The Philosophy of Computer Science

What did CompPhils do before AI?

Hera Brown

A TALK
presented for the exclusive purview OF the
OXFORD COMPSOC
on this humble night of

May 3, 2025

A common occurrence

There's a conversation I've had a lot over the last three years.

Q: So what do you study at university?

HB: Oh, I study computer science and philosophy!

Q: Oh, wow! Why'd you choose to do that?

HB: Looked fun.

Q: Neat.

Q: **So how do those subjects relate to each other**?

I've never been able to come up with a good answer [1].
Hopefully what I say today is a start.

[1] That doesn't mention AI.

What is a philosophy of computer science?

What might we want out of a philosophy of computer science, then? Well, much like what we'd want out of a philosophy of mathematics, we'd like:

1. A semantics of computer science: what does it mean to say a result is true?
2. An ontology of computer science: if it's a science, then what are we doing science on? What are the objects of our study?
3. An epistemology of computer science: how do we know what's true in computer science?

Let's consider each of these in turn.

A surely true result:

Let's start with a result (helpfully proved with computers) that's true:

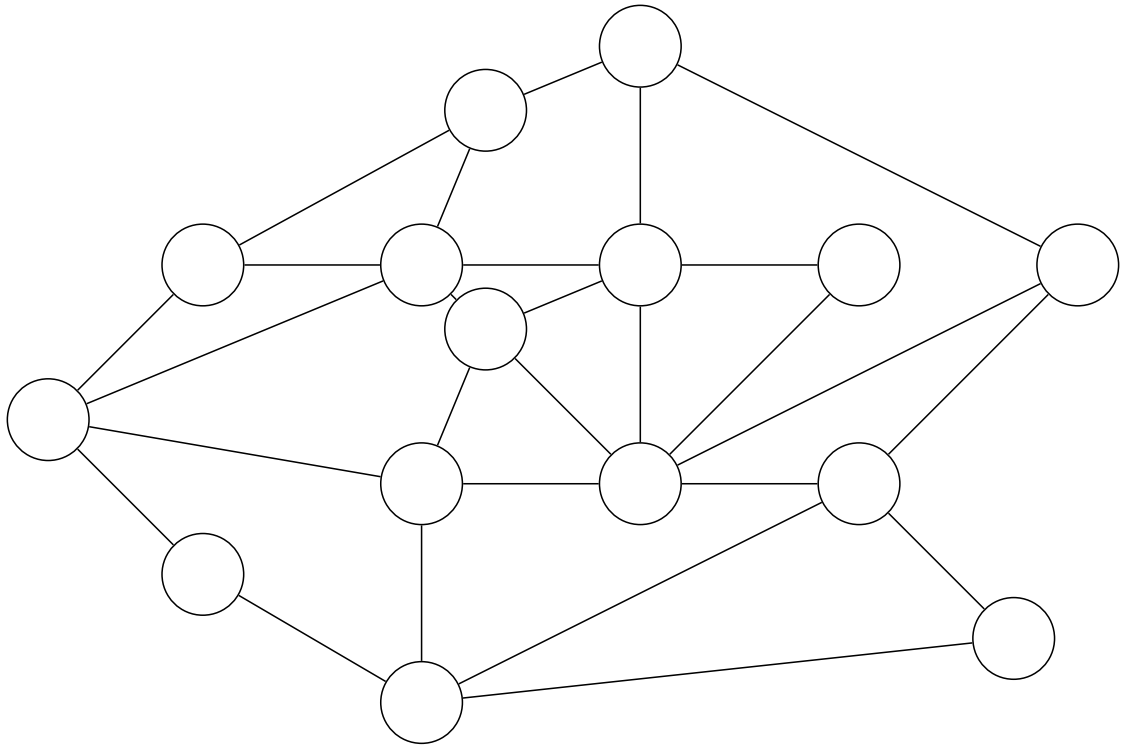
Theorem (The Four-Colour Theorem)

Every planar graph has a 4-colouring; that is, all the vertices of a planar graph can be given one of four colours such that no neighbouring vertices have the same colour.

