AI, Science, and Intellectual Processes: Preliminary Remarks and Arguments

By J. Terry Nutter

TR 90-38

AI, Science, and Intellectual Progress: Preliminary Remarks and Arguments*

Extended Abstract

J. Terry Nutter
Department of Computer Science
Virginia Polytechnic Institute & State University
Blacksburg, VA 24061

This paper argues that trying to answer questions like the precise relationship of AI to existing disciplines (psychology, philosophy, linguistics, etc.) is both premature and potentially harmful to all concerned. This is *not* to say that we cannot say anything useful on the underlying questions which plague those who question its status; but the answers that can currently be given will probably fail to satisfy critics and proponents alike. This dissatisfaction -- and indeed much of the debate -- results from a view of science which takes as its model mature, developed sciences, and ignores facts about necessary phases in their development. This paper also argues that the question whether AI is a science is usually standing surrogate for concerns that have nothing whatever to do with science, and which should be addressed on their own grounds. The fundamental thesis here is that understanding the current status of AI, and so understanding the relationship between the various proposed approaches, requires adopting a more sophisticated approach to the status and development of intellectual disciplines, and that such an approach can contribute substantially to a broad area of current disputes, including most notably the "traditionalist/connectionist" controversy.

The paper is organized around four questions: (1) "What is AI?"; (2) "Is AI science?"; (3) "Why do we care?"; and (4) "What does all this say about AI, cognitive science, and art forms?" The discussion of the fourth question will deal, among other things, with consequences concerning the relationship between "traditional" and connectionist AI. I outline my positions on these four questions below.

1. "What is AI?"

At this point, this question does not have a well defined answer. The best we can do is take a sociological approach (AI is what the AI community says it is). Under that approach, it depends on who you ask. There are at least four separable views current in the AI community. Under the first of these, AI is a discipline which tries to find ways to get computers to solve problems which, at least ostensibly, require intelligence for humans to solve them. We usually (but not always) restrict these to problems which are not strictly numeric (a statistical package is not an AI program, though we would normally say that it would take intelligence for a person to perform the same calculations). In these terms, AI has two aspects. One is a theoretical aspect, but has little or nothing to do with cognition. It is a branch of computer science

* A talk based on this abstract was presented at the AAAI-sponsored Workshop titled "Artificial Intelligence: Emerging Science or Dying Art Form?" , 21-23 June 1990, Binghamton, New York.

which studies algorithms for certain classes of problems, just as any other applied branch of the field does (operating systems, compiler design, etc.). As with those other branches, AI also has an engineering side, concerned with pragmatic issues in implementing the algorithms and methods studied by its theoretical side. It is no more a science-less engineering field than, say, aerospace or electrical engineering, both of which have perfectly respectable underpinnings in physics. But it can be viewed as computer engineering. This view I will call algorithmic AI (A-AI).

The second view holds that AI aims at producing models of intelligence. On this view's terms, it does not matter whether the intelligence in question is human-like or not. All that matters is that we would consider that the things modeled are legitimate aspects of cognition. This view I will call nonspecific-cognitional AI (NC-AI).

The third view holds that AI tries to develop adequate models of aspects of human cognition. As a rule, these models are intended to be extendible to all of human cognition, but -- as in any normal academic discipline, scientific or not -- they are generally approached incrementally. This view is deeply concerned that the models formed are valid models of cognition as it takes place in humans, and not just intelligent ways to do things. This view I will call human-cognitional AI (HC-AI).

Finally, there is the view that AI aims at producing artificial agents. This can be approached from either of two directions. One could claim that it is not enough to simulate cognition: we must actually reproduce it, and this is something above and beyond modeling it (or simulating it) in just the same way that producing a car engine is something over and above modeling or simulating one (a model of an engine cannot be put under a hood and used to propel a car). Or one can take the view (normally called "hard AI" these days) that because cognition is computational in nature, a computational model/simulation of intelligence *is* intelligent.

This view has two distinct versions, depending whether it takes its view of cognitive agents from the concerns embraced in NC-AI or those reflected by HC-AI. That is, while the aim is to produce genuine agents, these agents may or may not be synthetic ***human*** intelligences. Whichever of these versions we embrace, and whichever argument we produce in favor of embracing either one, the goal is to develop systems that *actually* see, thing, feel, and so on, in the same *sense* (whether or not by the same mechanisms) as we do. This family of views I will call synthetic AI (S-AI).

