The Mathematicians' Misconception

Transcript of a talk given by David Deutsch at the International Centre for Theoretical Physics, Trieste, Italy, on the occasion of being awarded the Dirac Medal.

March 14, 2018

A *tabula rasa,* with David Deutsch answering questions from the audience after this talk
Photo by Artur Ekert

A couple of years ago, the mathematician Hannah Fry made a TV documentary about Ada Lovelace, the nineteenth-century computer-theory pioneer. It was about an episode in the history of ideas which would have been absolutely pivotal, if anybody had noticed it at the time. Or in other words if Lovelace hadn't died young.

Because, well, from the evidence in that documentary, I suspect that the first human being to *get,* the universality of computation, was actually Lovelace and not her colleague Charles Babbage, the designer *of* the universal computer that she was theorising about: Babbage's Analytical Engine. Never built, but like many of these computers the significance was in the design, the theory, rather than actual building. The thing is, it would have had *two kinds* of universality and Babbage was obsessed with one of them.

He had perhaps been the first human being to understand what one could call *arithmetical* universality. His previous design, the Difference Engine, could compute polynomials in one fixed-point variable. So, a very limited kind of universality – universal for those. But Babbage realised that if he added just a few more features, conceptually very simple, the machine would make the jump to universality, becoming the *Analytical* Engine, universal for *any* arithmetic function, of any number of variables, of *any* finite precision. Basically what we would today call: computable functions. So *this* was *arithmetical universality.*

What Lovelace understood, I think, was the significance of the Analytical Engine's ability to compute: not just any *arithmetic* but *anything – in the world.* In the physical world.

She envisaged all sorts of applications like computer *music* and art and chess and so on. But this wasn't just a matter of usefulness. The abilities of the Analytical Engine as a *physical* object depend on a momentous property of the laws of physics themselves. All of them.

Namely: while the Analytical Engine could instantiate an infinitesimal fraction of all mathematical objects and relationships (what we call the computable ones), it could also – apparently – instantiate (or simulate, emulate) *all* possible motions of all possible physical objects, and their laws. Not just a tiny subset. This *physical* universality is an intrinsic property of the laws of physics. It doesn't follow from Babbage's arithmetical universality. It has nothing to do with mathematics. In fact neither of the universalities follows from the other.

Yet it seemed that both of them were exhibited by the same machine. Why? Well, whatever the reason, it's in the laws of physics. It would make no sense to try to prove it – other than from the laws of physics.

This unity of the two universalities was also conjectured later, explicitly, by Alan Turing in the 20th century: it's just Turing's conjecture. Sometimes called the Church-Turing thesis – it has various names. But the usual way that this conjecture is described is not that it's 'The unity of the two universalities'.

Why not? Well, Turing's great paper presenting his conjecture had an 'application', as he put it, to a fundamental puzzle posed by the mathematician David Hilbert: basically: what is the relationship between a *true* mathematical statement and a *provable* one. Hilbert had hoped that one could define a system of proof, such that a mathematical statement was true if and only if it could be proved under that system.

In the 1930s, mathematicians converged from several directions on the realisation that that is impossible. Notably, Kurt Gödel proved that there can be no method of proof that identifies all true mathematical propositions.

Now, Turing's approach – did the same in *that* respect, but it had wider implications – as we now know, because of *these,* physical, objects: computers. The reason Turing's approach had this additional reach was that: *Gödel's* model of proof was a model inside the arithmetic of the integers. So nothing to do with computation. He simply defined proofs as finite sequences of symbols drawn from a finite set, and all that stuff. But there was no 'Gödel's conjecture'. It was Turing who realised that *that* notion of what proving something means, isn't self-evident: so he acknowledged it as a substantive conjecture, the Turing conjecture. The model of proof that *he* used was *computation.* And the model of computation that he used was *physical.* Strips of paper divided into squares with symbols and a finite set of discrete operations on them. The universal Turing machine. And when he conjectured that this machine was universal *for proofs*, the phrase he used, was that it could compute anything which "would naturally be regarded as computable". At the time, the word 'computer' meant a human being. It wasn't one of these things. Someone whose job was to manipulate symbols on sheets of paper. And the manipulators obeying the rules, human beings. are physical objects too. So, by 'anything that would naturally be regarded as computable', Turing meant: computable *in nature* – by physical objects. And by 'provable', he meant provable by physical objects.

Now, that conjecture, unlike Gödel's proofs, might have been false. But it turned out to be true *in nature*. Or rather, very nearly true. As Richard Feynman remarked: they thought they *understood* paper; but they didn't. And when I proved Turing's conjecture from quantum theory, in 1985, it was with the slight correction that the universal machine is not Turing's paper machine nor Babbage's brass-gear machine, but the universal *quantum* computer.

