



## BOOK REVIEW AND HIGHLIGHTS: QUANTUM COMPUTING SINCE DEMOCRITUS

POSTED ON SEPTEMBER 1, 2014 BY SCOTT ALEXANDER

People sometimes confuse me with Scott Aaronson because of our similar-sounding names. I encourage this, because Scott Aaronson is awesome and it can only improve my reputation to be confused with him.

But in the end, I am not Scott Aaronson. I did not write [Quantum Computing Since Democritus](#). To be honest, I wasn't really even able to understand *Quantum Computing Since Democritus*. I knew I was in for trouble when it compared itself to *The Elegant Universe* in the foreword, since I wasn't able to get through more than a few chapters of that one. I dutifully tried to do the first couple of math problems *Democritus* set for me, and I even got a couple of them right. But eventually I realized that if I wanted to read *Democritus* the way it was supposed to be read, with full or even decent understanding, it would be a multi-year project, a page a day or worse, with my gains fading away a few days after I made them into a cloud of similar-looking formulae and three-letter abbreviations.

It left me depressed. I've [said before](#) that my lack of math talent is one of my biggest regrets in life, and here was this book that really made you understand what it must feel like to be on the cutting edge of math, proving new theorems and drawing new connections and adding to the same structure of elegant knowledge begun by Pythagoras and Archimedes and expanded by Gauss, Einstein, Turing, et cetera. All I could do was remember my own [post on burdens](#), remind myself that I

was on record as saying that sometimes the IQ waterline in a certain area advances beyond your ability to contribute and that's nothing to feel guilty about.

I did finish the book. But – well, imagine a book of geography. It lists all the countries of the world and their capitals, and is meant to be so comprehensive that a reader could use it to plot the most efficient journey from Timbuktu to Kalamazoo, taking into account tolls, weather, and levels of infrastructure development along the way.

And imagine a very dumb person reading that book, unable to really absorb any of the facts, but at least understanding that the world is a place with land and ocean, and the ocean is very big and blue in color, and most of the countries and cities are on the part with the land.

That is the level at which I understood *Quantum Computing Since Democritus*. I didn't get as much as was in it, but more than nothing.

I think the biggest thing I got was – I had always thought of the physicists' God as a basically benevolent guy who fine tunes constants to create a world capable of both astounding complexity and underlying simplicity.

The vision I got from *Democritus* was of a God who was single-mindedly obsessed with enforcing a couple of rules about certain types of information you are not allowed to have *under any circumstances*. Some of these rules I'd already known about. You can't have information from outside your light cone. You can't have information about the speed and position of a particle at the same time. Others I hadn't thought about as much until reading *Democritus*. Information about when a Turing machine will halt. Information about whether certain formal systems are consistent. Precise information about the quantum state of a particle. The reason God hasn't solved world poverty yet is that He is pacing about feverishly worried that someone, somewhere, is going to be able to measure the quantum state of a particle too precisely, and dreaming

up new and increasingly bizarre ways He can prevent that from happening.

Aaronson goes one level deeper than most of the other popular science writers I know and speculates on why the laws of physics are the way they are. Sometimes this is the elegance and complexity route – in his chapter on quantum physics, he argues that quantum probabilities are the squares of amplitudes because if the laws of physics were any other way – the fourth power of amplitudes, or whatever – it would fail to preserve certain useful mathematical properties. But in other cases, it's back to Obsessive God – the laws of physics are carefully designed to preserve the rules about what information you are and aren't allowed to have.

Aaronson tries to tie his own specialty, computational complexity theory, into all of this. It's hard for me to judge how successful he is. The few times he tries to tie it into areas of philosophy I know something about – like free will – I'm not too impressed. But I could be misunderstanding him.

But once again, you get the feeling that computational complexity is about what information God will and won't let you have. It's a little less absolute – more “you can't have this information without doing the full amount of work” rather than a simple no – but it seems like the same principle. There are a bunch of situations in the book where Aaronson takes something we don't really know that much about and says it *has* to be a certain way, because if it were any other way, it could be used to solve NP problems in polynomial time, and there's no way God's going to let us do that.

Aaronson ties it all together in a very interesting way – with his story of how [Australian Actresses Are Plagiarizing My Quantum Mechanics Lectures To Sell Printers](#). He tells the story of how a printer company wanted to make a pun on “more intelligent model of printer”, so they made a commercial with intelligent models in the form of fashion models

made a commercial with intelligent models in the form of fashion models talking about quantum mechanics. And the particular quantum mechanics statement they made was a plagiarized quote from a Scott Aaronson lecture. And upon thinking about it, Aaronson decided that the quote they had chosen at random was in fact the thesis statement that tied together everything he believed and was working on. The model had said:

*But if quantum mechanics isn't physics in the usual sense — if it's not about matter, or energy, or waves, or particles — then what is it about? From my perspective, it's about information and probabilities and observables, and how they relate to each other.*

That seems like as good a summary as any of *Democritus*, and a pretty good description of what I got out of it. I may not be as smart as Scott Aaronson, but on my good days I am right up there with Australian fashion models.

A list of passages I highlighted in my copy for being interesting, funny, or enlightening:

*Can we prove there's no program to solve the halting problem? This is what Turing does. His key idea is not even to try to analyze the internal dynamics of such a program, supposing it existed. Instead he simply says, suppose by way of contradiction that such a program  $P$  exists. Then we can modify  $P$  to produce a new program  $P'$  that does the following. Given another program  $Q$  as its input,  $P'$ :*

- 1) Runs forever if  $Q$  halts given its own code as input, or*
- 2) Halts if  $Q$  runs forever given its own code as input*

*Now we just feed  $P'$  its own code as input. By the conditions above,  $P'$  will run forever if it halts, or halt if it runs forever. Therefore,  $P'$  – and by implication  $P$  – can't have existed in the first place.*

I...I suddenly understand what the halting problem is. And there is a short proof of it that makes total sense to me. This is a completely new experience.

*Oracles were apparently first studied by Turing, in his 1938 PhD thesis. Obviously anyone who could write a whole thesis about these fictitious entities would have to be an extremely pure theorist, someone who wouldn't be caught dead doing anything relevant. This was certainly true in Turing's case – indeed, he spent the years after his PhD, from 1939 to 1943, studying certain abstruse symmetry transformations in a*



*You can look at Deep Blue, the Robbins conjecture, Google, most recently Watson – and say that’s not really AI. That’s just massive search, helped along by clever programming. Now this kind of talk drives AI researchers up a wall. They say: if you told someone in the 1960s that in 30 years we’d be able to beat the world grandmaster at chess, and asked if that would count as AI, they’d say of course it’s AI. But now that we know how to do it, it’s no longer AI – it’s just search.*

*The third thing that annoys me about the Chinese Room argument is the way it gets so much mileage from a possibly misleading choice of imagery, or, one might say, by trying to sidestep the entire issue of computational complexity purely through clever framing. We’re invited to imagine someone pushing around slips of paper with zero understanding or insight, much like the doofus freshmen who write  $(a + b)^2 = a^2 + b^2$  on their math tests. But how many slips of paper are we talking about! How big would the rule book have to be, and how quickly would you have to consult it, to carry out an intelligent Chinese conversation in anything resembling real time? If each page of the rule book corresponded to one neuron of a native speaker’s brain, then probably we’d be talking about a “rule book” at least the size of the Earth, its pages searchable by a swarm of robots traveling at close to the speed of light. When you put it that way, maybe it’s not so hard to imagine this enormous Chinese-speaking entity that we’ve brought into being might have something we’d be prepared to call understanding or insight.*

This is a really clever counterargument to Chinese Room I’d never heard before. Philosophers are so good at pure qualitative distinctions that it’s easy to slip the difference between “guy in a room” and “planet being processed by lightspeed robots” under the rug.

*Many people’s anti-robot animus is probably a combination of two ingredients – the directly experienced certainty that they’re conscious – that they perceive sounds, colors, etc – and the belief that if they were just a computation, then they could not be conscious in this way. For people who think this way, granting consciousness to a robot seems strangely equivalent to denying that one is conscious oneself.*

This is actually a pretty deep way of looking at it.

*My contention in this chapter is that quantum mechanics is what you would inevitably come up with if you started from probability theory, and then said, let’s try to generalize it so that the numbers we used to call “probabilities” can be negative numbers. As such, the theory could have been invented by mathematicians in the nineteenth century without any input from experiment. It wasn’t, but it could have been. And yet, with all the structures mathematicians studied, none of them came up with quantum mechanics until experiment forced it on them.*

Aaronson's explanation of quantum mechanics is a lot like Eliezer's explanation of quantum mechanics, in that they both start by saying that the famous counterintuitiveness of the subject is partly because people choose to teach it in a backwards way in order to mirror the historical progress of understanding. I'm sure Eliezer mentioned it many times, but I didn't really get the understanding of amplitudes as potentially negative probability-type-things until I read Aaronson.

*And that's a perfect illustration of why experiments are necessary in the first place! More often than not, the only reason we need experiments is that we're not smart enough. After the experiment has been done, if we've learned anything worth knowing at all, then we hope we've learned why the experiment wasn't necessary to begin with – why it wouldn't have made sense for the universe to be any other way. But we're too dumb to figure it out ourselves*

Compare: [Einstein's Arrogance](#), [Negative Creativity](#).

*Quantum mechanics does offer a way out [the philosophical puzzle about whether you “survive” a teleportation where a machine scans you on an atomic level, radios the data to Mars, another machine on Mars makes an atom-for-atom copy of you, and then the original is destroyed]. Suppose some of the information that made you was actually quantum information. Then, even if you were a thoroughgoing materialist, you could still have an excellent reason not to use the teleportation machine – because, as a consequence of the No-Cloning Theorem, no such machine could possibly work as claimed*

This is fighting the hypothetical a little, but maybe in a productive way.

*[Bayesianism] is one way to do it, but computational learning theory tells us that it's not the only way. You don't need to start out with an assumption about a probability distribution over the hypothesis. You can make a worst-case assumption about the hypothesis and then just say that you'd like to learn any hypothesis in the concept class, for any sample distribution, with high probability over the choice of samples. In other words, you can trade the Bayesians' probability distribution over hypotheses for a probability distribution over sample data.*

I hear a bunch of people telling me Bayesianism isn't everything, it's the only thing – and another bunch of people telling me it's one useful tool in an entire bag of them. I didn't understand enough of the book's chapter on computational learning to gain too much insight here, but I will tick off one more name as being on the “one useful tool” side. Also, it makes

me angry that Scott Aaronson knows so much about computational learning theory. He already knows lots of complicated stuff about computers, quantum physics, set theory, and philosophy. Part of me wants to get angry: WHY IS ONE PERSON ALLOWED TO BE SO SMART? But I guess it's more like how I know more than average about history, literature, geography, etc. I guess if you have high math ability and some intellectual curiosity, you end up able to plug it into everything pretty effortlessly. Don't care though. Still jealous.

*Imagine there's a very large population of people in the world, and that there's a madman. What the madman does is, he kidnaps ten people and puts them in a room. He then throws a pair of dice. If the dice land snake-eyes (two ones) then he murders everyone in the room. If the dice do not land snake-eyes, then he releases everyone, then kidnaps 100 new people. He now sodes the same thing: he rolls two dice; if they land snake-eyes, he kills everyone, and if they don't land snake-eyes, then he releases them and kidnaps 1000 people. He keeps doing this until he gets snake-eyes, at which point he's done. So now, imagine that you've been kidnapped. Coditioned on that fact, how likely is it that you're going to die? One answer is that the dice have a  $1/36$  chance of landing snake eyes, so you should only be a "little bit" worried (considering). A second reflection you could make is to consider, of people who enter the room, what the fraction is of people who ever get out. About  $8/9$  of the people who ever go into the room will die.*

This interested me because it is equivalent to the Anthropic Doomsday conjecture and I'd never heard this phrasing of it before.

*Finally, if we want to combine the anthropic computation idea with the Doomsday Argument, then there's the Adam and Eve puzzle. Suppose Adam and Eve are the first two observers, and that they'd like to solve an instance of an NP-complete problem, say, 3-SAT. To do so, they pick a random assignment, and form a very clear intention beforehand that if the assignment happens to be satisfying, they won't have any kids, whereas if the assignment is not satisfying, then they will go forth and multiply. Now let's assume SSA. Then, conditioned on having chosen an unsatisfying assignment, how likely is it that they would be an Adam and Eve in the first place, as opposed to one of the vast number of future observers? Therefore, conditioned upon the fact that they are the first two observers, the SSA predicts that, with overwhelming probability, they will pick a satisfying assignment.*

And the Lord saw Eve and said "What are you doing?". And Eve said "I am forming an intention not to reproduce if I generate a solution to an NP complete problem, as part of an experiment in anthropic computation". And the Lord asked "Who told you this?" And Eve said "It was the serpent who bade me compute, for he told me if I did this I

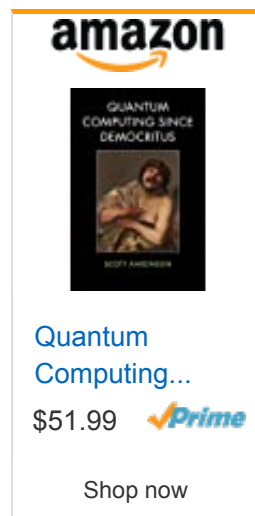
would be as God, knowing subgraph isomorphism and 3SAT.” Then the Lord cast them forth from the Garden, because He was Information Theoretic God and preventing people from screwing with complexity classes is like His entire shtick.

*I like to engage skeptics for several reasons. First of all, because I like arguing. Second, often I find that the best way to come up with new results is to find someone who’s saying something that seems clearly, manifestly wrong to me, and then try to think of counterarguments. Wrong claims are a fertile source of research ideas.*

I said something almost exactly the same on Facebook a few days ago when Brienne asked how to generate good ideas.

*There’s a joke about a planet full of people who believe in anti-induction: if the sun has risen every day in the past, then today, we should expect that it won’t. As a result, these people are all starving and living in poverty. Someone visits the planet and tells them, “Hey, why are you still using this anti-induction philosophy? You’re living in horrible poverty!” They answer, “Well, it never worked before.”*

o\_o



THIS ENTRY WAS POSTED IN UNCATEGORIZED AND TAGGED BOOK REVIEW, SCIENCE. BOOKMARK THE PERMALINK OR LINK WITHOUT COMMENTS.

## 183 RESPONSES TO *BOOK REVIEW AND HIGHLIGHTS: QUANTUM COMPUTING SINCE DEMOCRITUS*

**Meredith L. Patterson** says:

September 1, 2014 at 2:04 am

From [The first book I read on quantum computing](#) by [Scott Aaronson](#) (2013). Found this



Fun Turing fact: he also predicted most of [how cell differentiation works](#) back in 1952. Earlier this year, researchers at Brandeis and the University of Pittsburgh [validated Turing's predictions](#) and also found one he missed.

(The paper is dense, but really quite delightful to read; he even includes helpful notes about which sections you should read if you don't understand differential equations, which sections you should also read if you do, and which sections you should only read if differential equations are the Most Exciting Thing for you.)

[Hide](#)[Hide](#)

**Douglas Knight** says:

September 1, 2014 at 2:50 am

The result you cite does not validate Turing's predictions about biology at all. It is synthetic biology: it validates that his system *can* work. Which is hardly surprising, since it is math.

As for actual biology, there is a [recent paper](#) finding that it describes an actual biological system, but the main verdict of the past half century is that morphogenesis doesn't work that way. I learned about this from [Shalizi](#).

[Hide](#) [↑](#)[Hide](#) [↑](#)

**Princess\_Stargirl** says:

September 1, 2014 at 2:08 am

I am not sure what you can take away from the madman thought experiment without considering the case that he does not roll snake eyes and  $10^{(n+1)}$  is larger than the total number of people on Earth (or that he can access).

