

Artificial Intelligence: Its Scope and Limits, by James Fetzer, Kluwer Academic Publishers, Dordrecht, Boston, London.

Artificial Intelligence (AI) is the study of how to make machines behave intelligently, to solve problems and achieve goals in the kinds of complex situations in which humans require intelligence to achieve goals. AI has been studied from a number of points of view.

1. Since the best intelligence we know about is embodied in the human nervous system, we can study human and animal neuroanatomy and neurophysiology and try to build something sufficiently similar to be intelligent. Steady, slow progress has been made in this study, but it hasn't yet led to understanding of human problem solving, and there aren't yet any physiologically based theories of higher mental functions. Part of the problem is that neurons are universal computational devices, and therefore knowing more about how they work doesn't limit what complex structures can do computationally.

2. Some people, the connectionists, look for general principles of neural organization for intelligence and try to make intelligent systems out of simplified neurons.

3. Human intelligence is also studied on the psychological level, e.g., with experiments in which human subjects solve problems.

4. Expert system technology is based on the fact that much expert practice is based on large numbers of rules-of-thumb—“*if you see this do that*”. Thus Mycin proposes a diagnosis and recommends a treatment for a bacterial infection of the blood on the basis of symptoms and the results of tests. It knows nothing of bacteria as organisms that grow and reproduce and sometimes emit toxins. In fact Mycin knows nothing of processes occurring in time. It is not evident how general knowledge about events in general and bacteria in particular could be used in a Mycin-like program. Expert system technology has turned out to be useful for carefully selected (or just fortunately selected) problems. Some enthusiasts believe that human level intelligence is just a matter of having enough rules of thumb. Fetzer writes that the problem of expert systems is choosing the right expert. That's rarely the problem because expert system work almost always concerns areas where the experts agree. The problem is the fundamental limitation of collections of rules-of-thumb.

5. The logicist or formal reasoning approach to AI (I do that) studies the kind of “common sense informatic situation” in which intelligence solves problems. It expresses in mathematical logic facts about what information

and actions are available for achieving goals and uses logical inference to decide what to do. Common sense is the center of its domain, and it sees scientific theories, e.g. physics, as embedded in common sense knowledge and reasoning. Its needs have led to formalizing* nonmonotonic reasoning.

6. Contiguous with the logicist approach to AI are philosophical studies of epistemology and theories of mind. These concern what knowledge is, how it is to be obtained and how it is possible for (a human) to know. The many schools of epistemology and philosophy of mind take different views of artificial intelligence. These range from active cooperation in formalizing phenomena of AI interest (e.g., knowledge and causality), through Daniel Dennett's "*intentional stance*" that ascribes mental qualities in a way quite congenial to AI, to total disdain for the possibility of intelligence in non-biological systems.

Because mind has always been central to philosophy, philosophical schools have developed opinions about what constitutes knowledge, cause and mind. These opinions have led many philosophers to be boldly prescriptive about what AI researchers can and cannot do or should and should not do. Such prescriptions have been about as influential in AI as they have been in other sciences.

Philosophers also engage in conceptual analysis resulting in distinctions, and sometimes AI researchers find these distinctions useful. However, sometimes what purports to be merely a conceptual distinction turns out to make empirical claims, and sometimes these claims are false. One series of such inadvertent claims concerned the incorrigibility of reports of pain and was discussed by Daniel Dennett (1978). Another was made by John Searle, who asserted (1990) that a Chinese dialogue could equally well be interpreted as the score of a chess game or stock market predictions by a person unfamiliar with Chinese. This contradicts the experience of cryptography, the decipherment of ancient languages and (most likely) algorithmic information theory.

I bet you thought I'd never get around to talking about the present book.

Fetzer recounts the history of AI research, surveys the fields of AI makes some conceptual distinctions and makes some empirical claims. Although he discusses using logic for representing information, he ignores the logicist approach to AI—thus offending this reviewer.

There is also a considerable difference in emphasis between the AI community and Fetzer's "*... the most fascinating question about AI remains whether or not machines can have minds*". We would put the question as

“*What intellectual mechanisms are there, which are required for which tasks, and how can a computer be programmed to implement them?*” This difference is already apparent in the foreword, and Fetzer seems unaware that anyone would dissent from the way he structures the issues.

The book covers very many topics related to AI, and I find myself in disagreement in so many places that I’d have to write a whole book to deal with all of them. It’s especially hard because I mostly don’t like his formulation of the questions rather than merely disagreeing with the answers.

However, there’s a main contention, and I’ll deal with that. We have on page xiv:

“I shall argue that . . . digital machines are symbol systems, while humans are semiotic; that only semiotic systems are possessors of minds; and that the prevalent conceptions of AI are based on mistakes.”

I take this sentence as asserting that no digital machines are semiotic systems.

When someone asserts that AI cannot do something, there is usually an intuition behind it concerning what computer programs are like, perhaps about how they must represent information. If the person has paid some attention to the AI literature, then the intuition may, more or less, apply to the programs he has read about. Often the intuition can be taken as a challenge, however imprecisely expressed, to extend the capability of AI programs.

