Proof Verification Technology and Physics

Ernest Davis
Google Research NYC
December 20, 2017

Formal Mathematical Proof

- Rigorous mathematical proofs can be expressed as deductions in extensional logic (FOL or HOL).
- Deduction in formal logic can be characterized in terms of rules for symbol manipulation.

No human understanding; no intuition.

[Set theory suffices as foundation.]

(Frege, Peano, Whitehead & Russell, Tarski, etc.)

Proof Verification Software

has been successfully used to verify very complex and difficult mathematical proofs:

- The prime number theorem
- The Kepler conjecture
- Feit-Thompson theorem (All groups of odd order are solvable.)
- Most theorems in undergraduate math.

More or less, any rigorously proven theorem presumably can be verified if one wants to put in the work.

Can something similar be done for physics?

Outline:

- Discussion of math as point of comparison
- Doing this for physics
- Straw man: Tee-shirt model of physics
- Knocking down the straw man
  - Representation
  - Reasoning
- What *can* be done: Textbook word problems
- My own work [anti-climax]
- Earlier similar proposals and related work
- Looking forward and summary

Disclaimers

- It would be better if I knew more AI and logic.
- I don’t know nearly enough philosophy of science.
- My knowledge of physics is altogether inadequate.

However,

- At 61, you do research with the knowledge and skills that you have, not with those that you wish you had (Rumsfeld).

Value of the logical analysis for math

- A formal standard for rigorous proof.
  - Separate questions:
    - Is this all of what one means by “math” or by “proof”.
    - Historical: Is this what Euclid/Euler meant by “number” or “proof”?
    - Cognitive: Is this how mathematicians/lay people/rats think about mathematical concepts.?
- Metatheorems (Gödel , Turing, etc.)

Value of the proof verification technology for math and CS

- Increase confidence in difficult theorems
- Formal verification of mathematically-oriented software (e.g. floating point).
- Step toward automatically proving theorems
- Technology for other applications e.g. program verification, other kinds of reasoning, tutoring.

What hasn’t been done for math

- An AI that reads a journal article and translates it into a formal proof.
- A user interface that is inviting enough to tempt the “mathematician on the street” to use verification technology.

“Real world” word problems

Leaving aside the NLP problems, what math word problems do we know how to represent?

- For high school and college freshman math (algebra, geometry, calculus, combinatorics, discrete math)
- Perhaps for probability.
- Statistics is doubtful. [Statistics is non-monotonic: If you add more information, then conclusions become invalid.]

Word problem

"If a baseball and a bat cost \$1.10 together, and bats cost \$1.00 more than balls, how much does each cost?"

$\forall_{s:Set}\ Cost(s) = \sum_{x \in s} Cost(x)$

Ball(X). Bat(Y).

$\forall_{x,y}\ Ball(x) \wedge Bat(y) \Longrightarrow Cost(y) = Cost(x)+1.00$

$\forall_{x,y}\ Ball(x) \wedge Bat(y) \Longrightarrow x \neq y$

Cost({X,Y}) = 1.10.

Axioms of real addition and naïve set theory.

Word problem

Two trains 100 miles apart are flying toward each other. One is going 75 mph, the other is going 25 mph. A bird flies back and forth between them at 150 mph. How far does the bird travel before they collide?

$y(0) - x(0) = 100$

$\text{Until}(0, y(t) = x(t), x'(t) = 25)$

$\text{Until}(0, y(t) = x(t), y'(t) = -75).$

$C = \min(t \mid t > 0 \wedge y(t) = x(t))$

$z(0) = x(0)$

$\forall_t\ 0 < t < C \wedge z(t) = x(t) \Rightarrow$

$\quad \text{Until}(t, z(s)=y(s), z'(s)=150 \cdot \text{Sg}(y(t) - x(t))$

$\forall_t\ 0 < t < C \wedge z(t) = y(t) \Rightarrow$

$\quad \text{Until}(t, z(s)=x(s), z'(s)=150 \cdot \text{Sg}(x(t) - y(t))$

Evaluate: arclength(z,0,C)

Physics

Can we

- Represent physics — principles, measurements, observations, experiments —in a formal language?
- Characterize [some] physical argumentation, from principles to observables, in some formal theory of reasoning?
- Implement verification technology?

