Why Machines Will Never Rule the World: Artificial Intelligence without Fear
by Jobst Landgrebe & Barry Smith

Walid S. SABA*

Institute of Experiential AI (EAI) at Northeastern University

Journal of Knowledge Structures & Systems

December 2022
Vol.: 3 Issue: 4 Pages: 38-41

WHETHER IT WAS John Searle's Chinese Room argument (Searle, 1980) or Roger Penrose's argument of the non-computable nature of a mathematician's insight – an argument that was based on Gödel's Incompleteness theorem (Penrose, 1989), we have always had skeptics that questioned the possibility of realizing strong Artificial Intelligence (AI), or what has become known by Artificial General Intelligence (AGI). But this new book by Landgrebe and Smith (henceforth, L&S) is perhaps the strongest argument ever made against strong AI. It is a very extensive review of what building a mind essentially amounts to drawing on insights and results from biology, physics, linguistics, computability, philosophy, and mathematics.

Although the discussion in many sections of the book can be quite involved the overall argument is quite simple: Anything we engineer (a computer or any other machine) must ultimately be a system that can be modelled mathematically. That is, any engine we engineer is in the end a logical system that can be formally modelled and described by the mathematics available to us. The mind, however, and as L&S argue, is not a logical system but a dynamic complex system that no known mathematics can model or describe. L&S reject the mind-body duality and believe mental processes are themselves physical processes. Moreover, the complexity of modeling these mental processes is not simply a function of their complex temporal or stochastic behavior; rather, it is due to the fact that these processes are dynamic, adaptive, continuously evolving, and constitute systems whose behavior affects and is affected by the environment they function in. According to L&S this is the source of limitations of modern-day machine learning techniques: While one can "train" a deep network on a set of input-output pairs, beyond any narrow domain no set of training data can adequately predict the future environment since the state of that environment itself

*✉ walid.saba@gmail.com

is a function of the very system that we are training! For example, no amount of training data can be enough to train autonomous vehicles since the rules of the entire autonomous driving ecosystem might change based on the collective behavior of the autonomous agents themselves. This cyclical cause-and-effect behavior of complex systems cannot be modelled by any known mathematics.

Another interesting point L&S make is the notion of *granularity*. For L&S, it is complex systems all the way up: From specific components of the mind (e.g., the language faculty) to the mind itself, to the larger and more complex system, namely the animal living organism, and all these systems at various degrees of granularity are complex systems that no known mathematics can model. It is important to note here that L&S argue these systems are beyond engineering regardless of the underlying computing machinery. In fact, L&S argue that quantum computers, for example, while they may provide faster execution of some algorithms, they do not expand the space of computable functions, the limits of which were already set by the Church-Turing thesis.

Of particular interest to me was the somewhat detailed discussion of language and the complexities in open interactive dialogues. First, L&S consider language – and correctly so, in my opinion – to be a prerequisite to any AGI. (They therefore think Alan Turing got this right!) They then go on to argue that linguistic communication is itself a complex system that no mathematics can model, and thus, again, no AGI is possible. Like the situation of autonomous driving, in a real dialogue the interpretation of some utterance must be a function of previous utterances and the overall context that has been built so far. But since responses cannot be predicted in any meaningful way, the overall context is not well defined, and so the entire interaction cannot be mathematically modelled. L&S do not preclude the development of interactive bots in narrow domains, where responses and thus the overall context can be predicted to a certain extent. I personally do not subscribe to this argument: Language is an all-or-nothing process and no amount of circumscription of the domain will help us in predicting the next response/utterance a participant might make.

The argument of L&S mentioned above that the mind or some faculties of the mind (like language) are complex systems that are dynamic, adaptive, continuously evolving, and are systems whose behavior affects and is affected by the environment they function in, is very compelling. The argument that the behavior of such systems is beyond any *known* mathematics is also very convincing. I do, however, have two reservations as to the conclusions L&S make based on these two observations. First, "never" is a very long time, and if the behavior of such systems is beyond any known mathematics, I am not sure one can argue that AGI is thus impossible for there could be a *new* mathematics that mental processes require that is yet to be discovered. This, of course, is the Kantian point of view argued in his *Logic* as "everything in nature, as well in the inanimate as in the animate world, happens or is done according to rules, though we do not always know them" (Kant, 1819). Second, and while some complex behavior cannot be mathematically modelled, this also does not preclude the possibility of "building" such systems. In fact, we can, as the late John McCarthy once remarked, build systems that we cannot understand. Indeed, one can write a program in the *intensional* programming language LISP that would change its behavior at run-time. The behavior of such programs cannot be mathematically modelled (since they can introduce self-reference and thus will be reduced to the Halting function), yet such programs can be executed on our von Neumann machines. The fact that we are not able to model such programs with our known mathematics does not imply that we cannot build them. As Ord (2002) argues, if one considers hypercomputation one can conceive of a new mathematics that would change our view of the Church-Turing thesis and of what is computable.

Finally, I expected this book to have a brief discussion of *the frame problem in AI* (McCarthy & Hayes, 1969), a problem that touches on issues related to the nature of belief revision in complex systems, an issue that in my opinion has not yet been adequately addressed, at least mathematically or algorithmically, yet is a problem the solution of which will undoubtedly shed some lights on the whole discussion of what we can or cannot engineer.

In the end, and whether or not one believes we could one day develop the mathematics that would be needed to model the complex mental processes of our species, the arguments L&S make are very convincing and are, at a minimum, very sobering. At a time when unprecedented resources are being spent on AI systems that oversimplify and over-promise, resources that will be (and in some cases have been) thrown "down the drain," it is very important to consider issues related to AGI not from the vantage point of an (commercially) eager and overexcited engineer, but from the vantage point of what is physically, philosophically, and mathematically possible. In this regard, this book is a must for anyone working on, or interested in, building intelligent machines.

References

Kant, I. (1819). *Logic.* Transl. by J. Richardson. London: W. Simpkin & R. Marshall.

McCarthy, J. & Hayes, P. (1969). Some philosophical problems from the standpoint of Artificial Intelligence. In B. Meltzer & D. Michie (eds.), *Machine intelligence. Vol. 4* (pp. 463-502). Edinburgh: Edinburgh University Press.

Ord, T. (2002). Hypercomputation: Computing more than the Turing machine. Available at https://arxiv.org/abs/math/0209332

Penrose, R. (1989) *The emperor's new mind: Concerning computers, minds, and the laws of physics.* Oxford: Oxford University Press.

Searle, J. (1980). Minds, brains, and programs. *Behavioral and Brain Sciences*, *3*, 417-457.

Cite this article as:

Saba, W. S. (2022). *Why Machines Will Never Rule the World: Artificial Intelligence without Fear* by Jobst Landgrebe & Barry Smith. (Book Review). *Journal of Knowledge Structures & Systems*, *3*(4), 38-41.

EDITORIAL INFORMATION

Editor-in-chief:[a] Luis M. Augusto

[a]Book reviews are reviewed only by the editor-in-chief for appropriateness of content and style.