Proof tactics

Kyle Miller

Jan 20, 2017

In this note, I describe propositions and the rules ("tactics") you may use to prove them. The hope is that by understanding what constitutes a valid proof, you will feel more comfortable searching for one!

1 Creating propositions

The book defines a *proposition* as a declarative statement which is either true or false (but not both). One way to view mathematics is as an activity where we try to discover true propositions — though whether a proposition is "*interesting*" is very important to a mathematician and not captured by logic alone.

Telling whether a statement is a proposition can be tricky, but there are some rules you can follow to construct valid propositions. In particular, the rules avoid statements like "This statement is false."

First, a technicality: there might be *free variables* in the rules. A variable stands for some kind of object, and a free variable is a variable which is not yet defined. One way to *close* a free variable is to *bind* it with a universal or existential quantifier. Another way is to define the variable in the current context, which we will see in the section on inference rules. For a statement to have a definitive truth value, it should not have any free variables. For instance, if n is undefined then "n is even" is not a proposition, but "for all n, n is even" is a (false) proposition. Similarly, if n is an object and $n = 2$ is in the context, then "n is even" is a (true) proposition. We will write $p(x)$ to mean a proposition where with a free variable x (for convenience, we allow $p(x)$ to be something like "1=2" — after all, the constant function $f(x) = 2$ is a function of x).

- Basic propositions, like "x is cute", "x is the parent of y", "$1 = 2$", or "$x \in X$". These must be simple expressions of a property or relation in the world.
- Given a proposition p, the negation $\neg p$, also known as "not p."
- Given two propositions p and q, the conjunction $p \wedge q$ and the disjunction $p \vee q$. These are respectively also known as "p and q" and "p or q." The disjunction is the so-called *inclusive or*.
- Given two propositions p and q, the implication $p \to q$, which can be said in many ways, like "if p then q," "q if p," "p only if q," and so on.[1]
- Given a proposition p with x (possibly) a free variable, the universal quantifier $\forall x, p(x)$, said "for all x in the universe, $p(x)$." There is a cousin $\forall x \in X, p(x)$, but this is shorthand for $\forall x, x \in X \to p(x)$. The "possibly" is to allow for things like $\forall x, 1 + 1 = 2$.
- Given a proposition p with x (possibly) a free variable, the existential quantifier $\exists x, p(x)$, said "there exists an x such that $p(x)$. Similarly, there is a cousin $\exists x \in X, p(x)$, but this is shorthand for $\exists x, x \in X \wedge p(x)$.[2]

See Section 4 for a pseudo-implementation of these propositions.

[1] We take "if p then q" to mean that p is false or q is true (or both), which is known as the *material implication*. In fact, $p \to q$ is unnecessary notation and could be replaced by $\neg p \vee q$. We keep the arrow because it's more intuitive. From the Stanford Encyclopedia of Philosophy: The truth-functional theory of the conditional was integral to Frege's new logic (1879). It was taken up enthusiastically by Russell (who called it "material implication"), Wittgenstein in the Tractatus, and the logical positivists, and it is now found in every logic text. It is the first theory of conditionals which students encounter. Typically, it does not strike students as obviously correct. It is logic's first surprise. Yet, as the textbooks testify, it does a creditable job in many circumstances. And it has many defenders. It is a strikingly simple theory: "If A, B" is false when A is true and B is false. In all other cases, "If A, B" is true.

[2] The existential quantifier is also unnecessary, because $\exists x, p(x)$ is equivalent to $\neg\forall x, \neg p(x)$.

2 Rules of inference

Any given proposition is either true or false (by definition), but determining which of the two it is can be quite a challenge. Really the only tool we have at our disposal to tell one way or another is inference. A set of inferences which supports a proposition is called a proof. In this section, we will give a list of the main inference rules we use in proofs.

It should be said that inference is inherently limited. In addition to propositions being either true or false, a proposition is either provable or not provable, where a provable proposition is one for which there exists a proof. One might hope truth and provability were equivalent, but, while every provable proposition is a true proposition, unfortunately the opposite isn't the case! In Gödel's incompleteness theorem, he constructs a proposition G such that G is true if and only if G is not provable. If G were false, then G would be provable, and thus also true! So, assuming math is sound, G must be an unprovable true proposition.

A more optimistic result, on the other hand, is Gödel's completeness theorem, which is that any proposition which is true in all possible universes is provable. A consequence of this is that whenever you show that two truth tables are the same, there *must* be a corresponding proof that the corresponding propositions are equivalent!

