

# Arthur's Letter

- ARTHUR I have heard of simplexes before. The 3D simplex you describe is a tetrahedron. It is one of the five Platonic solids. Plato thought that these solids governed the harmony of the celestial spheres and made everything happen. You seem to be saying that one of the Platonic solids, the tetrahedron, can compute everything. Do you think that is an original idea?
- JAMES Like all public expressions of an idea, my writing draws on the ideas of others. But, so far as I know, the idea of using a simplex to encode all computations is original.
- ARTHUR Do you think that a Platonic simplex that encodes all computations existed prior to your expression of the idea, so that it is proper to say that you *discovered* it, or did it have no prior existence, in which case it is proper to say that you *invented* it?
- JAMES Like many philosophical questions, that one is many faceted. Let's deal first with physical existence. In my view anything that exists and is capable of change is physical.
- ARTHUR Surely not? Light exists; I can see it and if I sunbathe for too long it can even burn me. Light changes too; the light in the sky changes in intensity from night to day. So light exists and changes, as you require, but it is not physical, because it is not matter.
- JAMES Indeed light is not what physicists call *condensed matter*, but photons of light do have mass, so light is physical in the sense of being material. But that is not the sense in which I mean *physical*. For me a physical thing is anything that exists and which can change in some way.
- ARTHUR Why do you insist that physical things are capable of change? Surely a giant rock that never changes is physical?
- JAMES Suppose a physical observer in the physical universe comes across such a rock. The rock is unchanging, so the observer cannot make any material, energetic, or indeed any physical measurement of it. So far as the observer is concerned it makes no difference whether such a rock exists or not. So, obeying Ockham's razor - the maxim of accepting the simplest explanation that fits the facts - the physical observer should properly conclude that the rock does not exist.
- ARTHUR Suppose I stub my toe on such a rock and it hurts!
- JAMES Assuming that you are physical, then you have falsified my hypothesis that such a rock is non-physical. The fact that such a falsification is possible makes my hypothesis *scientific* in the sense that Karl Popper used the word. That is, my hypothesis can be tested.
- ARTHUR Does tomorrow exist?
- JAMES According to Einstein, yes. So does yesterday. By the way, if they did not exist then time travel in space-time would be impossible. But my impression of the world is that tomorrow does not exist yet, that yesterday no longer exists, and that time travel is impossible. But *now* exists and changes all of the time - excuse that tautology! So *now* is physical in either view.
- ARTHUR Do hobbits exist?
- JAMES No. An idea of hobbits exists in my head and in your head. A similar idea existed in Tolkien's head, but he is now dead. I have seen pictures of hobbits, and soft toys that fit my idea of hobbits. All these exist in the physical universe, but hobbits themselves do not exist. At least, we have not yet found any hobbit fossils dated to the Second Age of Middle Earth, which is when hobbits are supposed to have existed.
- ARTHUR Does the number  $\pi$  exist?

JAMES I believe that the idea of the number  $\pi$  exists in the physical composition of my brain, it certainly exists as marks on paper in mathematical texts, and as bit patterns in digital computers. But  $\pi$  is like a hobbit. It does not exist in itself. Nothing mathematical exists in itself.

ARTHUR Did numbers exist prior to the emergence of mankind?

JAMES Let's talk now about biological existence. In my view this is a subset of physical existence, so I will repeat the above arguments in a new guise.

Chimpanzees can count in the fairly obvious way of being able to make a sign to indicate how many things are present. They are also able to order the signs for numbers in increasing and decreasing order. So, given that Chimpanzees can count, it seems likely that apes that existed before *homo sapiens* could count too.

Venus fly traps can also count. These delightful plants do not close when a sensory hair is bent once, as might happen when wind blown debris touches the plant, but only when a hair is bent a second time, which is most likely to occur when an insect is crawling about on them. The ability to count often saves the Venus fly trap from expending effort on closing when there is nothing to eat on its leaves.

It might be that Chimpanzees learn to count, or that the idea of counting is innate in them. Certainly the idea of counting is innate in Venus fly traps. A Venus fly trap can be grown from seed without ever having an opportunity to learn how to count, yet it does so when catching flies.

Now if the first Chimpanzee that ever counted anything did so while learning to count then it would be proper to say that it *invented* counting. Or, perhaps, *re-invented* counting if some other species had learned the trick earlier. By contrast, if a super smart Venus fly trap were to write an original treatise on counting, by making explicit its innate ability to count, then it would be proper to say that it *discovered* counting, but *invented* the explanation of counting contained in the treatise.

