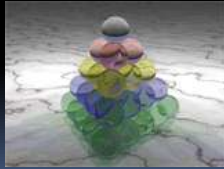
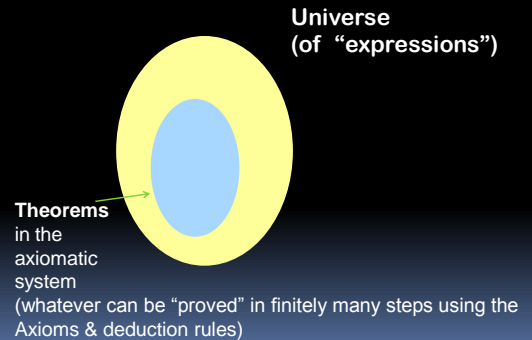


Proofs



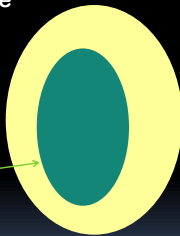
Recap: Axiomatic systems



Recap: Truths

Universe

Truths



Truths = a subset
Of course,
"interesting" subsets
correspond to
"meaningful" notions
of truth
(eg. Tautologies in
Propositional logic)

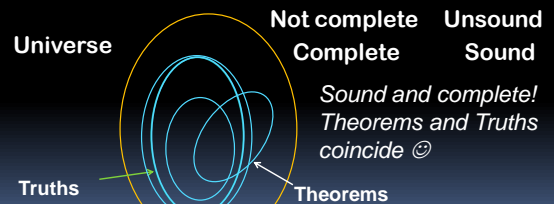
Recap: Soundness & Completeness

Axiomatic system is sound for some truth concept:

"all theorems are truths"

System is complete for some truth concept:

"all truths are theorems"



Proofs

Bits of Wisdom on Solving Problems,
Writing Proofs, and Enjoying the Pain:
How to Succeed in This Class

No specific topic covered today;
General 'fun' lecture

2. What is a proof?
1. How do I find a proof?
3. How do I write a proof?

2. What is a proof?
1. How do I find a proof?
3. How do I write a proof?



Typical philosophy for working in math:

**Small progress per day,
for many days.**

251 HMWK version: 15% progress per day for 7 days.


I don't have any magical ability. I look at a problem, and it looks something like one I've done before; I think maybe the idea that worked before will work here. When I was a kid, I had a romanticized notion of mathematics, that hard problems were solved in 'Eureka' moments of inspiration. [But] with me, it's always, 'Let's try this. That gets me part of the way, or that doesn't work. Now let's try this. Oh, there's a little shortcut here.... It's not about being smart or even fast. It's like climbing a cliff. If you're very strong and quick and have a lot of rope, it helps, but you need to devise a good route to get up there. Doing calculations quickly and knowing a lot of facts are like a rock climber with strength, quickness and good tools. You still need a plan — that's the hard part — and you have to see the bigger picture.



Terence Tao

2006 Fields Medalist,
winner of 10+ international math
prizes worth over \$5 million

10 tips for finding proofs

1. Read and understand the problem.
2. Try small or special cases.
3. Develop good notation. 
4. Understand why the problem seems hard (Put yourself in the mind of the adversary)
5. Clarify, abstract out, summarize pieces. Record partial progress.

10 tips for finding proofs

6. Use blocks of ≥ 1 hour, or at least 30 minutes.
7. Take breaks.
8. Use plenty of paper, and draw pictures if possible.
9. Collaborate, bounce off ideas.
10. A crisp write-up is important (both for scoring points, and checking that argument is airtight).

251 Homework Problem, Spring 2010:

The kitchen for a cookie baking contest is arranged in an m by n grid of ovens. Each contestant is assigned an oven and told to make as many cookies as possible in three hours. Prizes are awarded in the following manner: in each row the p people who produced the most cookies receive a prize. Likewise, in each column the q people who produced the most cookies receive a prize. Assume $p \leq n$, $q \leq m$, and that no two people produced the same number of cookies. Prove that at least pq people received two prizes for their cookie-baking performance.

