Note 1: Logic

CS 70, Summer 2024

1 Propositions

Mathematics is a search for truth. Unlike other disciplines, wherein a lack of precise truth is where they find their beauty and joy, we have precise truths in mathematics. To this end, we'll start by establishing exactly what kinds of things we'll be studying in our study of mathematics: statements that are true or false. We will slowly build our way up to proving such statements.

A statement which is either true or false is called a **proposition**. The following are examples of propositions:

- "$\sqrt{3}$ is irrational";
- "$1 + 1 = 5$"; and
- "Julius Caesar had two eggs for breakfast on his 10^{th} birthday."

Each of the above statements is either true or false. At the present moment, we don't know whether the first statement is true or false, but it has to be one or the other. The second statement that $1 + 1 = 5$ is false. And the final statement about Julius Caesar is again either true or false—although it's doubtful we'll ever know the answer to that one, through mathematics or not.

Here are some non-examples of propositions:

- "$2 + 2$";
- "$x^2 + 3x = 0$";
- "Cats are awesome";
- "Henry VIII was unpopular."

The first statement has no truth value. Something like "$2 + 2 = 4$" would indeed be a proposition, but "$2+2$" alone is not. Similarly, "$x^2+3x = 0$" is not a proposition as we don't know what x is. The statement is true when $x = 0$, but it's false when false $x = 1$. We'll revisit such kinds of statements later, but they are not propositions. The last two are not propositions because they use vague or subjective terms. The third statement is more an opinion, unless "awesome" is explicitly defined to be a set of characteristics which cats share. The same is true of the fifth statement: "unpopular" is ill-defined.

Propositions form the domain of interest for mathematics. They are the statements that are true or false, and we will work towards proving or disproving the truth of such statements. Thus we have defined the domain of interest for our search for truth in mathematics.

2 Propositional Logic

Now that we have looked at what kinds of things "have truth," our next order of business is to think about how truth behaves. In particular, how can we combine different truth values, and what rules govern such combinations? This line of thinking brings us to **propositional logic** (sometimes abbreviated PL), which is the study of how truth behaves under different functions which combine propositions.

2.1 The Boolean Connectives

We will usually let capital letter variables represent propositions, such as P, Q, R, etc. For example, we could let P be the proposition that "3 is odd." Such variables, which represent arbitrary propositions, are called **propositional variables**.

The simplest way of joining propositions together is through the use of the **connectives** "and," "or," and "not." Let P and Q be any two propositions.

- **Conjunction**. $P \wedge Q$, said in English as "P and Q." The conjunction is true only when both P and Q is true.

- **Disjunction**. $P \vee Q$, said in English as "P or Q." The disjunction is true only when at least one of P and Q is true.
- **Negation**. $\neg P$, said in English as "not P." The negation is true only when P is false.

These connectives allow us to combine propositions to get more complex ones. For example we can consider the proposition

$$\text{“}\sqrt{3} \text{ is irrational} \wedge 1 + 1 = 5,\text{”}$$

which is false since the second clause of the conjunction is false. Note that we didn't need to know whether $\sqrt{3}$ was irrational to determine this falsity of the proposition.

Statements like P, $P \vee Q$, and $P \wedge \neg Q$, which are created by combining propositional variables with connectives are called **propositional formulas**. While propositions like "$\sqrt{3}$ is irrational" and "$1 + 1 = 5$" are either true or false by themselves, the truth value of a propositional formula depends on the assignment of truth values to its propositional variables. For example, the propositional formula $P \wedge Q$ could be true if both P and Q are assigned to be true, or it could be false if P is assigned to be false. By plugging in actual propositions into a propositional formula, we get a proposition.

Propositional logic is not equipped with the tools to determine whether a particular proposition like "$\sqrt{3}$ is irrational" is true or not. Instead, propositional logic studies propositional formulas. By studying propositional formulas, which involve propositional variables rather than actual propositions, we are able to study how truth behaves without worrying about the truth values of the specific propositions we are using. That way, when we plug in actual propositions for the propositional variables, we'll know that whatever result we showed about the propositional formula still holds even after it's been converted into a proposition.

2.2 Tautological Truth and Equivalence

We will consider a notion of truth which depends entirely on the structure of the propositional formula (that is, the connective structure which forms it). This makes the name "propositional logic" a bit of a misnomer, since it doesn't actually care about whether the propositions themselves are true or false; rather, it cares about which propositional *formulas* are true or false.

We will start with an example. The propositional formula $P \vee \neg P$, known as the **law of the excluded middle**, says that any proposition P is either true or it is not. It is quite intuitive that this should always be true; in fact, we defined a proposition to be any sentence that was either true or false. This is our first example of finding truth in mathematics.

Thus we introduce our first notion of truth. A propositional formula which is *always* true, regardless of the truth values of its propositional variables, is called a **tautology**. We will say that such a propositional formula is **tautologically true**. This means that we can plug in *any propositions* for the propositional variables of such propositional formulas and get a true proposition. For example, we know that "$\sqrt{3}$ is irrational or $\sqrt{3}$ is rational" is true *regardless* of whether or not "$\sqrt{3}$ is irrational" because the propositional form we substituted it into is a tautology.

