intuitive sense, the propositions in the "if" statements usually need to have some sort of quantifier involved. We'll get to this soon, but you may have noticed that our current definition of "proposition" doesn't even cover the statement "if an integer is divisible by 4, it is divisible by 2." This is because our current definition of proposition is better-suited for introducing these symbols but doesn't actually quite cover all the ground we need it to. For that, we'll need to talk about formulae.

If and Only If ($\leftrightarrow$)**:** The statement $p\leftrightarrow q$ is equivalent to $(p\rightarrow q)\wedge(q\rightarrow p)$. Note that $p\leftrightarrow q$, $q\leftrightarrow p$, $\neg q\leftrightarrow\neg p$, and $\neg p\leftrightarrow\neg q$ are equivalent, and we can also use $p\leftrightarrow\neg q$ to function as an exclusive or.

Contradiction: If at any point in a proof we are forced to conclude that the statement $p\wedge\neg p$ holds, we have reached a contradiction. This is a problem, because if we can prove $p\wedge\neg p$ for a proposition p, we can prove *any* proposition, and logic will cease to have meaning. See exercise 1.2.

That does it for the simple logical relation symbols. Before moving onto quantifiers, as I hinted before, we need to first update our notion of "proposition" to allow statements like "an integer is divisible by four" without specifying what that integer is.

Formula: A formula is like a proposition, but it can depend on one or more free variables. For instance, the statement "x is divisible by four" is a formula dependent upon the free variable x, and we could write it p(x). Have we just completely stolen the notation and word for "formula" and used it in a very different way? Yes. Welcome to mathematics.

Notably, a formula dependent on one or more free variables is either true or false, but only depending on what the variable actually is. For instance, p(x) as above has p(4) true and p(3) false.

A formula does not necessarily have to depend upon a free variable. If it doesn't, that formula is also a proposition, although from here on out we'll call it a "sentence."

Free Variable: A variable inside a formula that isn't "bound" to any quantifier. That is, a variable that doesn't yet have any context telling it what it should be.

What's a quantifier? Well, it's a symbol that sort of "specifies" what we're looking for in a particular free variable (you'll see what I mean in a second). In first-order logic, we have two (really one) quantifiers:

For All ($\forall$): This quantifier says that we're considering all possible values of the variable. For instance, $\forall x(p(x))$ means exactly what it seems like it should; it's true when, for all x, the formula p(x) is true.

There Exists ($\exists$): This quantifier says that we're considering if something is true for any value of the variable. For instance, $\exists x(p(x))$ means exactly what it seems like it should; it's true when, for at least one x, the formula p(x) is true. Note that $\exists x(p(x))$ is equivalent to $\neg(\forall x(\neg p(x))$, so the quantifier is technically redundant.

When we quantify, we are always quantifying "over" something (the thing that goes right after the symbol). In first-order logic, that something is always variables; we are not allowed to make statements like "there exists a subset."

Also, if you're familiar with this sort of notation, you might think that for something like "$\exists x$" you'd need to specify where you're allowed to take x from; for instance, you might think of $\exists x \in \mathbb{Z}$ to mean "there exists x in the integers." However, in first-order logic, you often want to be thinking of the "universe" you're working in as something you're taking into account when you're evaluating the truth of a proposition, not in writing the proposition itself.

For instance, consider the proposition p(x) as "x^2 is not equal to x+1." Instead of saying something like "$(\forall x \in \mathbb{Z}(p(x)))$ is true" and "$\neg(\forall x \in \mathbb{R}(p(x)))$ is true," one would rather say "$\forall x(p(x))$ is true over the integers and false over the reals."

In fact, it is probably better to think of formulae as being constructed in their universe, and therefore the set you're taking the variable from as being by nature constrained to whatever place you made the formula in. This is because otherwise one could ask "what is the truth value of 'x^2 is not equal to x+1' when x is taken from the set of cat breeds." This is clearly nonsense, and the entire structure of logic breaks down in cases like this where something is permitted to be undefined. But if we're defining formulae as being part of their "universe," then we'd never even be able to construct p(x) in the universe of cat breeds unless we've defined addition and multiplication over cat breeds anyways.

All this to say: when we have a sentence like "$\exists x p(x)$," there is no ambiguity in what set x is taken from because we have (presumably) already defined that set when we created p(x).

Got it? Alright, let's do some exercises.

For the following exercises, p, q, r, s… are any propositions.

1.1) As rigorously as possible, show that $(((p \wedge \neg q) \rightarrow r) \wedge \neg r) \rightarrow (p \rightarrow q)$. This is proof by contradiction, essentially.

1.2) Show that $(p \wedge \neg p) \rightarrow q$ without assuming $\neg(p \wedge \neg p)$. You may, however, assume that $(\neg(p \wedge q)) \wedge p \rightarrow \neg q$ for any propositions p, q (this is known as the disjunctive syllogism).

1.3) Convince yourselves that $(((p \wedge q) \vee (r \wedge \neg p)) \wedge (r \leftrightarrow p)) \rightarrow (q \wedge r)$. You don't have to do this rigorously, it's just good for intuition.

1.4) Show that $\forall x(p(x)) \leftrightarrow \neg(\exists x(\neg p(x)))$ in any universe.

1.5) Find a counterexample to the claim $\forall x \exists y(p(x)) \leftrightarrow \forall y \exists x(p(x))$.

