AI Magazine Volume 8 Number 1 (1987) (© AAAI)

WORKSHOP REPORT

Derek Partridge

Workshop on the Foundations of AI: Final Report

The Workshop on the Foundations of AI (WFAI) was held at the Holiday Inn, Las Cruces, New Mexico, on 6, 7, and 8 February 1986. Financial support for the workshop came from the National Science Foundation; the American Association for Artificial Intelligence (AAAI); and the Computing Research Laboratory (CRL) at New Mexico State University, which also hosted the meeting.

My original vague idea for this workshop was backed enthusiastically by CRL right from the start, first by Roger Schvaneveldt as acting director and later by Yorick Wilks when he took over as director. Andrew Ortony played a leading role in both casting and production for this workshop; he claims that he doesn't love telephoning people, just doesn't mind it. These three and the rest of the program committee, as well as a number of other people, reviewed the considerable number of submitted papers. Production support was provided by CRL, particularly Dorothy DeLena, Jeannine Sandefur, Azzie Partridge, and Patty Lopez. Within the computer science department, Art Karshmer and Don Dearholt assisted me in numerous ways.

WFAI was attended by approximately 40 nonlocal invitees from both academic and industrial institutions. In addition, approximately 40 local attendees listened to the sessions. Each day was organized around five half-hour presentations interspersed with discussion periods of much the same duration.

Derek Partridge is principal investigator in the Computing Research Laboratory and a professor at New Mexico State University, Las Cruces, New Mexico 88003. After May 1987 Dr Partridge will be at the Department of Computer Science, University of Exeter, Exeter EX4 4PT, United Kingdom This report supersedes all previous reports; any traces of which should be totally erased from memory

A proceedings for this workshop was not generated and speakers were not required to produce anything more than an extended abstract in hard copy form Copies of the WFAI preprints, which range from complete papers to somewhat underextended abstracts, and a list of speakers with the titles of their papers are available from Computing Research Laboratory, 3CRL, New Mexico State University, Las Cruces, New Mexico 88003, or from CSNET: Rebecca@nmsu In addition, many workshop participants are contributing to *A Sourcebook on the Foundations of AI*, edited by Partridge and Wilks, which is to be published by Cambridge University Press

Foundations of AI: Why We Might Want to Dig for Them

Throughout the three full decades of its history, artificial intelligence (AI) has been a contentious topic, and with the recent explosive growth of commercial AI, many basic issues have ceased to be of purely academic concern. Thus, in addition to its intrinsic academic value, an exploration of the foundations of AI might uncover aspects of the field that will have significant economic impact.

Within the AI community, the perceived need for an examination of the foundations of AI is widespread. For example, John McCarthy in his president's message to AAAI in 1984, entitled "We Need Better Standards for AI Research," maintained that the criteria for evaluating AI research are not in very good shape. On the other side of the Atlantic, the *AISB Quarterly* throughout 1983 and 1984 was home for a debate between Alan Bundy and Stephan Ohlsson on various aspects of AI methodology; despite their differences, both agreed that the methodology is a mess. If research standards and methodology are in serious trouble, then it appears to be none too early to initiate an examination

Abstract This report makes a case for the need to examine the methodological foundations of AI. Many aspects of AI have not yet developed to a point of general agreement. The goals of AI work, the methods for acheiving these goals, the presentation of results, and the assessment of claims are all highly contentious issues. All aspects of AI methodology are subject to debate The Workshop on the Foundations of AI was conceived as a forum in which such a debate could proceed This report presents the rationale behind the event, the details of the program, and finally some afterthoughts.

of the foundations of AI. The field is rife with basic problems, and this situation in itself is the major problem.

Generally, we are witnessing wholesale adoption of the label AI for all things computer related. The term AI, which is meant to proclaim that the system so described possesses exciting new qualities, is in danger of becoming a hackneyed label which has lost all currency. AI does have something to offer the commercial world, but it would be to nearly everybody's benefit if this something were clarified; for example, which AI techniques hold a promise of immediate or near-term commercial viability, and which aspects of AI are still perceived to be difficult research topics. Perhaps it does not matter that the exact nature of AI is ill defined, but it seems important both for the integrity of the discipline of AI (if it is one) and for a healthy applications domain that the current scope and limitations of commercial AI be spelled out to some degree (definition would be too much to hope for). However, prior to any hope of significantly clarifying the commercial possibilities of AI, we must have some agreement on the basics of AI itself.