The "opposite" of the law of the excluded middle is the propositional formula $P \wedge \neg P$, which states that P is both true and not true. Such a propositional formula is *always* false, and we call it a **contradiction**, or say that it is **tautologically false**.

Concept Check 1. Reason whether each of the following propositional formulas is a tautology, a contradiction, or neither.

- $\neg(P \vee \neg P)$.
- $P \wedge Q$.
- $\neg(P \wedge \neg P)$.
- P.

While we state that it is intuitive that $P \vee \neg P$ is always true and that $P \wedge \neg P$ is always false, such intuition does not suffice a tautological proof. Instead, to prove this, we will introduce a tool known as a **truth table**. A truth table is an algorithm which we can use to verify whether any given propositional formula is a tautology. As such, it is an algorithm for discerning truth.

Truth tables are like function tables. A propositional formula takes as input the truth values of its atoms, so the truth table enumerates all possible inputs, along with the resulting truth values of the propositional formula of interest.

Below we provide the truth tables for the conjunction, disjunction, and negation connectives.

P	Q	$P \wedge Q$
T	T	T
T	F	F
F	T	F
F	F	F

P	Q	$P \vee Q$
T	T	T
T	F	T
F	T	T
F	F	F

P	$\neg P$
T	F
F	T

Let's use these truth tables to determine the truth tables for $P \vee \neg P$. We'll fill out the truth table one connective at a time.

We can start by filling in the truth values of P, since those are given as the truth table inputs.

P	P	$\vee$	$\neg P$
T	T	?	?
F	F	?	?

Next, we'll fill in the truth values of $\neg P$, which we can find using the negation truth table.

P	P	$\vee$	$\neg P$
T	T	?	F
F	F	?	T

Finally, we can find the truth value of the entire expression $P \vee \neg P$ using the disjunction truth table. The truth values we found for P and $\neg P$ are the inputs to the disjunction. We highlight the column in green to indicate that it corresponds to the truth value of the entire propositional formula. The other columns correspond to the truth values of the atoms of the propositional formula.

P	P	$\vee$	$\neg P$
T	T	T	F
F	F	T	T

Thus we have shown that no matter what the truth value of P is, the propositional formula $P \vee \neg P$ is always true. Hence it is tautologically true.

Concept Check 2. Use a truth table to prove that the propositional formula $P \wedge \neg P$ is tautologically false.

2.3 The Implication Connective

Another extremely important connective is the connective for reasoning: implication. Let P and Q be any two propositions.

- **Material Implication**. $P \rightarrow Q$, said in English as "if P, then Q.

Here, P is known as the **hypothesis** or **antecedent**, and Q is known as the **conclusion** or **consequent**.

To understand what the truth table of the material implication should look like, let's consider a couple of examples in English.

- "If you stand in the rain, you'll get wet." This says that whenever you're standing in the rain, you'll get wet.

  Note that you could get wet in other ways than standing in the rain, so this statement is different than the statement "If you get wet, then you're standing in the rain."

- "If you eat your dinner, then you can have dessert." This says that when you eat your dinner, you'll get dessert. It's possible you could get dessert another way, maybe by sneaking it from the fridge when no one is looking.

  So this statement is different from the statement "You can have dessert only if you eat dinner," where in this case the only way for you to get dessert is by eating dinner.

It's clear that whenever the antecedent P is true, then the consequent Q must also be true for $P \to Q$ to be true. If you were standing under the rain but you did not get wet, that would render the first material implication false.

But what about when P is false? In the first material implication, it might false that you're standing in the rain and false that you're wet. But the implication is still true. Similarly, it might be false that you're standing in the rain but true that you're wet, e.g., you jumped into a pool. But the implication is still true.

So the material implication is false only when P is true but Q is false. This yields the following truth table for the material implication.

P	Q	$P \to Q$
T	T	T
T	F	F
F	T	T
F	F	T

Now consider the propositional formula $\neg P \lor Q$. The truth table is worked out below.

P	Q	$\neg P$	$\lor$	Q
T	T	F	T	T
T	F	F	F	F
F	T	T	T	T
F	F	T	T	F

Note that the propositional formulas $P \to Q$ and $\neg P \lor Q$ have the same truth table. We say that two propositional formulas P and Q are **tautologically equivalent** if they have the same truth table, and we write $P \equiv Q$. We have just proved that $P \to Q$ and $\neg P \lor Q$ are tautologically equivalent. Two propositional formulas are tautologically equivalent when they are saying the same thing. In particular, they combine truth values in precisely the same way as one another. This fact will be useful for us to prove certain statements by transforming them into tautologically equivalent statements which are easier to prove.

The propositional formula $P \to Q$ can be said many ways in English.

- "if P, then Q"
- "Q if P"
- "P only if Q"
- "P is sufficient for Q"
- "Q is necessary for P"
- "Q unless not P"

Concept Check 3. Explain, in English, why each of the above statements is the same as $P \to Q$.