Potential value

- Philosophical understanding of the nature of physics. We'll come back to this.
- More powerful reasoning about physical systems: Beyond simulation
- Verification of software controlling physical systems (airplanes, robots, nuclear reactors)
- Step toward the super-AI-scientist, who will understand all of science.

Bayesian/MDL formulation (as framework and foil)

Space of scientific theories Φ.

Experiments/observations E.

Outcomes D_E

$$\text{argmax}_{H\in\Phi}\, P(H \mid D_E) =$$

$$\text{argmax}_{H\in\Phi}\, P(D_E \mid H) \cdot P(H)$$

$$P(H) \propto 2^{-|H|} \text{ (say).}$$

In what language can one express all possible theories in Φ, *and all possible experiments* E?

Is there a theory-neutral language of experiments?

Straw man: Tee-shirt model of physics

We encode the 100 top equations of physics into Coq/Isabelle. And then we're pretty much done.

Tee shirt model

- In math, all you need is ZFC + definitions.
- Doesn’t work at all for the other sciences (Chemistry, Biology, Cognitive Science, Social Sciences)
- Physics is a borderline case.

The equations themselves are more complicated than on the tee-shirt

Tee shirt version of Newtonian gravity

$$m\frac{d^2x}{dt^2} = F. \quad F = Gm_i m_j / r^2$$

Actually, for point objects

$$i \neq j \rightarrow \vec{F}_{j,i}(t) = G m_i m_j \frac{\theta(\vec{x}_j(t) - \vec{x}_i(t))}{|\vec{x}_j(t) - \vec{x}_i(t)|^2}$$

$$m_i \frac{d^2 \vec{x}_i(t)}{dt^2} = \sum_{j \neq i} \vec{F}_{j,i}(t)$$

Extended Objects: Particle model

Rigid: $\mathrm{C}(p_i, p_j) \rightarrow |\vec{x}_i(t) - \vec{x}_j(t)| = d_{i,j}$

Elastic: $\mathrm{C}(p_i, p_j) \rightarrow$

$$\vec{F}_{j,i}(t) = k_{i,j}(\left|\vec{x}_j(t) - \vec{x}_i(t)\right| - d_{i,j}) \cdot \theta(\vec{x}_j(t) - \vec{x}_i(t))$$

$$\sim\mathrm{C}(p_i, p_j) \rightarrow \vec{F}_{j,i}(t) = G m_i m_j \frac{\theta(\vec{x}_j(t) - \vec{x}_i(t))}{|\vec{x}_j(t) - \vec{x}_i(t)|^2}$$

$$\vec{F}_{i,j}(t) = -\vec{F}_{j,i}(t)$$

$$m_i \frac{d^2 \vec{x}_i(t)}{dt^2} = \sum_j \vec{F}_{j,i}(t)$$

What the straw man is missing: Grounding

- You have to understand the manifestation of gravity in experiments and observations:
  - Objects on spring scales and balance scales
  - Falling objects
  - Solar system
  - Tides
  - Stars

Really basic stuff is not on the tee-shirt

“What [scientific] statement … contain[s] the most information in the fewest words? … The atomic hypothesis … that *all things are made of atoms*.’’ *Feynman Lectures on Physics* chap. 1

But: Atoms and their interactions are not on the tee shirt. They are a class of solutions to Schrödinger’s equation given the characteristics of atomic nuclei, electrons, and the EM force, in a certain temperature range.

Boundary conditions

A lot of physics

— most of |H| —

is a characterization what the boundary conditions look like in various settings, across a range of scales.

Experimental equipment

- Experimental and observational measurements are not direct perceptions; they rely on technology of ever-increasing complexity.
- So the measurement technology itself has to be validated, theoretically and empirically (when possible — with gravitational lenses, there's only so much you can do.)