Anyway, let us begin the examining the rules of inference. We will give them formally, but in a math textbook or elsewhere they usually appear much more informally, out of consideration of making the proof flow (and at risk of introducing errors). You might consider taking proofs from a textbook and rewriting them more formally to understand their structure better. When you write proofs of your own, you might use a list like this one for inspiration. And, do not take these as prescriptions: discovering a proof for a proposition is not always straightforward and involves no small amount of searching.

When evaluating a proof, there are three pieces of context that the inference rules manipulate. The first piece of context is a background set of *assumptions*, sometimes known as *hypotheses*. At the beginning, the set of assumptions starts with the *axioms*, the propositions we agree to hold self-evidently true, like that 0 is a number, or that adding 1 to a number gets a new, bigger number. The set of assumptions also starts with all previously proved propositions (called variously *theorems*, *lemmas*, *corollaries*, or *propositions*). The second piece of context is a set of defined variables. All that "defined" means is that you can speak of it, so if $n \in \mathbb{N}$ is a variable, do not get the impression that you actually know what n is.[3] The third piece of context is the set of goals. When proving a proposition, the initial goal is that proposition. A proof is complete when the set of goals becomes empty.

You might hear people say to work on a proof both forwards and backwards. Proving forwards means to take assumptions and deduce new assumptions from them, and proving backwards means to replace goals with goals which imply them. At some point, these might meet in the middle, and you can pull out a proof.

Assumption. If p is both an assumption and a goal, then we may remove p as a goal. In a proof:

> We assumed p.

Modus ponens. If p and $p \to q$ are both assumptions, we may also assume q. In a proof:

> Since p and $p \to q$, q.

Deduction. To prove $p \to q$, we may prove q while temporarily assumping p. In a proof:

> Assume p. (or "Suppose p.")
>
> > q is a goal
>
> Thus, $p \to q$.

What this means is that we assume p for the duration of the indented region, and we may not leave the indented region until we remove q as a goal.

This is also known as a *conditional proof.*

Instantiation. To prove $\forall x, p(x)$, we introduce an "arbitrary" x into the context. In a proof:

> Let x be arbitrary.

[3] Unless, say, there is an assumption which says something like $n = 22$. Then you know what n is in that context.

$p(x)$ is a goal

Thus, $\forall x, p(x)$.

What this means is that we may refer to x in propositions in the indented region, and we may not leave the indented region until we remove the $p(x)$ goal. Do not misunderstand what "let x be arbitrary means": you must imagine your adversary is trying to find some worst-case x which exposes a flaw in your proof, so it represents the idea of "what if someone gave us an x." You do not know what x is unless some assumption lets you deduce something.

A common compound is $\forall x, p(x) \to q(x)$, which might be said as "for all x such that $p(x)$, $q(x)$." In a proof:

Let x be such that $p(x)$.

$q(x)$ is a goal

Thus, for all x such that $p(x)$, $q(x)$.

The first line simultaneously introduces an object x and an assumption $p(x)$ for the the indented region.

Invocation. If $\forall x, p(x)$ is an assumption and a is an object, then $p(a)$ is a new assumption.

Since $\forall x, p(x)$, and a is a thing, $p(a)$.

Or the compound:

Since $\forall x, p(x) \to q(x)$ and $p(a)$ is true, $q(a)$ is true.

Deduction is to modus ponens as instantiation is to invocation.

Everything else is just an elaboration of these five inference rules of Assumption, Modus ponens, Deduction, Instantiation, and Invocation. Strictly speaking, only three of them are needed, but the goal isn't to create a minimal system of formal reasoning, it's to understand all of the ways to prove things and roughly where they come from!

What follows are some examples of the forms in which they are commonly used. By tactics, I mean any technique which reduces to a series of rules of inference. Of course, a rule of inference is itself a tactic.

3 Tactics

Negation. The proposition $\neg p$ is the same as $p \to F$, so we may use a conditional proof:

Assume p.

F is a goal (i.e., the goal is to prove a contradiction)

Thus, $\neg p$.

Strengthening. If $p \to q$ and q is a goal, we may replace the goal.

Since $p \to q$, replace goal q with goal p.

This is equivalent to

p is goal

Since $p \to q$, q.

Thus q.

Similarly, $p \leftrightarrow q$ can be used to rewrite a goal q.

Contradiction. This is related to negation. To prove p, we can try to prove $\neg\neg p$.

Assume $\neg p$.

F is a goal

Thus, p.

This relies on propositions being either true or false: if p being false implies a contradiction, it must have been true.

