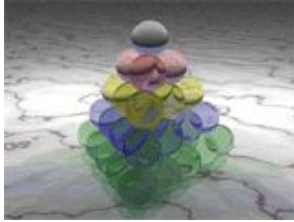


Problem Solving: Where Does the Aha Come From?



What is a proof?

How do I find a proof?

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.

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

Proofs — prehistory



Euclid's *Elements*
(ca. 300 BCE)
consisting of 13 books

It is a collection of definitions, axioms, theorems and mathematical proofs.

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

Proofs — 19th century



Bertrand Russell



Alfred Whitehead

Principia Mathematica, 3 vols, 1910, 12, 13

showed that math is really boiled down to **logic**;
defined at the very basic level what math was.

Vol. 2, page 80: "1+1=2"

Proofs — 19th century

The goal of mathematics is to abolish all thought.



Alfred Whitehead

The ultimate goal of mathematics is to eliminate all need for intelligent thought



Donald Knuth

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

conjectured in 1852

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.

Four Color Theorem

conjectured in 1852

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 that 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" to derivation

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

Progress started in late 19th century.

100's of papers, ~20,000 pages ...

1983: Gorenstein announces proof is complete.

However, one piece was missing.

2004: Aschbacher & Smith finished a 1221-page paper, announced proof is complete.

Classification of finite simple groups

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 was found.

1994: Bug fixed, 100-page paper.

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

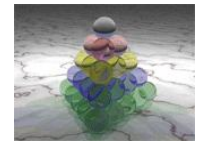
1999: Gaoyong Zhang, *Annals of Mathematics*:
proved "n=4 case of Busemann-Petty".

Kepler Conjecture



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

about cannonballs packing
in 3D
Euclidean space



physicists knew,
mathematicians believed

Kepler Conjecture

2005:
Tom Hales: 120 page proof



Plus code to solve 100,000
distinct optimization
problems...

The journal 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: "I will code up a
completely formal axiomatic
proof,
checkable by computer."



Open source "Project FlysPecK"

Hales estimates that producing a complete
formal proof will take around 20 years of work.

Computer-assisted proof

Proof assistant software:

HOL, Mizar, Coq, Isabelle

1. Checks that a proof encoded in an axiomatic system is valid.
2. Helps user code up such proofs.

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

Computer-formalized proofs

Fundamental Theorem of Calculus (Harrison)

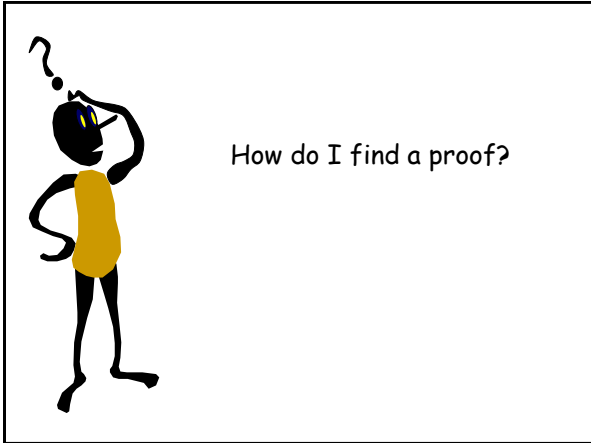
Fundamental Theorem of Algebra (Milewski)

Prime Number Theorem (Avigad @ CMU)

Jordan Curve Theorem (Hales)

Fixed Point Theorem (Harrison)

Four Color Theorem (Gonthier)



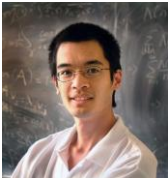
Typical philosophy for research level math:

Small progress each day, for many days.

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

I don't have any magical ability. ... 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.

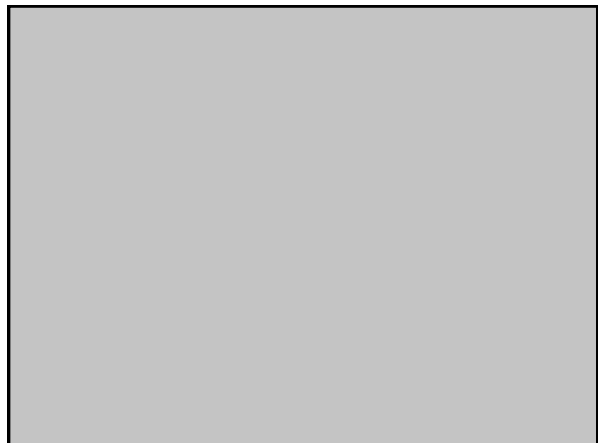


Terence Tao
2006 Fields Medalist,
winner of 10+ international math
prizes worth \approx \$2,000,000

- ### Some tips for finding proofs
1. Read and understand the problem.
 2. Try small or special cases.
 3. Use plenty of paper, and draw pictures
 4. Collaborate, bounce off ideas.
 5. Record partial progress.
 6. Use blocks of ≥ 1 hour, but not long
 7. Take breaks.

251 Homework Problem

The kitchen for a cookie baking contest is arranged in an $m \times n$ 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 $p \leq n$ people who produced the most cookies receive a prize. Likewise, in each column the $q \leq m$ people who produced the most cookies receive a prize. Assume that no two people produced the same number of cookies. Prove that at least $p q$ people received two prizes.





