

BOOK REVIEWS

Artificial Intelligence: The Very Idea. John Hauge-land. Cambridge, Massachusetts; The MIT Press, 1985. 287 pp.

In his introduction, the author states three goals for his book: "to explain, clearly and with an open mind, what AI is really all about; second, to exhibit the philosophical and scientific credentials behind its enormous appeal; and finally, to take a look at what actually has and has not been accomplished." Readers who are willing to accept the author's definition of AI will find that these three goals have been met quite well. AI is not viewed in this book as a particular collection of tools and techniques, nor is it seen as an effort to make machines efficiently perform certain tasks which are done well at present only by people. The author ignores definitions of the field which do not promote the assumption that one is grappling with fundamental questions having vast implications. He defined artificial intelligence as the attempt to create "machines with minds, in the full and literal sense." Fortunately, this most dramatic of definitions is not used as the excuse to weave an intricate tangle of abstruse speculations from which few ideas and fewer readers would emerge. To the contrary, the author's discussion is generally very clear, concrete, accurate, and interesting. My main regret is that the author does not take (or was not given by his editors) the space to more fully explore certain ideas.

In the first chapter, AI (as defined by the author), is placed in historical perspective as the most recent of investigations into the relations which hold between our mental and physical universes. Overviews of the work of Copernicus, Galileo, Hobbes, Descartes, and Hume take the reader from the ancient commonplace that things are not always what they seem to the more modern view that there is no intrinsic connection between thoughts and their alleged objects. Thus, we are faced with two questions which lie at the heart of AI: What makes a notation suitable for symbolizing some subject matter? What makes a suitable notation actually symbolize that subject matter? The author mainly addresses the second question, by discussing the meaning of meaning at various points throughout the book.

The next three chapters exhibit some of AI's credentials by discussing "interpreted automatic formal systems," one instance of which is purposefully programmed computers. The only programs discussed at any length are SHRDLU and GPS. As implied by the author's choice of goals for the book, AI has many credentials apart from its repertoire of working programs. Instead of focusing closely on existing programs, the discussion of AI's credibility revolves around the plausibility of certain assumptions. Perhaps the most obvious of these is "medium independence," the assumption that "essentially the same formal system

can be materialized in any number of different media, with no formally significant difference whatsoever." The media presently of concern to AI researchers, of course, are neurons and transistors. The author discusses other assumptions as well, including the very interesting one that "just as our smooth visual experience is somehow based in a 'grainy' retina, perhaps our own easy, flexible good sense is ultimately based in (sufficiently fine grained) rules and stereotypes." By the way, the author is neither strongly "pro-AI" nor "anti-AI"; rather, he takes the stand that AI is based on some very good ideas which may or may not be correct.

I was disappointed to find no references to Schank or any other natural language understanding researcher in the chapter "Semantics." I also wish the ideas of a "semantic division of labor" and reinterpretation of symbols had been more fully explored. On the other hand, the overview of Babbage's, Turing's, Von Neumann's, McCarthy's, and Newell's virtual machines in the chapter "Computer Architecture" could not be better. This chapter concludes that even though, for reasons of convenience, most AI programs are written in LISP, "the mind could have a computational architecture all its own. In other words, from the perspective of AI, mental architecture itself becomes a new theoretical "variable," to be investigated and spelled out by actual cognitive science research." This conclusion is a natural lead-in to a chapter discussing current work on knowledge representation, but there is no such chapter. Instead, the author turns to his third goal of looking at what has actually been accomplished in AI. He discusses early work on machine translation of natural languages, and describes the behavior of the programs GPS and SHRDLU. Here again, the reader could have benefited from more discussion of current AI work, although some references are given in the footnotes. Knowledge representation is discussed in a very general way. Schank's scripts, Minsky's frames, Bartlett's schemata, and Husserl's noemata are all thrown together in the single footnote relating to actual AI work; these are referred to collectively in the text as "linked stereotypes." There is, however, an interesting look at what the author calls the "frame problem." The example given is the task of correctly updating a knowledge base which represents a simple real-world physical situation after one of the objects being represented is moved.

In the final chapter, the author examines some of the fundamental differences between people and programs. The chapter is full of little gems. For example, the author points out that people follow an "ascription schema." That is, we ascribe beliefs, goals, and faculties so as to maximize a system's overall manifest competence. If someone says "Careful! That chair is hot!" and then sits on the chair himself, we will conclude that he lied to get the chair for himself, that he enjoys hot seats, or make

some similar conclusion which maximizes our opinion of his competence. The author points out that "the ascription schema constrains mental ascriptions once a system is specified; but it puts no limit on which systems should have mental states ascribed to them." He postulates a "Supertrap" which strikes matches in the presence of gas-soaked mice, topples dictionaries on mice, and, of course, snaps shut whenever mice nibble its bait "These habits betray a common malevolent thread, which is generalizable by (and only by) ascribing a persistent goal: dead mice." When we see other Supertrap behaviors, such as failure to harm cats that reek of gasoline, we become involved in a "semantic intrigue," an effort to understand how mental ascriptions cohere and interact. Whimsical examples aside, ascription is important for AI because it provides one more way to detect patterns that might otherwise go unnoticed. The ascription schema is proposed during the author's discussion of people's pragmatic sense. The final chapter also examines other fundamental differences between people and programs: our use of mental images, feelings, and ego involvement. Even if this chapter were not as thought-provoking and enjoyable as it is, it would be worth reading simply to remind oneself how extremely difficult problems in AI can (should?) be.

