What is a Proof?

1 What is a Proof?

A proof is a method of ascertaining truth. There are many ways to do this:

jury trial Truth is ascertained by twelve people selected at random.

word of God Truth is ascertained by communication with God, perhaps via a third party.

word of Boss Truth is ascertained from someone with whom it is unwise to disagree.

experimental science The truth is guessed and the hypothesis is confirmed or refuted by experiments.

sampling The truth is obtained by statistical analysis of many bits of evidence. For example, public opinion is obtained by polling only a representative sample.

inner conviction "*My* program is perfect. I know this to be true."

"I don't see why not..." Claim something is true and then shift the burden of proof to anyone who disagrees with you.

Mathematics also has a specific notion of "proof" or way of ascertaining truth.

Definition A *proof* is a verification of a *proposition* by a chain of *logical deductions* from a base set of *axioms*.

The three key ideas in this definition are highlighted: proposition, logical deduction, and axiom. Each of these terms is discussed in a section below. The last section contains some examples of complete proofs.

2 Propositions

Definition A *proposition* is a statement that is either true or false.

This definition sounds very general, but it does exclude sentences such as, "Where art thou Romeo?" and "Give me an A!".

Proposition 2.1 $2 + 3 = 5$

This proposition is true.

Proposition 2.2 $\forall n \in N, n^2 + n + 41$ is a prime number.

The symbol $\forall$ is read "for all" and will be discussed further in tutorial. The symbol $\mathcal{N}$ stands for the set of natural numbers: 0, 1, 2, 3, ... (See your TA for a complete list.) A *prime* is a natural number greater than one that is not divisible by any other natural number other than 1 and itself: 2, 3, 5, 7, 11, ...

Certainly the preceding proposition is true for many natural numbers n:

$$\begin{array}{cc}
n = 0 \rightarrow 41 & (prime) \\
n = 1 \rightarrow 43 & (prime) \\
n = 2 \rightarrow 47 & (prime) \\
n = 3 \rightarrow 53 & (prime) \\
\ldots & (more\ primes) \\
n = 20 \rightarrow 461 & (prime) \\
\ldots & (still\ more\ primes) \\
n = 39 \rightarrow 1601 & (prime)
\end{array}$$

But if $n = 40$, then $n^2 + n + 41 = 40^2 + 40 + 41 = 41 \cdot 41$, which is not prime. Since the expression is not prime *for all* n, the proposition is false!

Proposition 2.3 $a^4 + b^4 + c^4 = d^4$ has no solution when $a, b, c, d \in N$.

Euler (pronounced "oiler") conjectured this 1769. But the proposition was proven false 218 years later by Noam Elkies at the liberal arts school up Mass Ave. He found the solution $a = 95800, b = 217519, c = 414560, d = 422481$.

Proposition 2.4 $313(x^3 + y^3) = z^3$ has no solution when $x, y, z \in N$.

This proposition is also false, but the smallest counterexample has more than 1000 digits!

Proposition 2.5 Every map can be colored with 4 colors so that adjacent regions have different colors.

This proposition is true and is known as the "four-color theorem". However, there have been many incorrect proofs, including one that stood for 10 years in the late 19th century before the mistake was found. An extremely complicated proof was finally found recently with the help of a computer. There is still debate whether a proof is valid if no human can work through the whole argument!

Proposition 2.6 The original Pentium chip divided properly.

Intel's "proofs" by authority and by sampling turned out to be invalid. The proposition is false.

Proposition 2.7 Every even integer greater than 2 is the sum of two primes.

No one knows whether this proposition is true or false. This is the Goldbach Conjecture, which dates back to 1742. See the Globe handout and note the silly error.

3 Axioms

Definition An *axiom* is a proposition that is assumed to be true.

There is no proof that an axiom is true; you just assume it is true because you believe it is reasonable. Here are some examples:

Axiom 1 If $a = b$ and $b = c$, then $a = c$.

This seems very reasonable! But sometimes the right choice of axiom is not clear.

Axiom 2 (Euclidean geometry) Given a line l and a point p not on l, there is exactly one line through p parallel to l.

Axiom 3 (Spherical geometry) Given a line l and a point p not on l, there is *no* line through p parallel to l.

Axiom 4 (Hyperbolic geometry) Given a line l and a point p not on l, there are *infinitely many* lines through p parallel to l.

No one of the three preceding axioms is better than the others; all yield equally good proofs. Of course, a different choice of axioms makes different propositions true. Still, a set of axioms should not be chosen arbitrarily. In particular, there are two basic properties that one would want in any set of axioms; it should be consistent and complete.

Definition A set of axioms is *consistent* if no proposition can be proven to be both true and false.

This is an absolute must. One would not want to spend years proving a proposition true only to have it proven false the next day! Proofs would become meaningless if axioms were inconsistent.

Definition A set of axioms is *complete* if it can be used to prove or disprove every proposition.

Completeness is very desirable; we would like to believe that any proposition could be proven or disproven with sufficient work and insight.

