Proofs

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/Mysticism "*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.

"Cogito ergo sum" Proof by reasoning about undefined terms.

> This Latin quote translates as "I think, therefore I am." It comes from the beginning of a famous essay by the 17th century Mathematician/Philospher, René Descartes. It may be one of the most famous quotes in the world: do a web search on the phrase and you will be flooded with hits.

> Deducing your existence from the fact that you're thinking about your existence sounds like a pretty cool starting axiom. But it ain't Math. In fact, Descartes goes on shortly to conclude that there is an infinitely beneficent God.

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

Definition. A *formal proof* of a *proposition* is a chain of *logical deductions* leading to the proposition 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.

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, "Wherefore art thou Romeo?" and "Give me an A!".

Proposition 2.1. *2 + 3 = 5.*

This proposition is true.

Proposition 2.2. *Let* $p(n) ::= n^2 + n + 41.$

$$\forall n \in \mathbb{N} \; p(n) \text{ is a prime number}.$$

The symbol $\forall$ is read "for all". The symbol $\mathbb{N}$ stands for the set of *natural numbers*, which are 0, 1, 2, 3, ... ; (ask your TA for the 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, for example, 2, 3, 5, 7, 11, ...

Let's try some numerical experimentation to check this proposition: $p(0) = 41$ which is prime. $p(1) = 43$ which is prime. $p(2) = 47$ which is prime. $p(3) = 53$ which is prime. ... $p(20) = 461$ which is prime. Hmmm, starts to look like a plausible claim. In fact we can keep checking through $n = 39$ and confirm that $p(39) = 1601$ is prime.

But if $n = 40$, then $p(n) = 40^2 + 40 + 41 = 41 \cdot 41$, which is not prime. Since the expression is not prime *for all* n, the proposition is false! In fact, it's not hard to show that *no* nonconstant polynomial can map all natural numbers into prime numbers. The point is in general you can't check a claim about an infinite set by checking a finite set of its elements, no matter how large the finite set. Here are two even more extreme examples:

Proposition 2.3. $a^4 + b^4 + c^4 = d^4$ *has no solution when* a, b, c, d *are positive integers. In logical notation, letting* $\mathbb{Z}^+$ *denote the positive integers, we have*

$$\forall a \in \mathbb{Z}^+ \forall b \in \mathbb{Z}^+ \forall c \in \mathbb{Z}^+ \forall d \in \mathbb{Z}^+ \; a^4 + b^4 + c^4 \neq d^4.$$

Strings of $\forall$'s like this are usually abbreviated for easier reading:

$$\forall a, b, c, d \in \mathbb{Z}^+ \; a^4 + b^4 + c^4 \neq d^4.$$

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* $x, y, z \in \mathbb{N}.$

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[1] regions have different colors.*

---

[1]Two regions are adjacent only when they share a boundary segment of positive length. They are not considered to be adjacent if their boundaries meet only at a few points.

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 laborious proof was finally found about 15 years ago by a Mathematician named Haaken who used a complex computer program to categorize maps as four-colorable; the program left a couple of thousand maps uncategorized, and these were checked by hand by Haaken and his assistants—including his 15-year-old daughter. There was a lot of debate about whether this was a legitimate proof: the argument was too big to be checked without a computer, and no one could guarantee that the computer calculated correctly, nor did anyone have the energy to recheck the four-colorings of thousands of maps that was done by hand. Finally, about five years ago, a humanly intelligible proof of the four color theorem was found (see http://www.math.gatech.edu/ thomas/FC/fourcolor.html).

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 (Goldbach). *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.

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 3.1. If $a = b$ and $b = c$, then $a = c$.

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

Axiom 3.2 (Euclidean geometry). Given a line l and a point p not on l, there is exactly one line through p parallel to l.

Axiom 3.3 (Spherical geometry). Given a line l and a point p not on l, there is *no* line through p parallel to l.

Axiom 3.4 (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 an attractive property; 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 6.042 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 $p \longrightarrow q$ is true, then q is true. The expression $p \longrightarrow q$ is read "p implies q" or "if p, then q." A truth table for $\longrightarrow$ is shown below:

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

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

Rule.

$$\frac{p, \quad p \longrightarrow q}{q}$$

When the statements above the line, called the *antecedents*, are true, then we can infer that the statement below the line, called the *conclusion* or the *consequent*, is also true. There are many other natural inference rules, for example:

Rule.

$$\frac{p \longrightarrow q, \quad q \longrightarrow r}{p \longrightarrow r}$$

Rule.

$$\frac{p \longrightarrow q, \quad \neg q}{\neg p}$$

Rosen describes additional standardized inference rules useful in proofs. 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 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).

- 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 and *correct*!

- It has a nice structure, like a good program. It is broken up into separate parts that define and 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.

- Important conclusions are not "justified" by being "left to the reader," nor by intimidating phrases like "it is obvious that . . . " or "any moron can see that . . . ." These phrases save the writer's time, but consume the reader's time. Mistakes in proofs are typically found in these parts "left to the reader."

- 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. A good proof is strong evidence that no bugs exist.