Solution write-up

Proof by induction on $n+m$.

$P(k)$ = claim true when $n+m=k$ for all $(p,q) \in \{1,2,\dots,n\} \times \{1,2,\dots,m\}$

$P(2)$ is true ($n=m=p=q=1$)

Assume $P(k)$ is true. Let's prove $P(k+1)$. Suppose $n+m=k+1$.

If everyone who wins a prize wins two prizes, we are done, since at least $(mp+nq)/2 \geq pq$ people win prizes.

So there is someone who receives just one prize. Among those, pick the person, say X , who made the most cookies. Either X is not among top p in her row or not among the top q in her column.

Without loss of generality, assume the latter. (Why's this okay?)

Remove X 's column. By induction hypothesis, the remaining $m \times (n-1)$ grid has at least $(p-1)q$ people receiving two prizes (since every row has at least $(p-1)$ prize winners in new grid). Add to this set the q winners in X 's column, who by choice of X , all win two prizes (otherwise X wouldn't have been the largest single prize winner). This gives pq two-prize winners in all. QED.

If you just read the solution, it's frustrating:

Writeup is short: 3 short paras.

Seems to have some "aha!" moments (eg. choice of X)

Hides cognitive process behind discovery of "aha!"-like step(s).

But you need to set yourself up for making such a step.

For the write-up, you can step back and try for the clearest possible explanation (which often is also succinct, but some intuition is nice to include, especially in difficult proofs).

2. What is a proof?

1. How do I find a proof?

3. How do I write a proof?

What is a proof?

In math, there are agreed-upon rigorous rules of deduction. Proofs are right or wrong.

Nevertheless, what constitutes an acceptable proof is a social construction.

(But computer science can help.)

Proofs — prehistory



Euclid's *Elements*
(ca. 300 BCE)

Canonized the idea of giving a rigorous, axiomatic deduction for all theorems.

Proofs — 19th century

True rigor developed.

Culminated in the understanding that math proofs can be formalized with First Order Logic.



Bertrand Russell



Alfred Whitehead

Principia Mathematica, ca. 1912

Developed set theory, number theory, some real analysis using **formal logic**.

page 379: " $1+1=2$ "

It became generally agreed that you **could** rigorously formalize mathematical proofs.

But nobody wants to!
(by hand, at least)

But are English-language proofs sufficient?

Four Color Theorem

1852 conjecture:

Any 2-d map of regions can be colored with 4 colors so that no adjacent countries get the same color.



Four Color Theorem

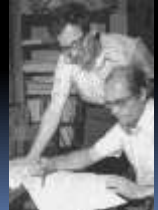
- 1879: Proved by Kempe in *Amer. J. of Math*
- 1880: Alternate proof by Tait in *Trans. Roy. Soc. Edinburgh*
- 1890: Heawood finds a bug in Kempe's proof.
- 1891: Petersen finds a bug in Tait's proof.

Kempe's "proof" was widely acclaimed.

Four Color Theorem

- 1969: Heesch showed that the theorem could in principle be reduced to checking a large number of cases.
- 1976: Appel and Haken wrote a massive amount of code to compute and then check 1936 cases (1200 hours of computer time).

Claimed this constituted a proof.



Four Color Theorem

Much controversy at the time. **Is this a proof??**

Arguments against:

- No human could ever hand-check the cases.
- Perhaps there's a bug in the code.
- Perhaps there's a bug in the compiler.
- Perhaps there's a bug in the hardware.
- No "insight" is derived – still don't know "why" the theorem is true.

Nevertheless, these days, pretty much everyone accepts that it counts as a proof.

"Simpler" proof: Roberston, Sanders, Seymour, Thomas 1997

Classification of finite simple groups (the "prime numbers" of group theory)

Theorem: Every finite simple group is one of [4 families or 26 special cases].

Progress started in late 19th century.

100's of papers, 10,000–20,000 pages later...

1983: Gorenstein announces proof is complete.