Surprisingly, making a complete, consistent set of axioms is not easy. Bertrand Russell and Alfred Whitehead tried during their entire careers to find such axioms for basic arithmetic and failed. Then Kurt Gödel proved that no set of axioms can be both consistent and complete! This means that any set of consistent axioms (an absolute must) can not be complete; there will be true statements that can not be proven. For example, it might be that Goldbach's conjecture is true, but there is no proof!

In this class, we will not worry about the precise set of axioms underpinning our proofs. The requirements are only that you be upfront about what you are assuming, that the background knowledge of math that you assume is self-consistent, and that you do not try to avoid homework and exam problems by declaring everything an axiom!

4 Logical Deductions

Logical deductions or *inference rules* are used to combine axioms and true propositions to construct more true propositions.

A fundamental inference rule is *modus ponens*. This rule says that if p is true and $p \Rightarrow q$ is true, then q is true. The expression $p \Rightarrow q$ is read "p implies q" or "if p, then q" and is also sometimes written simply $\rightarrow$. A truth table for $\Rightarrow$ is shown below:

p	q	$p \rightarrow q$
T	T	T
T	F	F
F	T	T
F	F	T

In this table, T stands for "true" and F stands for "false". Note that if p is false, then $p \Rightarrow q$ is always true. For example, any statement beginning, "If pigs fly, then ..." is true!

Inference rules are sometimes written in a funny notation. For example, modus ponens is written:

$$\frac{\begin{array}{c} p \\ p \Rightarrow q \end{array}}{q}$$

This says that if the statements above the line are true, then we can infer that the statements below the line are also true. There are many other natural inference rules, for example:

$$\frac{\begin{array}{c} p \Rightarrow q \\ q \Rightarrow r \end{array}}{p \Rightarrow r}$$

$$\frac{\begin{array}{c} p \Rightarrow q \\ \neg q \end{array}}{\neg p}$$

As with axioms, we will not be too formal about the set of legal inference rules. Each step in a proof should be clear and "logical"; in particular, you should state what previously proved facts are used to derive each new conclusion.

5 Examples of Proofs

The first two examples are simple; the last two are more tricky.

5.1 Direct Proofs

Theorem 5.1 (DeMorgan's Law) For any sets A and B, $\overline{A \cup B} = \overline{A} \cap \overline{B}$.

The usual way to prove that two sets are equal is to show that each is a subset of the other. This means that the proof will have two parts.

Method To prove $R = S$, show $R \subseteq S$ and $S \subseteq R$.

Proof. First, we will show $\overline{A \cup B} \subseteq \overline{A} \cap \overline{B}$. We do this by showing that any element of the left side set is also in the right side set.

Let x be any element of $\overline{A \cup B}$. This implies that x is not an element of the complementary set, $A \cup B$. Therefore, x is contained in neither A nor B. Since x is not contained in A, it must be contained in the complementary set, $\overline{A}$. Similarly, since x is not contained in B, it must be contained in $\overline{B}$. We now know $x \in \overline{A}$ and $x \in \overline{B}$. These facts imply that x is contained in the intersection $\overline{A} \cap \overline{B}$. We have shown that any element of the left side set is also in the right side set. This proves that $\overline{A \cup B} \subseteq \overline{A} \cap \overline{B}$.

Second, we will show $\overline{A} \cap \overline{B} \subseteq \overline{A \cup B}$. We will use the same style of argument for the second part of the proof; in fact, we will argue in reverse.

Let x be any element of $\overline{A} \cap \overline{B}$. This implies that x is an element of $\overline{A}$ and an element of $\overline{B}$. Since x is an element of $\overline{A}$, it can not be an element of the complementary set A. Similarly, since x is an element of $\overline{B}$, it can not be an element of B. We now know that $x \notin A$ and $x \notin B$. These facts imply that x is not contained in the union $A \cup B$. This implies that x is contained in the complementary set $\overline{A \cup B}$. Again, we have shown that any element of the left side set is contained in the right side set. This proves that $\overline{A} \cap \overline{B} \subseteq \overline{A \cup B}$.

Since each of the two sets is contained in the other, the two sets must be equal. ■

This is tedious! But this is the correct way to prove statements of this form.

Theorem 5.2 For any Boolean variables x and y,

$$(x \Rightarrow y) \Leftrightarrow (\neg y \Rightarrow \neg x)$$

The expression on the right is called the "contrapositive" of $x \Rightarrow y$. A Boolean variable is one that is true/false or, equivalently, 0/1. The symbol $\Leftrightarrow$ is read "if an only if" and is sometimes also written "iff". A statement like $p \Leftrightarrow q$ holds provided p is true exactly when q is true. A common mistake when writing proofs with iff statements is to forget the only if part and to only go one way, for example, to show only $p \Rightarrow q$.

Method (truth table method) To prove a statement about a small number of Boolean variables, make a truth table and check all the possible cases.