For the record, my intuition is that no human being has been born with the innate idea of encoding all possible computations with a simplex, so I believe that, in this biological sense, I invented the idea.

ARTHUR Let me ask you now about mathematical existence. Mathematicians write things like, "there exists an  $x$  such that  $x + 1 = 3$ ". The answer is that  $x = 2$ , but you deny that anything mathematical, such as  $x$  or 2 exists. In what sense then, is it proper for mathematicians to say that  $x$  exists?

JAMES When mathematicians say that something exists they may properly mean that the idea of the thing exists, as a state of affairs in their brains, as a proof written on paper, as bits in a computer, or whatever, but the thing itself does not exist.

ARTHUR What does a mathematician mean when saying, " $y$  is the set of numbers 1, 2, 3"?

JAMES The mathematician may properly mean that there are ideas of 1, 2, 3 and an idea of a set  $y$  that contains these things.

ARTHUR How can a mathematician have an idea about numbers which do not exist in themselves and then have an idea about the idea of numbers, such as that they can occur in sets?

JAMES In exactly the same way the Tolkien can have the idea of hobbits that do not exist in themselves, and then go on to have an idea about hobbits, such as that they live in houses cut into hillsides. The chain of reasoning can go on, with hobbits holding parties in the houses, being visited by the wizard Gandalf, and setting off on a journey to defeat the Lord of the Rings. Authors and mathematicians can be amazingly creative.

ARTHUR Does mathematical creativity bear on the question of discovery and invention?

- JAMES Yes. A mathematician might discover innate mathematical concepts and make them explicit. Alternatively a mathematician might invent original concepts based on imagination, learning, and innate concepts. I imagine that all mathematical creativity is a mixture of discovery and invention. To borrow a phrase from biologists, I believe that all intellectual creativity is a mixture of nature and nurture.
- ARTHUR That is just a paraphrase of your claims about biological discovery and invention. But I want to know if mathematics itself allows discovery and invention.
- JAMES Mathematics can be created by manipulating symbols, but it makes no sense to say that any particular string of symbols is discovered or invented in any non-physical sense.
- ARTHUR Suppose a mathematician creates a proof based on pre-existing axioms, is it proper to say that the proof has been discovered? Conversely, if a mathematician creates some new mathematical axioms, is it proper to say that the axioms have been invented?
- JAMES A random symbol generator could equally well generate such proofs or axioms, so proofs and axioms are not necessarily different in the manner of their creation that would justify labelling the one discovery and the other invention. And as the proofs and axioms are new in each case, neither had a prior physical existence that would justify describing them as discovered. All mathematics is invented, unless it has some physical pre-existence in which case it can be discovered.
- ARTHUR Let's explore the *potential* for existence. Philosophers have the axiom of *logical omniscience*. That is, if some being knows something, the axiom of logical omniscience means that it knows all logical consequences of the thing. Do the consequences exist? In other words, does a proof exist before a mathematician derives it from pre-existing axioms?
- JAMES Logical consequences do not exist in a being unless they are innate or it thinks of them.
- ARTHUR Can the *possibility* of something exist?
- JAMES If it is thought of then yes, otherwise it does not exist. To say that it is possible that  $x = 2$  is to state that  $x$  might or might not be equal to 2. But this is no different from saying that Frodo might or might not attend Bilbo's birthday party. The possibility does not exist, except in so far as it has been thought of. The thought might then be recorded, perhaps in a book, so that its author or others might more easily think of it.
- ARTHUR Why do you insist on something existing for it to be physical? Isn't it enough just to require that anything that can change is physical?
- JAMES A curve can have a changing slope as defined by differential calculus. The ideas of curves and slopes and differentials are all written down and exist in many people's heads, so the ideas are physical, but a mathematical curve itself does not exist, only the idea of the curve exists, so the curve is not physical. There must be both existence and change for a thing to be physical.
- ARTHUR If the objects of mathematics do not exist in themselves, and the possibility of them does not exist in itself, how can you use mathematical things, such as the simplex, to create feelings and consciousness?
- JAMES I can write down a mathematical formula in a book to measure sensations and act on them. This formula exists in a physical form as symbols in the book, but the formula itself is not physical. For example, the formula cannot change in response to sensations. By contrast, I can write the formula as a program in a computer. Now the program exist in a physical form and can change. For example, the program can change in response to sensations. So it can have feelings and responses to them.
- ARTHUR Let me see if I understand the nature of physicality here. If a mathematician writes on paper the formula for the construction of a house, then the formula exists as physical writing on paper. The paper, and the ink that makes up the formula, could be used as a blanket, a fairly minimal form of

