How to Prove?

Parham Phoulady

Methods of Proof

1 Statements and Predicates

Below, we will fist define and talk about the **mathematical** statements.

1.1 Simple Statements

A **mathematical** statement is a statement that is either true or false but not both simultaneously.

Commands, questions, opinions, etc. are not (mathematical) statements: they are not either false or true.

Subjective statements are not mathematical statements either: they are true or false based on a person's opinion/feeling, the context, etc. Moreover, they basically cannot be properly **measured** or **evaluated**.

For example, P: "the weather is cold outside" is not a mathematical statement because it can be true or false based on a person's opinion. P is a **subjective** statement and therefore, it is not a mathematical statement. However, Q: "the weather outside is 70 degrees Fahrenheit" is a mathematical statement because the weather can be **objectively** measured and the truth value of the statement can be established without considering a specific person's opinion, etc. The statement Q is an objective and mathematical statement. From now, by "statements" we mean "mathematical statements".

1.1.1 Predicates

A predicate is a declarative sentence whose true/false value depends on one or more variables and becomes a statement when specific values are substituted for the variables. A predicate $P(x)$ assigns a value true or false to each x in domain D depending on whether the property holds or not for x.

The true set of $P(x)$, denoted as $\{x \in D | P(x)\}$, is the set of all elements of D that make $P(x)$ true when they are substituted for x.

For example, consider the statement $P(x)$: "x is greater than 10". Based on different subjects (values of x), $P(x)$ is either true or false. The true set is $\{x | x > 10\}$. When you choose a specific value for the variables in your predicate, you will have a proposition. For example, $P(5)$ and $P(50)$ are propositions. The former is false and the latter is true. Technically, $P(x)$ itself is not considered an statement. Because it is not either true or false: based on different values of x it will be either true or false but not before knowing the subject. So, propositions are statements but not predicates themselves.

1.2 Quantified Statements

Quantified statements are those which contain mathematical quantifiers. Quantifiers are used to quantify predicates to claim over which range the predicate is true or false.

1.2.1 Quantifiers

The two most important quantifiers are **universal**, $\forall$ (for all), and **existential** quantifier, $\exists$ (exists).

The general form of the most two important quantified statements are as below:

- Universal statement: $\forall x : P(x)$,
- Existential statement: $\exists x : P(x)$.

In the general form above, $P(x)$ itself can be either a predicate or quantified statement, etc.

We can use De Morgan's laws to negate quantified statements:

- Negating a universal statement: $\sim\forall x : P(x)$ is equivalent to $\exists x : \sim P(x)$,
- Negating an existential statement: $\sim\exists x : P(x)$ is equivalent to $\forall x : \sim P(x)$,

Instead of "$\sim$", "$\neg$" is used commonly as well.

For a more complex example, consider the statement R, $\exists x : P(x) \wedge (\forall y : Q(y))$. Then, $\sim R$ will be $\forall x : \sim(P(x) \wedge (\forall y : Q(y)))$ which is equivalent to $\forall x : \sim P(x) \vee (\exists y : \sim Q(y)))$.

1.2.2 Quantified Conditional Statements

Consider the statement of the form $\forall x \in D :$ if $P(x)$ then $Q(x)$. Then,

- Its **contrapositive** is the statement: $\forall x \in D :$ if $\sim Q(x)$ then $\sim P(x)$,
- Its **converse** is the statement: $\forall x \in D :$ if $Q(x)$ then $P(x)$,
- Its **inverse** is the statement: $\forall x \in D :$ if $\sim P(x)$ then $\sim Q(x)$.

A conditional statement is logically equivalent to its contrapositive.

Now, after knowing these basic definitions, we can talk about arguments and methods of proof. We will review a few most important type of arguments and methods of proof.

2 Arguments

There are several different forms of logical arguments. To be brief, we will discuss just a couple of important forms of valid arguments.