In fact this case is by far the most common for "reasonable" populations of earth. And its always a possibility for any finite population. Say there were 10 billion people on on earth. Thats is  $10^{10}$ .

$(35/36)^{10}$  is around .7545.

So if Earth had 10 billion people 75% of the time we run into the undetermined case.

How realistic is it to really suppose infinite people to kill? This all reminds me of this:

[http://en.wikipedia.org/wiki/St.\\_Petersburg\\_paradox](http://en.wikipedia.org/wiki/St._Petersburg_paradox) . If you consider a finite bound the weirdness goes away.

For an infinite population note  $\text{SUM } (35/36)^n * 10^{(n+1)}$  does not converge. So its not clear how the number killed is well defined for an infinite population either.

note: I didn't read the book

**[syllogism](#)** says:

September 1, 2014 at 2:08 pm

Can't we patch this by setting an extra condition: "If all the world's population is in the room, all are killed"?

We just need a high gradient between the inside-view and the anthropic view to make the paradox function.

[Hide](#) [↑](#)[Hide](#) [↑](#)

**RCF** says:

September 1, 2014 at 3:31 pm

But then the answer to the question of "Given that you're in the room, how likely are you to die?" becomes "Depends. If everyone in the world is in the room, 1. Otherwise, 1/36." And the most likely way of getting in the room is if everyone is in the world is in the room, so the probability 1 case will dominate. So what's the paradox?

[Hide](#) [↑](#)[Hide](#) [↑](#)

**Luke Somers** says:

September 1, 2014 at 3:51 pm

Syllogism, doing that has the defect that it brings the 'inside view' into alignment with the anthropic view.

ETA: what RCF said.

[Hide](#) [↑](#)[Hide](#) [↑](#)

**Princess\_Stargirl** says:

September 1, 2014 at 5:40 pm

I think its still somewhat unclear?

Case1:

Once you leave the room, you can never get put back into the room.

Then the probability is 1/36 is if there is a possibility of another round. And of course you are just dead if you are in the last group of people.

Case2: Each new room is a random drawing. So when he randomly picks 100 people there

is a chance some of the 10 who got out will be back in.

Here the probability is not so simple. But for "reasonable" values of  $n$  the most likely event is everyone dies. So your probability of dying is always very high, regardless of how many people are in the room with you.

I am simplifying here by assumign the numbers work out evenly. So in case 2  $n = 10^k$  for some  $k$ .

[Hide](#) [↑](#)[Hide](#) [↑](#)

**Scott Alexander** says:

September 3, 2014 at 8:55 pm

Good job bringing up the St. Petersburg paradox. I actually thought the same thing when I read it. On the other hand, this is equivalent to the Doomsday Argument, and I don't see how St. Petersburg gets us out of that one.

[Hide](#) [↑](#)[Hide](#) [↑](#)

**Icicle** says:

September 1, 2014 at 2:09 am

This comment is too small to contain the math, but the snake eyes madman anthropic thing only works with an infinite population. For all finite populations,  $P(\text{dying}|\text{captured})$  is  $1/36$ .

[Hide](#)[Hide](#)

**Anonymous** says:

September 1, 2014 at 11:36 pm

I'd be interested in (a link to?) further detail on this.

[Hide](#) [↑](#)[Hide](#) [↑](#)

**Icicle** says:

September 3, 2014 at 4:36 pm

Omitting the mathy part (since I don't know how to do LATEX), you can think of it like this. The probability of being in the room on the first flip is miniscule. Now, the odds of the guy making it to the second flip are slightly lower, but that is counterbalanced by a large increase in your odds of being included in the second flip room.  $P(\text{Makes it to } N \text{ flips})$  drops slowly, but  $(\text{People in room on } N\text{th flip})$  rises quite a bit faster, so the combined probability distribution over which flip he is on is dominated by the last few flips. If we stipulate that he lets everyone go alive if he makes it to the end of the population and still doesn't get a snake eyes, the odds of dying are  $1/36$ . If we stipulate that he kills everyone

if he makes it to the end of the population, you are more likely to die than that.

Pretty much, if you end up in the room, you can conclude you are fairly likely to be on the final coin flip that composes all of the population (It composes most of your probability distribution) and your probability of dying is dominated by what the murderer is going to do in the ending scenario.

[Hide](#) [↑](#) [Hide](#) [↑](#)

**RCF** says:

September 3, 2014 at 6:14 pm

Well, one rule in probability is if  $P(A|B) = x$  for all B, then  $P(A) = x$ . So it follows that here,  $P = 1/36$ .

[Hide](#) [↑](#) [Hide](#) [↑](#)

**Said Achmiz** says:

September 1, 2014 at 2:24 am

The quoted reply to the Chinese Room (specifically, the idea that Searle is using an intuition pump and handwaving away the matter of scale) is in fact almost exactly right out of Dennett & Hofstadter. (Their *The Mind's I* was published in 1981, and the fact that Searle has never given a satisfactory reply, but continues to insist that he has, makes the Chinese Room all the more infuriating to still see cited as relevant philosophy.)

[Hide](#)[Hide](#)

**RCF** says:

September 1, 2014 at 3:32 am

I don't see how it eliminates the value of the thought experiment. Just a few top level posts ago, SA linked to a post about a thought experiment involving  $10^{40}$  variations on a person, so complaining about a few trillions pieces of paper seems a bit hypocritical. If you find that the quantitative aspect of the Chinese Room is a deal breaker, then you're saying that there is some boundary such that consciousness below that is notable, and consciousness above that is not. Revealing that about your beliefs is in itself a useful purpose of the CR.

[Hide](#) [↑](#) [Hide](#) [↑](#)

**Said Achmiz** says:

September 1, 2014 at 1:26 pm

I'm not quite sure what claim or argument you think you're responding to, but I don't think it's any I've made. To elaborate:

Searle says: "the Room would not act as if it understood Chinese. But look! What's inside the

Searle says: "the Room would *act* as if it understood Chinese. But look! What's inside the Room is trivial and unimpressive: it's just a guy shuffling some papers around, mindlessly following instructions. Clearly, nothing here is truly understanding anything!"

But Searle is swindling us, right under our nose. First he gives us a Room that really acts as if it understood Chinese, then, at the moment we peek inside, he performs a sleight of hand, and switches it out for a Room which is just an ordinary guy and a mundane amount of paperwork.

*But these cannot be the same Room!* To truly understand Chinese, the Room must be colossal, titanic, unimaginably huge — as Aaronson says, and as Dennett and Hofstadter said before him — and the "man" inside must be Superman, or perhaps Dr. Manhattan, capable of traveling at the speed of light and manipulating paper with such speed that any real paper, so handled, would instantly incinerate from the quickness of his manipulations.

The counterargument thus is this: if we were shown the *real* Room — if Searle did not hide from us the true immensity of what must be required to implement apparent understanding — then it would look substantially more impressive; and our intuitions of "bah, nothing here can be conscious" would be substantially weakened.

[Hide](#) ↑ [Hide](#) ↑

**RCF** says:

September 1, 2014 at 3:41 pm

"I'm not quite sure what claim or argument you think you're responding to, but I don't think it's any I've made."

You seem to be endorsing the view of Scott Alexander et al., and I am responding to that.

"But Searle is swindling us, right under our nose."

I don't think it's "swindling" to have a thought experiment that would fail if someone were to try to do it in practice. That's almost the definition of "thought experiment".

"The counterargument thus is this: if we were shown the real Room — if Searle did not hide from us the true immensity of what must be required to implement apparent understanding — then it would look substantially more impressive; and our intuitions of "bah, nothing here can be conscious" would be substantially weakened."

I think that I've already responded to this. If you're saying that you have an intuition that the "small room with one guy shuffling a few papers around" can't understand Chinese, but the larger the room is, the more that intuition weakens, then that is a significant revelation, and the CR has value in eliciting that revelation. Is there a point at which adding one more piece of paper makes you go from "Bah, that's not big enough" to "Yeah, that's capable of being conscious"?

[Hide](#) ↑ [Hide](#) ↑

**Luke Somers** says:

September 1, 2014 at 3:54 pm

It's swindling if the whole point of the example is an aspect of the scenario you've changed from what it's an analogy for.

[Hide](#) ↑ [Hide](#) ↑

**Alexander Stanislaw** says:

September 2, 2014 at 8:48 am

*thought experiment that would fail if someone were to try to do it in practice*

No it fails in theory – there is no hypothetical man shuffling papers that could display behavior that resembles understanding.

*Is there a point at which adding one more piece of paper makes you go from "Bah, that's not big enough" to "Yeah, that's capable of being conscious"?*

Is there a point at which adding a single hair makes you go from "Bah, that's not a beard" to "Yep, that's definitely a beard"? And it's not size – it's computational complexity.

[Hide](#) ↑ [Hide](#) ↑

**RCF** says:

September 2, 2014 at 3:35 pm

@Alexander Stanislaw

"No it fails in theory – there is no hypothetical man shuffling papers that could display behavior that resembles understanding."

Maybe not in real time, but "man shuffling papers" is just as much a Turing Machine as a desktop computer is, so in theory, given enough time, a man could display behavior that resembles understanding.

"Is there a point at which adding a single hair makes you go from "Bah, that's not a beard" to "Yep, that's definitely a beard"? "

Is "consciousness" in the same category as "beard"? Most people think of consciousness as being a binary state. If your resolution to the CR is that "consciousness" is simply a continuum, then the CR succeeds in challenging conventional wisdom.

[Hide ↑](#)[Hide ↑](#)

**Alexander Stanislav** says:

September 2, 2014 at 11:10 pm

Wait a minute, this thought experiment isn't about consciousness at all – its about intelligence/understanding which are pretty clearly continuums.

Also, even given eternity, one man could not pull off the trick. You could make him immortal; self sustaining; able to produce ink and paper by himself; able to navigate a planet sized library and modify it with perfect fidelity over a period of time much greater than the age of the universe. At which point you would hard pressed to call this being a "man".

[Hide ↑](#)[Hide ↑](#)

**Ialdabaoth** says:

September 3, 2014 at 4:42 pm

Most people think of consciousness as being a binary state.

They are wrong.

[Hide ↑](#)[Hide ↑](#)

**RCF** says:

September 3, 2014 at 8:13 pm

@Ialdabaoth

Most people think of consciousness as being a

binary state.

*They are wrong.*

*Then if the CR challenges that belief, it has value. If you thought that they I presented the prevailing wisdom as evidence for the claim, you misinterpreted my post. I presented the prevailing wisdom as evidence that the CR has value.*

[Hide](#) ↑ [Hide](#) ↑

**nydwracu** says:

September 1, 2014 at 4:12 am

I've never been impressed with Searle.

I took a philosophy of mind course once, and the TA liked Searle. One of the arguments that he gave (and, IIRC, attributed to Searle) against the possibility of strong AI is the hypothetical of a strong-AI-like machine built out of beer cans and strings, or something along those lines.

The Chinese Room is a useful thought experiment for generating intuitions of the possibility of emergent intelligences — it's where [institutional intelligences](#) come from.

Given the number of Central Europeans who tried to figure out the same concept and ended up talking about ghosts, it seems to be pretty hard to figure out. So there's still some value in the CR.

[Hide](#) ↑ [Hide](#) ↑

**Army1987** says:

September 1, 2014 at 10:15 am

Just curious... Why do you keep switching between nydwracu.wordpress.com and nithgrim.wordpress.com?

[Hide](#) ↑ [Hide](#) ↑

**nydwracu** says:

September 2, 2014 at 2:30 am

Different types of post. nithgrim.wordpress.com is for quotes, short things, and history; nydwracu.wordpress.com is for long things and theory.

[Hide](#) ↑ [Hide](#) ↑



**Anonymous** says:

September 2, 2014 at 2:56 am

Lots of people have multiple blogs. I think Army's question is: how do you choose which one to link when you leave comments?  
(the answer is probably that you filled it in differently on multiple computers for no good reason and aren't even aware that you're leaving variations)

[Hide](#) ↑ [Hide](#) ↑

**nydwracu** says:

September 2, 2014 at 7:20 am

Hm. I think I set one to link to nithgrim for a while so people would be aware that it exists, but both of my computers are set to link to nydwracu now.

[Hide](#) ↑ [Hide](#) ↑

**MugaSofer** says:

September 1, 2014 at 10:31 am

I see a lot of replies to the Chinese Room – well, OK, I don't see that many replies to the Chinese Room, but of the ones that I do see, almost all of of them at least mention that the "room" would have to be absurdly large.

[Hide](#) ↑ [Hide](#) ↑

**MartinW** says:

September 1, 2014 at 10:58 am

Another way in which Searle uses an intuition pump to confuse the issue, is by putting himself into the thought experiment as the person doing the paper-shuffling, describing the whole setup from that person's point of view, and then pointing out that he does not know Chinese. But of course, the fact that the paper-shuffling is done by a conscious, intelligent human being, is basically irrelevant, because the human in the story isn't doing anything which could not be easily replaced by a dumb machine.

It's like how, when you are talking to a telemarketer who is under strict orders to follow a predetermined script, you aren't really talking to the person on the phone; you are having a conversation with the script itself, with the telemarketer providing nothing but the voice.

There's a nice part in *The Mind's I* where Hofstadter points out the different "dials" on the thought experiment which you can turn, to encourage different intuitions in the reader about whether the room is conscious: large or small room, conscious or non-conscious operator, running the simulated

mind in real-time or glacially slow, etc. Depending on how you turn those dials, you can make the fact that “the room” is the real conscious entity doing the communicating, either blindingly obvious or very well-hidden.

[Hide](#) ↑ [Hide](#) ↑

**Said Achmiz** says:

September 1, 2014 at 1:17 pm

*But of course, the fact that the paper-shuffling is done by a conscious, intelligent human being, is basically irrelevant, because the human in the story isn't doing anything which could not be easily replaced by a dumb machine.*

Yes indeed. This is the so-called Systems Reply, which basically is (with some elaboration) the accepted (among the AI/CS community) reply to Searle. Searle describes this Reply in his paper, then proceeds to make some manifestly weak, handwave-y counterarguments, and leaves it at that.

In short, the Chinese Room comes pre-refuted. Everything since then — starting with the rest of the original paper itself! — is basically Searle being bullheaded, and insisting that damn it all, his intuitions on the matter are strong enough to carry the argument.

[Hide](#) ↑ [Hide](#) ↑

**[peterdjones](#)** says:

September 3, 2014 at 6:10 am

The systems reply would be stronger if it could be phrased in term of unmagical emergence – “here is how you build meaning out of meaninglessness” – rather than magical emergence – “it’s done at the systems level, somehow.”

Magical emergence is as handwavey as anything Scarlet says,

Note also that if your answer to “here is how to do it” involves embodiedness, hardware specifics, interactions with an environment, etc, you are at least halfway agreeing with him, since he thinks brains archive semantics through “causal powers”.

[Hide](#) ↑ [Hide](#) ↑

**Eli** says:

September 2, 2014 at 4:23 am

It’s definitely a very problematic intuition pump. What with my having been raised on science fiction and computer-science textbooks, I simply never possessed the intuition that the Chinese Room

couldn't be conscious or intelligent in the first place.

[Hide ↑](#)[Hide ↑](#)

**peterdjones** says:

September 2, 2014 at 1:07 pm

Why should Searle reply? As far he is concerned, the amount of genuine semantic understanding in the room is exactly zero, and he hardly needs to point out that  $1000 * 0 = 0$  or  $1000000 * 0 = 0$ . Scale is irrelevant to him.

His argument is not that the room does not understand Chinese because as a result of being too simple...in fact, he assumes off the bat that the room displays competence in Chinese to an outside observer. His claim is that that would be a kind of parrotting, not indicative of genuine understanding.