Extreme example: BACON

"Discovered" Kepler's third law $T^2 \propto R^3$.
(Langley and Simon, IJCAI-81)

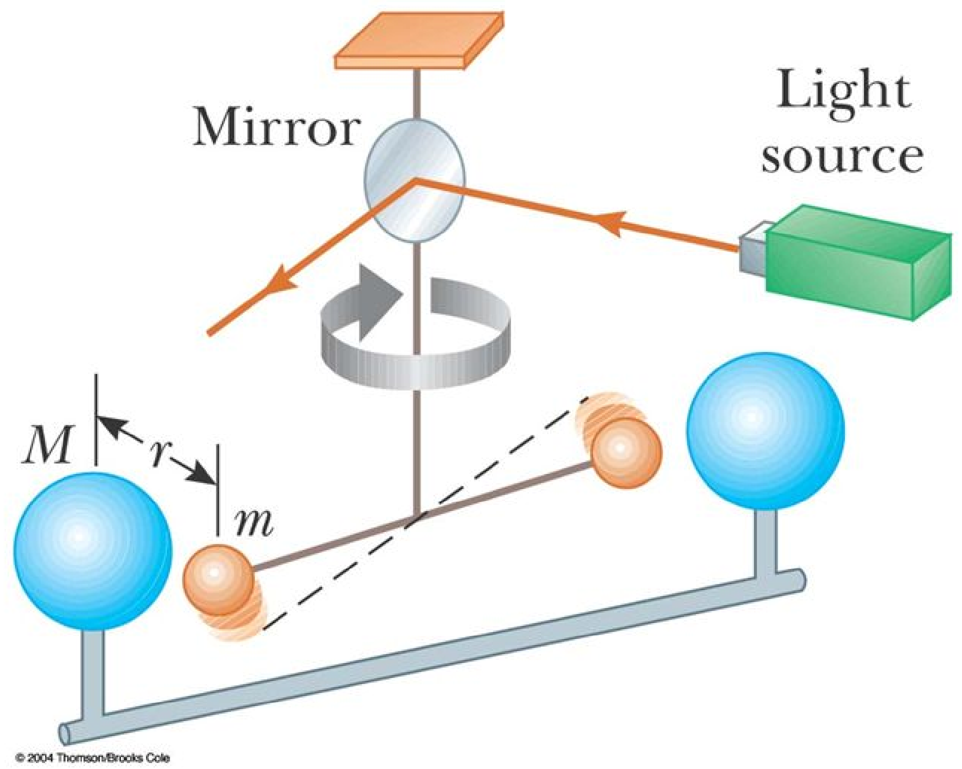
Input to BACON: A table of T and R for planets

Input to Johannes Kepler: Positions of planets and stars in the sky over years. (Projection of their 3D position onto the dome of earth's sky, which itself moves.)

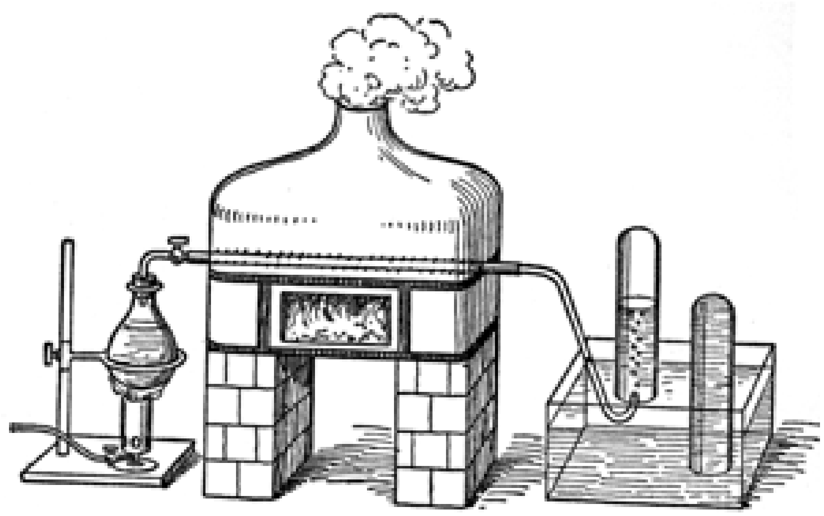
Measuring the gravitational constant

Measuring the Gravitational Constant (G)