2.1 Valid Arguments

Some of the most simple forms of valid arguments are listed below (taken from the book Discrete Mathematics with Applications by Susanna S. Epp):

Modus Ponens	$p \to q$ p $\therefore q$	**Elimination**	**a.** $p \vee q$ $\sim q$ $\therefore p$ **b.** $p \vee q$ $\sim p$ $\therefore q$
Modus Tollens	$p \to q$ $\sim q$ $\therefore \sim p$	**Transitivity**	$p \to q$ $q \to r$ $\therefore p \to r$
Generalization	**a.** p $\therefore p \vee q$ **b.** q $\therefore p \vee q$	**Proof by Division into Cases**	$p \vee q$ $p \to r$ $q \to r$ $\therefore r$
Specialization	**a.** $p \wedge q$ $\therefore p$ **b.** $p \wedge q$ $\therefore q$		
Conjunction	p q $\therefore p \wedge q$	**Contradiction Rule**	$\sim p \to c$ $\therefore p$

In each of the arguments listed above, the first statement(s) is the **premise** and the last statement (the statement after $\therefore$) is the **conclusion**. If we have n premises, $P_1, P_2, \ldots, P_n$, and argue that conclusion C is true, then an argument has the following logical form:

$$(P_1 \wedge P_2 \wedge \cdots \wedge P_n) \to C, \tag{1}$$

which means that if all premises are true, then the conclusion is true as well.

For example, Modus Tollens can be written as the logical form $(p \to q \wedge \sim q) \to \sim p$. This is a valid argument because a conditional statement and its contrapositive are equivalent. In other words, we know that if $p \to q$ is true, then so is its contrapositive $\sim q \to \sim p$. The other premise says that $\sim q$ is true. Therefore, $\sim p$ is true (which means that p is false).

2.2 Invalid Arguments

With an invalid reasoning or a **fallacy** we will have an invalid argument. It is important to note that while a reasoning can be a fallacy, it still does not necessarily mean that the conclusion is false. In many cases, the conclusion of an invalid argument might be true or false regardless of the fallacy.

Let's look at a couple of arguments:

1) You told me that you would buy pizzas for tonight if you left work early. You bought pizzas but you have not left work early. Liar!

2) If interest rates go up, stock market prices will go down. They didn't go up, so the stock market prices will not go down.

In the first argument, we made an invalid reasoning called "converse error". The argument is wrong because the conditional statement and its converse are not equivalent. Note that if s/he have said that "I would buy pizzas for tonight only if I leave work early", then the reasoning above would not have been invalid.

The second argument is an example of "inverse error". It assumes that the conditional statement and its inverse are equivalent while they are not. However, a conditional statement and its inverse are not equivalent. In this particular case, although the stock market prices may really not go down in near future, this argument is invalid nonetheless.

These were only a couple very simple examples invalid arguments. There are many fallacies and this website lists many of them: logicallyfallacious.com. It is important to know at least some of the most important formal fallacies.

3 Methods of Proof

We will discuss the following five fundamental methods of proof:

1) Direct proof,

2) Proof by counterexample.

3) Proof by contradiction,

4) Proof by contrapositive,

5) Proof by mathematical induction.

These methods are all different forms of **deductive reasoning** (as opposed to **inductive reasoning**).

Note that all the letter we use to denote different things, like P, Q, x, etc., are arbitrary and they may be represented by different letters/symbols in other cases.

3.1 Direct Proof

Generally, the statement, $P(x)$ we want to prove has a generic form of in (1). In direct proof, we prove C by a sequence of statement that are either taken from one of the premises or

are deducted from them, and finally we reach to the conclusion C. If we want to prove a statement for many different values of x, then we should not put any specific condition on x. The only assumption should be that x (n, etc.) satisfies the premise.

For example, let's say we want to prove the following statement

$$P(n) : \forall n \text{ if } n \text{ is even, then } n^2 \text{ is even.}$$

Note that we should prove the conclusion for **any** even value of n. Because the statement did not have any other condition for the value of n. So, we cannot put n equal to some specific even number and prove the conclusion. This might look like a really obvious thing in this simple statement but it is one of the common mistakes that many students do in relatively more complex statements.

Direct proof: we know that n is even. So, according to the definition of "even", $n = 2k$ for some integer k. Now, consider n^2:

$$n^2 = (2k)^2 \Rightarrow n^2 = 4k^2 = 2(2k^2) = 2k'.$$

Based on the assumption k is an integer, therefore, $k' = 2k^2$ is also an integer. Hence, n^2 is even.