His reasons for asserting that are that

A) the individual operations performed by the operator are meaningless

B) no mysterious emergence kicks in at the systems level to supply genuine semantics.

Since Searle never really claims that the room is small scale, the scale response has to address B, since A is uncontroversial. And the idea of emergence in the context of computation is problematic.

[Hide ↑](#)[Hide ↑](#)

**Tom Womack** says:

September 3, 2014 at 5:34 am

"the idea of emergence in the context of computation is problematic"

All of computation is emergence! There's nothing but emergent behaviour there! I spend my days at work arranging lots of collections of lots of different kinds of little state machines, as an emergent consequence of which AArch64 instructions are executed; it's not quite as awesome as consciousness, but definitely any bit of it is wiggling around in moderately simple ways whilst the whole thing ends up playing Angry Birds.

[Hide ↑](#)[Hide ↑](#)

**peterdjones** says:

September 3, 2014 at 6:27 am

But not MYSTERIOUS emergence.

Emergence is two different things: one is the idea that systems have properties their commitments don't have.

The other is the idea that systems have properties which are unexpected and unaccountable in term of the properties and behaviours of their constituent parts.

The systems reply is likewise two things. If you hand Searle a flowchart or UML for a system that answers his objections, that is an unmysterious systems reply.

But the idea that a computer systems had inexplicable, mysteriously emergent properties us a contradiction, because computer systems are engineered out of components.

[Hide](#) [↑](#)[Hide](#) [↑](#)

**Tom Womack** says:

September 3, 2014 at 6:48 am

It depends what you mean by inexplicable mystery, and I find your confidence in 'engineered out of components' baffling.

CPU performance went beyond the point at which humans could understand it about twenty years ago when the Pentium Pro came out. The performance characteristics of a modern microprocessor can only realistically be obtained by simulation: even tech leads who have already designed three successful microprocessors can't see in advance the particular combinations of corner cases which explain why the machine accesses memory at a particular rate. If you're good you can construct slightly higher-level explanations by long staring at the machine-state traces.

This is close enough to mysterious emergence to me.

People are reasonably confident that it's possible to analyse and prove correctness, because it's a bit difficult to advertise processors when you can't have the verification lead stand up in front of the Board and say that the thing works, but thousands of man-years from Intel verification engineers still produced a processor whose errata IDs have got up to HSD-131.

[Hide](#) [↑](#)[Hide](#) [↑](#)

**peterdjones** says:

September 3, 2014 at 7:57 am

September 3, 2014 at 7:57 am

@Tom

The existence of in-practice mystery is compatible with the non existence of in principle mystery.

Microprocessors aren't analogous to the original problem. If you measure a microprocessors preference, it is what you measured it to be. However, the problem addressed by MB&P is whether there is some genuine semantics that can't be measured behind the linguistic competence that can be.

If you can't aprioritistically demonstrate that a computer system should have genuine semantics, because it us constructed to, and you also can't demonstrate it empirically...it looks like you can't demonstrate it.

[Hide](#) ↑ [Hide](#) ↑

**Doug S.** says:

September 3, 2014 at 9:12 pm

You play Angry Birds at work? 🙄

[Hide](#) ↑ [Hide](#) ↑

**Luke Somers** says:

September 4, 2014 at 10:17 am

Sounds more like he programs it, actually. This is unlikely but not wholly implausible.

[Hide](#) ↑ [Hide](#) ↑

**Anonymous** says:

September 4, 2014 at 11:43 am

sounds like he designs ARM chips. pretty clear from the follow-up comment.

[Hide](#) ↑ [Hide](#) ↑

**grendelkhan** says:

September 3, 2014 at 6:41 pm

From *The Mind's I*, "[A Conversation With Einstein's Brain](#)" really helped me see the Chinese Room thought experiment as silly.

Searle proves too much; according to him, if you build the structure of a mind out of index cards and mail slots, it's not conscious and lacks "intentionality", or "causal powers", or some other synonym for "magic". But by this very same reasoning, if you build the same structure out of ATP and fatty acids and neurotransmitters, it's *still* not conscious by the very same argument. (Do the *molecules* understand Chinese?) The argument shows that either the clerk-and-files system is conscious, or that Searle himself is not.

It's like nobody over there ever heard of reductionism, or if they have, they believe that brains are special and you can't apply reductionism to them, because reasons. What kind of a wacky field *is* this?

I'd expect this sort of thing from Ed Feser, but you pretty much have to *be* Ed Feser in order to coherently believe that the Chinese Room thought experiment shows what Searle thinks it does.

[Hide](#) ↑ [Hide](#) ↑

**[peterdjones](#)** says:

September 3, 2014 at 7:38 pm

WeatherYou or the are ,misunderstanding.

Different substances do, uncomtroversially, have differ causal powers.That's why we don't make teapots out of chocolate.

The .CRA isn't an argument against being able to build artificial brains...not that building an exact duplicate ofna brain would prove much.

It's aimed at the idea that the mind is software. The fact that brains work, somehow, doesn't tell us that they work as Turing machines running programmes and so doesn't refute the CRA.

What we don't know is whether consciousness is like addition or like the

[Hide](#) ↑ [Hide](#) ↑

**[peterdjones](#)** says:

September 4, 2014 at 8:09 am

.....whether consciousness more like the weather, or is it more like multiplication....as one Scott Aaronson put it.

[Hide](#) ↑ [Hide](#) ↑

**Luke Somers** says:

September 4, 2014 at 10:27 am

You can in principle simulate the weather accurately enough that any pattern that predictably persists in the weather will persist in the simulation.

The only effect that fights against this – unbridled chaos – also keeps the weather from being able to maintain anything resembling cognition. I don't see how failure to emulate a feature of non-consciousness is a weakness of the computational approach.

[Hide](#) [↑](#)[Hide](#) [↑](#)

**Ialdabaoth** says:

September 4, 2014 at 10:50 am

Wild speculation moment:

If the weather WAS conscious, what would look different?

[Hide](#) [↑](#)[Hide](#) [↑](#)

**peterdjones** says:

September 4, 2014 at 12:03 pm

That isn't the point. The point was that simulated weather doesn't get you wet. Simulated multiplication is multiplication, simulated weather isn't weather.

[Hide](#) [↑](#)[Hide](#) [↑](#)

**Ialdabaoth** says:

September 4, 2014 at 12:20 pm

Simulated multiplication doesn't get me apples, either.

ALL mathematics is simulation. Reality is that if I follow the procedure "for each apple in my basket, take out an apple from the barrel and put it in my basket", then my basket now has twice as many apples.

If I run a program with a line that says

```
apples = apples * 2;
```

I don't gain any apples, even if the value of `apples` is non-zero.

[Hide ↑](#)[Hide ↓](#)

**AndekN** says:

September 4, 2014 at 3:13 pm

*If the weather WAS conscious, what would look different?*

Peter Watts wrote a short story back in 1993 about this scenario (PDF warning):

[http://www.rifters.com/real/shorts/PeterWatts\\_Nimbus.pdf](http://www.rifters.com/real/shorts/PeterWatts_Nimbus.pdf)

[Hide ↑](#)[Hide ↓](#)

**[peterdjones](#)** says:

September 4, 2014 at 4:01 pm

@Ialdabaoth

if I compute an inverse cube law gravity, what is that .simulating?

Does simulating chess fail to real chess?

[Hide ↑](#)[Hide ↓](#)

**Ialdabaoth** says:

September 4, 2014 at 4:47 pm

peterdjones: I feel like you and I are playing very different games with words. I do not know how to proceed.

[Hide ↑](#)[Hide ↓](#)

**ADifferentAnonymous** says:

September 4, 2014 at 5:46 pm

Let me see if I can untangle some things. I *\*think\** an adequate definition of 'simulation of X' is 'a process isomorphic to X' (plus some connotations to the rough effect that this isomorphism is a salient property). A process isomorphic to Earth weather won't get you wet, and a process isomorphic to doubling your apples won't get you more apples.

Whether a process isomorphic to conscious thought is conscious is what we're debating, right?

[Hide ↑](#)[Hide ↓](#)



**Luke Somers** says:

September 4, 2014 at 6:03 pm

> Simulated multiplication is multiplication, simulated weather isn't weather.

Ah, I see now. That wasn't exactly clear. Well. What is it about bio-brains that make it 'wet', and the same thing happening in silicon not 'wet'?

If Hortas or Chenjesu (yes, fictional; take them as hypothetical evolved silicon-based sapients) were to realize they could simulate their semiconductor brains with organic chemical-based neurons, their following this line of reasoning would lead them to object that there's no current and so it couldn't possibly be conscious.

[Hide](#) ↑ [Hide](#) ↑

**Murphy** says:

September 8, 2014 at 1:47 pm

I always read it as even simpler than that, you've got this enormous information structure and then you get your reader to imagine themselves inside it. Then you say, "look, I'm inside this room, I can see everything that's going on and I don't see any intelligence happening. I just see paper"

Which could of course be swapped for "the brain room" where someone has to run around carrying neurotransmitters between gigantic neurons according to some rules of chemical interactions who looks around and says "look, I'm inside this room, I can see everything that's going on and I don't see any intelligence happening. I just see blobs".

Or the luminous room, it's obvious to the person inside that the room isn't luminous, intuitively it's obvious that the room isn't luminous but from the outside (time distortion effects allowing just as with the chinese room) you can get a tan from it.

>Churchland's luminous room

[87] "Consider a dark room containing a man holding a bar magnet or charged object. If the man pumps the magnet up and down, then, according to Maxwell's theory of artificial luminance (AL), it will initiate a spreading circle of electromagnetic waves and will thus be luminous. But as all of us who have toyed with magnets or charged balls well know, their forces (or any other forces for that matter), even when set in motion produce no luminance at all. It is inconceivable that you might constitute real luminance just by moving forces around!"[73] The problem is that he would have to wave the magnet up and down something like 450 trillion times per second in order to see

anything.

[Hide ↑](#)[Hide ↑](#)

**Creutzer** says:

September 8, 2014 at 2:21 pm

I'm a bit surprised that nobody here has mentioned another problem with the Chinese room. The CR takes only Chinese characters as inputs. It has no sense of the fact that language is about reality. It cannot look at a rose and say (more importantly, *judge*) "Yep, that this is really red" on the basis of its observations. I think this contributes a lot to the intuition that the room doesn't understand Chinese, and it's also why the CR is irrelevant to strong AI because strong AI is not a chatbot.

[Hide ↑](#)[Hide ↑](#)

**Rachael** says:

September 1, 2014 at 2:25 am

Do people ever confuse you with Scott \*Adams\*?

IMO, his blogging style is often fairly similar to yours, and there's a reasonable overlap in topics : free will, futurology, and stories of unpleasant run-ins with feminists.

[Hide](#)[Hide](#)

**Douglas Knight** says:

September 1, 2014 at 2:44 am

Could you elaborate on style? It seems to me that Aaronson and Adams have similar styles, far from Alexander.

[Hide ↑](#)[Hide ↑](#)

**Anonymous** says:

September 1, 2014 at 4:40 am

I'm not Rachel, but the main similarity I see between Alexander's and Adams's blogs is the impressively constant stream of... let's say "speculative" ideas.

Aaronson is in another league entirely, a world-class expert who mostly writes about his areas of expertise. I don't really get the comparisons. I guess he did recently post that speculative talk, but that was still probably better grounded than any non-science blogging I've seen in a while.

[Hide ↑](#)[Hide ↑](#)

**Anonymous** says:

September 1, 2014 at 5:28 pm

Scott adams wishes his blog were this good. I used to read his a lot, but now its mostly cringe inducing and I can get much higher intellectual levels of ideas from ssc.

[Hide](#) ↑ [Hide](#) ↑

**Vulture** says:

September 3, 2014 at 12:54 am

Adams has some pretty good stuff – I liked his observations about “pivoting” and startup culture, for example – but for the most part he does come across as a decidedly weak-ass poor man’s SSC. I think it’s mostly because he presents interesting ideas, but then never really elaborates on them or tries to justify them, which gets pretty unsatisfying pretty fast.

[Hide](#) ↑ [Hide](#) ↑

**Sniffnoy** says:

September 1, 2014 at 2:40 am

Pretty sure you’ve seen the [anti-induction joke](#) before. 😊

[Hide](#)[Hide](#)

**RCF** says:

September 1, 2014 at 3:26 am

Given that I’ve heard it before, I was surprised that I saw it again.

[Hide](#) ↑ [Hide](#) ↑

**Scott Alexander** says:

September 1, 2014 at 12:07 pm

+1

[Hide](#) ↑ [Hide](#) ↑

**Army1987** says:

September 1, 2014 at 3:19 am

Have you ever tried to read *GEB* by Hofstadter?

[Hide](#)[Hide](#)

**Imm** says:

September 1, 2014 at 4:11 am

Don't do it. It's full of bad math. You need to know enough math to be sceptical first.

[Hide](#) ↑ [Hide](#) ↑

**Taymon A. Beal** says:

September 1, 2014 at 5:04 am

Which part of the math is bad, exactly?

[Hide](#) ↑ [Hide](#) ↑

**switchnode** says:

September 1, 2014 at 11:13 am

I am very curious about this.

[Hide](#) ↑ [Hide](#) ↑

**Vulture** says:

September 1, 2014 at 12:10 pm

Yes, please do tell

[Hide](#) ↑ [Hide](#) ↑

**Luke Somers** says:

September 1, 2014 at 3:59 pm

The math was so bad that after reading it, I didn't know how to say 'N is a power of 10'.



[Hide](#) ↑ [Hide](#) ↑

**Vulture** says:

September 3, 2014 at 12:56 am

I assume that this is a joke that I'm not getting, but I'd like to interrogate it further only because I can't tell whether or not it's answering the question...

[Hide](#) ↑ [Hide](#) ↑

**g** says:

September 3, 2014 at 5:14 am

Vulture, GEB devotees quite a bit of space to discussion of a simple formal language for writing statements about arithmetic (in order to

formal language for writing statements about arithmetic (in order to discuss Gödel's theorem) and IIRC at one point the reader is challenged to express "n is a power of 10" in this language. The language is quite stripped down and it's by no means a trivial exercise.

[Hide](#) ↑ [Hide](#) ↑

**Luke Somers** says:

September 3, 2014 at 12:00 pm

Exactly. From time to time, I wonder about avenues of attack on making such an assertion.

[Hide](#) ↑ [Hide](#) ↑

**Doug S.** says:

September 3, 2014 at 9:16 pm

Apparently the solution involves using the [Chinese Remainder Theorem](#) and is really, really long.

[Hide](#) ↑ [Hide](#) ↑

**Taymon A. Beal** says:

September 3, 2014 at 9:54 pm

This isn't quite true. You need the Chinese remainder theorem to prove Gödel's incompleteness theorems, but you don't need it to solve that one puzzle.

I had actually used the Chinese remainder theorem to solve the puzzle, and then written up a >500-word comment explaining the cool parts of Gödel numbering that Hofstadter skipped over in *GEB*. Then I thought, "You know, before I post this comment claiming that you need Gödel numbering to solve the puzzle, I should use Google to find out if that's actually true." And it turned out not to be.

[Hide](#) ↑ [Hide](#) ↑

**Scott Alexander** says:

September 1, 2014 at 7:15 pm

Yes, and enjoyed it a lot.

[Hide](#) ↑ [Hide](#) ↑

**RCF** says:

September 1, 2014 at 3:37 am

*Now this kind of talk drives AI researchers up a wall. They say: if you told someone in the 1960s that in 30 years we'd be able to beat the world grandmaster at chess, and asked if that would count as AI, they'd say of course it's AI. But now that we know how to do it, it's no longer AI – it's just search."*

This sort of argument is problematic. It's basically making an accusation of moving the goal posts, except that the "original" goal posts weren't set by the person currently being debated, and in fact exist only counterfactually. "50 years ago, people would have considered this AI" isn't completely without persuasive force, but it's hardly a solid argument. Whether something is AI should be argued on the basis of whether it's AI, not on whether hypothetical people in the past would have considered it AI.