. . . the intent . . . was to better expose and articulate some fundamental AI problems by closing in on them from several different directions . . .

The primary emphasis for the workshop was an exploration of the foundations of AI. Individual sessions were composed of both invited position papers and selected submitted papers. We also attempted to allow for ample discussion time.

The workshop was organized into three subtopics. Within each subtopic, a number of example issues were identified. This set was not intended to be exclusive; it was merely a partially organized catalog, a selection of specific issues of the type this workshop was designed to address.

The first subtopic was philosophical and logical foundations of AI. Presented in two sessions, topic discussions were chaired by Maggie Boden and Yorick Wilks, respectively. Questions for discussion included the following:

- What are AI's purposes, goals, and epistemological and ontological presuppositions?
- Is AI still maturing or already decomposing into subfields?
- In what sense are there AI theories, and what kind of theories are they—hacking and hypothetical deductive method?
- In what sense is AI scientific?

The second subtopic was relationships between foundations and programs. The two sessions were chaired by Ron Brachman and Roger Schvaneveldt, respectively. Discussion qestions included the following:

- How is an AI program to be evaluated?
- What kinds of goals can an AI program achieve?
- Can AI's goals be realized on single-processor computers?
- Does it make sense to say a program is a theory?
- Is it possible to describe AI programs?

The third subtopic was relationships between AI and other disciplines. Questions for discussion were presented in two sessions chaired by Len Uhr and Andrew Ortony, respectively. These questions are the following:

- What contribution can other disciplines make to progress in AI and vice-versa?
- How does cognitive science relate to AI?
- Is AI a separate discipline, or is it just a way of approaching many other disciplines?
- How can we make AI like a neighboring science?
- Should we try?
- Are AI discoveries rediscoveries of matters well known in central computer science or other neighboring disciplines; is AI just applied computer science?

Even these "specific issues" are still uncomfortably broad. The main danger for this workshop, as I saw it, was the fact that the foundations of AI are an ill-defined (some might say empty) topic; we risked generating either a series of high-level, rather vacuous exchanges or a collection of highly personal viewpoints that did not relate to one another. I trust that we avoided the worst outcomes. Nevertheless, although the dangers were clearly there, I believed that the importance of the problems merited an attempt to make some progress with some of them.

We had presentations from people who are clearly pro-AI and from others who have quite the opposite leaning, let us say, con-AI. Additionally, there were also people who might still be wondering what they are expected to do at this workshop, the non-AI group. Our intention was certainly not to pit pro against con and stand back to watch the sparks fly. The intent behind collecting this distinguished but obviously not homogeneous group of speakers was to better expose and articulate some fundamental AI problems by closing in on them from several different directions, including directions from which AI is viewed as something of a nonevent.

The three subtopic areas constituted our original wish list. As in all the best fairy stories, these wishes came largely true, although not without some coaxing and cajoling—destiny should not be left entirely to its own devices. In addition to the scheduled presentations, a number of attendees provided position papers that for one reason or another could not explicitly be included in the program. Where possible, I

reproduced these papers in the preprints in an effort to provide a comprehensive background to the discussions.

The workshop was organized around the three major topic areas, one topic per day. The first day was used to examine different aspects of the logical or philosophical underpinnings of AI. We had six very different but mainly critical views of basic AI issues.

On the second day, we tried to focus somewhat; we looked mainly at the role of programs in AI. It is perhaps the centrality of machine-executable constructions—loosely, programs—that distinguishes AI work (if it is to be distinguished) from related disciplines. Computer science might appear to have the strongest claim to these intriguing objects (computer programs), but the fundamental medium here is really machine-independent abstractions (algorithms). AI hinges on the workings of piles of code (despite the manifest dissatisfaction that this news causes, I think it is largely true). We questioned the role of programs in AI: Are there a number of rather distinct roles? Should we classify AI work on this basis? If a program is a major product of AI research, how can it be effectively communicated to and, hence, evaluated by others?

