

Are Minds Computable?

Carlos Gershenson

*Departamento de Ciencias de la Computación
Instituto de Investigaciones en Matemáticas Aplicadas y en Sistemas
Universidad Nacional Autónoma de México
A.P. 20-726, 01000 México D.F. México*

Abstract

This essay explores the limits of Turing machines concerning the modeling of minds and suggests alternatives to go beyond those limits.

Keywords: computability, mind, Turing, interactions

One of the main open questions in science and philosophy since ancient times has been how bodies relate to minds. René Descartes proposed his famous dualism to ontologically separate the physical (*res extensa*) from the mental (*res cogitans*). However, with Cartesian dualism one cannot relate bodies and minds. This has led many people to propose ways of reducing minds to the physical realm (e.g. Ashby, 1947).

Alan Turing (1950) was also interested in describing minds using only mechanical principles. Universal Turing machines (Turing, 1937) were proposed to decide whether a *mechanism* could produce a desired computation. Thus, one can transform the mind-body relation into the mind-mechanism relation (Johnson-Laird, 1987; Boden, 1988), avoiding the ontological part of the mind-body problem. This can be formulated with the question: *Can universal Turing machines compute a process similar to a mind?* Turing himself proposed his famous test to decide this (Turing, 1950; Proudfoot, 2011). This question is relevant for Artificial Intelligence, since a large part of A.I. research models aspects of human minds on machines with an architecture inspired by Turing machines.

Email address: cgg@unam.mx (Carlos Gershenson)

URL: <http://turing.iimas.unam.mx/~cgg> (Carlos Gershenson)

Stuart Kauffman (2012) has recently argued that Turing machines are not able to compute minds. After noting several limitations of Turing machines, Kauffman—following others such as Penrose (1989)—suggests an alternative in quantum mechanics. Here I defend that, even when Turing machines are not sufficient for computing minds, quantum mechanics is not required to provide a more viable explanation. The crucial difference between Turing machines and human minds lies in *interactions* (Gershenson, 2011).

Universal Turing machines (UTMs) have been a central concept in computer science and A.I. (Denning, 2010). UTMs are used to define computable functions. They are one of the theoretical foundations of modern computers. Still, they have limitations. Since A.I. has been working with computers based on the concept of a UTM, it is affected by the same limitations. Alternatives to classical A.I. (e.g. Rumelhart et al., 1986; Brooks, 1991) also suffer from these limitations, since these do not depend on representations (Gershenson, 2004).

The main limitations of UTMs—related to modeling minds—are two:

1. UTMs are closed.
2. UTMs compute—i.e. produce an output—only once they halt.

UTMs are closed because once an initial condition and a program are set, the result will be deterministic. However, there are many phenomena where data (in the tape or program) changes at runtime. One example of this can be seen with coupled Turing machines (Copeland, 1997; Copeland and Sylvan, 1999). Certainly, if the interactions and updating scheme between Turing machines (where the computation of one machine affect the data of another machine) are deterministic and defined precisely, these could be modeled by a UTM. Nevertheless, in most cases these are only known *a posteriori*, because of computational irreducibility (Wolfram, 2002). Thus, one cannot define the output of an “open” Turing machine, since the inputs during runtime are unknown *a priori*. As opposed to UTMs, minds are constantly receiving inputs.

Computations carried out by UTMs have to halt before an output can be extracted from them. However, there are many computations that do not halt (Wegner, 1998; Denning, 2011), for example, biological computation (Mitchell, 2011) is not about computing a function, but a continuous information processing that sustains living systems (Maturana and Varela, 1980). Minds also fall within this category. A mind does not compute a function and halts. A mind is constantly processing information.

From the above arguments, it can be concluded that UTMs cannot compute minds. Thus, A.I. as it stands now cannot model minds completely. Still, this does not imply that minds cannot be computed. Now the question is: *Are there (non-Turing) mechanisms capable of computing a process similar to a mind?* If we see computation as transformation of information (Gershenson, 2007), then human minds are already computing processes. Thus, the answer is affirmative—human minds can be seen as computers—but not very useful. Minds are computable, but how?. A more pragmatic framing of the question would be: *How can we describe computations of processes similar to minds?* This description should enable us to understand better cognitive processes and also to build more sophisticated artificial cognitive systems.