However, experts knew one piece still missing.

2004: Aschbacher & Smith finish a 1221-page paper, Aschbacher announces proof is complete.

Classification of finite simple groups

Some controversy: **Is the theorem proven?**

A concern:

Everyone who understands the proof may **die** before it's properly collated.

A ~5000 page, 13-volume series of books describing the proof is underway.

More anecdotes

1993: Wiles announces proof of Fermat's Last Thm. Then a bug is found.

1994: Bug fixed, 100-page paper.

1994: Gaoyong Zhang, *Annals of Mathematics*: disproves "n=4 case of Busemann-Petty".

1999: Gaoyong Zhang, *Annals of Mathematics*: **proves** "n=4 case of Busemann-Petty".

Kepler Conjecture



Kepler, 1611: As a New Year's present (???) for his friend, wrote a paper with this conjecture:

The densest way to pack spheres is like this:



Kepler Conjecture

2005:

Our neighbor Tom Hales:
120 page proof in
Annals of Mathematics



Plus code to solve 100,000 distinct optimization problems, taking 2000 hours computer time.

Annals recruited a team of 20 referees.
They worked for 4 years.

Some quit. Some retired. One died.

In the end, they gave up.

But said they were "99% sure" it was a proof.

Kepler Conjecture

Hales: "We will code up a completely formal axiomatic proof, checkable by computer."



Open source "Project Flyspeck":
2004 --- August 10, 2014

Computer-assisted proof

Proof assistant software like HOL Light, Mizar, Coq, Isabelle, does two things:

1. Checks that a proof encoded in an axiomatic system for First Order Logic (or typed lambda calculus theory) is valid.
2. Helps user code up such proofs.

Developing proof assistants is an active area of research, particularly at CMU!

Computer-assisted proof

Suppose, e.g., HOL Light certifies a formal proof.
Can you trust it?

- You don't need to trust the million-line proof.
- You don't need to trust the process used to generate that proof.
- You just need to trust HOL Light's 430-line program for verifying FOL deductions.

Computer-formalized proofs

Fundamental Theorem of Calculus (Harrison)

Fundamental Theorem of Algebra (Milewski)

Prime Number Theorem (Avigad++ @ CMU)

Gödel's Incompleteness Theorem (Shankar)

Jordan Curve Theorem (Hales)

Brouwer Fixed Point Theorem (Harrison)

Four Color Theorem (Gonthier)

2. What is a proof?
1. How do I find a proof?
3. How do I write a proof?

So as to get full points on the homework.

Your homework is not like the Four Color Theorem.

The TAs can correctly decide if you have written a valid proof.

Here is the mindset you must have.

Pretend that your TA is going to code up a formalized proof of your solution.

Your job is to write a complete English-language **spec** for your TA.

You must give a spec to your TA that they could implement with no complaints or questions.

Equivalently, you must convince your TA that you know a complete, correct proof.

Alternate Perspective



You: must present an airtight case.



Your TA

Possible complaints/points off from your TA:

- A does not logically follow from B.
- You missed a case.
- This statement is true, but you haven't justified it.

But also:

- I don't understand your proof.
- This explanation is unclear.
- Your proof is very hard to read.

Problem: Prove $n^2 \geq n$ for all integers n .

Solution:

We prove $F_n = "n^2 \geq n"$ by induction on n .
The base case is $n = 0$: indeed, $0^2 \geq 0$.
Assume F_n . Then
 $(n+1)^2 = n^2 + 2n + 1 \geq n^2 + 1 \geq n + 1$ (by F_n).
This is F_{n+1} , so the induction is complete.

Read the question carefully.

Some common induction mistakes

"The base case F_0 is true because [...].
For the induction step, assume F_k holds for all k .
We now show that F_{k+1} holds..."

You just assumed what you're trying to prove!

"The proof is by strong induction.
The base case F_0 is true because [...].
For the induction, assume F_k holds for all $k \leq n$.
We will now show F_{k+1} [...]"

What is k ? Where did n go?