It is good style to avoid proofs by contradiction or to limit their scope to as small a part of your proof as you are able to make it. Since a proof by contradiction involves counterfactual reasoning, it is difficult to be certain that the contradiction wasn't just a mistake!

Another form of proof by contradiction is

> Assume $\neg p$.
>
> > p is a goal
>
> Thus, p.

The first line of a contradiction proof is usually "assume for sake of contradiction that $\neg p$," to let the reader know what to expect. Then the contradiction is announced with "which is a contradiction, thus p."

Conjunction. To prove $p \land q$,

> p is a goal
>
> q is a goal
>
> Thus, $p \land q$.

In other words, once both p and q are valid assumptions, $p \land q$ is one, too.

Contrapositive. Rather than proving $p \to q$ directly, it might be easier to prove its contrapositive $\neg q \to \neg p$.

> Assume $\neg q$.
>
> > $\neg p$ is a goal
>
> Thus, $p \to q$.

Note that this is not a proof by contradiction! There are fewer assumptions in the context, so less room for error.

Implication by contradiction. The statement $\neg(p \to q)$ is equivalent to $p \land \neg q$, so $p \to q$ can be proven like this:

> Assume p and $\neg q$.
>
> > F is a goal
>
> Thus, $p \to q$.

The first line tends to be written as "assume for sake of contradiction that p but $\neg q$."

If you find that you do not use p as an assumption, consider rewriting the proof as a proof of the contrapositive instead.

Lucky disjunction. To prove $p \lor q$, you might be lucky and have p as an assumption:

> Since p, $p \lor q$.

Disjunction by implication. The statement $p \lor q$ is equivalent to $\neg p \to q$.

> Assume $\neg p$.
>
> > q is a goal
>
> Thus, $p \lor q$.

Disjunction by contradiction. The statement $\neg(p \lor q)$ is $\neg p \land \neg q$.

> Assume (for sake of contradiction) $\neg p$ and $\neg q$.
>
> > F is a goal
>
> Thus, $p \lor q$.

Again, if you did not play $\neg p$ and $\neg q$ off each other, you might consider a non-contradiction proof.

Conjuction by contradiction. The statement $\neg(p \wedge q)$ is $p \to \neg q$, so we may attempt:

Assume (for sake of contradiction) $p \to \neg q$.

F is a goal

Thus, $p \wedge q$.

Biconditional. To prove $p \leftrightarrow q$, there are a few options. The most elegant is to prove a sequence of biconditionals $p \leftrightarrow r_1$, $r_1 \leftrightarrow r_2$, and so on until $r_n \leftrightarrow q$. But proving both $p \to q$ and $q \to p$ is fine. The question then is whether to prove them directly, by the contrapositive, or by contradiction. For instance, here is one of those nine options:

Assume p

q is a goal

Thus, $p \to q$.

Assume $\neg p$

$\neg q$ is a goal

Thus, $q \to p$.

Therefore, $p \leftrightarrow q$.

Construction. To prove $\exists x, p(x)$:

Construct some object a (using other existential propositions)

$p(a)$ is a goal

Thus, $\exists x, p(x)$.

Existential by contradiction. The statement $\neg\exists x, p(x)$ is $\forall x, \neg p(x)$.

Assume (for sake of contradiction) $\forall x, \neg p(x)$.

F is a goal

Thus, $\exists x, p(x)$.

("If it's not true that $p(x)$ is false for all x, then there must be some x where $p(x)$ is true.")

Existential invocation. If $\exists x, p(x)$ is an assumption, we may introduce an object a and the assumption $p(a)$.

Since $\exists x, p(x)$, let a be such that $p(a)$.

This use of "let" is not to be confused with the "let" in Instantiation of a universal quantifier. Usually the latter uses the word "arbitrary" to differentiate them.

Universal quantifier by contradiction. Since $\neg\forall x, p(x)$ is $\exists x, \neg p(x)$, we may do something like

For sake of contradiction, let a be such that $\neg p(a)$.

F is goal

Thus, $\forall x, p(x)$.

This says, if there were a counterexample, then we would reach a contradiction, so there must not be a counterexample.

A similar case is $\forall x, p(x) \to q(x)$, whose negation is $\exists x, p(x) \wedge \neg q(x)$ (check this!), so a proof might be

For sake of contradiction, let a be such that $p(a)$ and $\neg q(a)$.

F is a goal

Thus, $\forall x, p(x) \to q(x)$.

For example, to prove "For all $x \in \mathbb{R}$ such that $x > 1$, $x^2 > x$" we might do

For sake of contradiction, let $x \in \mathbb{R}$ be such that $x > 1$ and $x^2 \leq x$.

