Why I Am Not a Super Turing Machine

Gert-Jan C. Lokhorst

lokhorst@fwb.eur.nl

Faculty of Philosophy

Erasmus University Rotterdam

UCL, 24 May 2000

Purpose

I will discuss the question: *Is Man a Super Turing Machine* (i.e., a machine which is stronger than any Turing Machine)?

I will be particularly interested in the Super Turing real-valued ARNNs described in Hava Siegelmann, *Neural Networks and Analog Computation: Beyond the Turing Limit*, Birkhäuser, Boston, 1999.

These are equivalent with the biologically more or less realistic "third generation analog neural networks with spiking neurons" studied by Wolfgang Maass and others.

> See W. Maass & C. M. Bishop, editors, *Pulsed Neural Networks*, MIT Press, Cambridge (Mass.), 1999.

I will present all objections to the thesis that Man (or the Human Brain) is a Super TM that I could think of, and I will try to determine whether these objections are valid.

Contents

1 The Bekenstein Bound **4**

2 Discreteness of Space-Time **4**

3 The Question of Empirical Meaningfulness **5**

4 Noise **6**

5 General Philosophical Worries: Strict Finitism **7**

6 Lucas-Type Arguments **9**

1 The Bekenstein Bound

THEOREM (Bekenstein 1981). A spherical region with radius R and energy E can contain only a limited amount of information I (in the sense of number of distinguishable (quantum) states):

$$I \leq 2\pi ER/\hbar c \ln 2$$

where $\hbar$ is Planck's constant and c is the speed of light. Alternatively,

$$I \leq kMR$$

where M is the mass in the region and $k \approx 2.57686 \cdot 10^{43}$ bits/(m kg).

COROLLARY. All physical systems are limited in power to that of finite automata. My body cannot contain a TM [c.q. ARNN with rational weights] or Super TM [e.g., advice TM or ARNN with real-valued weights]. Only a FA fits in.

OBJECTION. The Bekenstein bound was derived in the context of black-hole physics. But we still have no complete theory of quantum gravity.

REPLY. Bekenstein's derivation has withstood 20 years of scrutiny. No counterargument has survived.

2 Discreteness of Space-Time

Many physicists believe that space-time is ultimately discrete. But this issue has not yet been settled.

> See Philip Gibbs, *Event-Symmetric Space-Time*, Weburbia Press, Bristol, 1998.

3 The Question of Empirical Meaningfulness

We can make only a *finite* number of observations of *limited* precision of any physical system. This implies that the claim that system X is a Super TM or TM may not be *empirically meaningful.*

As Gregory Mulhauser wrote
[`http://www.labs.bt.com/people/mulhaug/docs/tutorials/biocognition/biocognition.htm`]:

> Given that the notion of universal computation is defined over an infinite space of functions, it would appear that no finite series of measurements could ever reject the hypothesis that a real physical system was only computationally equivalent to a finite automaton as opposed to some variety of universal computer. In other words, the hypothesis that a real physical system such as a human cognizer has universal computational power does not appear to be empirically meaningful.

NOTE. Even if we were able to make a finite number of measurements with *infinite* precision, this might not help. Let X be some ARNN. Assertion: For every finite running time T, you cannot exclude that X is a finite automaton. Sketch of proof: During T, the neurons assume only a *finite* number of activation values. A net which is the same as X except that it is explicitly stipulated that the neurons can assume *only these* activation values produces the same behaviour during T. But such a net is essentially a finite automaton. [Note that this depends on the assumption that computation proceeds in discrete steps.]

CONCLUSION. Even if the brain were a Super TM or TM, it might well be the case that there is no way to establish this fact.

OBJECTION. There may be *theoretical* reasons to prefer the analog approach (for example, elegance or simplicity). Cf. the description of the orbits of the planets as ellipses rather than epicycles. You can approximate ellipses with epicycles, but the ellipse theory is more elegant and easier to work with.

REPLY. Such theoretical reasons have not yet been presented.

4 Noise

There is a lot of noise in the brain: thermal noise, effects of radioactivity in the environment, unreliability of the neurons, and so on. This has a dramatic consequence:

> Wolfgang Maass & Eduardo D. Sontag, Analog neural nets with Gaussian or other common noise distributions cannot recognize arbitrary regular languages, *Neural Computation* 11 (1999) 771-782.

ARNNs which are subject to analog noise are, in general, just as powerful as "noisy finite automata", i.e., probabilistic finite automata with strictly positive transition matrices. They cannot be made resistant to realistic types of noise.