Henry Cavendish (1798)



Chemical reaction



Passing steam over heated iron filings, the iron rusts, and you generate hydrogen.

$$2Fe + 3H_2O \rightarrow Fe_2O_3 + 3H_2$$

Faraday, *The Natural History of a Candle*

Universality

Rubner's demonstration of conservation of energy in a dog is an important experiment for *physics;* it shows that the constraints of physics apply to animals.

Claim (Laplace): Newton's law can almost perfectly account for all planetary motions with very high accuracy.

Feynman: "All things are made of atoms."

Claim: Almost all terrestrial phenomena (except a few nuclear reactions) are consequences of quantum mechanics, EM, and gravity, as applied to a configuration of atomic nuclei and electrons.

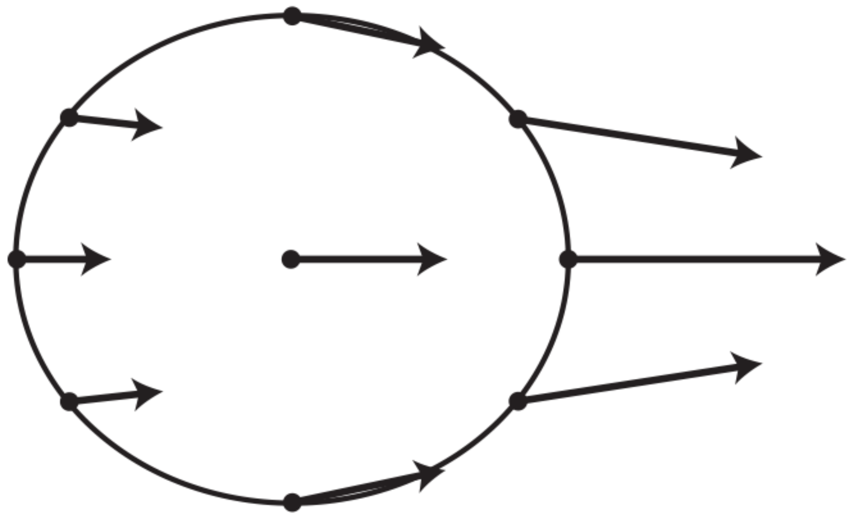
Hand-waving arguments

- Eliminable in mathematics
- Not eliminated in physics. In practice "arguments from first principles" include:
  - Idealization. "Assume a spherical cow."
  - Abstraction. Circuit diagram.
  - Approximate models: Continuum mechanics.
  - Ignore irrelevant issues, negligible quantities.
  - Argument by analogy
  - "Physical intuition" (perhaps analogy)

But the physicists largely agree on how to hand wave, and apparently are doing it right.

Explanation of the Tides

- On the side of Earth that is directly facing the moon, the moon's gravitational pull is the strongest. The water on that side is pulled strongly in the direction of the moon.
- On the side of Earth farthest from the moon, the moon's gravitational pull is at its weakest. At the center of Earth is approximately the average of the moon's gravitational pull on the whole planet. (NOAA, NASA, “SciJinks”)



Maxwell-Boltzmann distribution

Assume that the distribution of velocities of particles in a gas:

A. Is isotropic
B. The components of the velocity in two orthogonal directions are independent.

Then the distribution of velocities follows a Gaussian. (The only remaining degree of freedom is the variance, which is the temperature.)
Michael Strevens, *Tychomancy*

Covalent bonding

The reason is simply that when you allow an electron to wander over a larger space, the kinetic energy always goes down. If you double the size of the space in one direction, the kinetic energy in that direction goes down by a factor of 4. If you consider the two H-atoms as two boxes, doubling the x-size of the box keeping the y and z sizes the same, reduces the kinetic energy from X+X+X to X/4 + X +X or by a factor of 3/4, so the binding energy of two boxes end to end with non-interacting electrons is 1/4 the kinetic energy.