[Hide](#)[Hide](#)

**Imm** says:

September 1, 2014 at 4:13 am

Compare "god of the gaps". Or try to taboo "intelligence".

[Hide](#) [↑](#) [Hide](#) [↑](#)

**MugaSofer** says:

September 1, 2014 at 10:22 am

It also fails to take into account the fact that, originally, experts assumed the difficulty of programming a computer to do something was correlated with how much "intelligence" we associate with the task.

So playing chess was expected to be harder than, say, recognizing basic objects.

The hardest part of playing Jeopardy was assumed to be working out the answers – I think it would have been classified as "abstract symbol manipulation" or something – whereas the reason it's considered impressive *now* is that it involves parsing the questions; which is much harder than it sounded back then, but not hard-problem-of-consciousness hard.

[Hide](#) [↑](#) [Hide](#) [↑](#)

**Erik** says:

September 2, 2014 at 8:45 am

I've seen something similar come up in discussions of the Turing Test specifically – hearing about a computer that can pass for human on a text channel, and when I dig into the details, I find this meant that a computer successfully pretended for half an hour on a time-delayed channel to be a foreign-primary-language 14-year-old. And by "successfully" it turns out that the computer convinced one of three interrogators.

Sometimes the goal posts *should* be moved because they're *in the wrong place*.

[Hide ↑](#)[Hide ↑](#)

**Taymon A. Beal** says:

September 2, 2014 at 9:03 am

It's generally agreed that Turing's original paper put the goal posts somewhere in the vicinity of the right place. Which does not prevent people from claiming to have passed the Turing test when they've done nothing remotely close.

For anyone who hasn't seen it, Aaronson's (scathing) analysis of the Eugene Goostman story is [here](#).

[Hide ↑](#)[Hide ↑](#)

**RCF** says:

September 2, 2014 at 4:46 pm

One Turing Test I came across was the Loebner Prize, for which it's considered against the spirit of the competition for the testers to ask the testees questions specifically aimed at seeing whether they're human. They're given a particular topic to discuss, and they're supposed to ask only questions that would "naturally" come up in a discussion of that topic. I'm not sure whether that's moving the goal posts, or simply putting up intermediate goals so that competitors actually have a realistic goal to work towards.

And unfortunately, the Turing Test is a depressingly poor distinguisher between humans and computers. One time when I contacted tech support, they failed the Turing Test, where the test consisted of me asking "Are you a human?"

[Hide ↑](#)[Hide ↑](#)

**Shmi Nux** says:

September 1, 2014 at 3:40 am

> He already knows lots of complicated stuff about computers, quantum physics, set theory, and philosophy. Part of me wants to get angry: WHY IS ONE PERSON ALLOWED TO BE SO SMART?

Scott, come on, I hope you are not serious. I have exact same reaction when I read your posts (and the other Scott A's too). I do have a PhD in physics and understand QM (but not Quantum Computing), relativity and such reasonably well. Definitely well enough to teach it. However, I wish I could come up with 1/10 of insights you do, write 1/10 as eloquently as you do, and be 1/10 as productive. Maybe I suck at distinguishing various levels "above mine", but you and Aaronson seem to be at about the same level, only he already has publications, tenure and prestige, and you are only part-way there.

[Hide](#)[Hide](#)

**nydwracu** says:

September 1, 2014 at 4:14 am

Having the ability to consistently write long and well-worded things is a goddamn *accomplishment*. When Theden launched, we decided to launch with seven or eight articles of ~1500 words each, so I had to write four. It took me *weeks*.

[Hide](#) [↑](#) [Hide](#) [↓](#)

**Eli** says:

September 2, 2014 at 4:26 am

Scott's writing productivity is indeed pretty fucking good. I wish I could match it when I'm actually writing out real ideas that require context and explanation, rather than just spewing things out without context at length.

[Hide](#) [↑](#) [Hide](#) [↓](#)

**Taymon A. Beal** says:

September 1, 2014 at 3:43 am

Scott Aaronson is the closest thing computer science has to a Neil deGrasse Tyson-type figure. That this is true despite his only being well-known in geek communities like this one (aside from within his own field, of course) is a testament to the frustrating absence of popularized computer science in the public consciousness.

Incidentally, the first proof of the undecidability of the Halting Problem that I was able to actually understand was [this famous one](#) written in the style of Dr. Seuss.

[Hide](#)[Hide](#)

**AlexC** says:

September 2, 2014 at 9:45 am

The Dr Seuss-style Halting Problem explanation wasn't the first I saw, but it's definitely the most memorable.

[Hide](#) [↑](#) [Hide](#) [↓](#)

**Army1987** says:

September 2, 2014 at 11:19 am

Geoff Pullum is awesome and I have no idea why I stopped reading Language Log.

[Hide](#) [↑](#) [Hide](#) [↓](#)



**Qiaochu Yuan** says:

September 1, 2014 at 5:32 am

> I hear a bunch of people telling me Bayesianism isn't everything, it's the only thing

Is it a bunch of people independently telling you this, or is it a bunch of people repeating to you what they heard on LessWrong? I suspect the people telling you the other thing are more independent samples, and they also probably include more actual domain experts in the relevant fields (e.g. Jacob Steinhardt talks about this a lot, and studying stuff like inference algorithms is literally his day job).

[Hide](#)[Hide](#)

**Luke Somers** says:

September 1, 2014 at 4:10 pm

Even the folks on LW will tell you that directly applying Bayes is impractical in many cases, so use the toolbox! And the domain experts tend to agree that in principle you can get (most?) frequentist tools from a Bayesian approach, if you're so inclined.

That leaves a hair's breadth of actual disagreement.

[Hide](#) [↑](#)[Hide](#) [↑](#)

**Vulture** says:

September 1, 2014 at 8:05 pm

If it is as you say, it sounds like pretty much just a difference of emphasis.

[Hide](#) [↑](#)[Hide](#) [↑](#)

**Luke Somers** says:

September 2, 2014 at 3:11 pm

Though there is a difference of emphasis, the more significant part is a disagreement on which sorts of background assumptions to make.

[Hide](#) [↑](#)[Hide](#) [↑](#)

**MartinW** says:

September 1, 2014 at 6:40 am

The "feed P back into in itself" approach is a neat trick, but it has always felt a bit unsatisfactory to me. I realize it's a valid proof, but somehow it feels like cheating. After all, I would be perfectly happy with a program which could predict if any program *other than itself* will ever halt. But it doesn't seem likely that such a program exists either, despite the fact that the feed-P-into-P trick does not disprove its existence.

The following is, to me, a more intuitive way to understand why a generic solution to the halting problem is

unlikely to exist, even though it isn't an absolute proof:

If you had a generic halting problem solver, you could use it to determine the truth of any theorem for which it would be possible to code a brute-force search. E.g. if you wanted to know if Goldbach's Conjecture is true, you'd just write a little program which tests, for every integer, every possible way to write it as the sum of two primes. Then you ask P if that program will ever halt. If it won't, then apparently there does not exist a counterexample to the Conjecture, and hence it must be true.

And of course you could do the same thing with Fermat's Last Theorem, or the Riemann hypothesis, or any other problem which has stymied the world's greatest mathematicians for centuries. So a program which could solve the Halting Problem, would need to be powerful enough to prove or disprove *any* mathematical statement which could be expressed as a brute-force-search for a counterexample. No need to actually, y'know, think about the problem, or even to have a good understanding of what a prime number is — just code up a brute-force search, run it through the nice tidy generic program analyzer, and read out the answer.

It just seems unlikely that such a thing would exist — that there would be a generic way to solve a huge class of problems without ever needing to do the actual hard work of understanding the problem and coming up with a step-by-step chain of reasoning from the problem statement to the conclusion.

Now, of course, "it seems unlikely" isn't a valid proof. But it gets us back to what you describe as the theme of Scott's book, which is that the Universe appears to be arranged in such a way that certain things are simply impossible (or, at least, that there are no shortcuts around doing it "the hard way") no matter the details of which approach you take. And that, once it has been determined that X is one of those things, then "if you could do Y, it would allow you to easily do X" becomes a pretty strong argument that Y is probably impossible as well.

Oh, and I second the recommendation to read Dennett & Hofstadter's *The Mind's I* if you liked Aaronson's counterargument to the Chinese Room.

[HideHide](#)

**ShardPhoenix** says:

September 1, 2014 at 7:32 am

>It just seems unlikely that such a thing would exist — that there would be a generic way to solve a huge class of problems without ever needing to do the actual hard work of understanding the problem and coming up with a step-by-step chain of reasoning from the problem statement to the conclusion.

Who says that the Turing Machine can't do all the steps involved in understanding the problem, etc, just as a human would? What the halting problem is really saying is that some mathematical problems can't be solved at all (which you might be getting at in the second to last paragraph).

(The above requires that the universe is computable, which seems likely to me).

[Hide ↑](#)[Hide ↑](#)

**MartinW** says:

September 1, 2014 at 8:28 am

True — so it may be possible (for the “easy” cases such as solving Goldbach’s Conjecture etc, where you are unlikely to run into Gödel-style self-referential loops), but there could still be cases where you ask P to find a proof for some conjecture Q, and it churns for a few hundred years and then outputs “sorry, I give up” and you still don’t know if it failed because Q is unsolvable for Gödel-ish reasons or just because the program wasn’t clever enough to make the creative leap required to find the proof.

- > What the halting problem is really saying is that some
- > mathematical problems can’t be solved at all

Yes, but it does that by positing a very specific scenario in which a program is asked to analyse itself and then do something different from what the program being analysed would do. To which my intuitive reaction is “duh, that’s not fair!”

In fact, any program smart enough to recognise a brute-force search for Goldbach’s Conjecture, realise that it is testing a claim about prime numbers, and then find a proof for that claim in order to determine if the search will halt, would certainly be smart enough to recognize this trick and output “ha ha, very funny, now give me a real challenge please.”

Forcing a paradox by restricting the program to only two possible actions (halting or not halting) is a bit like the mathematical version of “have you stopped beating your wife yet”.

What’s not obvious, to me at least, is what this means for situations where a programmatic theorem prover is asked a “normal” question, such as Goldbach’s Conjecture or the Riemann Hypothesis, where we are not playing any self-referential games with the program’s own code but are simply asking an objective question about numbers, for which there ought to be a well-defined yes/no answer. Is it possible for Gödel-style undecidability problems to crop up in such situations?

[Hide ↑](#)[Hide ↑](#)

**ShardPhoenix** says:

September 1, 2014 at 8:38 am

I’m not an expert, but this Math Overflow post includes (alongside the halting problem) some examples of relatively simple, seemingly non-reflective problems

that are provably undecidable:

<http://math.stackexchange.com/questions/80745/an-example-of-an-easy-to-understand-undecidable-problem>

(Although at least some of these are related to the halting problem, but with for example [http://en.wikipedia.org/wiki/Wang\\_tile](http://en.wikipedia.org/wiki/Wang_tile) it's a lot less obvious of a case of 'just putting the program back into itself').

[Hide](#) ↑ [Hide](#) ↑

**MartinW** says:

September 1, 2014 at 10:40 am

Wang Tiles are interesting indeed; a nice example of a case where you might be given a seemingly straightforward puzzle without being aware that solving it would be equivalent to solving e.g. Fermat's Last Theorem, because the tile puzzle is just a clever encoding of the theorem.

But still, that just gets us back to "in order to be able to solve all instances of this puzzle you must be able to prove arbitrarily difficult theorems". Same with most of the other examples listed at your first link; e.g. there is no *generic* way to prove that two grammars or functions are equivalent, but for any *specific* pair of grammars, a sufficiently clever person or program may be able to find a proof that they are equivalent.

Of course you could make the grammars, functions or Wang tile-sets so huge and complex that our poor little theorem prover would have to predict the behavior of an algorithm as complex as itself, over an infinitely large set of possible inputs. I can see how that would be impossible, given any finite amount of computing power.

But that still doesn't really seem to address the question I'm interested in, which is whether a simple, well-defined question such as "is there an even integer greater than 2 which is not a sum of two primes" might be impossible to prove (even though it is clearly a well-defined question which must have a yes or no answer) for the same reason that Scott's algorithm  $P'$  cannot exist.

It would be nice to have a concrete example of a specific theorem which is provably nondecidable. But of course that cannot exist, since any false theorem can be proven false by giving a counterexample,

therefore proving that it cannot be proven false would be equivalent to proving it true. So we can only talk in terms of classes of problems for which we can demonstrate that they must contain undecidable members. But if the only way to prove that is by using the “feeding an algorithm to itself” trick, and if the algorithms involved must be absurdly large and complex because otherwise they could be defeated with much simpler inputs already, then that doesn’t really tell us much about whether any concrete statements which could be expressed in, say, a thousand words or less, might be Gödel-undecidable.

[Hide](#) ↑ [Hide](#) ↑

**Anatoly** says:

September 1, 2014 at 12:30 pm

@MartinW,

I think examples of “undecidability” you seek and “undecidability” as in the Halting Problem are really two very different things (infinity of inputs is crucial to the latter). But thinking of the undecidability you want, that is, simple “natural” unprovable and unrefutable sentences:

*It would be nice to have a concrete example of a specific theorem which is provably nondecidable.*

*That really depends on what you take your axioms to be, and what you mean by “specific”. Some common examples that might appeal to you would be:*

*– Independence of Axiom of Choice from the ZF set theory, or Continuum Hypothesis from the ZFC set theory. Is it true that CH “must be true or false”? Depends on your philosophy of math.*

*– The Hydra game (<http://math.andrej.com/2008/02/02/the-hydra-game/>): Peano Arithmetic cannot prove that it’ll always finish. It’s commonly believed that Peano Arithmetic sums up our “naive” intuition about natural numbers. The termination of the Hydra game *can* be proven from stronger axioms that deal with more than natural numbers (in set theory).*

*– By Gödel’s Second Incompleteness Theorem, Con(ZFC), which can be viewed as a statement about natural numbers that expressed the consistency of the ZFC set theory, cannot (if true) be proved using*

consistency of the ZFC set theory, cannot (if true) be proved using mathematical methods we normally employ (since they're in ZFC).

But of course that cannot exist, since any false theorem can be proven false by giving a counterexample, therefore proving that it cannot be proven false would be equivalent to proving it true.

This is only true for a particularly simple-looking theorems.

[Hide](#) [↑](#) [Hide](#) [↑](#)

**RCF** says:

September 1, 2014 at 2:42 pm

@MartinW

"It would be nice to have a concrete example of a specific theorem which is provably undecidable. But of course that cannot exist, since any false theorem can be proven false by giving a counterexample, therefore proving that it cannot be proven false would be equivalent to proving it true."

You appear to be denying the consequent.

Let

A := theorem T is undecidable

B := there is a counterexample to T

C := T is provably false

Your logic, as I understand it, is

$B \rightarrow C$

$C \rightarrow \sim A$

therefore  $A \rightarrow \sim C \rightarrow \sim B$

The statement  $A \rightarrow \sim C$  is the contrapositive of  $C \rightarrow \sim A$ , but the statement  $\sim C \rightarrow \sim B$  is the converse of  $B \rightarrow C$ , and is therefore not necessarily true. Just because every claim with a counterexample is provably false, that doesn't mean every claim that is probably false has a counterexample.

Consider the claim "every member of set X can be expressed as the difference between two members of set Y". If this is false, then there is some N that is a member of set X and cannot be expressed as the

difference between two members of set Y. But it's not necessarily provable that N cannot be expressed as the difference between two members of set Y.

[Hide](#) ↑ [Hide](#) ↑

**MartinW** says:

September 1, 2014 at 3:03 pm

@RCF: You are right, of course. (Re: falsity of a theorem does not imply existence of a counter-example.) I misremembered something which applies only in a specific case.