Objection. As Maass himself emphasizes, not much is known about the way in which information is actually processed in the brain. (His own third generation networks with spiking neurons are just an approximation.)

Reply. These results are very general.

5 General Philosophical Worries: Strict Finitism

TMs and Super TMs are defined in terms of unbounded tapes, infinite series, and so on. But such infinite entities are deeply suspect. There is a new movement in the philosophy of mathematics, Strict Finitism, which is even more distrustful of the infinite than Brouwer's intuitionism, and which tries to do mathematics without assuming any kind of infinite entities. Strict finitists accept *finite* sets of natural numbers, rational numbers and finite precision real numbers (i.e., numbers of the form $m.n$, where $m, n \in \mathbb{N}$), but they do not accept infinite sets and infinite series, real numbers, surreal numbers, transfinite numbers, and so on. For these people, TMs, advice TMs, analog neural nets, and so on, are very problematical constructs. Finite automata, on the other hand, are OK.

> Brian Rotman, *Ad Infinitum. The Ghost in Turing's Machine. Taking God out of Mathematics and Putting the Body Back In.* Stanford UP, Stanford, 1993.

6 Lucas-Type Arguments

Super Turing ARNNs have certain limitations. For example, no ARNN can compute the function $f : \mathbb{R}^2 \mapsto \{0, 1\}$ defined by $f(r_1, r_2) = 1$ if $r_1 = r_2$ whereas $f(r_1, r_2) = 0$ otherwise (Siegelmann, p. 147). Gifted human mathematicians do not have such limitations. Therefore they are more powerful than any Super Turing ARNN.

OBJECTION. The second premise looks suspect, to say the least.

Anyway, it is to be noted that ***Super TM status cannot be an exclusively human property***.

What makes Man different from the rest of the Nature? The human genome. But the human genome contains only a finite amount of information. If man has Super TM capabilities, that cannot be due to the human genome.

ILLUSTRATION in terms of real-valued ARNNs. A real-valued ARNN is equivalent with an advice TM (Siegelmann, ch. 4). An advice TM is a TM which receives an an advice string $\nu(n)$ for each input of length n. [Let $r = \nu(1)\nu(2)\cdots$. r may be encoded as a weight $\delta(r)$, which explains why advice TMs can be simulated by ARNNs.] An advice TM is stronger than any TM if ν cannot be generated from a finite rule. The human genome, which contains only a finite amount of information, obviously cannot contain the code for a function of the latter type. Similarly, it cannot contain a recipe for constructing an uncomputable weight $\delta(r)$.

So if the human brain is a Super TM, then this cannot be due to the Human Genome. The source of the uncomputability must lie elsewhere, for example in physics or chemistry. In other words, Super TM Status cannot be an exclusively human property.

Objection. The human brain may be the only device which harnesses the uncomputability which is to be found in Nature [Penrose?].

Reply. How? This sounds like Cartesian dualism, which says that the Human Brain is the only organ that is subject to the influence of an Immaterial Soul. Just as Cartesian dualism did not explain why the human brain is the only organ that is in contact with an Immaterial Soul, so it is hard to understand why the human brain should be the only device that is in touch with naturally occurring uncomputable processes.

Conclusion

There is no reason to suppose that Man is a Super Turing Machine or even a Turing Machine. The Finite Automaton Hypothesis is sufficient.

> A philosophy of mind built on this foundation is to be found in R. J. Nelson, *The Logic of Mind*, Reidel, Dordrecht, 2nd edition, 1989.

Objection. Lucas-type arguments. Man can reach certain insights that no TM, let alone any finite automaton, can reach.

Reply. Lucas's arguments have often been shown to be invalid. I have never understood why some people find them attractive. Minsky wrote as follows in his book *Computation: Finite and Infinite Machines* (1967), in a passage which is similar to the passage from Mulhauser we have quoted, and which is concerned with the computational capacities of discrete nets without circular signal paths:

> We have seen that, even with this restriction, we can realize any stimulus-response behavior function that is duration-bounded in the manner indicated in our discussion. [...] One should not conclude that the behavior patterns so realized are necessarily uninteresting or unimportant. We must admit that, at least in principle, one could so build a machine which would behave in every

> respect like some particular man, for the lifetime of that man. That is, one could (in principle) arrange things so that for each sequence of stimuli the "man" might ever receive, his proper response would be elicited (p. 55).

Even Lucas's own behavior could be accounted for in this way, so even in his case there is no reason to think of ultra-Turing intellectual capabilities.

These slides are available at
`http://www.eur.nl/fw/staff/lokhorst/hypercomputationUCL.html`