The kinetic energy of an electron in an H-atom is equal to the binding energy (this is the Virial theorem--- the kinetic energy cancels half the potential energy in a 1/r potential to make a binding energy), so you get 1/4 of 12 eV or 3eV of binding energy from this. This is a terrible approximation, because the electrons repel each other, and the H-atom is not a box, but it shows you that allowing the electron volume to spread gains you a lot of energy on the atomic scale, and it is now plausible that even with repulsion, the electrons will bind, and they do.

— Ron Maimon, Physics Stack Exchange

Things that are partially understood

- Historical example: the tides.

  Generally, the twice-daily tides were understood in terms of the moon's gravity by the early –mid 18^{th} century.

  Predicting the tides at a particular location from first principles could not be done until the late 19^{th} century.

Things that are partially understood

In the first chapter we spoke of the great strides that have been made since the early Greek observations of the strange behavior of amber and of lodestone. Yet in all our long and involved discussion, we have never explained *why it is that when we rub a piece of amber we get a charge on it* nor have we explained *why a lodestone is magnetized* ... So you see this physics of ours is a lot of fakery --- we start out with the phenomena of lodestone and amber, and we end up not understanding either of them very well.

Feynman Lectures on Physics, vol. 2 chap. 37

Textbook word problems

Ignoring the NLP issues:

- Choosing the idealization
- Setting up the equations
- Solving the equations

Textbook word problems

Ignoring the NLP issues:

- Choosing the idealization

  Problem wording gives clues. E.g. if no geometric constraints on the shape of an object is not mentioned, then treat it as a point mass.

- Setting up the equations

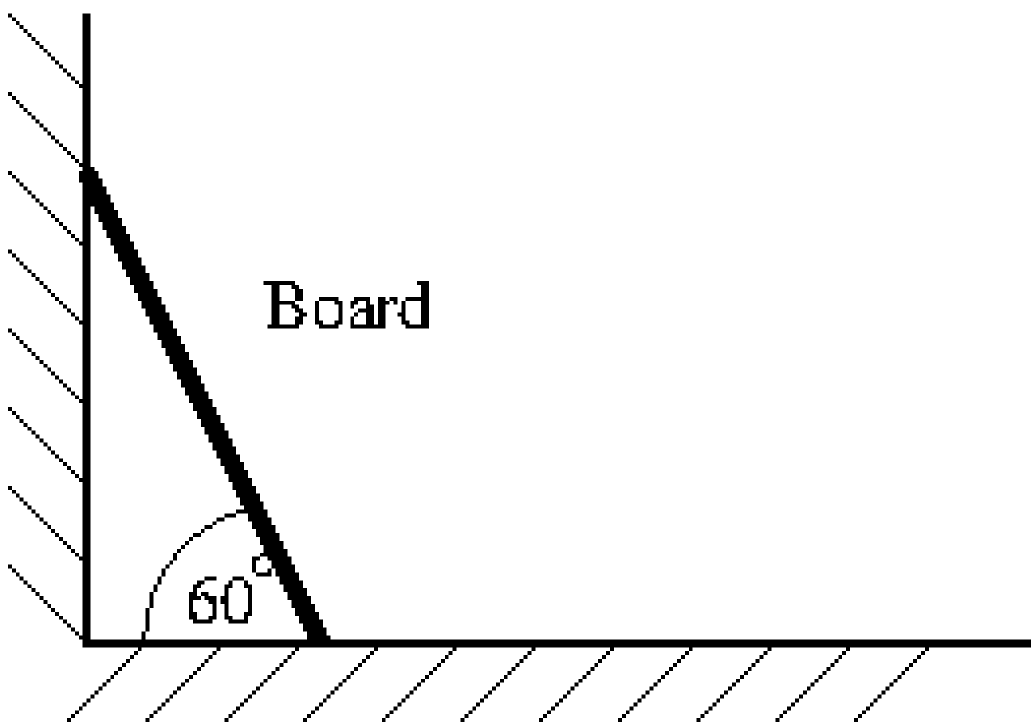
  Largely doable.