In this example, we derived the conclusion by directly starting with the premise. Later, after a few deductions we reached to the conclusion we wanted to show.

Direct proof, structurally, one of the most simple methods of proof. However, compared to other methods of proof, it is relatively more difficult to use the direct proof method to prove some statements. For example, we will see that some statements are much easier to prove using the proof by contrapositive than by direct proof.

3.2 Proof by Counterexample

This method of proof (which is considered, instead of a method of proof, a proof strategy, etc. sometimes) is usually used to prove negative statements (or in other words, to refute the validity of a statement). In this method, we generally need to show that a statement is incorrect for at least 1 value of x. For example the statement

$$P(n) = \sim\forall \text{ positive integers } n \text{ s.t. } 2 \leq n \leq 7, n \text{ is prime.}$$

The statement says that it is incorrect that all integers greater than or equal to 2, and less than or equal to 7 are prime. Therefore, it claims that there is **at least** one value of $n \in 2, 3, 4, 5, 6, 7$ for which n is not prime. There are only two values in the set of possible values for which n is not prime. Therefore, to prove by counterexample, we can say that: let's put $n = 4$. But $n = 4 = 2 \times 2$. Therefore, n is the product of two integers greater than 1 and therefore, it is not prime.

There is another value of n that makes n not-prime. However, we do not need to consider that as well because we already showed that is not true that for all of those values n is prime.

Proof by counterexample is usually used in combination with other methods of proof. Because usually the statements that we want to prove do not have the negative form of the statement above. So, usually, this method is used after we apply other modification to the statement as we will see later.

3.3 Proof by Contradiction

Proof by contradiction is one of the common methods of proof. In this method, generally, we assume that **while** the premises are true, the conclusion is false and show that this leads to a contradiction. We can then deduct that the first assumption (premises were true but the conclusion was false) was false. Which means that the original statement is true.

To make it more clear that how and why this method of proof works, consider proving the following statement

$$p \to q.$$

When is this (conditional) statement false? Whenever p is true but q is false. This is the only time that the conditional statement is false (remember truth tables?). For example, if p and q are both false, it does not invalidate the conditional statement.

As an example, assume I say "if I get A in algorithms class, I will throw a party". The only way that someone can prove that I lied is when "I get A in algorithms class but I do not throw a party" (here, "but" means "and"). On the other hand, if I want to prove that I did **not** lie, I need to show that this did not happen: "I got a good grade in algorithms class and I did not throw a party".

Proof by contradiction works in this way: we assume "p is true **and** q is false" and show that it leads to a contradiction (a contradiction is simply an **always-false** statement). After showing that, we can deduct that the original assumption (p is true and q is false) was incorrect itself, and therefore, the original conditional statement will be proven to be true.

Consider the following statement:

$$P(n) : \forall n \text{ if } n^2 \text{ is even, then } n \text{ is even.}$$

Proving this statement using the direct proof method is not as easy as proof by contradiction. To prove it by contradiction, we assume that for **some** n, n^2 is even but n is odd. Based on the definition of "odd", $n = 2k+1$ for some integer k. Then $n^2 = (2k+1)^2 = 4k^2+2k+1 = 2(2k^2+k)+1$. But because k is an integer, then $k' = 2k^2+k$ is also an integer and therefore, $n^2 = 2k'+1$ is odd. Which is in contradiction with the original assumption that n^2 was even. Hence, our original assumption was false (n^2 is even and n is false) and the statement is proven to be true.

3.4 Proof by Contrapositive

Proof by contrapositive works by proving the contrapositive of a conditional statement instead of the original statement.

As a brief background, remember that the conditional statement,

$$p \to q,$$

and its contrapositive,

$$\sim q \to \sim p,$$

are equivalent. In logic theory, whenever we say two statements are **equivalent** it means that they are both true or both false at the same time. In other words, instead of proving any statement, it is enough to prove any equivalent statement to that original statement.