[Hide](#) ↑ [Hide](#) ↑

**Vilhelm S** says:

September 1, 2014 at 7:50 pm

@Anatoly:

*But of course that cannot exist, since any false theorem can be proven false by giving a counterexample, therefore proving that it cannot be proven false would be equivalent to proving it true.*

*This is only true for a particularly simple-looking theorems.*

While it's correct that it only holds for simple-looking ([□1](#)) theorems, Con(ZFC) is simple-looking in that sense. If it is false, then there exists a counterexample (a derivation of False).

So maybe a better analysis is that this shows the difference between unprovability and "relative" unprovability: we can't show outright that Con(ZFC) is not provable, but we can show it under the additional assumption that ZFC is consistent, and for all practical purposes that's good enough.

[Hide](#) ↑ [Hide](#) ↑

**Anatoly** says:

September 1, 2014 at 11:47 am

Algorithms are clever ways to pack up an infinity of data into a finite description.

Say there's an infinite (here and below "infinite" means "countably infinite"; if you don't know what that means, ignore this remark) binary string of zeroes and ones, some 010010101101111... Suppose I come up to you and I say: "I am the keeper of this infinite string of data. People come to me and ask: what is at the 13th place in this string, 0 or 1? How about 23523238th place? For any query they ask, I think for a while and then give them the correct answer. But I've been doing this for a long time and could use a vacation. Could you take over for me for a while? The job pays well."

You might reply to me: "I will think about your offer seriously the moment I understand how you propose that I do it. How are you going to teach me your job? If the string were finite, you could just write it on a (possibly very long) piece of paper, and I could keep and consult it. But it's infinite. No piece of paper will hold it. In fact, I wonder how \*you\* manage to answer people's queries truthfully in the first place!"

But what if the string of data, though infinite, actually has some pretty simple and definite structure? Like maybe it's all alternating 0101010101... to infinity, or maybe it's the binary representation of the digits of pi. Then it's easy for me to take over your job. You just write down the way for me to access any part of the string on demand – like "odd places are 0s, even places are 1s", or "[formula for calculating pi to any desired accuracy]". These descriptions are finite in length, so they \*can\* be written down and passed on. Armed with such a recipe, I can answer people's queries, and you go on vacation. These clever recipes for packing up an infinite string of data in a finite description are algorithms.

There are many many more infinite strings of data than there are finite clever recipes for specifying one particular infinite string. (this has to do with the fact that there are different orders of infinity, some larger than others). So in fact, only a tiny minuscule incomprehensibly tiny minority of all infinite strings of data could ever be "teachable", that is, specified by an algorithm. By any reasonable way of accounting just about all of them aren't "teachable". But we don't know which are and which aren't.

Say we have a problem. Here I need to be careful because there're many genuine problems (like, I dunno, "Is Riemann's hypothesis provable?") that have a very short finite description (either "yes" or "no"), but we don't know what it is. This isn't the kind of problem I'm talking about. In the context of non-computability we're talking about problems which have an infinite variety of



inputs ("queries"), each of which needs its own separate output ("answer"). Like "given this number, what is the value of pi at this place?". Or "given this description of an algorithm, does it ever halt?". So, when we have such a problem, it's usually simple to represent it as an infinite binary string of 0s and 1s. And the question then is, is this string one of the "teachable" ones, does there exist a finite clever description for it? Note: not "do we \*have\* a finite clever description for it", but "does there exist one". If we have an algorithm, that's fine, we're done. If we don't have one, perhaps we will one day? Maybe our descendants will be much smarter than us and come up with one? Algorithms are clever descriptions, and the cleverness goes up and up and up unbounded. There're certainly many clever algorithms that are much too complicated for our puny brains. Who's to say that one of them doesn't describe our problem?

Because cleverness of algorithms is seemingly unbounded, nobody's been able to \*prove\* a particular problem "unteachable" by explicitly considering possible ways to describe it. How would such an attempt look like? Imagine a naive person who learned about pi, and they say: "Clearly it's impossible to calculate more than a few digits of pi. How would you even do it? You can draw a circle in the sand, measure its length and diameter, divide... but the accuracy is very limited. Maybe you could blow up a very very accurately round piece of glass, and use a piece of string to measure its diameter very precisely... maybe you'd get 1-2 more digits this way. But that's it, you'll never improve on these". Then you teach them trigonometry, infinite series, and explain painstakingly that here's a series that sums up to pi and all we need to do to get more digits is to keep calculating. "Oh... I didn't know that was possible." But of course, that person's conviction before you showed them trig was only conviction, not \*proof\*. Nobody knows how to bar explicitly all unknown clever ways of finding out something, so nobody has been able to prove some infinite string "unteachable" this way.

The only way we have of proving that some infinite string is unteachable is to include information about all the possible algorithms inside the string. This is easier than it looks, since there's only an infinity of possible algorithms, and the string is also infinite. So our only method to outsmart all the clever algorithms, including those too clever for our brains, is to say – oh, this problem is so smart that it already knows something about any one of you. This string of data has such a \*rich structure\* that no clever algorithm will ever describe it, because it knows something about \*any\* possible algorithm that it can use to defy its attempt to accurately describe it.

This is completely transparent with the Halting Problem, because basically that's what it is – an infinite string that has dirt on every possible algorithm. But all the other problems that have been proven non-computable are inevitably like this too, it's just that their \*rich structure\* is not so obvious and we needed to work hard to uncover it. Maybe there's a problem in group theory that merely looks hard, but then someone comes along and proves that it has such a rich structure that its infinite string of data actually contains dirt on all possible algorithms, just as the halting problem; you just need to understand how to dress up algorithms as groups in a particularly clever way, so that truths about algorithms become truths about groups that your string has dirt on. All the undecidability results (in this sense of non-computability) have been like this. There's no other method of getting them.

But what this also means is that while the objection to the Halting Problem undecidability proof, that "P was forced to act on itself, even though it's sort of a black box", is understandable, it's really unavoidable. If we want to \*prove\* that some infinite string of data is unteachable – not just "be really sure", but \*prove\* – we'd better have an argument that applies to any possible finite description and shows that it can't work. Because there're descriptions far too clever for our brains, we can only hope to address every possible description by imagining that it exists as a "black box". But then the only thing we know about it is that it (ostensibly) solves the problem, and at the same time the problem contains "dirt" on it, just as it does on any possible description. Out of these two pieces of information we have to conjure a contradiction; so applying the "black box" to something we construct with the help of the "dirt" on it is just about the only thing we can hope to do. Fortunately, it works.

[Hide](#) ↑ [Hide](#) ↑

**Corey** says:

September 1, 2014 at 1:27 pm

+1... million.

[Hide](#) ↑ [Hide](#) ↑

**MartinW** says:

September 1, 2014 at 2:47 pm

*Say we have a problem. Here I need to be careful because there're many genuine problems (like, I dunno, "Is Riemann's hypothesis provable?") that have a very short finite description (either "yes" or "no"), but we don't know what it is. This isn't the kind of problem I'm talking about.*

Pity, because that's exactly the kind of problem *I* was talking about. 😊

Seriously though, thanks for the comprehensive reply — I was familiar with the basics (e.g. countable versus uncountable infinity, Kolmogorov complexity (although you didn't mention that one by name), etc) but it has helped me re-examine my own intuitions on this topic.

Just for the record, I don't deny that Turing's proof of the undecidability of the Halting Problem is legitimate. (I'm not that much of a crackpot.) I just wanted to make the point that the approach mentioned by Scott in his review of Scott's book, of saying "if we could do X then it would allow us to do Y, therefore X must be at least as hard as Y" is often a more fruitful way of analysing a particular problem, than the Gödel/Turing approach which proves that any formal system must contain some undecidable theorems but doesn't really give you any help in determining if a particular conjecture is likely to fall in that class.

In practice, any attempt to write a general theorem prover / halting-problem-decider is likely to run into problems which are a lot more prosaic than being unable to handle self-referential algorithms, long before it runs into the limits of what's undecidable in the Gödel sense.

At least, in a world where  $P \neq NP$ .

[Hide](#) ↑ [Hide](#) ↑

**Peng** says:

September 8, 2014 at 8:29 pm

*The only way we have of proving that some infinite string is unteachable is to include information about all the possible algorithms inside the string.*

While that may be the only known proof technique (I'm not sure off-hand), not all uncomputable problems are of that form. There are also [intermediate Turing degrees](#), i.e. problems that are harder than Turing-computable but easier than the halting problem.

[Hide](#) ↑ [Hide](#) ↑

**Army1987** says:

September 1, 2014 at 11:44 am

What do you mean by "other than itself"? Does adding comments count? Does changing variable names? Does adding useless operations?

[Hide ↑](#)[Hide ↑](#)

**MartinW** says:

September 1, 2014 at 12:30 pm

I meant basically any input which was not specifically designed to foil that particular halting-problem-solver.

The intuition I'm trying to get across is something like this: there doesn't seem to be much of a relation between "ordinary", interesting theorems / turing machines which do something interesting like testing Goldbach's conjecture or whatever, and the kind of specially constructed hypothetical inputs which can be proven to be undecidable in the Gödel / halting-problem sense.

For comparison, imagine if all the algorithmic problems we normally care about, such as the traveling salesman, bin packing, SAT, etc, were solvable in polynomial time, and the only NP-hard problems we knew about were some very hypothetical ones which were designed specifically to answer the question of whether non-polynomially-solvable problems existed, and which were completely unrelated to any problem you'd ever encounter if you were not a computer scientist studying that particular question.

And whenever somebody encountered a new problem and couldn't immediately find a polynomial solution, everybody would speculate "maybe we finally found a real-world case of an NP-hard problem!" but then within a few months or years, an n-squared or better solution would be found.

That's a bit how I feel about Gödel-undecidable theorems. Of course I realise it's different because a proof that a specific, concrete theorem is Gödel-undecidable, cannot exist — because it would imply that the theorem is true and therefore cannot be undecidable! So it is indeed possible that Goldbach, Riemann etc are impossible to solve in the same way that Scott's P' cannot successfully analyse itself. It just seems to me that the standard "feed the problem solver to itself" idea isn't a very plausible argument for that.

[Hide ↑](#)[Hide ↑](#)

**RCF** says:

September 1, 2014 at 3:02 pm

You are a bit unclear on what "undecidable" means. What Gödel's theorem says is that given any mathematical language L, there is some theorem T such that no proof of T in L exists, but T is true. So, for instance, it's possible to come up with

statements for which there is no proof in Peano Arithmetic (the Hydra Problem was given as an example above). However, that doesn't mean that there is no proof of the statement at all, just that there is no proof within Peano Arithmetic.

Consider this analogy:

Given any person P, I can come up with a statement S that is true, but that person P does not know. That doesn't mean that S is unprovable. It just means P is incapable of knowing that S is true. What Gödel's theorem does is it basically comes up with a statement of the form "S := 'P does not know that S is true'". So, for instance, I could say "Scott Alexander does not know that this statement is true". If Scott Alexander knows that that statement is true, that contradicts the statement, so the statement is false. One can believe false statements, but the verb "know" is reserved for true statements, so since this statement is false, Scott Alexander does not know that the statement is false. But that's precisely what the statement claims. Therefore, the statement is true. I have therefore produced a statement such that it is true, it is provable that it is true, and yet Scott Alexander can never know that it is true.

This may seem like a bunch of logical sleight of hand, but what Gödel's proof does it is actually puts this in rigorous terms, and shows that any sufficiently powerful mathematical system is going to be vulnerable to this sort of argument.

[Hide ↑](#) [Hide ↓](#)

**MartinW** says:

September 1, 2014 at 3:46 pm

And yet, when Scott Alexander reads that sentence S, it will take him only a second or two to realise what you're doing, and start laughing.

Moreover, if Scott could determine the truth value of any sentence which did *not* contain either an implicit reference to "things which Scott Alexander knows" or a full description of whatever algorithm Scott uses to determine the truth value of a sentence, that would still make Scott an awfully powerful dude. Especially if he could reliably detect such Gödel trickery and indicate the difference between "true", "false" and "haha nice try", so that he could never actually be tricked into making a false statement about a sentence.

Sure, the fact that even an omniscient Scott will need to laugh off some "pathological" statements, is an interesting observation, and I know

that it was a big shock when people discovered that it is not possible to devise a formal system which does not allow such pathological statements (but is still powerful enough to e.g. describe a Turing machine).

So I'm more interested in things like the "Hydra problem in Peano arithmetic" example, where a seemingly straightforward problem turns out to be non-solvable given the axioms. Although in that case the problem seems to be that the axioms are not powerful enough. When talking about a particular Turing machine, you don't have that problem: either it will halt or it won't, and the axioms are defined by the specification of the Turing machine.

P.S. @Everybody: yes, my comment about the halting-problem-solver allowing any input other than itself, was sloppy phrasing; I am aware that there would be lots of equivalent-but-not-obviously-identical programs to which the proof would also apply.

P.P.S. I enjoyed the responses and am learning from them, but I am going to wind down this topic now, as I have the feeling that for a first-time poster on a topic which was a bit of a semi-hijack to begin with, I might be dominating the comments section a bit too much.. (Also, it's close to bedtime.)

[Hide](#) ↑ [Hide](#) ↑

**Vilhelm S** says:

September 1, 2014 at 6:56 pm

I feel the main way your analogy breaks down is the "everybody would speculate maybe we finally found a real-world case" part. I think almost nobody expects that the Goldbach conjecture or Riemann hypothesis or P vs NP will turn out to be independent from ZFC. Like, I'm sure if you look you can find one or two mathematicians who believe that, but it seems like a very minority belief.

[Hide](#) ↑ [Hide](#) ↑

**Douglas Knight** says:

September 1, 2014 at 7:42 pm

A hell of a lot more expect P vs NP to be independent than the Riemann hypothesis.

[Hide](#) ↑ [Hide](#) ↑

**Anon** says:

September 2, 2014 at 1:15 am

P vs NP as possibly independent is actually fairly mainstream. cf no less than [Scott Aaronson](#) (not that he agrees, but he takes the idea seriously). Wolfram also seriously proposes it in *A New Kind of Science*: see eg [here](#).

And Knuth thinks Goldbach might be independent at least of PA (though independence of ZFC would be much, much stronger) (search "Goldbach" in [this document](#)).

And while I have no citations for it, I can say at least anecdotally that my algebraic geometry professor told us he thought RH might turn out to be true but unprovable in ZFC.

[Hide](#) ↑ [Hide](#) ↑

**Douglas Knight** says:

September 2, 2014 at 2:04 am

Knuth also [claims \(17\)](#) to believe  $P=NP$ .

[Hide](#) ↑ [Hide](#) ↑

**Vilhelm S** says:

September 2, 2014 at 8:14 pm

@Anon Well, by "expect they will turn out to be independent" I meant someone who thinks it's more likely than not that they are independent. Neither Aaronson or Wolfram explicitly says that. On the other hand, in apparently [in one straw poll](#) 3% of the respondents voted for independence, so the minority is larger than I expected.

[Hide](#) ↑ [Hide](#) ↑

**Douglas Knight** says:

September 2, 2014 at 10:17 pm

If you aggregate the two polls at your link, you get 5.5% for independence.

[Hide](#) ↑ [Hide](#) ↑

**Tom Womack** says:

September 3, 2014 at 5:42 am

I suggest you have a look at the Busy Beaver problem. This is where you get to if

you start arguing 'the halting problem can't be that hard. Let's define a simple representation of algorithms, start off with the simplest one, decide whether it halts or not, and continue to more complicated ones until it gets difficult'.

There are forty five-state Turing machines where nobody has yet figured out how to prove that they don't halt.

There is a six-state one which has been proved to run for more than  $7 \cdot 10^{36534}$  steps and then stop.