As it was argued above, *interactions* (Gershenson, 2011) are the missing element in UTMs. Probably this is due to a reductionist bias in science. Classical science, since the times of Galileo, Descartes, Newton and Laplace, has tried to simplify and separate in order to predict (Kauffman, 2000). This is understandable, since traditional modeling is limited by the amount of information included in the model. This has naturally led to neglecting interactions. However, modern computers—a product of reductionist science—have enabled us to include much more information in our models, opening the possibility to include interactions.

Computing with interactions (Wegner, 1998) implies models of computation that are open and that do not halt. Interestingly, modern computers—while based on UTMs—are able to perform interactive computation. It should be noted that Turing-computability—which is theoretical—is different from practical computability (in a physical computer). For example, there are Turing-computable functions that are not computable in practice simply because there is not enough time and/or memory in the universe. On the other hand, computers can compute non-Turing-computable functions, such as halting functions with the aid of an “oracle” (a posteriori). Interactions are another example, since computers may receive inputs while processing information and may not halt. This is positive for A.I. since it implies that modeling minds with current computers does not imply a hardware challenge.

The remaining challenge is related to reductionism. By necessity, many A.I. systems rely on interactions, be it with humans or with other artificial systems. Examples include conversation systems (Pineda et al., 2007) and agents for the Semantic Web (Hendler, 2001). However, even these systems simplify modeling to reductionist processes, where allowed interactions are

defined beforehand, and computations are well defined. This contrasts with natural systems—humans included—where there is a constant interaction with the environment (Clark, 1997) that is not predefined. The ability to respond to unforeseen circumstances seems an obstacle for A.I. Still, there are several examples of adaptive systems that go beyond this limitation (e.g. Braitenberg, 1986; Bongard et al., 2006).

A.I. can describe minds in terms of computation only if this is dynamic (changing in runtime) and continuous (non halting). This will enable artificial systems to respond to interactions which can occur at any moment and do not stop occurring.

To achieve dynamic computation, systems should be able to be interrupted and modified. New states should be able to be created and old states should be able to be deleted.

To achieve continuous computation, systems should be programmed with the goal of constant interaction, as opposed to producing a set of outputs. Systems should not deliver results. Systems should deliver processes.

Dynamic, continuous computing will certainly be less predictable than traditional computing. Still, it will have the possibility of being more creative and adaptive, since new solutions will be constantly sought as novel situations arise.

References

- W. R. Ashby, The Nervous System as Physical Machine: With Special Reference to the Origin of Adaptive Behavior, *Mind* 56 (221) (1947) 44–59, URL <http://tinyurl.com/aqcmdy>.
- A. M. Turing, Computing Machinery and Intelligence, *Mind* 59 (236) (1950) 433–460.
- A. M. Turing, On computable numbers, with an Application to the Entscheidungs problem, *Proceedings of the London Mathematical Society* 2 (42) (1937) 230–65, doi:10.1112/plms/s2-42.1.230.
- P. Johnson-Laird, How could consciousness arise from the computations of the brain, in: *Mindwaves*, Blackwell, Oxford, 247–257, 1987.
- M. A. Boden, *Computer models of mind*, Cambridge University Press, 1988.