We introduce one final connective. Let P and Q be any two propositions.

- **Material equivalence**. $P \leftrightarrow Q$, said in English as "P if and only if Q." The material equivalence is true when both $P \to Q$ and $Q \to P$ are true.

Let's consider a couple of examples in English.

- "I wear a hat if and only if it's sunny." If it's sunny, that means I'll wear a hat. But this also means that if I'm wearing a hat, it must be sunny, since that's the only time I wear a hat. The two conditions can only be true at the same time.
- "If and only if it's his birthday, Chidi eats cake." If it's Chidi's birthday, he eats cake. But if he's eating cake, it must be his birthday, since that's the only time he eats cake. Again, the two conditions can only be true at the same time.

The truth table for material equivalence is given below.

P	Q	$P \leftrightarrow Q$
T	T	T
T	F	F
F	T	F
F	F	T

Concept Check 4. Prove that $P \leftrightarrow Q$ is tautologically equivalent to $(P \to Q) \land (Q \to P)$.

Note that in the truth table, $P \leftrightarrow Q$ is true only when P and Q have the same truth values. This means that two propositional formulas are tautologically equivalent if and only if their material implication is tautologically true!

Concept Check 5. Prove once more that $P \to Q \equiv \neg P \lor Q$ by showing that $(P \to Q) \leftrightarrow \neg P \lor Q$ is tautologically true.

Why do we care about tautological equivalence? Because it allows us to transform more difficult propositional formulas into equivalent ones which may be easier to work with.

For any two propositions P and Q, consider the propositional formulas $P \to Q$, $\neg Q \to \neg P$, and $Q \to P$. The propositional formula $\neg Q \to \neg P$ is called the **contrapositive** of $P \to Q$, and the propositional formula $Q \to P$ is called the **converse** of $P \to Q$. How are these propositional formula related to the original implication $P \to Q$?

Their truth tables are given below.

P	Q	$P \to Q$	$\neg Q \to \neg P$	$Q \to P$
T	T	T	T	T
T	F	F	F	T
F	T	T	T	F
F	F	T	T	T

Concept Check 6. Confirm that the above truth tables are correct.

In particular, observe that these truth tables prove that $P \to Q \equiv \neg Q \to \neg P$; they are tautologically equivalent. If we ever wanted to prove that P implies Q, we could instead prove it by proving that $\neg Q$ implies $\neg P$.

As an example, consider the implication "if you stand in the rain, you'll get wet" from earlier. Its contrapositive is "if you didn't get wet, then you didn't stand in the rain" and its converse is "if you get wet, then you stood in the rain."

The converse is a different claim than the original implication; in the original implication, you could have gotten wet through some other way (like jumping into a pool). The converse claims that the only way to get wet is to stand in the rain.

However, the contrapositive is the same statement as the original implication. If you didn't get wet, you might have done a number of things, but none of them could have been standing in the rain. This is the same as the original implication, which says whenever you stand in the rain, you get wet.

Concept Check 7. Write down the contrapositive and the converse of the implication "if you eat your dinner, then you can have dessert."

2.4 De Morgan's Laws

The class of tautological truths and equivalences in propositional logic is rich and varied. Together, they form what is known as the *Boolean algebra*, a set of base tautological truths and equivalences that can be used to prove other, more complex, tautological truths and equivalences. We have already seen one such tautology: the law of the excluded middle. Here, we will explore a pair of equivalences in the Boolean algebra.

Theorem 1. *De Morgan's Laws.* The following tautological equivalences are true.

$$\neg(P \wedge Q) \equiv \neg P \vee \neg Q$$
$$\neg(P \vee Q) \equiv \neg P \wedge \neg Q.$$

Proof. We prove that $\neg(P \wedge Q) \equiv \neg P \vee \neg Q$. The truth tables of $\neg(P \wedge Q)$ and $\neg P \vee \neg Q$ are given below.

P	Q	$\neg$	$(P$	$\wedge$	$Q)$	$\neg P$	$\vee$	$\neg Q$
T	T	F	T	T	T	F	F	F
T	F	T	T	F	F	F	T	T
F	T	T	F	F	T	T	T	F
F	F	T	F	F	F	T	T	F

Since the truth values of $\neg(P \wedge Q)$ and $\neg P \vee \neg Q$ are the same, the two are tautologically equivalent.

Concept Check 8. Prove the second tautological equivalence from De Morgan's laws.

2.5 Summary

Propositional logic studies the truth of propositional formulas, which are created by combining propositions together with connectives like $\wedge$, $\vee$, $\neg$, $\rightarrow$, and $\leftrightarrow$. In propositional logic, we were able to define tautological truth and tautological equivalence. A propositional formula is tautologically true if it is always true regardless of the truth values of its propositions. Two propositional formulas are tautologically equivalent if they always share the same truth values. To prove tautological truth or equivalence, we must use truth tables.