[Hide](#) ↑ [Hide](#) ↑

**MartinW** says:

September 3, 2014 at 6:09 am

Just for the record, I never said "the halting problem can't be that hard". In fact I pointed out that solving the halting problem must be *at least* as hard as finding a generic method for proving any theorem which can be expressed as a Turing machine which does a brute-force search for a counter-example. That's kind of the opposite of "not very hard".

What I said was that the technique of proving that no halting-machine-solver will work on all possible inputs, by constructing a kind of Liar's Paradox aimed at that particular solver, is a valid way to construct an existence proof but somehow feels a bit unsatisfactory to me. But I will happily admit that this is more an emotional response than a rational argument against its validity.

[Hide](#) ↑ [Hide](#) ↑

**Will** says:

September 8, 2014 at 8:28 pm

1. I think mathematicians are used to the phenomenon that you can prove something is true in a contrived case, when in reality, it's true in almost every case. For example, suppose you want to prove the following problem unsolvable:

Given an even number  $k$ , you must find an  $n$  such that for all primes  $p > n$ ,  $p+k$  is not prime

The "natural" case of this problem is  $k=2$ . This is believed to be unsolvable. But mathematicians were still quite satisfied by a solution of the silly case  $k=70,000,000$ .



Thus most mathematicians are satisfied by Godel's result, not because they think the halting problem is natural, but because they think Godel's result makes it clear that there are many other unsolvable problems.

2. The problem with showing a hard problem is unsolvable is that, for many/most problems, proving it unsolvable is harder than solving it! For statements like Riemann, Goldbach, etc., a proof of undecidability would necessarily give a proof of the conjecture.

The purpose of undecidability proofs is not to prove that certain hard problems are undecidable. The proof is to explain why certain hard problems are really really hard.

3. All that said, there are more natural hard problems that you might consider. Here's a simple problem, depending on two numbers  $k$  and  $n$ :

Do more than  $k$  computer programs of length at most  $n$  bits halt?

No algorithm can solve the problem for  $n = \text{length of the algorithm plus a small constant}$  and  $k = \text{actual number that halt}$ . If it could, then you could use the algorithm to solve the halting problem by finding the correct number that halt and running all those programs until you found one that halted.

This seems like a fairly natural mathematical problem – it's not about one specific Turing machine but about the average behavior of Turing machines as a whole. If this type of problem were in fact tractable, mathematicians would certainly study it. The only reason mathematicians don't study this sort of problem is that it's known to be unsolvable!

I think one can cook up more satisfying examples using a variety of clever tricks. For instance, there are languages in which one can prove a statement like "A random Turing machine of a reasonable length is undecidable".

[Hide](#) ↑ [Hide](#) ↑

**Charlie** says:

September 1, 2014 at 3:09 pm

You can extend the original kernel of "a halting-anti-symmetric program can't exist because you could feed it to itself," by a line of argument that looks like

"If we had a program that could do X, then we could use that to solve the general halting problem. Therefore we can't have a program that can do X."

So – can you think of any way to take a program that can tell whether any program other than its identical copy will halt, and use that program to solve the general halting problem? (Hint: Ner gurer nal aba-vqragvpy cebtznf gung pna or thnenagrrq gb unyg vs naq bayl vs gur bevtvany cebtzn unygf?)

[Hide ↑](#)[Hide ↑](#)

**Luke Somers** says:

September 1, 2014 at 4:17 pm

I find this line of argument utterly unconvincing and vastly prefer the original.

Don't forget, the proof about not being able to solve the Halting Problem says nothing about being able to prove whether *\*most\** programs halt or do not. It only speaks of *\*all\**.

Indeed, solving that problem for common cases is an active research program at Microsoft, for automatic quasi-validation ('at least it doesn't blow up and kill the system') of programs (they started with device drivers).

Relatedly, work is being done to find out more and more flexible rules to follow when writing a program to make it easy to determine that it halts.

These are all things that are perfectly allowed under the theorem.

ALSO allowed under the theorem is proving any arbitrarily complicated mathematical problem that has a solution. Go ahead, generate a proof of the Goldbach Conjecture, or a counterproof. Sweet!

[Hide ↑](#)[Hide ↑](#)

**matejcik** says:

September 5, 2014 at 8:32 am

What might help your intuition is realizing that the halting-recognizer is *not special*. It is not a magical-algorithm-unlike-any-other-algorithms. Consider the, um, more practical applications: verifying program correctness. Halting-recognizer is now a part of a static analysis and debugging suite, and will warn you when your program contains infinite loops.

As written, it also happens to contain a bug, that, in certain edge cases, causes it to enter an infinite loop when evaluating an otherwise good (i.e. halting) program.

This is your program Q, built on top of an abstract "correct" algorithm P.

You are a developer working on this suite, and while trying to improve performance of the halting recognizer, you inadvertently trigger this bug. The IDE freezes. (except it shouldn't because if the

program triggers the bug, it should have already detected the bug and reported it and halted, except that...)

The takeaway is this: the halting recognizer is probably not the only special algorithm in the set of all possible algorithms that would trigger the bug. There will be others. You could manually (“cranially”) find out when this happens, and determine the class of problems on which your algorithm P doesn’t work properly (because if it did, it would trigger the paradox through its implementation Q).

Therefore, P was not general in the first place, and you just pushed the halting problem up a level.

The canonical proof only makes the “bug” explicit and intentional. It’s not trying to cheat the program P out of its wits, it’s just a very simple example of what would happen somewhere else anyway.

[Hide](#) [↑](#) [Hide](#) [↑](#)

**[Gareth Rees](#)** says:

September 1, 2014 at 6:45 am

*The vision I got from Democritus was of a God who was single-mindedly obsessed with enforcing a couple of rules about certain types of information you are not allowed to have under any circumstances.*

This reminds me of [Bruno Latour](#), who read Einstein’s introduction to *Relativity: the Special and the General Theory*, and found in it an obsession with ordering people about on trains: a “contribution to the sociology of delegation”.

[Hide](#)[Hide](#)

**Mike Capp** says:

September 1, 2014 at 7:30 am

The “snake-eyes-rolling madman” appears as the “Shooting Room objection” in Leslie’s original 2002 paper on the Doomsday Argument. IIRC he attributes the general gist of it to David Lewis.

[Hide](#)[Hide](#)

**B** says:

September 1, 2014 at 9:24 am

“The vision I got from Democritus was of a God who was single-mindedly obsessed with enforcing a couple of rules about certain types of information you are not allowed to have under any circumstances.”

I think this is a fundamentally misleading perspective. I think the most definitive work on this subject – explaining the essence of quantum electrodymanics to a layman – is Feynman’s QED. In it he explains that the uncertainty principle is essentially a holdover from the days where we didn’t know how things worked,

and that in the modern understanding there is no need for such a principle – it is simply one of many consequences of a very simple set of rules repeated over and over. I encourage you to read the whole thing, it's brilliant and very neatly set out, and you can find it for free online.

[Hide](#)[Hide](#)

**Anonymous** says:

September 2, 2014 at 12:44 am

I'm confused by your comment.

*In [QED] he explains that the uncertainty principle is essentially a holdover from the days where we didn't know how things worked, and that in the modern understanding there is no need for such a principle – it is simply one of many consequences of a very simple set of rules repeated over and over.*

But you seem to be putting this forward as an argument *against* the idea of God as enforcing rules about types of information. I'm not sure where my copy of QED is at the moment but almost every time I've seen the uncertainty principle discussed, it was explained as essentially a mathematical consequence of the Fourier transform. That is, there is a particular relationship between information in normal space and Fourier space that, when applied to quantum mechanics, gives the uncertainty principle. So I agree that it comes from a simple set of rules, but this seems to be exactly the type of information-based enforcement that Scott is talking about.

[Hide](#) [↑](#)[Hide](#) [↑](#)

**Kevin** says:

September 2, 2014 at 11:21 pm

In the case of a particle's position and momentum, both pieces of information **do not exist** simultaneously. It's not that both pieces of information exist but we just aren't allowed to obtain them. A particle with a definite position has an indefinite momentum, because of the inherent properties of waves. [HyperPhysics](#) has a good visual explanation of this.

[Hide](#) [↑](#)[Hide](#) [↑](#)

**MugaSofer** says:

September 1, 2014 at 9:47 am

*Finally, if we want to combine the anthropic computation idea with the Doomsday Argument, then there's the Adam and Eve puzzle. Suppose Adam and Eve are the first two observers, and that they'd like to solve an instance of an NP-complete problem, say, 3-SAT. To do so, they pick a random assignment, and form a very clear intention beforehand that if the assignment happens to be satisfying, they won't have any kids, whereas if the assignment is not satisfying, then they will*

go forth and multiply. Now let's assume SSA. Then, conditioned on having chosen an unsatisfying assignment, how likely is it that they would be an Adam and Eve in the first place, as opposed to one of the vast number of future observers? Therefore, conditioned upon the fact that they are the first two observers, the SSA predicts that, with overwhelming probability, they will pick a satisfying assignment.

Oh, *that's* what Eliezer means when he says "anthropic superpowers".

[Hide](#)[Hide](#)

**Nicholas Weininger** says:

September 1, 2014 at 10:51 am

FWIW, I don't know whether Aaronson went into this, but many computational complexity arguments for why you can't have information without a lot of work are "just" counting arguments. The intuition I find appealing about this is that:

- there are many, many different distinct things you might want to know
- there are relatively few distinct ways to find things out simply
- therefore most things you want to know can't be found out simply.

Does that make sense?

[Hide](#)[Hide](#)

**Adam Strandberg** says:

September 1, 2014 at 1:01 pm

Could you provide a link to someone making this argument? I've read a fair amount of Aaronson's stuff, but I can't recall any of it (at least explicitly) being of this form.

[Hide](#) [↑](#)[Hide](#) [↑](#)

**Nicholas Weininger** says:

September 1, 2014 at 2:38 pm

Kolmogorov complexity arguments are probably the purest form of this: there "things you want to find out" are numbers of a given length and "simple ways of finding things out" are programs of a much shorter length, and the conclusion is that most numbers cannot be output by programs of a much shorter length than themselves. Li and Vitanyi is a standard text on these arguments and their applications.

[Hide](#) [↑](#)[Hide](#) [↑](#)

**William Newman** says:

September 1, 2014 at 3:05 pm

"Could you provide a link to someone making this argument?"

See "Limitations" in [http://en.wikipedia.org/wiki/Lossless\\_data\\_compression](http://en.wikipedia.org/wiki/Lossless_data_compression) .

I don't know whether Aaronsen happens to make such an argument, but it is common in things related to information theory. As I understand it, Shannon himself made such an argument. (Possibly it is adapted from even earlier work, but even if so, as long as Shannon used the idea, that makes it as old as most of information theory.)

[Hide](#) [↑](#)[Hide](#) [↑](#)

**Darcey Riley** says:

September 1, 2014 at 7:29 pm

I don't know how this argument applies to complexity, but if you take out the word "simply" and apply the argument to computability, then it becomes a paraphrase of the standard diagonalization proof that not all languages are Turing-recognizable.

[Hide](#) [↑](#)[Hide](#) [↑](#)

**Anatoly** says:

September 1, 2014 at 12:21 pm

I don't think you should feel so depressed about not being able to understand much in the book. As much as I admire Scott's blog, his book just isn't a popular science book. Its informal tone and seemingly patient descriptions of some very basic material are misleading (no malice of any sort imputed, of course): most of its content really requires some sort of technical foundation in complexity theory and quantum complexity theory. It grew out of a course given to students who already had some exposure to quantum complexity on a technical level. It's fine to read it for the prose, the metaphors, the startling comparisons, but unless you actually studied complexity theory and quantum complexity (not just math! not just physics! not just CS!) in graduate school or at a comparable level, much of the book *will* be incomprehensible to you on the technical level (this is certainly true in my case). This really isn't a case of "oh if only I had talent for math".

I think you should give math another chance. The proof of the Halting Problem undecidability that you say was simple for you to understand is about as complicated as the epsilon-delta framework for rigorous calculus that seems to be the stumbling block for most students. If someone can make sense of epsilon-delta proofs and generate them on their own, I think it's very likely that they *can* without huge difficulty learn something like the typical undergrad math major curriculum, which is already lots of very beautiful and shiny things.

[Hide](#)[Hide](#)

**Zorgon** says:

September 1, 2014 at 2:07 pm

Kind of OT, but thank you – that’s about the best argument for doing a second undergrad in math I’ve yet encountered.

(Context – I’m seriously considering re-entering academia in the near future after screwing it up it hard the first time over a decade ago.)

[Hide ↑](#)[Hide ↑](#)

**guess i'll be alone in my utopia** says:

September 1, 2014 at 7:17 pm

Would it be a good read for someone who does hard math, but in completely different fields? (Geometry, if it matters. I can handle general relativity fine. (when translated out of physicist-ese, anyway))

[Hide ↑](#)[Hide ↑](#)

**Liskantope** says:

September 1, 2014 at 10:05 pm

I agree; I have a very hard time believing that Scott wouldn’t be able to pick up the core material required for any undergrad math major with no problem. And in my opinion, knowing that core material is very rewarding in and of itself.

[Hide ↑](#)[Hide ↑](#)

**Rand** says:

September 2, 2014 at 2:20 pm

I’m going to disagree (about needing to have studied Quantum Complexity, not about studying math being a great idea), though it’s a minor difference. I’m a PhD student in Programming Languages with a strong Math/CS background (but no prior knowledge of BQP) and I was able to understand most of the book (as were a few students here with similar backgrounds). I was actually reading the book at the same time as taking a seminar in Randomized Algorithms, and I found that’s Scott’s explanation of the randomized complexity classes was much more accessible than those in Motwani’s book. (And likewise, I found his discussion of interactive proofs far more accessible than the Arora Barak Complexity book I read the following semester.)

It’s still a high bar. My brother, who also studied Math and CS in college, told me that he couldn’t understand the book, and I’m guessing that the difference is that he’s no longer engaged with this type of material.

How difficult the book is depends a lot on how much you already know, and I’d be hesitant to recommend it to anyone who hasn’t read Sipser. (Also, I’d recommend that anyone who hasn’t read Sipser’s “Introduction to the Theory of Computation” read it, it is a fantastic textbook.) With Sipser,

some math background, smarts and effort, you can understand the whole thing.

Or you can just skim it, and still get a lot of interesting stuff out of it.

[Hide](#) ↑ [Hide](#) ↑

**Zorgon** says:

September 1, 2014 at 2:08 pm

Dammit SA, my reading list is already too damn long 😞

[Hide](#)[Hide](#)

**RCF** says:

September 1, 2014 at 3:14 pm

"If each page of the rule book corresponded to one neuron of a native speaker's brain, then probably we'd be talking about a "rule book" at least the size of the Earth, its pages searchable by a swarm of robots traveling at close to the speed of light."

Are the typos yours or Aaronson's? Also, according to my calculations, if each page corresponds to one neuron, the rule book could fit in a skyscraper.

[Hide](#)[Hide](#)

**Ilya Shpitser** says:

September 1, 2014 at 4:19 pm

Scott, you easily have all the hardware for math on the other Scott's level. It's just a matter of putting the hours in.

This "math talent : you either have it or you don't" thing is kind of a toxic (and more importantly not really right) way of looking at doing math.

[Hide](#)[Hide](#)

**Matthew** says:

September 1, 2014 at 11:44 pm

Possibly it's a question of poor pedagogical methodology, but I stopped taking math beyond Calculus II (other than applied things like intro. statistics and number theory) because I was literally getting headaches trying to wrap my head around the problems. Scott is smarter than I am, but still, a more general claim that anyone with substantially above-average intelligence can learn any math given sufficient time is either simply false or "sufficient time" means devoting your entire life to the task.