The final day was concerned with the relationship between AI and other disciplines. Is there just an AI branch, say, to psychology or linguistics? Is AI a discipline apart, a subject in its own right? Is the structuring of AI better managed in terms of related disciplines than in terms of, say, the role of the program?

A major current that I found in most of these papers was guarded optimism (as opposed to outright condemnation) for what AI has already achieved as well as what it might achieve in the future if only certain problems can be overcome. However, there was not a great deal of agreement about what these problems are, except perhaps that they are all basic issues—problems with the foundations of AI. A second theme which appeared repeatedly was that AI really resembles the conventional sciences much more than it appears at first sight. The apparently discrepant features of AI are just exaggerations of similar features in the conventional sciences. As such, they serve to highlight the realities of the classical sciences, realities that are often overlooked or disregarded. In sum, AI with its apparent idiosyncrasies can be used as a magnifying glass to facilitate examination of the realities of the conventional sciences.

Reflections on WFAI

After the event I collected a few musings from some members of the program committee.

Andrew Ortony

I suppose my main reaction to the workshop was that it turned out much as I expected it would and as most such workshops do. That is, the informal interactions and discussions seemed to me to be more valuable and interesting than the formal presentations. Such a turn of events is not necessarily a bad thing; after all, researchers from some area will always need some occasion or context in order to conglomerate, and it might be that workshops, conferences, and the like serve no real intellectual function beyond providing the occasion for such a conglomeration.

I suspect that you noticed at least a hint of disappointment in the last paragraph. I guess there is a hint because I had hoped that some interesting and important issues might be discussed in the formal part of the proceedings. However, my disappointment was very mild; hoping is not expecting. Specifically, I think that there are a number of key issues having to do with the status of AI (science versus engineering), its relation to theories (can programs be theories and if so when), and the nature of experiments in AI that could have been discussed to the benefit of the field.

To take just one of these issues, I believe the science versus engineering issue is a red herring and only arises because some people take an elitist view that only pure science is worth doing. Two interesting things, one pragmatic and one mythical, mitigate this pompous view. Oxford University practices AI in a department called Engineering Science (that's where Brady is). One presumes that after Ryle and others, Oxford University knows about category mistakes and that it doesn't consider engineering science to be one. Second, prophets of doom for AI are like those who, having seen Icarus's failure at heavier-than-air flight, might have given up because of the lack of real progress in the subsequent 5, 10, 50, 500, or 2000 years. They all would have been wrong: Following the engineering triumphs came the development of the sophisticated science of aeronautics. We should liken AI to heavier-than-air flight and aeronautics, not to chemical engineering as Roger Needham suggested in his talk.

Although missing what I considered to be intelligent presentations on these issues, I must say that I was impressed by the degree and depth of some of the informal discussions. Clearly, the most frequently and deeply discussed topic was connectionism. This topic is filled with interesting issues for those intrigued by foundations (and anything else). It was probably a mistake of ours in planning WFAI to overlook it. Part of the problem is that there clearly is a need for education on the topic. Many experts in the field appear to misunderstand exactly what connectionism is and what motivates it, at least according to the claims of its protagonists. There are big differences of opinion about whether connectionism is simply a style of implementation, whether and in what ways it is fundamentally different from symbolic representations, and so on. I learned a lot (although still not enough) about the topic just by listening to such debates, which alone made WFAI a thoroughly worthwhile event for me.

Leonard Uhr

I'll concentrate on several issues that were raised briefly (but largely ignored) which I think are of crucial importance.

. . . Perhaps when we have AI systems operating at the level of Cro-Magnon man doodling on the cave wall in his spare time, it will be appropriate to add a veneer of logical reasoning capability . . .

First, what is AI, and what is it about? Taken broadly, AI is the study of how to get intelligence into computers. This statement means that AI must look at and try to understand human and animal intelligence, the only examples of intelligence around. Without them, we wouldn't even have come up with the idea. AI, then, is the science of intelligent information processing. Just as physics has its natural side (meteorology, acoustics, and so on) and its constructive side (engineering), AI has its natural and constructive sides. Ideally then, AI should be developing the theoretical tools and concepts for psychology and much of the cognitive sciences and the neurosciences as well as establishing principles and techniques for achieving intelligent artifacts. Such is clearly not the case today probably because AI is still in a prescientific, pretheoretical stage. Hacking and bludgeoning away with bigger computers and an AI flavor to things might produce some useful programs but it might slow down the development of a general understanding.