John W. L. Ogilvie
Modula Corporation
Provo, Utah

Heuristics: Intelligent Search Strategies for Computer Problem Solving. Judea Pearl. Addison-Wesley Publishing. 1984. 382 pp.

The view of AI science offered by Judea Pearl is thoroughly traditional and standard, and therein lie both its strengths and its weaknesses as a monograph, a reference, or a textbook in its field. As a graph-theoretic analysis of search strategy that clearly conforms to well-established AI methods and techniques, it expands upon these to incorporate probabilistic performance analysis principles, thus providing a (partial) formal framework of search strategy, evaluation criteria, and decision methods that are all firmly grounded in operations research.

To those readers for whom mathematical logic and probability calculus represent the most promising theoretical foundations of AI science, especially if understood in terms of graph theory and standard probabilistic models, this book will be quite useful and illuminating for the purposes of a textbook and as a reference. Pearl's survey of search strategies with respect to various probabilistic features of "heuristic information" provides valuable insights for general readers, students, and practicing researchers alike. From this perspective, the strength and value of Pearl's work will not be questioned here. For the purposes of teaching and promoting the general aspects of that the-

oretical approach, his book is clearly worthwhile and even innovative. Granting all of this, the only complaint that might be raised is altogether excusable, if not also entirely minor, *i.e.*, that the material presented might not be so easily grasped by the "casual reader" as the author supposes.

Discursively considered, however, and especially for the purposes of AI research, these very same strengths can be seen as weaknesses from the viewpoint of at least two alternative approaches: (1) nonformalist or antiformalist theories, which completely reject standard mathematical logic and traditional probability theory; or perhaps, (2) nonstandard or alternative formal theories, which can displace those views as prevailing paradigms. Now it clearly was not Pearl's aim to forestall alternative theories or to justify his own approach in contrast to other views. The comments that follow are not being offered as criticisms *per se*. They should instead be regarded as advice for those who may wish to pursue such alternative approaches, but who could benefit from a survey of precisely that direction in AI science they might ultimately choose to oppose, for reasons of their own.

Advocates of nonformalism and antiformalism in AI science tend to regard "heuristics" as their last line of defense, so to speak, against formal encroachment upon their research territory, as Pentland and Fischler (1983) or Bierre (1985) stand opposed to Nilsson (1983), for example, or as the notorious "Great Debate" runs, in general. Pearl's functional analysis of heuristics as the (somewhat arbitrary) catalyst for algorithmic procedures does not yield "heuristics" at all on this view, it seems, since these are "formally ineffable" by virtue of being exactly that which algorithms are not. The objection that Pearl's analysis is pervasively algorithmic, however, has some merit after all, if the "algorithmic properties" of "heuristic methods" (*i.e.*, those of completeness, admissibility, "dominance," and "optimality" in Chapter 3) are just the kinds of properties that "heuristics," by definition, cannot have. But it should come as no surprise to any antiformalist that these are the kinds of properties any formalist would seek to identify and establish, even under the name of "heuristics." Yet this does not count against the analysis itself, nor does it diminish the usefulness (in its particular domain) of the search strategies, evaluation criteria, and decision methods provided by Pearl's account.

Pearl's conception of heuristics as rules of thumb, intuitive judgments, educated guesses, and common sense hints at their subjective character as inferential guidelines that are defeasible in light of new information. In particular, he defines these techniques as "strategies using readily accessible though loosely applicable information to control problem-solving processes in human beings and machine(s)"¹ (p. vii). As such, Pearl's heuristics that

¹On the notion of a scientific paradigm, see Kuhn (1970)

can be used in some cases but not in every case, where their utility in any particular case could be diminished, strengthened, or even overruled as new and better information becomes available. The problem, therefore, is to provide a framework within the use of heuristics that can be objectively validated, since the fact that a given heuristic has been useful in the past, for example, affords no guarantee that it will continue to be useful in the future. What is not at all clear, in other words, is whether or not the various fragments of mathematical logic and probability calculus adopted by Pearl can provide an objective justification for reliance upon the techniques he recommends.