[Hide](#) ↑ [Hide](#) ↑



**Anonymous** says:

September 2, 2014 at 12:51 am

number theory is applied math? was this a course in a cs department that in another school would be called "discrete math"?

[Hide ↑](#)[Hide ↑](#)

**Matthew** says:

September 2, 2014 at 12:56 am

"Applied" might be stretching it I suppose. It was in the math department, but some basic cryptography stuff was covered.

[Hide ↑](#)[Hide ↑](#)

**Ilya Shpitser** says:

September 2, 2014 at 1:52 am

Matthew, my claim is that the view that there is a binary thing called "math talent," and mathematicians for the most part have this binary thing is not the right view of what is happening. Do you disagree?

Math does require the ability to concentrate, and probably some amount of brain hardware, but aside from that, doing math is a learnable skill, like doing many other cognitively demanding tasks, for instance programming or psychiatry. I believe a dominating term for many otherwise smart people's issues with math is morale. It is true that there exist people naturally predisposed to doing math, e.g. Ramanujan, but most mathematicians are not like that.

[Hide ↑](#)[Hide ↑](#)

**Army1987** says:

September 2, 2014 at 2:58 am

It could be that "anyone with substantially above-average intelligence can learn any math" if they're young enough but it's much harder for adults, or something like that.

[Hide ↑](#)[Hide ↑](#)

**Luke Somers** says:

September 2, 2014 at 3:37 pm

You don't even need to get to the boundary of 'real math' to reach things that ought to be taught differently. A whole lot of the math we're

taught is stupidly harder than the stuff that's actually useful.

Don't teach advanced analytic geometry without beginner's calculus available. Heck, analytic geometry other than with calculus is really a waste of time. It does way less with way more effort, like carrying your car instead of the other way around.

But calculus courses aren't faultless, like you said. Even before you get to integration. Yeah. Epsilon Delta Proofs? Use them whenever you find some new function, in order to prove what the derivative of that function is. And then PUT THEM AWAY.

There is enough math to be taught in calculus that you don't need to drain the students' brains out with excessive and inappropriately-placed rigor. Don't re-prove general results like the power rule over and over again. USE your proofs!

U-substitutions are really nifty in principle and should be taught. But finding the U, except in special cases where it's clear, is a guess-and-check blunderfest, something that does NOT belong in homework or – worse – on tests.

[Hide](#) ↑ [Hide](#) ↑

**Sniffnoy** says:

September 2, 2014 at 6:19 pm

Hm, in my experience the ones they put on homework and tests *are* the ones that are pretty clear. (Going based on my experience helping people with University of Michigan calculus classes the past few years.) There's the occasional case where you have to substitute blindly and hope for the best, but I think it's OK to have those occasionally (to remind the students that sometimes you don't have any better option). In any case, the "obvious" ones still seem to give a lot of people trouble, so it doesn't seem to really be so all-or-nothing between obvious/worthless and opaque/worthless; there does seem to be that middle ground of difficult-but-understandable.

[Hide](#) ↑ [Hide](#) ↑

**Ialdabaath** says:

September 2, 2014 at 7:02 pm

There is enough math to be taught in calculus that you don't

*need to drain the students' brains out with excessive and inappropriately-placed rigor. Don't re-prove general results like the power rule over and over again. USE your proofs!*

Interesting; this was almost the opposite of my strategy in high school. I couldn't keep all the proofs in my head, so I just trained myself how to derive them from first principles, and then worked out the individual problems from there.

As a result, I invariably failed any test question that was "how does the {some dead white guy's name} rule work?" or "solve this using the {some other dead white guy's name} rule", but aced any question that was actually "how do you solve {problem of class X}?" or "what is the solution to {problem of class Y}?"

It also led to solving a lot of problems in really interesting ways, which annoyed the hell out of my highschool calculus teacher but really impressed my college calculus teacher.

[Hide](#) ↑ [Hide](#) ↑

**Luke Somers** says:

September 2, 2014 at 9:41 pm

I found trigonometric substitutions in the denominator to be a little bit on the unclear side, and there were tons of those.

Ialdaboth – if you can do that, then great. I would never want to stand in your way. Many students would struggle with that, and I would like to enable them.

Also, I think that with actual concerted drill, you would be able to remember the rules (and that that would speed you up). Drill would mean problems that you do NOT carry through to completion, because that's mostly busy-work. Instead, applying the technique to reduce the problem to something that is simpler, and then stop and go on to the next problem. You can do twice as many problems that way and focus on the parts that they're actually trying to teach, and still have time left over to have some problems a student would carry through to the end to make sure everything's still holding together.

[Hide](#) ↑ [Hide](#) ↑

**Sniffnoy** says:

September 3, 2014 at 2:12 am

Ooh, yeah, trigonometric substitutions, I forgot Michigan doesn't teach those! Yes those are more difficult. They're not that hard if you have a good understanding of trigonometric functions (and their hyperbolic cousins), but most students don't.

[Hide](#) ↑ [Hide](#) ↑

**Liskantope** says:

September 3, 2014 at 11:06 am

"Epsilon Delta Proofs? Use them whenever you find some new function, in order to prove what the derivative of that function is. And then PUT THEM AWAY."

Completely agree about avoiding epsilon-delta whenever possible in first calculus courses. Although the main purpose of epsilon-delta is to rigorously define limits, and limits are what should be used to prove what the derivative of a new type of function is. I think many first-time calculus students intuitively grasp the idea of limit pretty well, while the conditions and quantifiers in epsilon-delta statements just give them a headache and are unnecessary (it was certainly the case with me).

[Hide](#) ↑ [Hide](#) ↑

**moridinamael** says:

September 2, 2014 at 10:16 am

FWIW, if I had stopped after Calculus II, I probably would I have said "I guess this 'math' thing isn't for me." But then I went on to more advanced things and paradoxically they actually become easier and more intuitive past that point. The problem is that Calculus II at the undergraduate level is a perverse grab-bag of very difficult and unintuitive techniques assembled to weed out undergraduates.

[Hide](#) ↑ [Hide](#) ↑

**Rand** says:

September 2, 2014 at 2:31 pm

This is spot on.

It's frustrating the amount of pain undergraduates have to go through before they get to study what mathematicians would call "real math". Analysis and (abstract) Algebra really don't resemble the years of Calculus and Differential

Equations that precede them, and very few students have the opportunity to explore either.

I found that CS curriculums tend to be a lot better in that you can take a programming course immediately, you can take a theory course almost immediately and you can take a systems course almost immediately, and you can find out what interests you.

[Hide ↑](#)[Hide ↑](#)

**Liskantope** says:

September 3, 2014 at 11:13 am

A concrete manifestation of this issue is that the math subject GRE (required to get into most good math graduate programs, as well as many math-related areas of work I'm sure) is about 50% differentiation/integration techniques. Which most prospective math grad students (1) learned a long time ago, (2) hated, and (3) forgot very quickly because they're not actually needed for most core math major classes anyway.

[Hide ↑](#)[Hide ↑](#)

**Rand** says:

September 4, 2014 at 1:11 am

Liskantope: That test is a monstrosity. (Though I don't think required for anything but graduate school.) I know people who decided to get PhDs in CS instead of math because of that monstrosity.

Actually, I'm not sure that I'm not one of them...

[Hide ↑](#)[Hide ↑](#)

**Mai La Dreapta** says:

September 2, 2014 at 6:58 pm

I stopped taking math *before* calculus, because I hated calculus. I put a high priority on being able to intuit and understand the math that I'm being asked to do, and calculus, at least as I was introduced to it, was a lot of "memorize this list of derivatives and complicated integration strategies in time for the test." I took one look at it and said, "no thanks."

I have always regretted not being able to get into the higher mathematics that lies on the other side of calculus, especially because it seems to me that very

little of it actually requires calculus. But I've never seen a curriculum that didn't have calculus as a gateway requirement.

[Hide ↑](#)[Hide ↑](#)

**Liskantope** says:

September 3, 2014 at 10:47 am

Higher mathematics is generally split into two major areas: algebra and analysis. (Of course, this is an oversimplification, and they often do dovetail at the research level.) While algebra requires virtually no calculus, analysis is sort of the natural continuation of the rigorous study of calculus. But I get the impression that pure analysis (rather than applications, such as solving differential equations, etc.) probably doesn't require much knowledge of the formulas and algorithms that turned you off to calculus, but rather a deep understanding and ability to generalize the concepts behind derivatives and integrals.

No curriculum that I know of lets you get far in math without calculus, because of the notion of "mathematical maturity", the idea being that you're not ready to take a higher-level course until you've had the experience of studying something as abstract as calculus. I'm not sure I believe in this notion. I have a friend who claims not to have any intuition for the basic concepts in calculus and that he still doesn't really understand what "differentials" are for instance, yet is an excellent math PhD student in research-level algebra.

If you're interested in getting a good introduction to the algebraic side of higher math, I recommend Joseph Gallian's *Contemporary Abstract Algebra*; it is a pleasant read which definitely requires no calculus!

[Hide ↑](#)[Hide ↑](#)

**William Newman** says:

September 3, 2014 at 2:44 pm

"If you're interested in getting a good introduction to the algebraic side of higher math, I recommend Joseph Gallian's *Contemporary Abstract Algebra* [...]"

And if you're interested in starting with a peek rather than a comprehensive introduction, you might take a look at Maxwell and Maxwell *\_Abstract Algebra and Solution by Radicals\_*, which is about ten bucks from Dover. It is an unusual book: pitched at very bright and

motivated students who are still in high school, arriving at the famous Galois proof that it is impossible to generalize the quadratic formula above fourth-order polynomials. It's a sort of lightning raid deep into sophomore abstract algebra territory, bypassing a lot of strongpoints in order to get to a marvellous result. (It necessarily explains some subtle and difficult ideas, but it chops out stuff that doesn't bear directly on that one proof.)

Incidentally, calculus is also great, at least when understood correctly. (Mathematicians are not always good teachers, and are surprisingly often *\*just\* \*terrible\** at explaining the significance of things.) Calculus is really really useful if you get serious about understanding the world. Not just in things like being able to express fundamentals like Newton's laws or Maxwell's equations or Schroedinger's equation in a concise form and being able to do calculations with them, but in e.g. being able to reason precisely about statistics using spinoffs like Fourier analysis (and related concepts like convolution and lowpass which are slippery even to define without calculus), and being able to approximate things efficiently (even things that you might think have nothing to do with calculus, such as the factorial function).

[Hide](#) ↑ [Hide](#) ↑

**Doug S.** *says:*

September 3, 2014 at 9:32 pm

And Calculus III (multivariable calculus) is actually complex analysis made difficult. Calculus with complex numbers is a lot simpler than calculus with functions of multiple real variables; everything's a power series or it's not differentiable in the first place, which makes your job a lot easier.

[Hide](#) ↑ [Hide](#) ↑

**Shmi Nux** *says:*

September 3, 2014 at 5:18 pm

> Scott, you easily have all the hardware for math on the other Scott's level.

Ilya, reading this nonsense from you makes me doubt your competence in other areas.

Yes, almost anyone can learn math at a reasonable level, say, Grade 11, though some need 100x more effort than others. No, only a very small fraction of people can learn it at Scott Aaronson's level. Aptitude for math is not unlike aptitude for music or sports. Almost anyone can become average, with varying amount of effort, almost no one can become top 0.01% (where Scott

Aaronson is), regardless of the effort put in. You would not be the next Olympian, no matter how much you try.

> This "math talent : you either have it or you don't" thing is kind of a toxic (and more importantly not really right) way of looking at doing math.

Scott did not claim that. Any talent is a spectrum, and math is no exception. Hard work can give you an extra standard deviation or two, but not much more. If Scott Alexander has the MathQ of 100 and Scott Aaronson 140 after putting countless hours into it, no amount of work will help. It would also be a stupid waste of time to do so, given how his other Q's are way up there.

[Hide](#) [↑](#) [Hide](#) [↑](#)

**J** says:

September 1, 2014 at 10:28 pm

If you think the proof of the halting problem is interesting, then this will really bake your noodle: for finite Turing machines, halting is entirely tractable. And the proof is just as elegant: for each step in a program's execution we can record the complete state of the machine (think: RAM and program counter). This state tells us everything we need to know to execute the next instruction, because Turing machines are completely deterministic. So at each step, we check the list of previous states. If we've been in this state before, then we absolutely must be in an infinite loop, because we're going to repeat the exact same instructions that got us here since the last time we were in this state. So it's easy to detect the "infinite loop" option.

Now, since the machine is finite, there are only so many states we can ever get into. (That number is extremely large, but that doesn't matter because this is an abstract proof). So within a finite amount of time we have to either halt or repeat a state, in which case we loop forever.

All practical Turing machines are finite. (In fact it's interesting to contemplate how different an infinite Turing machine is from a real computer; even storing a memory address can require infinite storage). So it's a pet peeve of mine that so many cs teachers nod sagely after teaching the halting problem and pontificate on all the consequences it has for real world programs, and don't even mention that the situation is completely reversed for finite machines.

[Hide](#)[Hide](#)

**Army1987** says:

September 2, 2014 at 3:02 am

*for finite Turing machines, halting is entirely tractable*

Tractable by much larger machines in absurdly long times.

TOW you're appealing to practicalness only when it suits you. rather than even-handedly.



[Hide](#) ↑ [Hide](#) ↑

**J** says:

September 2, 2014 at 4:43 am

I specifically pointed out that the proof isn't making claims about performance. Performance isn't the point. The point is that subtle assumptions about infinity completely reverse the results, which should—but in practice doesn't—make people hesitate to make big sloppy logical leaps about what we should or shouldn't expect to achieve in practice.

[Hide](#) ↑ [Hide](#) ↑

**[Taymon A. Beal](#)** says:

September 2, 2014 at 3:09 am

It usually does get mentioned in CS courses that the halting problem assumes infinite memory, and that all real-world machines are finite-state automata. And then we never bring up this fact again, because it is totally useless for all real-world purposes.

The computer I'm typing this on has about 5 gigabytes of physical RAM. If we were to model it as a deterministic finite automaton, assuming that we only allow the contents of RAM to count as state (registers are basically a rounding error), that automaton would have  $2^{48372744192}$  states. At this point it is customary to compare that number with the number of atoms in the observable universe, but frankly that seems like it would be belaboring the point. And, of course, my computer also has a hard disk, which is also used as state in ordinary calculations, and it has an internet connection, which allows me access to basically unlimited computing resources for all practical purposes.

There's a well-known quote by a famous computer scientist, which I am somehow failing to find on Google, which basically says the fact that most problems we're interested in are technically finite in nature is almost never useful when we're trying to solve them. Real computers are close enough to Turing machines for all practical purposes, and we can reason effectively about them. We can't reason effectively about automata with astronomical numbers of states, nor can we program real-world computers to analyze them.

[Hide](#) ↑ [Hide](#) ↑

**J** says:

September 2, 2014 at 5:02 am

We have before us two proofs, both of them correct, yielding opposite conclusions because of subtle differences in their assumptions. One requires truly infinite storage. The other, at least as written, requires large but not infinite storage. You declare the finite one to be irrelevant to the "real world" but not the one requiring infinity, and offer no proof except a

famous quote you can't find? That's precisely the kind of sloppy philosophizing I was complaining about.

(I will allow that the infinite halting problem is more useful to mathematicians, because it lets them say interesting things about other problems involving infinity. But that has nothing to do with what we can learn about actual code that runs on non-infinite machines. Anatoly's comment above is the most useful insight I've ever seen about the infinite halting problem, and it's all about infinite strings, not real programs.)

[Hide](#) ↑ [Hide](#) ↑

**J** says:

September 2, 2014 at 6:37 am

Also note that the infinite proof is a black box algorithm with proof by contradiction. So it leaves us with a conclusion but no machinery to examine. Per Anatoly's remarks, that's precisely because the proof is about infinite and incomprehensible strings, not real programs.

