

A Broader, More Inclusive Definition of AI

Peter Stone

PSTONE@CS.UTEXAS.EDU

*Department of Computer Science
 University of Texas at Austin, Texas*

Editors: Dagmar Monett, Colin W. P. Lewis, and Kristinn R. Thórisson

Pei Wang’s article “On Defining Artificial Intelligence” (Wang, 2019) is a thoughtful and well-written argument in favor of a particular working definition of AI and an associated research project called NARS, for Non-Axiomatic Reasoning System.

It is structured in three parts. First, it argues in favor of the need to define Artificial Intelligence. Second, it argues in favor of Wang’s particular definition. Third, it explains how his definition leads to NARS.

I focus my comments here on the first and second parts. While there is much to agree with in the article, in the interest of discourse, I further focus my comments on the points with which I disagree. That being said, I fully acknowledge that it is much easier to criticize than it is to write an article with no room for criticism!

1. A Much Broader Definition

From the highest level perspective, I agree with Wang’s exposition of the values of specifying one’s working definition, and commend him for acknowledging on more than one occasion that there is room for different definitions. But despite this acknowledgment, I note that on more than one occasion he seems to argue for the need to converge on a single definition or the superiority of his own definition, neither of which I endorse. Personally, I hold strongly to the “big tent” view of AI that allows, and even encourages, multiple perspectives and agendas, and thus working definitions, to co-exist within the same field. It is with this view that I prefer a broad definition such as the one we put forth in the 2016 report of the One Hundred Year Study on AI (Stone et al., 2016):

“Artificial Intelligence (AI) is a science and a set of computational technologies that are by inspired by—but typically operate quite differently from—the ways people use their nervous systems and bodies to sense, learn, reason, and take action.”

2. The need for a Definition

Section 1.3 of the article presents a very useful discussion on what a definition is, and Section 2 lays out an interesting classification and generalization of the various types of AI definitions.

However in my opinion, the article over-reaches in a few ways. For example:

“Though a well-defined concept is not easy to obtain, its benefits are hard to overstress. It will prevent implicit assumptions from misleading a research project.”

While having a definition may indeed help focus or guide one’s research, I would not go so far as to say that research is “misguided” if it is not tied closely to a particular definition of AI.

Even for research that does start from a definition, Wang writes:

“In particular, the definition distinguishes the features of human intelligence that need to be reproduced in an AI system from those that can be omitted as irrelevant.”

In my opinion, this statement leads directly to the need for a plethora of working definitions, so that as a field we can investigate a broad range of the features of artificial intelligence. In fact, the statement requires different working definitions over time as knowledge and tools progress. Just as there are now considered to be different types of human intelligence (e.g. spatial intelligence, emotional intelligence, etc.), the field of AI has room for, and indeed requires, investigations of machine intelligence from various perspectives.

A few more minor points that bear mentioning follow.

- One justification raised for needing a single definition is so that policy makers can assess what AI systems will be able to do in the future. On the contrary, I think it is incumbent on AI researchers to stress that AI is *not one thing* and should therefore not be regulated as such. Policies ought to be developed sector by sector with regards to specific AI-based technologies that are relevant to that sector (see the AI100 report for further discussion on this point).
- Another justification put forth for a definition of AI is so that we will know “how to build one.” I disagree that AI is a “thing” to be built, and again it is certainly not *one* thing.
- The definition of Capability-AI takes an applications-oriented perspective, but then seemingly limits AI research in this paradigm to *matching* human performance. It ought to leave room for superhuman performance being realized by AI-based systems, as we have seen from recent game-playing systems.

In summary, I agree with the author regarding the usefulness of working definitions for helping focus one’s research. But I caution that definitions can also be exclusionary, and object to attempts to use narrow definitions as justification for limiting the field by dictating what “counts” as AI. The author is correct that the inclination to coin terms such as “AGI” has arisen to differentiate from AI research that is more narrowly focused. However, I disagree with the need to differentiate in this way. The term AI, and the field of AI can, and do, encompass both narrow and broad research foci and applications.

3. Wang’s Definition

As for Wang’s working definition itself, I think it is perfectly fine as “a” definition of AI. However I do not endorse it as “the” definition.

Actually, I do not find that the definition stands alone. Rather, to fully understand it requires reading its explanation throughout the 2.5 pages of Section of 3.2. For example, the phrase “adapt to its environment” does not necessarily lead to any requirements over beliefs, actions, tasks, or problems. And the meaning of “insufficient knowledge and resources” only becomes clear through the prose that follows. This need for extensive explanation violates at least the “exactness” desideratum of a good definition.

In any case, the author's definition is very well-paired with his research program and vice versa. It is an elegant coupling that is indeed commendable and worthy of emulation. However the leap from there to statements such as the following goes too far.

“The current field of AI is actually a mixture of multiple research fields, each with its own goal, methods, applicable situations, etc., and they are all called AI mainly for historical, rather than theoretical, reasons.”

In my opinion, the field of AI can tolerate, and in fact actively benefits from, research projects and perspectives that arise from a variety of working definitions, or that are even not directly tied to working definitions at all. It is for that reason that for the purpose of defining the field, I much prefer the AI100's much broader definition, as quoted above.

References

- Stone, P.; Brooks, R.; Brynjolfsson, E.; Calo, R.; Etzioni, O.; Hager, G.; Hirschberg, J.; Kalyanakrishnan, S.; Kamar, E.; Kraus, S.; Leyton-Brown, K.; Parkes, D.; Press, W.; Saxenian, A.; Shah, J.; Tambe, M.; and Teller, A. 2016. Artificial Intelligence and Life in 2030. One Hundred Year Study on Artificial Intelligence. Stanford University, Stanford, CA. <http://ai100.stanford.edu/2016-report>.
- Wang, P. 2019. On Defining Artificial Intelligence. *Journal of Artificial General Intelligence* 10(2):37.