Proof.

x	y	$x \rightarrow y$	$\neg y \Rightarrow \neg x$
T	T	T	T
T	F	F	F
F	T	T	T
F	F	T	T

By inspection of the four cases, $x \Rightarrow y$ is true exactly when $\neg y \Rightarrow \neg x$ is true. ■

5.2 Proofs by Contradiction

The two preceding examples are *direct proofs*. In a direct proof, axioms and previously proved theorems are combined until a new proof is done. Sometimes an *indirect proof* is easier. This is also known as a *proof by contradiction*. The idea is to assume the opposite of what we are trying to prove and try to derive a falsehood or contradiction. Then we can conclude that the assumption was false. More formally, we have the following method.

Method (Proof by Contradiction) To prove that p is true, we prove that $\neg p \Rightarrow False$.

We can verify the validity of this technique by the truth table method:

p	$\neg p$	$\neg p \Rightarrow False$
T	F	T
F	T	F

The table shows that p is true iff $\neg p \Rightarrow False$.

Theorem 5.3 $\sqrt(2)$ is an irrational number.

Remember that an irrational number is a number that can not be expressed as a ratio of two integers.

Always note when your proof is by contradiction and make clear what statement you are assuming for the purpose of obtaining a contradiction.

Proof. The proof is by contradiction. Assume for purpose of contradiction that $\sqrt{2}$ is rational. Then we can write $\sqrt(2) = a/b$ where a and b are integers and the fraction is in lowest terms. Squaring both sides gives $2 = a^2/b^2$ and so $2b^2 = a^2$. This implies that a is even; that is, a is a multiple of 2. As a result, a^2 is a multiple of 4. Because of the equality $2b^2 = a^2$, $2b^2$ must also be a multiple of 4. This implies that b^2 is even and so b must be even. But since a and b are both even, the fraction a/b is not in lowest terms. This is a contradiction, and so the assumption that $\sqrt{2}$ is rational must be false. ■

Remember that the intermediate statements in a proof by contradiction may very well be false. This is because they were derived from a proposition that was assumed to be true only for the purposes of contradiction. A common mistake is to forget this and to regard these statements as true!

5.3 A False Proof

The next theorem is incorrect. There is an erroneous step in the proof. Can you find it?

False Theorem 5.4 1 = -1

Proof.

$$\begin{aligned} 1 &= \sqrt{1} \\ &= \sqrt{(-1)(-1)} \\ &= \sqrt{-1}\sqrt{-1} \\ &= \sqrt{-1}^2 \\ &= -1 \end{aligned}$$

■

The step $\sqrt{(-1)(-1)} = \sqrt{-1}\sqrt{-1}$ is invalid. Generally, $\sqrt{uv} = \sqrt{u}\sqrt{v}$ does not hold over complex numbers.

The lecture in Fall 97 ended here.

6 Good Proofs and Bad Proofs

An estimated 1/3 of all mathematical papers contain errors. Even some of the world's most famous mathematicians have botched proofs. Here are some famous examples.

- Andrew Wiles recently announced a proof of Fermat's Last Theorem. It was several hundred pages long. It took mathematicians months of hard work to discover it had a fatal flaw (so Wiles produced another proof of several hundred pages; this one seems to have convinced people for now).
- Gauss's 1799 Ph.D. thesis is usually referred to as being the first rigorous proof of the Fundamental Theorem of Algebra (every polynomial has a zero over the complex numbers). But it contains quotes like

  > "If a branch of an algebraic curve enters a bounded region, it must necessarily leave it again. ... Nobody, to my knowledge, has ever doubted [this fact]. But if anybody desires it, then on another occasion I intend to give a demonstration which will leave no doubt."

  Fields Medalist Steve Smale writes about this, calling it an "immense gap" in the proof that was not filled in until 1920, more than a hundred years later.

- In 1900 Poincare carelessly claimed a certain very simple topological characterization of the 3-dimensional sphere. Later realizing it was not so obvious, he demoted the claim to the status of a "conjecture" in 1904. The Poincare Conjecture is now one of the biggest open questions in mathematics (two Fields Medals have been given out for partial progress on it).

Here are some of the characteristics of a good proof:

- It is clear, brief, –and correct!
- It is elegant; surprisingly short or clever. Like a "nice hack" in code.
- It has a nice structure, like a good program. It is broken up into separate lemmas that prove key intermediate properties. This makes it easy to understand the reason the whole thing works. It also makes it more likely that pieces can be reused.
- The pieces are general and abstract. This avoids the clutter of unnecessary hypotheses, useless restrictions, etc. Again, the analogy to programming holds; a subroutine should be as generally applicable as possible.
- Inobvious conclusions are not justified by shortcuts such as "it can be shown that..." and "It is obvious that..." and "any moron can see that...". These phrases save the writer's time, but create a chance for error and consume the reader's time.
- Like a scientific experiment, someone else must be able to "replicate" (i.e. understand) your proof.

Proofs are important. They permit you to convince yourself and others that your reasoning is correct. The insights gained can help you understand why something is true and whether it will stay true when other things change. Proofs are particularly important in computer science and electrical engineering. Bugs have proven costly for Intel, AT&T, and Airbus. An good proof is strong evidence that no bugs exist.