3 First-Order Logic

While it's nice to prove that certain propositional forms are always true or equivalent to one another, the theory above is somewhat boring. Everything can be decided by comparing truth tables, and the theory provides us nothing to actually prove the propositions themselves. For example, we saw above that we could prove $P \rightarrow Q$ by instead proving $\neg Q \rightarrow \neg P$, but we still have no idea how we would to prove either statement.

We've also seen that the $\rightarrow$ connective behaves a little strangely because it's based only on the truth values of its inputs rather than any meaningful notion of implication. For example, a proposition like "$0^2 = 0 \rightarrow$ every square is a rectangle" is true, despite the fact that it doesn't seem like $0^2 = 0$ has much to do with the fact that every square is a rectangle.

The first issue—the inability of propositional logic to prove that a proposition is true—is because it is not an expressive enough system of logic. Because of how it abstracts away every statement to just some proposition, the theory of propositional logic does not allow us to reason about propositions the way we want to.

Consider the below statements.

- "For any natural number n, $n^2 + n + 41$ is prime."
- "If an integer n is odd, so is n^3."
- "There is an integer k which is both even and odd."

These statements assert something about lots of propositions, all of which exist within some greater structure of numbers. For example, the first statement is asserting that $0^2 + 0 + 41$ is prime, $1^2 + 1 + 41$ is prime, and so on. Moreover, we all agree on some rules about the way the integers work which would make a statement about $n^2 + n + 41$ either true or false. In propositional logic, all of this structure is stripped away, since the entire statement is just reduced to some proposition P.

Let's consider an example. We said earlier than $x^2 + 3x = 0$ was not a proposition, since we don't know what the variable x is. However, this becomes a proposition if we plug in a specific value of x: "$0^2 + 3(0) = 0$" is a true proposition, while "$1^2 + 3(1) = 0$" is a false proposition. Such statements are called **predicates**. A predicate is a function which takes as input some element from a domain (such as the real numbers, matrices, animals, quantum fields, etc.) and outputs a proposition.

Concept Check 9. Determine which of the following are predicates.

- "$n^2 + n + 41$ is prime."
- "p is an electron $\wedge$ p has a negative charge."
- "x is a cool number."
- "b is studying data science $\rightarrow$ b is studying statistics."

As we saw in the examples at the beginning of this section, we're able to transform a predicate into a proposition by quantifying the inputs to the predicate which return true propositions. We'll use a very coarse degree of quantification: either saying that there is at least one input which returns a true proposition, or that every possible inputs returns true propositions. By quantifying a predicate, we're able to create a **first-order sentence**. A first-order sentences is just a proposition which uses quantifiers, like "there is a real number x such that $x^2 + 3x = 0$" or the other examples from the beginning of this section. The study of the truth of these kinds of propositions is known as **first-order logic**, sometimes abbreviated as FOL.

As with propositions, we will typically let capital letter variables represent predicates. For example, we could let $P(x)$ be the predicate "x is odd." We call a variable representing an arbitrary predicate a **predicate variable**.

We quantify over predicate variables in two ways. The two **quantifiers** of first-order logic are the existential and universal quantifiers. Let $P(x)$ be any predicate.

- **Existence**. $\exists x P(x)$, said in English as "there exists an x such that $P(x)$." An existential statement is true only when there is an object satisfying the predicate.
- **Universal**. $\forall x P(x)$, said in English as "for all x, $P(x)$." A universal statement is true only when every object satisfies the predicate.

There are rules to how these quantifiers can be applied. A variable x, y, z, etc. can only be quantified over if it is not already bound by a quantifier. For example, consider the binary predicate "$x < y$." If I create a new unary predicate "$\exists y(x < y)$", the variable y is bound by the existential quantifier and we can no longer quantify over it. Statements like $P(x)$, $\exists x P(x)$, and $\forall x P(x)$, which are created by combining predicate variables and quantifiers are called **first-order formulas**. In a valid first-order formula, all variables must be bound to a quantifier.

Concept Check 10. Classify each of the following as either a predicate, a first-order formula, or neither.

- $\exists x \exists x P(x)$.
- $\exists y \exists x P(x, y)$.
- $\forall x P(y)$.
- $\forall x \forall y P(y)$.
- $\exists x \forall y P(x, y)$.

Unlike a propositional formula, we aren't able to determine the truth of a first-order formula by checking the truth values of its input, since the predicate variables could represent *any* predicate. Instead we will consider *all possible predicates*.

For any set of first-order formulas, a **model** consists of the following two pieces.

(1) A nonempty set of objects $\mathfrak{D} \neq \varnothing$, typically referred to as the **domain**, along with the rules that govern them.

(2) An interpretation for each of the predicates which appear in the first-order formulas.

Consider the first-order formula $\exists \forall x P(x)$. Here are some possible models for this formula.

- The real numbers ($\mathfrak{D} = \mathbb{R}$) with their usual arithmetic (addition, multiplication, subtraction, and division) with an interpretation $P(x)$: "x is rational."

  In this model, the formula is false, since $x = \pi$ is not rational.