- D. Proudfoot, Anthropomorphism and AI: Turings much misunderstood imitation game, *Artificial Intelligence* 175 (5-6) (2011) 950 – 957, ISSN 0004-3702, doi:10.1016/j.artint.2011.01.006, URL <http://www.sciencedirect.com/science/article/pii/S000437021100018X>.
- S. Kauffman, Answering Descartes: Beyond Turing, in: S. B. Cooper, A. Hodges (Eds.), *The Once and Future Turing: Computing the World*, Cambridge University Press, 2012.
- R. Penrose, *The Emperor's New Mind: Concerning Computers, Minds, and the Laws of Physics*, Oxford University Press, Oxford, 1989.
- C. Gershenson, The Implications of Interactions for Science and Philosophy, Tech. Rep. 2011.04, C3, UNAM, Mexico, URL <http://arxiv.org/abs/1105.2827>, 2011.
- P. J. Denning, Ubiquity symposium 'What is computation?': Opening statement, *Ubiquity* 2010, ISSN 1530-2180, doi:10.1145/1880066.1880067, URL <http://doi.acm.org/10.1145/1880066.1880067>.
- D. E. Rumelhart, J. L. McClelland, the PDP Research Group (Eds.), *Parallel Distributed Processing: Explorations in the Microstructure of Cognition*, MIT Press, 1986.
- R. A. Brooks, Intelligence without representation, *Artificial Intelligence* 47 (1-3) (1991) 139 – 159, ISSN 0004-3702, doi:10.1016/0004-3702(91)90053-M, URL [http://dx.doi.org/10.1016/0004-3702\(91\)90053-M](http://dx.doi.org/10.1016/0004-3702(91)90053-M).
- C. Gershenson, Cognitive Paradigms: Which One is the Best?, *Cognitive Systems Research* 5 (2) (2004) 135–156, URL <http://dx.doi.org/10.1016/j.cogsys.2003.10.002>.
- B. J. Copeland, The Broad Conception of Computation, *American Behavioral Scientist* 40 (6) (1997) 690–716, doi:10.1177/0002764297040006003, URL <http://abs.sagepub.com/content/40/6/690.abstract>.
- B. J. Copeland, R. Sylvan, Beyond the universal Turing machine, *Australasian Journal of Philosophy* 77 (1) (1999) 46–66, doi:10.1080/00048409912348801, URL <http://www.tandfonline.com/doi/abs/10.1080/00048409912348801>.

- S. Wolfram, A New Kind of Science, Wolfram Media, ISBN 1-57955-008-8, URL <http://www.wolframscience.com/thebook.html>, 2002.
- P. Wegner, Interactive foundations of computing, Theoretical Computer Science 192 (2) (1998) 315 – 351, ISSN 0304-3975, doi:10.1016/S0304-3975(97)00154-0, URL [http://dx.doi.org/10.1016/S0304-3975\(97\)00154-0](http://dx.doi.org/10.1016/S0304-3975(97)00154-0).
- P. J. Denning, Ubiquity symposium: What have we said about computation?: closing statement, Ubiquity 2011 (2011) 1:1–1:7, ISSN 1530-2180, doi:10.1145/1967045.1967046, URL <http://doi.acm.org/10.1145/1967045.1967046>.
- M. Mitchell, Ubiquity symposium: Biological Computation, Ubiquity 2011, ISSN 1530-2180, doi:10.1145/1940721.1944826, URL <http://doi.acm.org/10.1145/1940721.1944826>.
- H. Maturana, F. Varela, Autopoiesis and Cognition: the Realization of the Living, D. Reidel Publishing Co., Dordrecht, 2nd edn., 1980.
- C. Gershenson, The World as Evolving Information, in: Y. Bar-Yam (Ed.), Proceedings of International Conference on Complex Systems ICCS2007, URL <http://uk.arxiv.org/abs/0704.0304>, 2007.
- S. A. Kauffman, Investigations, Oxford University Press, 2000.
- L. A. Pineda, V. M. Estrada, S. R. Coria, J. F. Allen, The obligations and common ground structure of practical dialogues, Inteligencia Artificial, Revista Iberoamericana de Inteligencia Artificial 11 (36) (2007) 9–27, URL <http://tinyurl.com/3wlao2r>.
- J. Hendler, Agents and the Semantic Web, Intelligent Systems, IEEE 16 (2) (2001) 30–37, ISSN 1541-1672, doi:10.1109/5254.920597, URL <http://dx.doi.org/10.1109/5254.920597>.
- A. Clark, Being There: putting brain, body, and world together again, MIT Press, Cambridge, MA, 1997.
- V. Braitenberg, Vehicles: Experiments in synthetic psychology, MIT Press, 1986.

J. Bongard, V. Zykov, H. Lipson, Resilient machines
through continuous self-modeling, Science 314 (5802)
(2006) 1118–1121, doi:10.1126/science.1133687, URL
<http://dx.doi.org/10.1126/science.1133687>.