It's true, right? And when asked why it's true, you'd probably say that it's because we have a proof of the result.
This would be the *intuitionist* position: mathematical statements are true just when we have proofs or witnesses of their truth.
The four-colour theorem, then, is secretly a meta-proof; given a graph, the four-colour theorem provides you a method to construct a four-colouring of that graph.

A problem

The intuitionist account of truth seems (true to their word) intuitive. But we like to make claims about specific computational objects, not just about general classes of them.
For instance, we'd like to say that the following graph is four-colourable:



But to do so, the intuitionist has to provide a witness. So it seems that we don't know whether this graph is four-colourable until we've got a four-colouring!
That doesn't seem right; we shouldn't need to construct an explicit witness. The graph just *does* have a four-colouring, regardless of whether or not we've proved it so.

Still a problem

- The intuitionist can respond by saying that we only need a proof *in principle*. Given enough time and ink, we can of course apply the proof of the Four-Colour Theorem to that graph, and so get a witness to its truth.
- But something still seems wrong. When talking about ordinary language statements — say, that there are three sheep in the field outside, or that there are thirteen compphils in my year — those facts are true just because there *are* three sheep out there, there *are* thirteen third-year compphils.

Figure 1: The Problem.

- For sure a proof of these facts (say, a photograph of the sheep, or census of the compphils) would be useful. But they're not true *because* there are proofs of these statements. Rather, they're true because of facts about objects out in the world.
- This leads us to our second position.

Numbers are real?!

- The *platonist* position holds that computational objects actually exist, and so statements about computer science are just like statements about anything. "There are three sheep outside" is true because there are, indeed, three sheep outside. "There are infinitely many prime numbers" is true because there are, indeed, infinitely many prime numbers.
- To the platonist, numbers aren't like all the other objects of our day-to-day lives, though. They're not like tables and chairs, and they have neither a place nor time at which they exist. Rather, the platonist wants to say that they're *abstract objects*.
- So the Four-Colour theorem was a discovery to the platonists; the planar graphs were all already out there, and we just discovered that they had this nice property.

Figure 2: Your host at leisure.

Trouble in Paradise

- The problem with platonism is how we come to know mathematical facts. We usually come to know about things by perceiving them, or inferring from what we've already perceived. So I know that there are three sheep in a field because I can see that there are three sheep in a field. But how do I know that the four-colour theorem is true, if it's about graphs that I can't see or touch or smell or so on?
- It's an especially worrying problem, since computer science is applicable to the real world. It's no use using an algorithm if you don't know it'll solve the problem you have. So, we care about the truth of statements in computer science; it's a worry that platonism can't explain how we know that they're true!

Many Questions

That's just a brief introduction to the questions we might ask, and the answers we might give. There's a lot more to be considered:

- Are we even asking the right questions about, say, the existence of graphs? Do *graphs* have to exist, or just sets of elements with the right relations to each other? If computer science can be reduced to (complicated) set theory, do we just need to suppose that sets exist? [This is the position of *structuralism*]
- If, in an axiomatised theory, we accept things as true because they're derived from the axioms, then how do we know that the axioms are true? How do we know when to reject an axiom?
- Is it all just in our heads? Is computer science a social construct? Then why do computer scientists all agree on what is and isn't true?
- Why should I care?

These are all important questions.

Figure 3: You.

- In any case, thanks for listening.
- If you're one of my aunts, uncles, or otherwise relatives, then hopefully this answers your question about what it is I do, next time you ask it at Christmas.
- If you're a computer science/maths and cs student, know that it's never too late to switch.
- If you're a compphil and wondering where the hell in the course I learned all this, then you should take the philosophy of maths module. It will be eight brutal, confusing weeks and then you shall be enlightened.

Thanks for listening!

Any questions?