Since $x > 1 > 0$, we can divide both sides of $x^2 \le x$ by x to get $x \le 1$.

But $x > 1$ and $x \le 1$ together are a contradiction!

Thus for all $x \in \mathbb{R}$ such that $x > 1$, $x^2 > x$.

(Note: contradiction was not necessary here.)

Destructuring a conjunction. If $p \wedge q$ is an assumption, then both p and q follow.

Since $p \wedge q$, p.

Cases. If $p \vee q$ is an assumption and r is a goal, we may proceed by cases.

Case I. Assume p.

r is a goal

Case II. Assume q.

r is a goal

Thus, $(p \vee q) \to r$.

This shows up often for a natural number. It is true, for instance, that $n = 0 \vee n = 1 \vee n \ge 2$, so we may proceed by three cases to prove something about the number.

Contradiction within cases. Sometimes one of the cases cannot ever happen, and to deal with this you may try to prove a contradiction to deal with that case. For instance, what if we wish to prove something about a pair of numbers n and m which are already present in the context, but you are able to prove $n \neq m$.

Case I. Assume $n < m$.

p is goal

Case II. Assume $n = m$.

F is the goal

Case III. Assume $n > m$.

p is goal

Thus, p, since $n < m \vee n = m \vee n > m \to p$ and $n < m \vee n = m \vee n > m$ is a tautology.

Induction. This is a structural inference rule which comes from the axioms about natural numbers, namely that 0 is a number and that every number n has a unique successor $n + 1$. Suppose we want to prove $\forall n \in \mathbb{N}, p(n)$.

Base case. $p(0)$ is a goal.

Inductive case. Let $n \ge 0$ and assume $p(n)$.

$p(n + 1)$ is a goal.

Since $p(0)$ and $\forall n \in \mathbb{N}, p(n) \to p(n + 1)$, then $\forall n \in \mathbb{N}, p(n)$ by induction.

The intuition is that if $p(0)$ and $p(n) \to p(n + 1)$, then $p(0) \to p(1)$, $p(1) \to p(2)$, $p(2) \to p(3)$, and so on. You can show $p(n)$ is true for any number you can think of, so we take it to be true for all numbers.

One application is a different proof of Euclid's theorem, that there are infinitely many prime numbers. Let $p(n)$ be the statement that there are at least n distinct prime numbers.

Base case I. $n = 0$. There are evidently at least 0 distinct prime numbers.

Base case II. $n = 1$. Since 2 is prime, there is at least 1 prime number.

Inductive case. Assume there are at least n distinct prime numbers.

Let $p_1, \dots, p_n$ be some distinct prime numbers, which are presumed to exist.

Let $m = p_1 p_2 \cdots p_n + 1$.

This number m when divided by any p_i has remainder 1.

Thus, the prime factorization of m must contain a prime p_{n+1} distinct from $p_1, \ldots, p_n$.

Thus, there are at least $n+1$ distinct prime numbers.

By induction, for any n there are at least n distinct prime numbers.

By definition, there are infinitely many distinct prime numbers.

Well-ordering. This can sometimes make nicer proofs than by induction. To prove $\forall n \in \mathbb{N}, p(n)$, we observe that if the set $\{n \in \mathbb{N} : \neg p(n)\}$ is nonempty, there is a smallest number in the set, and then we make use of the minimality to arrive at a contradiction.

Assume $n \in \mathbb{N}$ is the smallest such that $\neg p(n)$.

F is a goal

Thus, $\forall n \in \mathbb{N}, p(n)$.

Alternatively, we may use the "proof by infinite descent" form, where the contradiction we show is that there is yet a smaller number in the set with the same property.

Assume $n \in \mathbb{N}$ is the smallest such that $\neg p(n)$.

$\exists m \in \mathbb{N}, m < n \land p(m)$ is goal

Thus, $\forall n \in \mathbb{N}, p(n)$.

Example: every number has a prime factorization.

Assume n is the smallest number without a prime factorization.

Case I. Assume n is prime.

Then $n = n$ is a prime factorization, contradicting the fact n has no prime factorization.

Case II. Assume n is not prime.

Then $n = ab$ for some $a, b > 2$.

Both $a < n$ and $b < n$, so by assumption they have prime factorizations.

Then n has a prime factorization obtained by juxtaposing these prime factorizations.

Thus, $\neg(n \text{ prime} \lor (n \text{ not prime}))$, a contradiction.

Thus, every $n \in \mathbb{N}$ has a prime factorization.

Example: $\sqrt{2}$ is irrational.

