Consciousness as Turing Machines

by Sven Nilsen, 2020

In this paper I argue that the only physical existence of approximate Turing complete machines is in the form of consciousness, which suggests the hypothesis that Turing complete machines is the mathematical language of consciousness, which is not necessarily complex information networks.

A pure function in mathematics is deterministic. It contains no noise. When extending pure functions with sources of randomness, one can create non-deterministic functions that are probabilistic.

However, quantum logic requires replacing probabilities with complex probability amplitudes. Such functions are possible to analyze and predict, but only if some pure function, possibly extended with sources of randomness, is introduced to “observe” the quantum system.

The complex probability amplitudes “sums over” the input states of the pure function that corresponds to different outputs. Without a pure function to “observe” the quantum system, there is no way to predict how the quantum system evolves.

The universe itself corresponds to the `id` function that sums over a single complex probability amplitude per output. This preserves the structure of the universe wave function. Well, this picture is not entirely correct, as the complex probability amplitudes rotate over time. This forms a group of physical states. However, for purposes of analyzing the relationship to consciousness, one can ignore the rotation and use `id` instead, since time is not important except for generating the group of states.

So, although the “original” quantum non-deterministic function of the universe is unknown and impossible to construct, one can reason about it in terms of pure functions. It is at the same time crystal clear and confusing. Where did pure functions originate from?

It turns out that pure functions are constructed out of building blocks that relies on the existence of observations. Without observations, there are no building blocks to construct pure functions, hence no way to make new observations, which in turn results in no physical experience of the universe.

Observations are entangled with the state of the observer. Pure functions are only possible, because at least one observer exists, which can assemble pure functions out of previous observations to generate new observations.

Instead of assuming that pure functions are abundant in the universe, one can assume that they are very rare in existence compared to the vast wilderness of quantum systems.

One might assume that pure functions are only found where observers are found, which in turn interprets the world from a view point such that pure functions look abundant. This is because new experiences are generated from previous experiences, such that the only thing observer see is observations. Of course, they look abundant because that is all the observers observe!

If one can imagine pure functions are like natural numbers, then one can imagine quantum non-deterministic functions are like real numbers.

A Turing complete machine is built entirely out of pure functions. It is deterministic, behaving exactly like a pure function, capable of simulating any pure function. This relationship makes it easy to connect the dots from pure functions to Turing complete machines.

A universe filled with quantum systems, with the rarely occurrence of pure functions, will contain approximate Turing complete machines only where there are pure functions. Why? Because Turing complete machines are constructed out of pure functions. This is a mathematical tautology.

This suggests the hypothesis that Turing complete machines is the mathematical language of consciousness. By this I do not mean that they are the same, because Turing complete machines are an idealization. I mean that approximate Turing machines can only exist in the form of consciousness, nowhere else. There is simply not enough pure functions in existence to put them elsewhere.

When observing the world, it looks like it is full of pure functions, therefore one might mistake reality as a place simulated by a Turing complete machine. However, quantum non-deterministic functions are a generalization of pure functions extended with random sources, such that whatever things the observer observes, are quantum systems seen through the lens of pure functions.

This does not imply that every consciousness is an approximate Turing complete machine. Nor does it imply that every approximate Turing complete machine, that the observer sees, is conscious. To make this statement precise, one can use first-order logic:

$$\forall x \{ \text{turing\_complete}(x) \Rightarrow \text{conscious}(x) \}$$

Which is equivalent to the following:

$$\forall x \{ \neg\text{conscious}(x) \Rightarrow \neg\text{turing\_complete}(x) \}$$

The second version is a little more intuitive. If something is not conscious, like a chair or a rock, then it is also not Turing complete. However, a digital computer can be thought of as not conscious, which would conclude it is not Turing complete! This seems like a contradiction, until you remember that digital computers are observed through the lens of pure functions that make up observations.

Since quantum non-deterministic functions can easily simulate approximate Turing complete machines, but only when observed, it is not strange that the world can contain such machines, without implying that digital computers are necessary conscious. They are quantum systems, observed by the observer. The apparent paradox of existence of such machines, is not decidable, but not necessarily a paradox.

Neither does this imply that the observer must consist of complex information networks. A measure apparatus, which has a function, can act as an observer on behalf of a scientist, as long the recorded information, in principle, is possible to extract later. Hence, complex information networks might have nothing to do with observations. For an approximate Turing complete machine, complex information networks are not special. The argument that consciousness arrives in complex information networks is kind of like the argument “I do not understand this, so it must be magic”. I accept no such arguments.

In my opinion, it is far better approach to analyze carefully the equations of quantum physics and interpret its semantics in mathematical logic. However, why are our brains complex information networks? Because that is how nature evolved approximate Turing complete machines into existence!