There has been a great deal of dispute over which of these we are willing to countenance as AI, or which of these are legitimate areas of research. At this point, most of that dispute is misguided, and the rest is premature. What is important at this stage is that enterprises coming under any of these headings know which one they are under, and that they are evaluated relative to their particular goals and intentions.

2. "Is AI science?"

What do we mean by "science"? If we go to philosophy of science (the best place to look for definitions of "science"), we find a plethora of accounts of what science is, with the most widely accepted views probably still those of Popper and Kuhn. Popper's definition is, roughly speaking, that a discipline is scientific provided that its theories are in some manner and to some degree empirically testable (falisfiable).

By this test, three out of four of the views above succeed (all but S-AI), and that the problems S-AI has are so deep that to rule it out probably also rules all all enterprises part of whose domain is people.

Kuhn's test is less explicit and more complex. It has two parts: a sociological judgment, and a life cycle that follows a particular sort of pattern. Under the first test, yes, of course AI is science (ask the NSF; etc.). Under the second and more sensitive test, AI is a *young* science, specifically a preparadigmatic science. This preparadigmatic state has crucial consequences, including a strong suggestion of potentially disastrous consequences for cutting off *any* line of investigation too early.

3. "Why do we care?"

What is this debate *really* about? The answer to this probably has nothing whatever to do with whether something is a science, and equally has little to do with AI in particular. This same debate has taken place over the social sciences, and in one form or other over most new discipline. The real issue is usually a dispute over resources, and that dressing it up as a dispute over what is and is not science is misleading at best and harmful at worst. This is not to say that resource allocation should not be questioned and examined. It *is* to say that questions and examinations should be explicit and open, and not buried behind other issues. We should also remember that not everything that deserves study (and resources) is a science: few reputable scholars would want to terminate all work in the humanities. Rather than asking whether AI is science, engineering discipline, or something else, we should be asking what are the consequences of those categorizations, how they shoul affect resource allocation, how they would promote intellectual development, and why the answers matter. But whatever AI is, it is *not* an engineering discipline with no corresponding scientific ground. *All* engineering disciplines have corresponding scientific grounds (perhaps in one discipline, perhaps in more); and if the ground of engineering AI is *not* (called) scientific AI or cognitive science or some similar title, then it must be something else. It cannot be nothing.

4. "What does all this say about AI, cognitive science, and art forms?"

First, there is no evidence that AI is dying, or that in a reasonable sense it should be. If AI *does* die (or stop growing) in the near future, it will not be because it ran out of results or made no progress, and it will certainly not be because it was found not to be a science. Compared to other disciplines at corresponding points in their life cycles, AI has made astounding progress. The contrary appearance results from a combination of impatience and unrealistic expectations, most significantly the unexamined belief that every science springs forth from the forehead of Zeus with a full blown, fully developed paradigm attached. Arguments in favor of a different view will be brought from the history of science.

Second, there are good reasons for treating AI and cognitive science as different but closely related fields, and nurturing both. This does not commit us to any one of the four views of AI above: no matter which view is taken, there are still reasons to separate the two in principle, so that while many cognitive scientists may turn out also to be AIers, and vice-versa, there will always be some scientists in each field who are not in the other. This again is a perfectly normal situation with regard to other sciences, and should occasion no particular remark here.

Third, preparadigmatic sciences are not like normal science (in the Kuhnian

sense; some arguments will also rely on N.R. Hanson's analysis). This fact has consequences with regard to the current view of the relationship between "traditional" AI and connectionist models. Briefly, there are two different credes of connectionism. One view holds that these models have a great deal to say about certain aspects of cognition, and that their study will prove interesting and revealing in ways that other current models do not. The other is the crede according to which connectionism is the one true model which should replace all others. It is crucial to the development of AI that the field as a whole both embrace the first crede and reject the second. The argument for rejecting connectionism-as-the-one-true-faith is probably the more controversial (though these days, it's hard to tell). It will include three subsidiary arguments: (a) an argument in favor of viewing cognition as multi-layered in ways that connectionism at least today cannot model, but combinations of connectionist and traditional AI may be able to; (b) an argument against the view that even S-AI necessarily requires building artificial *brains*; and (c) an argument that the attempt to force a single model on a science in its preparadigmatic phases is probably the fastest and most effective way to kill it.