Assume for sake of contradiction that $\sqrt{2}$ is rational.

Let $p, q \in \mathbb{N}$ be such that $\sqrt{2} = \frac{p}{q}$ and q is the smallest such denominator.

Then $p^2 = 2q^2$.

By divisibility, p is divisible by 2, so $p = 2m$ for some $m \in \mathbb{N}$.

Then $(2m)^2 = 2q^2$, which is the same as $2m^2 = q^2$.

By divisibility, q is divisible by 2, so $q = 2n$ for some $n \in \mathbb{N}$.

Then $2m^2 = (2n)^2$, which is the same as $m^2 = 2n^2$.

But this means $\sqrt{2} = \frac{m}{n}$ and $n < q$, contradicting q being smallest.

Thus $\sqrt{2} \neq \frac{p}{q}$ for all $p, q \in \mathbb{N}$.

Therefore, $\sqrt{2}$ is irrational.

Zorn's lemma. This is an advanced method you can safely ignore for this course. This is an inference rule which relies on the so called "Axiom of choice," which standard mathematics assumes. Suppose X is a set with a partial order, which we will write as $\subset$ (so imagine X is a set of sets; also we take the convention that $U \subset U$ for any set U). Under a certain condition, we can show X has a *maximal* element, an element $M \in X$ such that for all $V \in X$, $M \subset V \to V = M$. This is not necessarily a *maximum* element, which is an element $M \in X$ such that for all $V \in X$, $V \subset M$. In the following, a set X is *totally ordered* if for all $U, V \in X$, $U \subset V \vee V \subset U$, and $U \subset V \wedge V \subset U \to U = V$.

Let $Y \subset X$ be an arbitrary totally ordered subset.

$\exists U \in X, \forall V \in Y, V \subset U$ is a goal (i.e., there is an upper bound for Y in X)

Thus, by Zorn's lemma, X has a maximal element.

For example, let us prove that every vector space V has a basis.

Let X be the set of all linearly independent subsets of V.

Let Y be an arbitrary totally ordered subset of X.

Let $J = \bigcup_{I \in Y} I$ be the union of the linearly independent subsets in Y.

Let $v_i \in J$ and $c_i \in \mathbb{R}$ be arbitrary such that $c_1 v_1 + \cdots + c_n v_n = 0$.

Each v_i is in one of the linearly independent subsets of Y. Call it $I_i \in Y$.

Since Y is totally ordered, there a j such that $I_i \subset I_j$ for all i.

Thus, $v_i \in I_j$ for all i.

Since I_j is an independent set, $c_1 = \cdots = c_n = 0$

Thus, J is an independent set. That is, $J \in X$.

Since $I \subset J$ for all $I \in Y$, J is an upper bound for I in X.

Thus, by Zorn's lemma, there is a maximal independent set $M \in X$.

For sake of contradiction, suppose M does not span V.

Let $v \in V$ be an element which is not in the span of M.

The set $M' = M \cup \{v\}$ is an independent set by the basis extension theorem.

Since $M \subsetneq M'$, this contradicts M being maximal.

Therefore, M spans V.

Thus M is a basis of V, since a basis is a linearly independent spanning set.

4 The grand interpreter

To give meaning to the symbols (that is, to give semantics to the syntax), we can imagine a computer which can perform infinitely many operations in finite time to evaluate the truth of a proposition. We will translate the logical symbols given in the rules for producing valid propositions into pseudo-Python. In a way, this explains why it is that those rules produce statements which are actually propositions, since each of these pseudo-Python statements will give a defined truth value.

For $\neg p$,

```
return not p()
```

For $p \wedge q$,

```
return p() and q()
```

For $p \vee q$,

```
return p() or q()
```

For $p \to q$,

```
if p():
  return q()  # Then q ought to be true for the conditional to be true
else:
  return True # Then the conditional is considered to be true (since it's not false)
```

For $\forall x, p(x)$,

```
for x in universe:
  if not p(x):
    return False
return True
```

For $\exists x, p(x)$,

```
for x in universe:
  if p(x):
    return True
return False
```

The statement “This statement is false” can be thought of as

```
def p():
  return not p()
```

which is an infinite recursion. This is disallowed by the proposition construction rules because there is nothing which introduces the ability to refer to statements themselves, much less to refer to the constructed statement itself.

This is not exactly obvious that by not being able to refer to a statement you cannot cause a contradiction. For instance, the (untyped) lambda calculus has issues like the following:

```
(lambda x: x(x))(lambda x: x(x))
```

There is no self-reference, yet somehow it evaluates to itself *ad infinitum*.