Spring '11 homework 2:

How many ways to arrange $c \geq 0$ ♣'s and $d \geq 0$ ♦'s so that all ♣'s are consecutive?

Solution:

You can have any number between 0 and d ♦'s, then the string of ♣'s; then you must have the remainder of the ♦'s. Hence there are $d+1$ possibilities.

Fallacious if $c = 0$: there is only 1 possibility.

**Handle all edge cases!
Don't have any missing parts in your spec.**

Problem: Prove $2^n > n$ for all integers $n \geq 1$.

Solution:

$F_n = "2^n > n"$

$F_1 = "2 > 1"$

$F_n \Rightarrow F_{n+1}$

$2^{n+1} = 2 \cdot 2^n > 2 \cdot n$ (induction) $\geq n+1$

because $n \geq 1$

Therefore proved.

This is not a full sentence.

This is not written in English!

Is this a definition? A claim?

What does this check mark mean?

Is this a claim? An assumption?

Oh, apparently you're doing an induction? [sarcasm]

Is it? Why?

Spring '11 homework 2, #3a:

There is a circle of 15,251 chips, green on one side, red on the other. Initially all show the green side. In one step you may take any four consecutive chips and flip them. Is it possible to get all of the chips showing red?

Intended solution:

No. If g of the 4 flipped chips are green, then after flipping $4-g$ of them are green. Note that g and $4-g$ have the same parity; hence the parity of the number of green chips will always remain odd.

Solution:

No it is not possible. Let's assume for contradiction we converted all 15,251 chips to red. But this means in the very last step there must be 4 consecutive green chips and the remaining 15,247 must be red. Repeating this k times for $1 \leq k \leq 3812$, we get three consecutive red chips, with the rest green. But we started from all green, contradiction.

If asked to show something is impossible, it does not suffice to show that one particular method does not work.

Spring '11 homework 2, #3b:

There is a circle of 15,251 chips, green on one side, red on the other. Initially all show the green side. In one step you may take any **seven** consecutive chips and flip them. Is it possible to get all of the chips showing red?

Intended solution:

Yes. Number the chips $0 \dots 15,250$. Flip the sequence $[0, 1, \dots, 6]$, then $[1, 2, \dots, 7]$, then $[2, 3, \dots, 8]$, etc., up until $[15,250, 0, 1, \dots, 5]$. Now each chip's been flipped exactly 7 times, an odd number. Hence each chip is now red.

Solution:

At any given time, let g be the number of chips showing green and r the number of chips showing red. The possible remainders when a number is divided by 7 are 0, 1, 2, 3, 4, 5, 6, 7. A flip that involves 6 red and 1 green increments the current modular class of g by 5 while the move that involves 1 red and 6 green decrements the current modular class of g by 5. Originally, with the number 15,251, the modular class of $g \bmod 7$ is 5. Thus, it is possible to make all chips red.

In short: this proof does not make sense. Do not just write a bunch of random facts.

A software company interview question



Four guys want to cross a bridge that can only hold two people at one time.

It is pitch dark and they only have one flashlight, so people must cross either alone or in pairs (bringing the flashlight).

Their walking speeds allow them to cross in 1, 2, 5, and 10 minutes, respectively.

Is it possible for them to all cross in 17 minutes?



Intuitive, But False

" $10 + 1 + 5 + 1 + 2 = 19$, so the four guys just can't cross in 17 minutes"

"Even if the fastest guy is the one to shuttle the others back and forth – you use at least $10 + 1 + 5 + 1 + 2 > 17$ minutes"

Vocabulary Self-Proofing

As you talk to yourself, make sure to tag assertions with phrases that denote degrees of conviction

Keep track of what you actually know – remember what you merely suspect

" $10 + 1 + 5 + 1 + 2 = 19$, so *it would be weird if* the four guys could cross in 17 minutes"

~~even~~ if we use the fastest guy to shuttle the others, they take too long."



If it is possible, there must be more than one guy doing the return trips:
it must be that someone gets deposited on one side and comes back for the return trip later!