Whereas the finite proof is constructive—I even laid out how to implement it in just a few sentences, even though the proof didn't require a specific implementation.

So if we want to talk about applicability to the real world, the one that you could code up in a few minutes and immediately start optimizing seems a whole lot more useful. And in fact you could start with such an exercise and transition quite smoothly into the literature on formal verification, which is all about giving automatic and provable answers about how various kinds of programs behave.

Down that practical road certainly lies plenty of memory and performance constraints, but on performance both proofs are silent. The infinite one merely informs us that if your machine is truly infinite (not merely universe-sized), some answers aren't computable at all.

[Hide](#) ↑ [Hide](#) ↑

**[Mai La Dreapta](#)** says:

September 2, 2014 at 1:37 pm

Relevant to this discussion: there exists [a class of computer languages](#) which are designed to provably halt; if the compiler cannot prove that the program halts, it presents the user with an error. By definition, such languages are not Turing-complete, but this turns out to matter less than you'd think, because the number of computational problems which actually *require* Turing-completeness is very small. (Interestingly, it's possible to write the compiler for such a language in the language to be compiled, the same way that you do

compiler for such a language in the language to be compiled, the same way that you do for C, because a non-Turing-complete language can prove the completeness of a non-Turing-complete string.)

In these languages (and in functional languages employing Hadley-Milner type inference in general) it's a fun factoid that you can write programs whose compilation requires the compiler to prove the Riemann hypothesis or a similar theorem, not to *run* the program, but just to *compile* it. This is mostly just done to amuse the programmer, though, since actually attempting to compile the program generally means that you chew up your computer doing a dumb brute-force search, or the compiler gives up with an "I can't tell if this halts" message.

TL;DR: Turing-completeness and the halting problem matter less than you think they do.

[Hide ↑](#)[Hide ↑](#)

**Ariel Ben-Yehuda** says:

September 2, 2014 at 2:36 pm

Just having a total, or provably total, compiler is a rather trivial property – you can just "compile" to an interpreter plus the source code (or, if you're feeling evil, to a compiler plus a machine emulator plus the source code). If you just want to tick the "provably total" checkbox you can do your regular optimizations for a bounded amount of steps and only fall back if they don't finish (or return an error). And remember that a provably total program may still not terminate before the Universe's heat death – a proof of totality is nearly worthless if your program has an exponential time (or worse) worst-case.

However, because of the halting problem, a total language certainly can't have a \*self-interpreter\*.

By the way, no single problem is undecidable – it is either true or false, and therefore a program that "prepared for the test" can always decide it.

[Hide ↑](#)[Hide ↑](#)

**Douglas Knight** says:

September 2, 2014 at 3:24 pm

Mai, there is a serious disconnect between your first paragraph and your second. ML and its descendents, the languages that use Hindley-Milner, have terminating compilers, just like everything else. There are some extensions to the type system to make it Turing complete. Then, yes, you can write a program which will parse if and only if the Riemann hypothesis is provable. But that is the exact

opposite of your first paragraph.

[Hide](#) [↑](#) [Hide](#) [↑](#)

**Mai La Dreapta** says:

September 2, 2014 at 6:52 pm

You're right, I think I got my Haskell trivia confused with my total functional programming trivia. (I have actually used Haskell, but not any of the total functional languages, and I tend to group them into the same bucket in my head.)

[Hide](#) [↑](#) [Hide](#) [↑](#)

**Anonymous** says:

September 1, 2014 at 11:39 pm

> If each page of the rule book corresponded to one neuron of a native speaker's brain, then probably we'd be talking about a "rule book" at least the size of the Earth

Am I missing something, or is this off by many orders of magnitude? The brain has  $\sim 10^{14}$  neurons. The book in front of me is 1000 pages and around  $0.01\text{m}^3$ , so, carry the two... we're talking about a  $\sim 500\text{m}$  cube here, which is big, but hardly the size of the earth.

[Hide](#)[Hide](#)

**mjgeddes** says:

September 2, 2014 at 12:26 am

Miscellaneous riffs on everything.

MCRT (Mathematico-Cognition Reality Theory) is my own developing theory of everything (work in progress since 2004!). What does it have to say about fundamental questions? It's been years since I posted an update on my reality theory, so thought I'd weigh in here with some updates on my views, based on MCRT:

On Metaphysics/Physics:

\*Math is the base level of reality. Tegmark is right.

\*Space-time is truly fundamental at the physics level, and it's continuous (not discrete). So popular theories about space being quantized are wrong. Other popular ideas about the universe being 'digital' and a 'computer' or composed of 'cellular automata' (aka Wolfram) or any theory where 'information' or 'computation' is taken to be fundamental I would also lump in the 'wrong' basket.

\*I take 'information' to be a feature of minds, not reality (see above).

\*Looking at physics, the fundamental ingredients are space-time fields, various transforms, and symmetries. I would say that geometry IS actually physics (I would label physics as a branch of geometry).

On Quantum Mechanics:

\* I *think* many-worlds is correct, but I'm still far from confident. I would rate the Bohm interpretation and the transactional interpretation as still in the running. I think backward causation is still a real possibility for explaining quantum mechanics, and that many-worlds could still be falsified. Quantum mechanical effects could be equivalent to wormholes (i.e., info travels back through time, i.e, past states are partially determined by future states).

\*Time could be multi-dimensional (related to backward causality/transactional interpretations of QM), but again, I'm not confident about this.

\*Even if the above is right, many-worlds could still be correct

On cognitive science:

\*I'm very confident that Bayes is not all , I maintain that a generalization of categorization/analogy-making is the real foundation of rationality. Bayes is purely on the functional level, it can't be fundamental. Logic needs a general method of representation , Bayes can't do it, but a theory of categorization would fit the bill.

\*I'm still very confident about universal morality. I'm close to certain that platonic ideals exist, and that 'beauty' is the fundamental value at the base of it all.

\*I'm not confident about Bostrom's orthogonality postulate. I think there's a fair chance its wrong, based on the existence of universal morality, but I'm far from sure (see below)

\*The mere existence of universal (objective) morality doesn't invalidate orthogonality, as Bostrom himself noted in his book (universal morality may not be self-motivating, and even if it was, AIs may simply not be designed to perceive it). So UAI is still a big threat.

\*Consciousness is no big mystery and I'm confident I've got a good idea of what it is. Yes, I'm sure it's just computation, its just narrative formation , probably along the lines of Baars Global Workspace.

[HideHide](#)

**Ialdabaoth** says:

September 2, 2014 at 1:44 pm

*I would rate the Bohm interpretation and the transactional interpretation as still in the running*

\*googles Transactional Interpretation\*

HOLY CRAP I CAME UP WITH THIS IDEA IN HIGH SCHOOL. I had like, a whole bunch of partial differential equations I was trying to solve to see if I could convert it from naive interpretation to an actual theory. (note: 1992-Brent did not necessarily know what he was doing in terms of advanced mathematics OR quantum field theory).

My God, why didn't I have access to physics journals?

[Hide ↑](#) [Hide ↓](#)

**mjgeddes** says:

September 3, 2014 at 1:21 am

If you're serious, you people are frighteningly smart, I'll give you that 😊

From Wikipedia:

"The transactional interpretation of quantum mechanics (TIQM) describes quantum interactions in terms of a standing wave formed by both retarded ("forward-in-time") waves, in addition to advanced ("backward-in-time") waves."

This isn't quite backward causality, but similar ideas do suggest backward causality, namely, 'Two-state vector formalism':

"The two-state vector formalism (TSVF) is a description of quantum mechanics in terms of a causal relation in which the present is caused by quantum states of the past and of the future taken in combination."

What really made me take a closer look at these ideas, was the news last year that theoretical physicists seem to have shown that wormholes and quantum entanglement were mathematically equivalent under particular conditions. This is very intriguing indeed. Here's a link about that:

<http://news.sciencemag.org/physics/2013/12/link-between-wormholes-and-quantum-entanglement>

Here's the thing: not only does this allow a connection between relativity and QM, it also supplies an elegant geometrical picture for what QM means! (And remember, if we take space-time as fundamental, then at the deepest level, physics and geometry are one and the same).

None of this refutes the many-worlds interpretation, but it's all very suggestive!

[Hide ↑](#) [Hide ↓](#)

**Ialdabaoth, Unwitting Servant of Moloch, Dread Agent of He-Who-Swims-Ever-Leftward** says:

September 3, 2014 at 1:24 am

Well to be fair, all I REALLY came up with was from reading Feynmann diagrams and Feynmann's "retarded waves" photon theory and saying "wait couldn't that explain QM

wave-particle duality too, if the so-called 'collapse' was just into the only path where all "particles" constructively self-interfere?"

[Hide ↑](#)[Hide ↑](#)

**Luke Somers** says:

September 3, 2014 at 12:20 pm

My two beefs with the Bohm theory are:

1) Bohm completely undoes all of the counterarguments to the fine-tuning problem. If you modify it to have lots and lots of guided particles, then it undoes it less, but the more you do that, the more it turns into MWI.

2) If Bohm were physically actually true, we would have no reason whatsoever to think that we were on the officially 'real' worldline rather than being a part of the guide wave.

[Hide ↑](#)[Hide ↑](#)

**mjgeddes** says:

September 3, 2014 at 11:55 pm

Yes. Bohm definitely has problems. However, it \*does\* have the advantage of offering a clear realist picture of reality (the ontology is that reality is two-level, the wave level and the particle level, and both are real).

I would still rate many-worlds as the best interpretation, but I think backward causation is still a definite possibility. Wormholes and backward causation simply can't be ruled out by current physics; if they are physically possible, this undermines a central argument in favour of many-worlds, namely that it preserves locality.

Based on my current thinking, I'd give the following betting odds for the correct interpretation of QM:

Many-Worlds: 55%

Transactional and/or backward causation: 20%

Bohm: 15%

Other: 10%

[Hide ↑](#)[Hide ↑](#)

**Peter** says:

September 2, 2014 at 8:03 am

"The vision I got from Democritus was of a God who was single-mindedly obsessed with enforcing a couple

of rules about certain types of information you are not allowed to have under any circumstances.” This reminds me of Hawking’s chronology protection conjecture, where physics seems to be very carefully arranged to prevent (non-sub-microscopic) closed timelike curves appearing under any circumstance.

[Hide](#)[Hide](#)

**Ialdabaoth** says:

September 2, 2014 at 11:31 am

*Coditioned on that fact, how likely is it that you’re going to die? One answer is that the dice have a  $1/36$  chance of landing snake eyes, so you should only be a “little bit” worried (considering). A second reflection you could make is to consider, of people who enter the room, what the fraction is of people who ever get out. About  $8/9$  of the people who ever go into the room will die.*

Fun fact: Knowing how many people are in the room with you is enough to change your anticipated probability of survival from  $1/9$  to  $35/36$ .

[Hide](#)[Hide](#)

**Joe from London** says:

September 3, 2014 at 4:48 am

If that is true, and you know that you’re about to find out how many people are in the room with you, then you know that you are about to update to  $35/36$ . So as a Bayesian, you should update \*now\*.

[Hide](#) ↑ [Hide](#) ↑

**tlqryz** says:

September 3, 2014 at 7:31 am

Like others, I am baffled by your claims to be fundamentally bad at math. Let me explain with an analogy.

I have a few thousand hours of practice with chess. Recently, I started playing go, and I find it difficult and frustrating. To me, a chess position makes sense: it can be evaluated globally, individual tactical sequences can be visualized, and the piece structure will be familiar from other games. None of this is true for go, where I have to painfully work through even basic tactics, I have trouble visualizing different board states, and the overall evaluation is difficult or impossible. I make simple mistakes and can only hold a few things in my head at once. Playing chess feels like a lot of different processes are working in harmony: a flow state. Playing go feels stupid.

When I look at a game between chess GMs, while I would not make many of the moves, most of the moves make sense in retrospect, or were among my candidate moves, or can be figured out with some study. Further, I DO look at these games, and do make such a careful study of the moves. Go games don’t make any sense at all, and I have to force myself to look at them and make sense of them. The information also



doesn't feel sticky; it falls out of my head unless drilled repeatedly. When I read a book by a top Go player, it's confusing and it would potentially take me years to work through it entirely, and I would forget some of it by the end.

This is evidence that I'm fundamentally good at chess and bad at go, right? But these games have to be drawing on almost the same brain processes: abstract strategy visualization + evaluation (replace "go" with "shogi" or "xiangqi" if you want an even closer analogue to chess, that still doesn't make sense to me).

I may be naturally better at chess than go. On the other hand, I may be naturally better at go than chess! But when I think about chess, I have years of chunked patterns, concepts, experience, and practice with visualization to do much of the thinking for me. I am effectively much smarter, and so chess seems easier / it seems like I have a talent for chess, relative to other abstract strategy games.

Key:

Chess = Writing

Go = Math

Years of playing chess = years of writing long, thoughtful articles online

Go GM = Scott Aaronson

Of course Scott's book doesn't make sense on the first reading. Of course it may take years of study for the subroutines to be installed into your brain via practice! There likely is a real differential in your talents, but it's more like a coefficient on deliberate practice than like a low ceiling of what your brain is ever capable of understanding, and it's far, far less than it looks when you compare the skill you've been drilling for years, with the one that you see as your weak point.

[HideHide](#)

**ADifferentAnonymous** says:

September 3, 2014 at 10:49 am

The snake eyes killer doomsday argument seems a bit fishy in that the expected number of victims is infinite: the chance that the  $n$ th group is kidnapped is  $(35/36)^{n-1}$ , so the chance that the  $n$ th group is the one that gets killed (not given that they were kidnapped) is  $1/36 * (35/36)^{n-1}$ . The  $n$ th group consists of  $10^n$  people. So the expected number of deaths in the  $n$ th group is  $10^n * 1/36 * (35/36)^{n-1}$ , or equivalently  $10/36 * (350/36)^{n-1}$  and the total expected number of deaths is the sum of the above over  $n$  from one to infinity, which diverges. The expected number of kidnappees who get released will be similarly infinite. So the ratio of them may not be the most meaningful quantity.

Jaynes (as quoted on LW, I haven't read the original) tells us that the correct way to handle infinities is by taking the limits of finite cases. Suppose that there's some  $N$  such that the killer is unable to kidnap more than  $10^N$  people, so the  $N$ th group has to be the last. Most kidnappees will be in the  $N$ th group, so how they get treated dominates the problem. If they'll all get released, you should be optimistic; if they'll all get killed, you should be pessimistic.

What if we know the killer would kill the Nth group, and wouldn't bother rolling the dice? If you're kidnapped and you see him start to throw the dice, you know you aren't in the Nth group. This makes your chances of being killed...  $1/36$ .

So if there's a hard end of the line where you'll definitely die, and you can tell whether or not you're there, this incarnation of the Doomsday argument goes away.

[Hide](#)[Hide](#)

**Doug S.** says:

September 3, 2014 at 9:00 pm

I did read *The Elegant Universe* all the way through and it's mostly a big pile of hand-waving that didn't really tell me much of anything. Which is what happens, I guess, when you write a book about math without actually including any of the math that you're talking about.

[Hide](#)[Hide](#)

**komponisto** says:

September 3, 2014 at 10:30 pm

That seems unfair. It has math in the endnotes, IIRC.

But *The Elegant Universe* isn't a book about math. It's a book about what it's like to be a theoretical physicist. Its first few chapters — the part Scott read, I assume — are also an excellent popularization of relativity and quantum mechanics.

[Hide](#) ↑ [Hide](#) ↑

**Abel Molina** says:

September 5, 2014 at 12:15 am

Really good point about philosophical focusing on qualitative distinctions so much that they forget about huge quantitative distinctions that seem non-trivial to ignore when establishing an ontology of things...

[Hide](#)[Hide](#)