Second, after thirty years of published papers about AI programs, AI has not even begun to develop the necessary methodology and canons for presenting, evaluating, and comparing programs. Maybe this statement is too extreme, but a field where people keep questioning whether the program that a paper describes actually can handle anything other than the one or two examples given in the paper, even runs, or has even been coded needs better methods. Clearly, the program, as well as the results that have been achieved and how these results compare with those of other programs, should be described unambiguously. The total domain of behavior the program is asserted to handle should also be described along with reasons for thinking why it should be able to handle these things. The crucial piece of information that makes it worth reading about the program is almost always a comparison, one that shows how and why this program is better than some other program of interest, or possibly some animal, at a significant task. We also need to understand how general the program is, how difficult its tasks are, and why they are significant.

Roger Schvaneveldt

At the time, I was somewhat surprised to find so much of the conference devoted to the connectionist program when the primary topic was foundations. In retrospect, however, it appears that the connectionist approach does conflict with the standard AI symbolic program. From a psychological perspective, the advancement of the connectionist view is not particularly surprising. Elements of the idea have been around for years, particularly in the area of pattern recognition. In recent years, the view that concepts are represented by collections of "exemplars," rather than some abstract entity such as prototypes or schemata, has received increasing support from the empirical literature and from theories based on exemplar storage. The phenomena that we formerly thought of as requiring abstract representations are explained by exemplar models of processes that accumulate information from collections of exemplars. This accumulation produces effects that are much like the effects expected from abstract representations.

The point is that several of the connectionist ideas were brewing within psychological models for some time. What might be new is the brash rejection of the need for any kind of symbolic or abstract representations. (In weaker moments connectionists give maybe 10 to 15% to symbolic computation). Traditional AI approaches might well be threatened by such claims.

Other impressions from the conference included a general feeling that within the traditional approach, the major foundational issues for AI have not changed very much over the course of AI history. (Maybe that's why they are foundational!) My own reaction to arguments about foundations is that we should be aware of them, but they are unlikely to be settled by argument. The definition of the field of AI will accompany its development, the problems it attacks, and its successes and failures. Compare the history of experimental psychology: The very objectives of the field have changed over the phases of introspectionism, behaviorism, and cognitivism.

Conclusion

Quite shamelessly (as Yorick would say), I'll give myself the last word. It seemed to me that truth and proof got a drubbing at the workshop and about time too. Absolute truth and associated abstractions are not notions that are usually associated with the fundamentals of intelligence. Perhaps when we have AI systems operating at the level of Cro-Magnon man doodling on the cave wall in his spare time, it will be appropriate to add a veneer of logical reasoning capability (assuming that it has not already emerged as an epiphenomenon). Whatever else it might be, the essence of intelligence is not prolonged, step-by-detailed-step logical reasoning, which, of course, is not to say that intelligent reasoning cannot be implemented in terms of some yet-to-be-discovered system of logic. It's the surface-level behavior that I'm taking issue with; I have nothing to say about how this behavior might be realized (at

least not until the next paragraph). When asserting that formal problem solving is a bad paradigm for AI research, David Marr pointed out that when we are solving formal problems we are certainly doing something well; however, what we are doing well is not the formal problem solving itself because that part of the process we are very poor in dealing with. George Boole, for example, with his misnamed "Laws of Thought" has (inadvertently, I suppose) led AI into a cul-de-sac from which it has yet to back out.

Connectionism and its second coming is the next unavoidable issue, as several of the commentators indicated. In proposing that the language-of-thought hypothesis supported a Turing-von Neumann architecture for cognition at the expense of a network representation, Jerry Fodor stirred up a hornet's nest of connectionists. If subsymbolic networks are necessary to support intelligent systems, then the conceptual transparency of the resultant AI systems is likely to be on a parr with that of the brain—somewhere close to zero. Homogeneous networks of subsymbolic elements will severely test our ability to understand our models; so, if subsymbolic network architectures are in some sense necessary, then this information might be bad news for AI.