Where Does the Aha Come From?

The 'Aha' comes from:

Recall the chocolate breaking problem
(induction on the total number of ovens)

Recall the problem of all students of the same gender
(remove a person and then insert it back)

Solution write-up

Proof by induction on $n+m$.

$P(k)$ = claim true when $n+m=k$

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

Assume $P(k)$ is true. Let's prove $P(k+1)$.

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. WLOG, assume the latter.

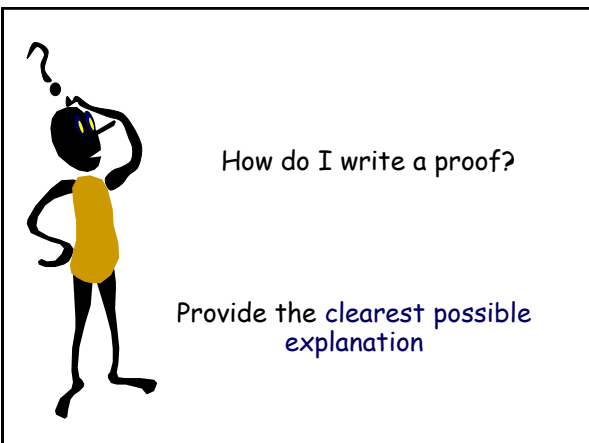
Remove X 's column. By induction hypothesis, at least $(p-1)q$ people receiving two prizes. Add to this set the q winners in X 's column. This gives pq two-prize winners in all.

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

Writeup is short

Hides cognitive process behind discovery of "aha!"
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).



Your homework is not like
the Fermat Last 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 specification 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.

Example:

How many ways to arrange $c \geq 0$ 0's and $d \geq 0$ 1's so that all 0's are consecutive?

Solution:

You can have any number of 1's, followed by all 0'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:

$$F_n = 2^n > n$$

$$F_1 = 2 > 1$$

$$F_n \Rightarrow F_{n+1}:$$

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

because $n \geq 1$

Therefore proved.

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]

This is not a full sentence.

This is not written in English!

Problem Solving:

Where Does the Aha Come From?




I am just going to ask you a software company interview questions

There is a rectangular field, with a circular lake somewhere inside of it.

You know the dimensions of the field, and the radius and location of the lake.

Your task is to find a straight line such that cuts the grass in half.

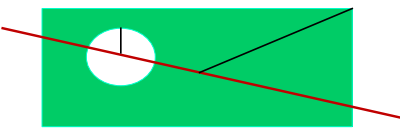





There is a rectangular field, with a circular lake somewhere inside of it.

You know the dimensions of the field, and the radius and location of the lake.

Your task is to find a straight line such that cuts the grass in half.



A flashlight problem



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.


How long does it take them to cross the bridge?

Intuitive, but ...


$10 + 1 + 5 + 1 + 2 = 19$

Even if the fastest guy is the one to shuttle the others back and forth - you use at least 19 minutes


The Aha Moment?



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!



1, 2, 5, 10



1, 2, 5, 10

$\underline{1, 2}, 5, 10$
 $5, 10$

$1, 2$

$\underline{1, 2}, 5, 10$
 $5, 10$

$1, \underline{2}$

$\underline{1, 2}, 5, 10$
 $5, 10$
 $2, 5, 10$

$1, \underline{2}$
 1

$\underline{1, 2}, 5, 10$
 $5, 10$
 $2, \underline{5}, 10$

$1, \underline{2}$
 1

$\underline{1, 2}, 5, 10$
 $5, 10$
 $2, \underline{5}, 10$
 2

$1, \underline{2}$
 1
 $1, 5, 10$

$\underline{1, 2}, 5, 10$
 $5, 10$
 $2, \underline{5}, 10$
 2

$1, \underline{2}$
 1
 $\underline{1}, 5, 10$

<u>1, 2</u> , 5, 10	
5, 10	1, <u>2</u>
2, <u>5, 10</u>	1
2	<u>1</u> , 5, 10
1, 2	5, 10

<u>1, 2</u> , 5, 10	
5, 10	1, <u>2</u>
2, <u>5, 10</u>	1
2	<u>1</u> , 5, 10
<u>1, 2</u>	5, 10

<u>1, 2</u> , 5, 10	
5, 10	1, <u>2</u>
2, <u>5, 10</u>	1
2	<u>1</u> , 5, 10
<u>1, 2</u>	5, 10
	1, 2, 5, 10

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

The future belongs to the computer scientist who has

- **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

Another reverse flipping problem

Assume that n is an arbitrary positive integer. The n cards of a deck are labeled $1, 2, \dots, n$. Starting with the deck in any order, we repeat the following operation: if the card on top is labeled k , reverse the order of the first k cards.

362154 → 263154 → 623154 → 451326 → 315426
→ 513426 → 243156 → 423156 → 132456

Prove that eventually the first card will be 1, at which state the algorithm terminates.



Study Bee

Solving problems:

- Understand problem
- Try small cases
- Use enough time & paper
- Put yourself in the mind of adversary

Writing proofs:

- Write a complete, correct spec
- Put yourself in TA's shoes
- Use good English!