Here’s an outline of Fetzer’s main argument. He introduces a distinction between *symbol* systems and *semiotic* systems. Symbol systems are taken to be what Newell and Simon call physical symbol systems. Next he argues that intelligence requires semiotic systems. He then seems to declare that symbol systems are not semiotic systems. We understand symbol systems, e.g., programs that compute with symbols and also their extensions that generate symbolic expressions from inputs and affect the external world by their outputs. So what are semiotic systems and how are they different?

Semiotic systems involve symbols, the things they denote and the being for whom they denote it. The same symbol can denote different things for different people. When a symbol in a machine denotes something, Fetzer requires that we distinguish between its denoting the thing for the user of the machine and denoting it for the machine itself. I gather he contends the latter doesn’t happen and can’t happen. An example is that “chair” denotes chairs for Fetzer and me but doesn’t really denote chairs *for* the database system of the chair manufacturer.

Fetzer doesn't make the idea more precise, but I would conjecture that he is concerned with this fact that when a person uses "chair" to denote chairs he has another concept of chair to which he relates the word "chair". A pocket calculator may have no other concept of 7 to which it relates the 7-key and the 7-display, although it has an internal representation of numbers which is distinct from these.

I agree that this is a worthwhile distinction, but I think we can make computer programs do it also. It will be important when we make computers introspective in significant ways, i.e., make them use sentences about their own knowledge or lack thereof. Maybe it can be said to happen already in a computer vision or robotic construction program that relates chair to recognition and construction criteria.

On p. 50, Fetzer says

"Moreover, there appear to be several unexamined alternatives with respect to Newell and Simon's conception, since other arguments might be advanced to establish that symbol systems properly qualify either as semiotic systems of Type I or Type II or else that special kinds of symbol systems properly qualify as semiotic systems of Type III, which would seem to be an important possibility that has yet to be considered."

By p.52 Fetzer seems to have rejected the possibility, saying; *"By combining distinctions between different kinds (or types) of mental activity together with psychological criteria concerning the sorts of capacities distinctive of systems of these different kinds (or types), the semiotic approach provides a powerful combination of (explanatory and predictive) principles, an account that, at least in relation to human and non-human animals, the symbol-system hypothesis cannot begin to rival. From this point of view, the semiotic-system conception, but not the symbol-system conception, appears to qualify as a theory of mind."* I can't find anything between these pages that even purports to supply an argument.

Fetzer (following C.S. Peirce (1839-1914)) puts semiotic systems in three categories—iconic, indicative and symbolic. A statue is iconic, because it looks like the object; a smell of burnt coffee is indicative, because it indicates that someone has left an almost empty coffee pot on the burner; and "chair" is purely symbolic. He points out that Newell and Simon don't treat these iconic and indicative cases in their definition of physical symbol system, and therefore argues that their concept doesn't cover semiotic systems in general. It seems to me that Newell and Simon could easily extend their concept to cover those two cases, because they are even easier to treat than the purely

symbolic. It is hard to believe that Fetzer relies so much on what seems to be a quibble.

Fetzer makes another distinction between symbolic and semiotic systems.

Symbolic systems can simulate, i.e., match the i-o behavior of another system whereas semiotic systems replicate, i.e., match the internal behavior as well. This has no obvious relation to the previous distinction.

Moreover, semiotic systems can make mistakes whereas symbolic systems can only malfunction. Distinguishing mistakes from malfunctions is sometimes worthwhile. Indeed any system doing nonmonotonic reasoning can make mistakes, i.e., reach a wrong conclusion because it failed to take into account an important fact. However, nonmonotonic reasoning systems are symbolic.

Another quibble concerns symbols. Newell and Simon use strings of characters as symbols and not the characters themselves. Fetzer seems to think symbols should be elementary.

Finally, let's consider the extent to which AI is ready for Fetzer's implicit challenge to represent the Peircian *semiotic* relation among a symbol, an external object and a person. To deal with persons, we propose to use the notion of formal context now being intensively studied in AI. There isn't yet an exposition of what has been discovered, but (McCarthy 1987) describes a preliminary approach. R. V. Guha's forthcoming Stanford dissertation will do much more. The idea is to write $ist(c, p)$ to assert that the sentence p is true in the context c . Contexts can be associated with persons in order to handle the third term in the semiotic relation, but they have many other uses.

One basic approach to the formal logic of semiotics involves axiomatizing a term $meaning(term, context)$, where *context* in general involves a person and a situation. In many cases, however, many of the axioms hold in contexts covering many persons and situations. In those (very common) cases, a more conventional meaning theory can be used.

I think AI isn't very far along in meeting the challenge which Fetzer's book partly expresses. Unfortunately, there isn't yet a clear understanding of what the semiotic challenge is, whether from Fetzer or from within the AI community.

With all this, even completely refuting Fetzer's argument that digital machines can't have minds wouldn't constitute much progress. Progress requires consideration of specific mental abilities and how to program the computer to use them. **References**

Dennett, Daniel (1978): “Why You Can’t Make a Computer that Feels Pain”, in *Brainstorms*, Bradford Books.

McCarthy, John (1987): “Generality in Artificial Intelligence”, *Communications of the ACM*. Vol. 30, No. 12, pp. 1030-1035. Also in *ACM Turing Award Lectures, The First Twenty Years*, ACM Press, 1987.

Searle, John (1990):