For the following exercises, we are working in the integers. Don't worry about formal proofs, just state whether they are true or false.

1.6) $\forall x \forall y \exists z(x+z=y)$
1.7) $\forall x \forall y \exists z\ (x*z=y)$
1.8) $\exists x \forall y \exists z\ (x*z+x+z=y)$
1.9) $\forall x\ (\exists y(4y=x*x) \vee \exists y(4y+1=x*x))$
2.0) $\exists x \forall y\ (\exists z\ x*z+1=y) \rightarrow (\forall z\ \neg(z+z+x=y))$

Section 2: Arithmetic

We've been discussing a lot of abstract logical concepts, but now let's see how it works in practice. We'll start by setting up the first-order "peano axioms" of arithmetic in the nonnegative integers--the sentences we assume are true. Here "s" is the successor function, which is the function that adds 1. The other symbols non-first-order symbols we will need are the functions + and *, and the constant symbol "0" (a constant symbol just refers to a fixed constant in our "universe"). We call this the set of these symbols a **language**, consisting of **function symbols** s, +, and *, **constant symbols** 0, and no **relation symbols**. Don't worry about this now, but we'll come back to it in a bit.

1) $\forall x\ \neg(s(x)=0)$
2) $\forall x\ ((\neg(x=0)) \rightarrow (\exists y\ (s(y)=x)))$
3) $\forall x\ x+0=x$
4) $\forall x \forall y\ x+(s(y))=s(x+y)$
5) $\forall x\ x*0=0$
6) $\forall x \forall y\ x*(s(y))=(x*y)+x$

And then for every formula p(a, b^), where b^ is the vector of free variables in p besides the first, the sentence

7) $\forall x\hat{}\ [[p(0, x\hat{}) \wedge \forall y(p(y, x\hat{}) \rightarrow p(s(y), x\hat{})] \rightarrow \forall z(p(z, x\hat{}))]$.

Why is this allowed? Well, saying $\forall x\hat{}$ isn't a problem, because it's just shorthand for $\forall x_1 \forall x_2 \forall x_3 \dots \forall x_n$, where n is the number of terms in b^, noting that formulae must be finite.

And secondly, there is nothing stopping us from having infinite axioms, so we are, in fact, allowed to do this for every formula. The formulae themselves are composed of first-order syntax, so however weird this is, it technically works.

2.1) Explain what the set of axioms encoded in 7) actually means.

Also note that we are allowed to use the symbol = in formal logic anywhere, regardless of the "universe" we are in. The symbol = has three critical "axioms"--really just assumptions at this point--attached to it, and in addition has the ever-important substitution property. We won't go into more detail on this, but it is very important to know we can do this.

2.2) From the axioms of Peano arithmetic, prove that $\forall x(s(0)*x=s(0))$.
2.3) From the axioms of Peano arithmetic, prove that $\forall x(0*x=0)$.

Section 3: Into the Deep End (Gödel's completeness theorem and consequences)

Some definitions:

A **theory** T is a set of sentences. More precisely, an **L-theory** T is a set of sentences made up of sentences from a **language** L, which as you might remember consists of some symbols for functions, relations (such as "greater than" or a relation determining being in the same equivalence class), and constants.

We say a theory is **inconsistent** if we can prove a **contradiction** assuming everything in the theory; otherwise it is consistent. Remember that a **contradiction** is a sentence of the form $p \wedge \neg p$.

We say a theory is **satisfiable** if it has a model. Don't worry about what a "model" is here; it's basically just a way of saying there's some mathematical structure out there where everything in the theory is true.

Gödel's Completeness Theorem states that if T is a theory and p is a sentence, then we can prove p from T if and only if p is true in all models of T.

Since no mathematical models will ever have $p \land \neg p$ true for any sentence p (an existing mathematical model with a live contradiction would lead to some serious problems), if a theory is inconsistent, it cannot have a model by Gödel's completeness theorem.

Meanwhile, if a theory is unsatisfiable, it has no models, meaning that for each sentence p, $p \land \neg p$ is true. Therefore, by Gödel's completeness theorem, we can prove $p \land \neg p$ from T, so T is inconsistent.

So we have just proven that a theory T is consistent if and only if it is satisfiable.

3.1) Use the above claim and the fact that proofs are finite to show that a theory T is satisfiable if and only if every finite subset of T is satisfiable. (Compactness)
3.2) Use 3.1 to prove that there is an uncountable model in which every first-order sentence true of the natural numbers is also true in that model. Remember: theories can contain uncountably many sentences.

Section 4: Conclusions

What does this all mean? Well, to start, first-order logic is a strange place. Because of the lack of quantification over subsets, first-order logic allows Gödel's Completeness Theorem, which gives rise to a very strong tool in the form of the Compactness Theorem but also has some truly strange results.

In general, first-order logic does a decent job of describing (some) mathematical structures, but its strengths show where either its results can be generalized or where you need to illuminate the inner workings of some elementary proof. Logic, although not necessarily first-order, is the last step you get to in mathematics by repeatedly asking "why" at a proof (past logic, you start running into philosophy very quickly). As a result, it is extremely fundamental and therefore can be applied almost everywhere in mathematics.

4.1 (if time): Describe the language and axioms of graph theory. What about linear orders?