- The natural numbers ($\mathfrak{D} = \mathbb{N}$) with their usual arithmetic (addition and multiplication) with an interpretation $P(x)$: "$x \geq 0$."

  In this model, the formula is true, since all natural numbers greater than or equal to 0.

- A group of people $\mathfrak{D} = \{\text{Akemi}, \text{Benoit}, \text{Chidi}\}$ where Akemi is a data science major, Benoit is a computer science major, and Chidi is a computer science and data science double major, with an interpretation $P(x)$: "x is a data science major."

  In this model, the formula is false, since Benoit is not a data science major.

Concept Check 11. Come up with another model for the first-order formula $\forall x P(x)$.

We can quickly see that there are infinitely many models for any first-order formula.

In a model with a finite domain, we can express quantified propositions using just conjunctions and disjunctions. For example, if we have a model with the domain $\mathfrak{D} = \{1, 2, 3, 4\}$ and some predicate P, then

$$\forall x S(x) \equiv S(1) \wedge S(2) \wedge S(3) \wedge S(4) \qquad \text{and} \qquad \exists x S(x) \equiv S(1) \vee S(2) \vee S(3) \vee S(4).$$

If our domain $D = \{a_1, a_2, \ldots\}$ has infinitely many elements, we can think of these like infinite conjunctions and infinite disjunctions:

$$\forall x P(x) \text{ is like } \bigwedge_{i=1}^{\infty} P(a_i) \qquad \text{and} \qquad \exists x P(x) \text{ is like } \bigvee_{i=1}^{\infty} P(a_i).$$

However, first-order logic does not permit infinitely long propositions, so this interpretation is not strictly true. However, the intuition provided by thinking about our quantifiers this way is quite precise, and can help in thinking about how these things work.

3.1 Logical Truth, Implication, and Equivalence

While we have strengthened the expressiveness of our language, we have lost our earlier notions of truth and equivalence. In particular, we can no longer use truth tables to prove that something is tautologically true or that two propositions are tautologically equivalent. As we saw in the previous section, the truth of a statement like $\forall x P(x)$ depends on the model we are working with. This brings us to discussing *logical* truth and equivalence, which are weaker forms of tautological truth and equivalence.

We say that a statement of first-order logic is **logically true** if every model makes it true. This is akin to how a statement of propositional logic is tautologically true if every possible truth assignment makes it true, but much, much more general. To reason about whether a first-order logic statement is true, we must reason about every possible model we could create of our statement. For example, we saw in the previous section that the statement $\forall x P(x)$ is not logically true. There are some models which make it true: for example, the natural numbers with the predicate $P(x)$: "$x \geq 0$." But there are also models which make it false: for example, the real numbers with the predicate $P(x)$: "x is rational."

Concept Check 12. Determine which of the following are logically true.

- $\forall x(x = x)$.
- $\exists x(x = x)$.
- $\exists x(x \neq x)$.
- $\forall x \forall y(x = y \rightarrow y = x)$.

We say that two first-order formulas are **logically equivalent** if every model which makes one formula true makes the other formula true. Again, this is similar to the idea of tautological equivalence, except we now must consider everything in an infinite class of models rather than a finite set of truth assignments. For example, the statements $\forall x P(x)$ and $\forall y P(y)$ are logically equivalent. Any model of $\forall x P(x)$ must satisfy the predicate P for all elements in its domain; the same is true for any model of $\forall y P(y)$.

Finally, we return to the idea of implication. We earlier discussed the material implication $\rightarrow$, and talked about its truth table. We saw that it behaves a little differently from the way we typically think about implication, since it depends only on the truth values of its two arguments rather than any notion of the hypothesis implying the conclusion. For first-order logic we introduce a new operation, $\implies$, which we call the **logical implication**. For φ and ψ two first-order formulas, we say that the implication $\varphi \implies \psi$ is true if it's the case that any model which makes φ true also makes ψ true. We use the same terminology: φ is known as the hypothesis, and ψ is known as the conclusion. Note that in this class, we will use the logical implication $\implies$ in place of the material implication $\rightarrow$ without concern.

This is the most natural extension of the idea of implication to the much more expressive world of first-order logic, and it aligns nicely with our own intuitions about how implication works. When we say an if/then statement like "if you stand under the rain, then you'll get wet," we're not saying anything about how the sentence behaves under an assignment of truth values to the propositions "you stand under the rain" and "you'll get wet." Rather, we're saying that any world where you're standing under the rain is also a world where you'll get wet.

By comparing the definitions of logical truth and logical equivalence, we can see that two first-order formulas φ and ψ are logically equivalent by showing that they logically imply one another: $\varphi \implies \psi$ and $\psi \implies \varphi$. If two first-order formulas φ and ψ are equivalent, we write $\varphi \iff \psi$.

3.2 Quantifier Rules

It seems that we have given ourselves quite a daunting task. To prove whether things are logically true, equivalent, or implied, we need to reason about all possible models—of which there are infinitely many! Fortunately, like our Boolean connectives, the quantifiers are also governed by rules which allow us to manipulate them. These are the instantiation and generalization rules. We will first discuss the universal insantiation and the existential generalization rules.