Now, as an example of proof by contrapositive, consider the following statement

$$P(n) : \forall n, \text{ if } n^2 \text{ is odd, then } n \text{ is odd.}$$

Again, try to prove this statement directly (by direct proof: start with the premise that for some n, n^2 is odd and try to reach to the conclusion that n is odd). You will see that it is not very easy to prove it by direct proof. However, it is very easy to prove is by contradiction or contrapositive. To prove it by contrapositive, instead of the original statement, we prove its contrapositive: if n is not odd, then n^2 is not odd. Because "not odd" means "even", then we need to prove that for any n, if n is even, then n^2 is even. This is very easy to prove (we proved it by direct proof).

3.5 Proof by mathematical induction

If we want to prove a statement for a sequence of values of its input, generally, we can prove it by mathematical induction. The proof by mathematical induction is done in two steps: the basis step, and the inductive step. In the basis step, we prove the statement for an initial value and in induction step we prove that if it is true for any value greater than or equal to the initial value, then it is true for the next value.

For example, suppose we want to prove the following statement:

$$P(n) : \forall n \geq 1, 1 + 2 + \cdots + n = \frac{n(n+1)}{2}.$$

We will prove it by induction.

Basis step: if $n = 1$, we should prove that $1 = \frac{1(1+1)}{2}$. This is clearly true: $\frac{1(1+1)}{2} = \frac{1\times 2}{2} = 1$.

Inductive step: suppose $P(n)$ is true for some $n \geq 1$. In other words, we want to prove that if

$$P(n) : \forall n \geq 1, 1 + 2 + \cdots + n = \frac{n(n+1)}{2}$$

is true, then

$$P(n+1) : \forall n \geq 1, 1 + 2 + \cdots + (n+1) = \frac{(n+1)(n+2)}{2}.$$

is true. Here, $P(n)$ is called the induction hypothesis. Note that in prove by mathematical induction, generally, you should try to somehow break down the conclusion in the statement

you want to prove, $P(n+1)$, in different parts such that one or more of the parts become similar to the induction hypothesis, $P(n)$. After that, you apply the hypothesis to that part. After that, we combine the remaining parts to reach to the conclusion.

To see it more clearly, look at the proof below. We start by the left side of the equation we want to prove and then we separate a part of it that is the same as the left part of our hypothesis:

$$\begin{aligned}
1+2+\cdots+(n+1) &= [1+2+\cdots+n]+(n+1) \\
&= \frac{n(n+1)}{2}+(n+1) && \text{because of induction hypothesis} \\
&= \frac{n(n+1)+2(n+1)}{2} \\
&= \frac{(n+1)(n+2)}{2}, && \text{after factoring the common factor, } (n+1)
\end{aligned}$$

which is what we wanted to prove.

The version of the induction explained above is sometimes called the **weak** induction. We also have **strong** induction. The different between these two kinds of induction is between their induction hypotheses. In strong induction we assume that the statement is true for all $k \leq n$ (and greater than or equal to the initial value) and then we prove it is true for $k+1$. For example, consider the following statement:

$$P(n) : \forall n \geq 2, n \text{ can be factored into primes.}$$

Basis step: $n = 2$: because 2 is prime itself, the prime factorization of 2 is itself, 2.

Inductive step: suppose $P(k)$ is true for all k such that $2 \leq k \leq n$ (induction hypothesis). We prove that $P(k+1)$ is true. We consider two cases for the integer $k+1$:

Case 1: $k+1$ is prime. Then $k+1$ is a prime factorization of itself and we are done.

Case 2: $k+1$ is not prime. Because $k+1$ is composite, it can be written as the product of two positive: $k+1 = p \times q$ for two integers $2 \leq p, q \leq k$. Now, consider $P(p)$ and $P(q)$. Based on induction hypothesis both p and q can be factored into primes. By multiplying their prime factors we can get a prime factorization of $k+1$.

Mathematical induction can be applied in many different forms. For example, sometime we may need to prove a statement for some initial value (basis step, for example, for $n = 0$) and then prove that if $P(n)$ is true, then $P(n-1)$ is true (inductive step). In that case, we would prove the statement for all $n \leq 0$.

Also, sometimes we need to take larger steps in the inductive step and go from $P(n)$ to $P(n+2)$. In that case, we need to prove the statement for two sequential initial values (why?).

A combination of the last couple of different versions explained above (and other similar versions) are used for relatively more complex statements. We skip that for simplicity.