- Solving the equations

  Piece of cake

Word problem example

Assuming that the board slides without friction, when does it lose contact with the wall?



Word problem

- Solid(Board). Solid(Wall). Solid(Ground).
- Fixed(Wall). Fixed(Ground). Rigid(Board).
- RightFace(Wall) = $0 \times [0,1]$.
- UpperFace(Ground) = $(-\infty,\infty) \times 0$
- Place(Board,0) = LineSeg(<0,sin(60º)><1/2,0>)
- Isolated(Board,{Wall, Ground}).
- NoInterpenetrate. NoFriction.
- UniformGravity(<0, −g>).

Idealization

Pendulum: Galileo, Cavendish, Foucault, Yo-yo

Setting: 2D/3D. Attachment fixed/rotating with earth.

Bob: Point mass, extended.

String: Fixed distance from attachment. Inelastic. Elastic

Bends along its length. Twists along its axis.

Massless/massed.

Can interfere with itself or with bob.

My own research

- Developing a formal language in which physical behavior can be described at the commonsense (mesoscopic) level.
- Support for qualitative reasoning.
- First-order language with naïve set theory, real arithmetic, real-valued time, Euclidean space
- Solids, liquids, gasses.
- Containers.
- Towards representing experiments like Faraday, Cavendish

Some sample inferences

- If a container remains closed, then matter cannot go from inside to outside.
- Liquid can be carried carefully without spilling in an open container.
- If you put rocks into a pail of water, the level of the water will rise.
- The reaction $2H_2 + O_2 \rightarrow 2H_2O$ consumes twice as many moles of H_2 as of O_2.

Similar proposals

Hilbert's 6th problem: *Mathematical Treatment of the Axioms of Physics.*

The investigations on the foundations of geometry suggest the problem: *To treat in the same manner, by means of axioms, those physical sciences in which already today mathematics plays an important part; in the first rank are the theory of probabilities and mechanics.*

Similar proposal

Work on axiomatizing physics by:

Hilbert, Russell, Carnap, Montague, etc.

Clarifying the relation of thermodynamics to statistical mechanics.

Clarifying the relation of continuum mechanics to particle mechanics.

Quantum logic.

Similar thought

- “The Master Algorithm is the germ of every theory: all we need to add to it to obtain theory X is the minimum data necessary to induce it. In the case of physics, this would be the results of perhaps a few hundred key experiments.”

Pedro Domingos, *The Master Algorithm*

The Vienna Circle and Logical Positivism

Am I reinventing, badly, an 80-year old project by much smarter people, that completely failed?

I hope not.

- The refutations of logical positivism establish that it can't be the whole story. But it still could be part of the story, or a useful approach.
- We can somewhat punt on the question of the ultimate foundations

Hostile reactions

- Refutations of logical positivism: Wittgenstein, Popper, Quine, Kuhn
- General non-interest of physicists and mathematicians
- Jack Schwartz, "The pernicious influence of mathematics on science"

Philosophical insights, revisited

Bayesian/MDL: $\text{argmax}_{H \in \Phi} P(H | D_E) =$

$$\text{argmax}_{H \in \Phi} P(D_E | H) \cdot P(H)$$

- Characterize physical theories (Φ).
- Characterize physics arguments ($P(D_E | H)$).
- Characterize inductive bias ($P(H)$): Symmetry? Locality? Mechanical theories? Simple foundational theories and complex boundary conditions?
- Establish external validity?

Going forward

- Develop representational languages, at many levels of scale and abstraction.
- Push on deductive reasoning, from various starting points.
- Work on characterizing the “hand-waving”. Incorporate work on analogy etc. (Forbus & Gentner)

Final comments

- Carrying out this project for college physics would be orders of magnitude larger than formally verifying college math.
- Justifying equipment may require more advanced physics.
- We are far from the general AI scientist.
- But even doing some of this might be very valuable.