- Universal instantiation (UI). $\forall x P(x) \implies P(c)$.

  Suppose we have established that $\forall P(x)$. Then for any element c of our model, we can infer that $P(c)$.

- Existential generalization (EG). $P(b) \implies \exists x P(x)$.

  Suppose we have established that some element b of our model satisfies $P(b)$. Then we can infer that $\exists x P(x)$.

  It is important to note that if we existentially generalize different objects in our domain, we must use different quantifier variables for each one. For example, if we know $R(b, c)$, we can infer $\exists x R(x, c)$ and then infer $\exists y \exists x R(x, y)$. But we cannot infer $\exists x R(x, x)$.

Let's try and use these rules to prove some logical truths.

Example 1. $\forall x S(x) \implies \exists x S(x)$.

Consider any model which makes the first-order formula $\forall x S(x)$ true.

Let c be any element in the domain of our model. By universal instantiation, $S(c)$.

Therefore, by existential generalization, $\exists x S(x)$.

Thus we have shown that any model of $\forall x S(x)$ is also a model of $\exists x S(x)$, so we have shown that $\forall x S(x) \implies \exists x S(x)$.

Concept Check 13. Show that the implication $\forall x \exists y R(x, y) \implies \exists x \exists y R(x, y)$ is logically true.

Now we discuss the existential instantiation and universal generalization rules.

- Existential instantiation (EI). $\exists x P(x) \implies P(b)$ for some new b.

  Suppose we have established that $\exists x P(x)$. Then there is at least one object in the domain which satisfies P, and we can give it a name, b. So we can infer $P(b)$.

  It is important to note that we must use a *new* name. We cannot use the same name as another element of the domain, since we don't know which element it is in the domain which satisfies the predicate. For example, if we know $\exists x R(x, b)$, we cannot infer $R(b, b)$. Rather, we must create a new name c and infer $R(c, b)$.

- Universal generalization (UG). $P(c)$ for c arbitrary $\implies \forall x P(x)$.

  Suppose we introduce a new name c to stand for a completely arbitrary member of the domain and are then able to prove $P(c)$. Then we can conclude $\forall x P(x)$.

  Universal generalization can be difficult to conceptualize. The key part of this inference rule is that we make no assumption about c. This is what allows us to universally generalize to the entire model, since our object c could have been any element of the domain.

Let's try out each of these rules. Here's an example using the existential instantiation rule.

Example 2. $\big(\exists x P(x) \wedge \forall x (P(x) \implies Q(x))\big) \implies \exists x Q(x)$.

Consider a model which makes $\exists x P(x) \wedge \forall x (P(x) \implies Q(x))$ true.

By existential instantiation, let b be the object satisfying $\exists x P(x)$, so $P(b)$.

By universal instantiation, we have that $P(b) \implies Q(b)$.

But then, since the model makes $P(b)$ true and $P(b) \implies Q(b)$ true, it must also make $Q(b)$ true.

So we can existentially generalize to $\exists x Q(x)$.

Thus we have shown that any model of $\exists x P(x) \wedge \forall x(P(x) \implies Q(x))$ is also a model of $\exists x Q(x)$, so we have shown that $\left(\exists x P(x) \wedge \forall x(P(x) \implies Q(x))\right) \implies \exists x Q(x)$.

Note that in the example, our hypothesis has an existential and universal statement, and we used existential and universal instantiation. Our conclusion has an existential statement, and we used existential generalization. This is a common pattern—instantiation rules are how quantified statements are used, and generalization rules are how quantified statements are demonstrated.

The next two examples use the universal generalization rule.

Example 3. $\left(\forall x P(x) \wedge \forall x(P(x) \implies Q(x))\right) \implies \forall x Q(x)$.

Consider any model which makes $\forall x P(x) \wedge \forall x(P(x) \implies Q(x))$ true.

Consider an arbitrary element c of that model. By universal instantiation, $P(c)$ and $P(c) \implies Q(c)$.

So the model must make $Q(c)$ true.

But c was arbitrary, so this is true for any element of the domain: by universal generalization, we have that $\forall x Q(x)$.

We have shown that any model of $\forall x P(x) \wedge \forall x(P(x) \implies Q(x))$ is a model of $\forall x Q(x)$, so we have shown that $\forall x P(x) \wedge \forall x(P(x) \implies Q(x)) \implies \forall x Q(x)$.

Again, observe that the claims in the hypothesis are utilized through universal instantiation, and the claim in the conclusion is shown using universal generalization.

Example 4. Consider a unary predicate S and a binary predicate R satisfying $\forall x \forall y(S(x) \wedge S(y) \implies R(x, y))$. Show that this logically implies that $\forall x(S(x) \implies R(x, x))$.

Consider any model which makes $\forall x \forall y(S(x) \wedge S(y) \implies R(x, y)$ true.

Consider an arbitrary element a of the domain.

By universal instantiation, $\forall y(S(a) \wedge S(y) \implies R(a, y))$.