Suppose we leave 1 for a return trip later

We start with 1 and X and then X returns
Total time: 2X

Thus, we start with 1,2 go over and 2 comes back....



1 2 5 10



1 2 5 10



1 2 5 10
5 10 2 1



1 2 5 10
5 10 2 1

$$\begin{array}{r} \underline{12}510 \\ 510 \\ 2 \quad 510 \end{array}$$

$$\begin{array}{r} \underline{2}1 \\ 1 \end{array}$$

$$\begin{array}{r} \underline{12}510 \\ 510 \\ 2 \quad \underline{510} \end{array}$$

$$\begin{array}{r} \underline{2}1 \\ 1 \end{array}$$

$$\begin{array}{r} \underline{12}510 \\ 510 \\ 2 \quad \underline{510} \\ 2 \end{array}$$

$$\begin{array}{r} \underline{2}1 \\ 1 \\ 1510 \end{array}$$

$$\begin{array}{r} \underline{12}510 \\ 510 \\ 2 \quad \underline{510} \\ 2 \end{array}$$

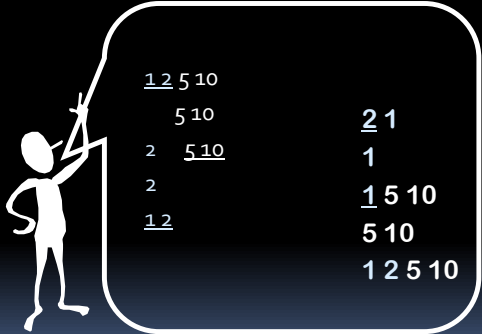
$$\begin{array}{r} \underline{2}1 \\ 1 \\ \underline{15}10 \end{array}$$

$$\begin{array}{r} \underline{12}510 \\ 510 \\ 2 \quad \underline{510} \\ 2 \\ 12 \end{array}$$

$$\begin{array}{r} \underline{2}1 \\ 1 \\ \underline{15}10 \\ 510 \end{array}$$

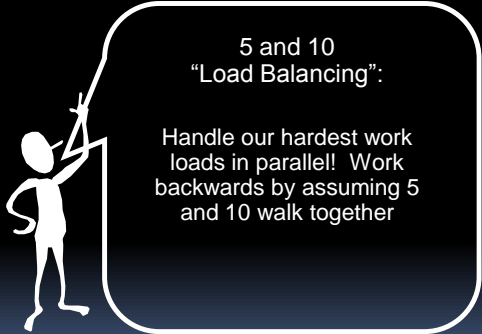
$$\begin{array}{r} \underline{12}510 \\ 510 \\ 2 \quad \underline{510} \\ 2 \\ \underline{12} \end{array}$$

$$\begin{array}{r} \underline{2}1 \\ 1 \\ \underline{15}10 \\ 510 \end{array}$$




$$\begin{array}{r} \underline{1\ 2} \ 5\ 10 \\ 5\ 10 \\ 2\ \underline{5\ 10} \\ 2 \\ \underline{1\ 2} \end{array}$$

$$\begin{array}{r} \underline{2} \ 1 \\ 1 \\ \underline{1\ 5} \ 10 \\ 5\ 10 \\ 1\ 2\ 5\ 10 \end{array}$$



5 and 10
 "Load Balancing":
 Handle our hardest work loads in parallel! Work backwards by assuming 5 and 10 walk together

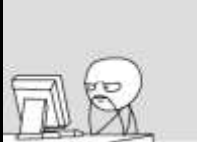
That really was an interview question



Why do they ask such questions, as opposed to asking for a piece of code to do binary search?

Success in computer science requires:

- Content: An up to date grasp of fundamental concepts and problems
- Method: Principles and techniques to solve the vast array of unfamiliar problems that arise in a rapidly changing field



Study Guide

Solving problems:
 Understand problem
 Try small cases
 Use enough time & paper
 put yourself in the mind of adversary

Writing proofs:
 like designing a complete, correct spec
 put yourself in the TA's shoes
 use good English!