DID SOMETHING GO WRONG?

by John McCarthy, Stanford University

Abstract

The premise of this symposium seems to be that something went wrong with AI. Otherwise things would be better. What would be better isn't stated. From my point of view what would have been decisively better is to have achieved human-level AI. However, I suspect some people of being disappointed that the AI companies didn't make enough money.

Here are some considerations.

1. Understanding intelligence well enough to make a human-level artificial system is a very difficult scientific problem. That's not surprising a priori. The physicists haven't solved turbulence after more than 100 years, and it was 100 years from Mendel till the genetic code was understood. The biologists still can't make a cell, although they apparently can make a virus.

2. Still, AI might have been done better if someone smarter and more energetic than us AI researchers had worked in the field, say an Einstein.

3. Here are some specific things that went wrong.

Almost all AI researchers concentrated on themselves deciding what phenomena to take into account and building systems that reasoned about exactly those phenomena. This was true of all the applied systems and most of the basic research systems, maybe all that were actually implemented.

Here's an example of the limitations. Suppose a system is following a script based on a housewife shopping for dinner. She or it suddenly asks itself how many steaks to buy. It depends on whether her son is coming to dinner, which in turn depends of whether his airplane flight will arrive on time. She or it switches temporarily to a script involving phoning the airline's flight information number and interacting with its voice operated system.

4. My own recipe for reaching human-level AI still, after 50 years, involves expressing commonsense knowledge and reasoning in languages of mathematical logic. I see many more requirements than I saw 50 years ago. These include formalized nonmonotonic reasoning, reasoning with approximate entities without if-and-only-if definitions, contexts as objects, individual concepts and propositions as objects, heavy duty axiomatic set theory, and the ability to reason about and improve its own reasoning methods.

We also need the ability to understand human language well enough to make systems that can learn by reading texts on the Internet. The difficulty is much more a matter of semantics than of grammar.