By universal instantiation again, $S(a) \wedge S(a) \implies R(a, a)$. This is the same as $S(a) \implies R(a, a)$.

Since a was arbitrary, we can universally generalize: $\forall x(S(x) \implies R(x, x))$.

We have shown that any model $\forall x \forall y(S(x) \wedge S(y) \implies R(x, y)$ is a model of $\forall x(S(x) \implies R(x, x))$.

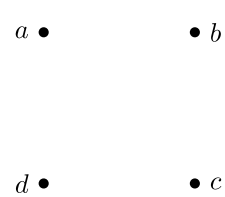
As one final example, we consider the logical implications of $\forall x \exists y R(x, y) \wedge \exists y \forall x R(x, y)$.

Example 5. Which of the following are logically implied by $\forall x \exists y R(x, y) \wedge \exists y \forall x R(x, y)$?

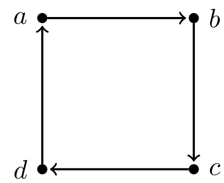
- $\exists y R(y, y)$.
- $\forall x \forall y R(x, y)$.

An extremely helpful tool for any proof is to play around with small examples. In this context, that means we must create models of the two hypotheses and check whether the conclusions also hold in our models. Very simply models usually work—dots that are labelled by unary predicates P and Q, or dots that have arrows between them for binary predicates R.

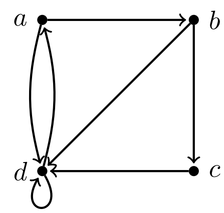
We begin to construct a model of $\forall x \exists y R(x, y) \wedge \exists y \forall x R(x, y)$. Our domain $\mathfrak{D} = \{a, b, c, d\}$ will be the set of dots below. Our interpretation of $R(x, y)$ will be "there is an arrow pointing from x to y."



The first condition $\forall x \exists y R(x, y)$ requires that every dot points to some dot. A simple way to make this condition true is to create a cycle of arrows.



The second condition $\exists y \forall x R(x, y)$ requires that there is some dot which every dot points to. Let's have that be dot d. We add arrows so that every dot (including d) points to d.



This model makes the first-order formula $\forall x \exists y R(x, y) \wedge \exists y \forall x R(x, y)$ true. We check if it makes each of $\exists y R(y, y)$ and $\forall x \forall y R(x, y)$ true.

Our model indeed makes $\exists y R(y, y)$ true, since it makes $R(d, d)$ true. This doesn't prove that $\forall x \exists y R(x, y) \wedge \exists y \forall x R(x, y) \implies \exists y R(y, y)$ is true, since we would need to show that it's true for *any* model, and not just the one specific model that we constructed above.

Our model makes $\forall x \forall y R(x, y)$ false. For example, it is not the case that $R(a, c)$ in our model. This suffices to disprove the implication, because we've found at least one model which makes $\forall x \exists y R(x, y) \wedge \exists y \forall x R(x, y)$ true but makes $\forall x \forall y R(x, y)$ false.

Let's return to $\exists y R(y, y)$. We saw that it might be true. Let's think about why it's true in our model. The second condition required that there were some dot such that every dot pointed to it, including that dot itself. We chose to use d to make that existential claim true, which is why we had $R(d, d)$. But it seems like this should indeed hold in general. Let's prove it.

> Consider any model which makes $\forall x \exists y R(x, y) \wedge \exists y \forall x R(x, y)$ true.
>
> By existential instantiation on the second condition, there is some object u in the domain such that $\forall x R(x, u)$.
>
> By universal instantiation on $\forall x R(x, u)$, this must also be true for our object u: $R(u, u)$.
>
> By existential generalization on $R(u, u)$, we have that $\exists y R(y, y)$.
>
> We have shown that any model which makes $\forall x \exists y R(x, y) \wedge \exists y \forall x R(x, y)$ true must also make $\exists y R(y, y)$ true.

Note that we used a name u other than a, b, c, or d in our existential instantiation, since we have to use a *new* name.

Concept Check 14. Prove or disprove whether each of the following are implied by $\forall x \exists y R(x, y) \wedge \exists y \forall x R(x, y)$.

- $\exists x \exists y R(x, y)$.
- $\exists t \forall s R(s, t)$.
- $\forall x \forall y (R(x, y) \vee R(y, x))$.

- $\exists x \exists y(R(x,y) \wedge R(y,x))$.

3.3 Restricted Quantification

Everything we have proved up until now have been fundamental logical truths about predicates and logic. But what if we want to talk about a statement like "for any natural number n, n^2+n+41 is prime"? If we write that in first-order logic as

$$\forall n(P(n^2+n+41)),$$

it won't make any sense, and it certainly won't be logically true. A model might assign a different interpretation to P than "is prime," or the model might not even be over numbers!