shelter, by someone sleeping rough. But if the formula is executed by a team of builders, or a robot, the result would be a house for our erstwhile rough sleeper to sleep in. I'll grant you that I, and the average rough sleeper, would recognise the difference between the formula on paper and the house. But the house had to be built by skilled builders using building materials. It seems to me that you have not said what physical things constitute feelings, analogous to bricks, and what actions should follow from having a feeling, analogous to laying bricks.

JAMES A digital computer has inherent properties that can be sensed. For example, different parts of computer memory run at different speeds. A computer might measure this difference by fetching identical copies of data from different parts of memory. So speed of access is something that exists in the computer and might be felt by it. In fact, many operating systems do measure things like the amount of time that programs spend in different parts of memory, or the number of errors that have occurred when accessing memory.

ARTHUR I can see that being able to measure such things would have an evolutionary advantage for a robot. It might arrange its programs to respond to the environment exactly as fast as needed, thereby leaving as much processing power as possible to satisfy its internal motives, such as designing better robots. Or it might have some way of repairing faulty memory, or of copying important programs to reliable parts of memory at the expense of less important programs. But surely the ability to measure something is not the same as feeling it?

JAMES In general no. If I measure the distance to the moon using lidar, that is not the same as feeling the distance to the moon. But speed of access exists in the computer, and speed of access is the physical content of the computer's measurement. In an abstract definition of a program, such as a Turing machine, this physical content would not be present, so a Turing machine cannot have feelings, but a physical computer can.

ARTHUR According to Turing, Turing machines carry out one operation in one unit of time, so a Turing program could simulate the delays of accessing memory in a real computer, so the delays would be the content of a Turing machine's feelings.

JAMES No. A Turing machine can specify the delays, but the delays do not occur unless the program is executed on a physical computer, or in a mathematician's head. The computer or mathematician may feel the delays imposed by executing the delaying instructions, but the Turing machine does not exist in itself and cannot have access to any sort of physical feeling.

ARTHUR Can Mary, a fully sighted psychologist who has a theory that describes everything about colour vision, experience the colour red if he lives in a completely black and white part of the world?

JAMES That depends. If Mary is able to use the theory to generate appropriate inputs to his sensors then, yes, he can experience the colour red, or, for that matter, what it feels like to be a woman, or even a brain in a bottle. Alternatively if Mary is human and not suffering from synaesthesia then the theoretical knowledge cannot cause the physical content of red in his retina or visual cortex, so he cannot experience red.

ARTHUR Can a computer see red?

JAMES If the computer has a colour camera as an input device mapped into its memory, then certain patterns in that memory will occur only in the presence of certain combinations of wavelengths of light in particular spatial distributions on the sensor. These patterns will not be able to occur in any other way, so they are the physical content of the computer's sensation of red. Camera designers go to a great deal of trouble to arrange that these physical patterns correspond as closely as possible to human colour perception, so, to the extent that camera designers have succeeded, a computer is able to see and feel red.

ARTHUR Can a computer be conscious of red?

JAMES        If a computer can measure a patch of an image as red, and can attach the word “red” to it, then it is able to obtain linguistic information from the image. Hence it can linguistically articulate its feeling of red. Secondly, if it can also map from the word “red” to red patches of the image, then it can feel what the word “red” means. This ability to map both ways between an image and a label constitute visual consciousness. Thirdly, if it can communicate the existence of red patches to another observer, then it might agree a public meaning for the word “red” that is grounded in its own feeling of red. Fourthly, the computer might surmise that humans have a different feeling of red to itself, based on observing differences in its own and human descriptions of coloured objects. These third and fourth abilities are extraneous to the consciousness of the feeling of red, but are a public indicator of consciousness of the feeling of red and are a private refinement of the meaning of the word “red”. I believe that a computer that can do the first two things is conscious of red. I expect you to assent to that if a computer can do the latter two things. Will you assent to that?

ARTHUR        Maybe. Build such a computer and I’ll test it.

### **Contributors**

James A.D.W. Anderson [Author@bookofparagon.btinternet.co.uk](mailto:Author@bookofparagon.btinternet.co.uk)