But I soon found out that not everyone saw it that way. I also had referee problems. The referee of the paper in which I presented that proof insisted that Turing's phrase "would *naturally* be regarded as computable" referred to mathematical naturalness – mathematical intuition – not *nature.* And so what I had proved wasn't Turing's conjecture. It was about physics. So I asked some mathematicians what mathematical intuition is. Turned out, it was as much of a mystery to them as to me. Some of them said it was *meta*mathematical intuition. Fair enough, but they couldn't tell me what that was either. Some kind of mathematical mysticism, I think. But one thing they were all adamant about nevertheless, was that Turing's conjecture was about whether his *mathematical* model of proof matched – not the physical world – but something else. Like *mathematical* intuition or something.

Now, Turing's basic insight, was that proof is computation, and computation is physical, and hence proof is physical. That it isn't physical – seemed to me a philosophical absurdity. But it was an absurdity that all the mathematicians I asked insisted on. And most (not all) – most *non*-mathematicians who'd thought about computation, didn't. So I called it the Mathematicians' Misconception. (The denial that proof is physical – is one way of putting it.)

By the way, Rolf Landauer (Charles Bennett's old boss) had been campaigning for years with the slogan *computation is physical* – and proof also.

Just to be clear: Mathematical *facts* – like Fermat's last theorem – *aren't* physical. That there's a difference between truth and provability was the main point of all those 1930s discoveries.

Still, in my paper, I had to defer to prevailing usage, so I changed it, to define '*Turing's* conjecture' as that vague *metamathematical* idea. And the referee at least agreed to let me call my result a 'proof of the Turing *Principle'* to distinguish it from the conjecture. The Principle that there can be a physical object whose motions contain those of all other objects. Nevertheless, now people sometimes call that the Church-Turing-Deutsch Principle. And that's how, the Mathematicians' Misconception ended up giving *me* credit for something Alan Turing did, and arguably, Ada Lovelace did.

A few years later, I gave a talk in Oxford, arguing that it makes no sense to regard Turing's conjecture, in *any* form, as something one might hope to prove one day from logic, like Fermat's last theorem. But that it could be proved to be a property of quantum mechanics.

Sitting in the front row was Robyn Gandy, who'd worked with Turing. He got a bit agitated, and at the end, he stood up, and declared (with good humour but very emphatically) "I've never heard such a load of rubbish in my life".

I tried to explain further, but he seemed implacable. However, he'd also given a talk at the same event and at the dinner afterwards, he came over to where I was sitting, and said "You know, I think there might have been a grain of *truth* in there somewhere. Let's talk about it later". And we did discuss it later, but unfortunately we did not reach a resolution.

He was a mathematician. He had the misconception.

Unfortunately in the bigger picture, the Mathematicians' Misconception has done more than just cause amusing anecdotes. It expresses the idea, acknowledged or not, that somewhere out there, in the world of mathematical abstractions, or in some supernatural world of mathematical intuition, there is *the* authentic, official, though ineffable (now we know that Hilbert was wrong), *definition of proof.* And if some physical process that doesn't conform to that definition, turns out to allow us to *know* some new, necessary truth, that process wouldn't constitute a proof of that truth. There's the misconception.

It so happens that a quantum computer's repertoire *of integer functions* is the same as the Turing machine's. They differ only in speed. So some people view this as vindicating the Mathematicians' Misconception. But no. First of all, we only know that 'they only differ in speed', from physics, from quantum theory. And second, quantum theory won't be the final theory in physics – and even if it is, *you can't prove that either, from mathematical intuition*. In reality, we only have physical intuition: never provable, always incomplete and full of errors.

The misconception also affects thinking about *information*. For example, a *quantum* cryptographic device may perform a *classical* information-processing task, that is provably impossible classically. So the misconception makes people say 'well, quantum cryptography isn't an *information-processing* task;

it's just an engineering task, like building a washing machine'. Why? Just because Turing machines couldn't perform it! Again, they think there's a mathematical definition of *information* somewhere, independent of physics.

The same holds for probability, by the way.

Similarly again, the answer to Eugene Wigner's famous question about why *mathematics* is 'unreasonably effective in science', is not that the physical world is actually being *computed*, on a vast computer – belonging to God. Or to super-normal aliens – Snailiens. Because there's no reason, other than the Misconception, why the Snailiens' computer should itself generate that particular tiny piece of mathematics *we* call 'computable'.

Purely mathematical intuition will never reveal anything about proof, or computation, or probability, or information. If you want to understand any of those fundamentally, you *must* start with laws of physics. And in particular with what is currently the most fundamental theory in physics: quantum theory. It won't always be the most fundamental. But its replacement will not come from mathematics, logic, or the supernatural.