So what can we do? One solution is to, in some sense, bake the model into the predicate form:

$$\forall n(n \in \mathbb{N} \implies P(n^2+n+41)).$$

This way, the statement is true in models which aren't the natural numbers, since the hypothesis is false in such models. The only models which will make the hypothesis true are models of the natural numbers, so we have effectively restricted the models we are interested in to those of the natural numbers. There is the issue that P could yet be interpreted as some other predicate under different models of the natural numbers, but this can be fixed by explicitly defining our predicate in terms of first-order logic:

$$P(n) : n \neq 1 \wedge \forall y \forall z(y \cdot z = n \implies y = 1 \vee z = 1).$$

Thus the question of whether the proposition

$$\forall n\Big(n \in \mathbb{N} \implies \big(n^2+n+41 \neq 1 \wedge \forall y \forall z(y \cdot z = n^2+n+41 \implies y = 1 \vee z = 1)\big)\Big)$$

is truly limited to the familiar structure of the natural numbers and the rules they obey.

Of course, it can be a little annoying to have to explicitly write out every predicate we wish to use in first-order logic. We will usually simply explain what the predicates we are working with are in English, and assume that there is an appropriate first-order logic translation.

Let's consider our earlier statement "there is some x such that $x^2+3x=0$." Again, we can't simply write this as

$$\exists x(x^2+3x=0)$$

since the discussion of its logical truth would have to include any possible model—but we're interested specifically if this is true of the real numbers. A tempting option is to mimic what we did in the universal quantification case and use

$$\exists x(x \in \mathbb{R} \implies x^2+3x=0).$$

Note that we would read this in English as "there is some element x such that if it is a real number, then $x^2+3x=0$" Something like this won't work—statements of the form $\exists x(P(x) \implies Q(x))$ are quite strange. Note that such a claim is true in any model containing any object which does not satisfy P. Call such an object b. Then $P(b) \implies Q(b)$ is vacuously true, since $P(b)$ is not true. Thus $\exists x(P(x) \implies Q(x))$.

Instead, for restricting existential quantifiers, we instead write

$$\exists x(x \in \mathbb{R} \wedge x^2+3x=0).$$

This says that there is some object x which is a real number and satisfies $x^2+3x=0$. That's what we were intending to say.

It can be notationally bulky to carry around these additional terms inside our quantified expressions. Instead, we will commonly use the following notational shorthand: for a set A, we will use

$$(\forall x \in A)P(x) \text{ to mean } \forall x(x \in A \implies P(x))$$
$$(\exists x \in A)P(x) \text{ to mean } \exists x(x \in A \wedge P(x)).$$

3.4 Negating Quantifiers

Similar to propositional logic, there is a De Morgan's law for quantifiers.

Theorem 2. *De Morgan's Laws.* The following logical equivalences are true.

$$\neg\forall x P(x) \iff \exists x \neg P(x)$$
$$\neg\exists x P(x) \iff \forall x \neg P(x).$$

Proof. We are not currently equipped to prove these statements. Instead we will reason that they are true informally. Consider the first logical equivalence $\neg\forall x P(x) \iff \exists x \neg P(x)$.

Consider any model which makes $\neg\forall x P(x)$ true. If it is not the case that every single object satisfies P, there has to be at least one which doesn't. So $\exists x \neg P(x)$. Therefore $\neg\forall x P(x) \implies \exists x \neg P(x)$.

Now consider a model which makes $\exists x \neg P(x)$ true. If there is some object which does not satisfy P, it cannot be that every object satisfies P. So $\neg\forall x P(x)$. Therefore $\exists x \neg P(x) \implies \neg\forall x P(x)$.

Concept Check 15. Provide an informal proof of the second logical equivalence in De Morgan's laws.

Another way to informally justify these is as an extension of De Morgan's laws for the disjunction and conjunction connectives. In particular, for some domain $\mathfrak{D}$, we have that

$$\neg\exists x P(x) \text{ is like } \neg \bigvee_{d\in\mathfrak{D}} P(d) \text{ is like } \bigwedge_{d\in\mathfrak{D}} \neg P(d) \text{ is like } \forall x \neg P(x).$$

Again, such arguments are informal since we may not have infinitely long conjunctions and disjunctions in first-order logic.

3.5 Numerical Quantification

We only added two new quantifiers to our language: one to say that there is at least one object satisfying a predicate, and one to say that the predicate is satisfied by all objects. This is a very coarse understanding of what it means to "quantify" something. What if we want to say that there are at exactly two objects satisfying a predicate? Or no more than three?

The first-order formula for "there are two objects satisfying P" cannot be $\exists x\exists y(P(x) \wedge P(y))$, since it may be that $x = y$. Instead we have

$$\exists x\exists y(x \neq y \wedge P(x) \wedge P(y)).$$

Similarly, we can get "there are no more than two objects satisfying P" as follows:

$$\exists x\exists y\forall z(P(z) \implies z = x \vee z = y).$$

Thus we are able to write "there are exactly two objects satisfying P" as

$$\exists x\exists y\big(x \neq y \wedge P(x) \wedge P(y) \wedge \forall z(P(z) \implies z = x \vee z = y)\big).$$

Concept Check 16 Translate the following two sentences into first-order logic.

- There exists exactly one object.
- There exist exactly three objects.