From the perspective of formal architectures for applied epistemology (e.g., Ringle (1979)), however, the theoretical view presupposed by Pearl exhibits intrinsic weaknesses that can only be overcome by adopting a fundamentally different formal language and logic for acquiring, representing, and utilizing knowledge, where the kind of knowledge in question can be either ordinary, expert, or scientific. If we suppose, for example, that logical reasoning and theorem-proving are indeed amenable to an AND/OR graphical analysis (as he suggests, p. 27), then deduction and induction are both limited to natural or axiomatic formal systems that will be exclusively extensional in character. There is convincing evidence against this approach, however, revealing that extensional analyses will be inadequate for representing knowledge regardless of its kind, since the following problems remain unsolved, concerning the representation of scientific knowledge, in particular:

- (a) the analysis of counterfactual and subjunctive conditionals;
- (b) the distinction between lawlike and accidental generalization; and,
- (c) the discovery of adequate criteria for scientific explanation.²

Each of these three problems is a crucially important issue in AI science, we may assume, especially in light of recent advances in “nonmonotonic reasoning” as a distinctively heuristic mode of knowledge processing. By contrast, monotonicity and consistency are claimed to be one and the same mathematical property of heuristic information (3.1.5), clearly suggesting that (a)-(c) could indeed be underlying problems not directly addressed within the theoretical framework of Pearl’s analysis.³

²(a)-(c) and (d)-(f) below are identified as the inherent shortcomings of extensional or “weakly intensional” analyses in Fetzer (1981), pp 167-8, 289-90. Similar difficulties are also encountered with respect to contexts of ordinary knowledge, as Nute (1980) suggests, especially regarding his conception of “hypothetical deliberation” in ordinary reason and discourses.

³Concerning nonmonotonicity, see Gabbay (1982), McCarthy (1980), McDermott and Doyle (1980), and Reiter (1980), to name a few. Given Pearl’s characterization of heuristics (p. vii), this conflation of monotonicity and consistency as being *identical* properties of “heuristic information” appears to be a major flaw in his approach.

Even if we further suppose that his account is based upon (what Fetzer (1981) identifies as) a “weakly intensional” framework by emphasizing the hypothetical, potential, or possible features of “heuristic information” as Pearl characterizes them, this does not prevent the following additional difficulties from arising:

- (d) explicating the meaning of single-case dispositional predicates;
- (e) providing an ontological foundation for nomological conditionals; and
- (f) supplying nonteleological solutions to (a)-(c).⁴

We may assume that these are also matters of crucial importance for AI science, which emphasize the need to clarify the ontic presuppositions upon which any knowledge processing system ultimately depends, whether it is artificial or natural in kind. The fundamental point to observe, therefore, is that, to the extent to which his analysis represents either a “classical” or a “neo-classical” instrumentalism embracing a purely extensional or even a weakly intensional language and logic (Fetzer, 1981, pp. 161-171) for the processing of “heuristic information,” Pearl’s view of heuristics either leaves open or begs all the questions that correspond to problems (a)-(c) and (d)-(f).

In all fairness, Pearl apparently intends to set aside these sorts of problems and commit to a classical framework and method afforded by standard mathematical logic and traditional probability theory. As suggested earlier, therefore, (a)-(c) and (d)-(f) are offered less as criticisms than as defining the limitations of his work. Unfortunately, his extensional approach represents what is evidently a widespread paradigm in scientific methodology and epistemology, despite its inability to overcome those very limitations. Insofar as AI science is, as many have suggested, “new,” “young,” or “immature” in the sense of having no prevailing paradigm at all, it would indeed be a mistake to adopt this one if (a)-(c) and (d)-(f) are problems that must be solved if AI is to succeed in attaining its goals. These are alternative theories, after all, to first- and second-order predicate calculi and standard theories of probability (where some combination of these appears to be the basic foundation of Pearl’s view), not all of which are nonformal or antiformal, where at least one of these (the probabilistic causal calculus *C* appearing in Fetzer (1981), Chapter 3, and in Fetzer and Nute (1979)) is sufficient to overcome the problems delineated above.⁵

Another difficulty with this book appears to be the author’s failure to provide precise formal definitions of some

⁴See Note 2. Also, see McCarthy (1977) and (1980), for example, concerning some of the ways in which epistemological and ontological problems are significant in AI science. Finally, (f) refers to teleological explanation, in particular, and since an AI artifact should be capable of explaining its inferences, its decisions, its valuations, and soon, nonteleological modes of explanation are desirable goals for AI science to pursue.

⁵While *C* is yet to be implemented, a nonmonotonic logic of subjunctive conditionals is being developed by Nute.

of the crucial terms involved in his analysis, such as “probability” and “likelihood,” without which it is difficult to appraise the suitability or the utility of his probability-based heuristic models. In this respect, what Pearl seems to have accomplished sometimes looks like a formalism in search of an interpretation without which the truth or the falsity of his claims is often impossible to assess. If the conceptions upon which his view is based do indeed conform to one or another of the traditional Bayesian models, moreover, then the very idea of a probability-based heuristic confronts a number of difficult problems of its own with respect to the distribution of probabilities to sets of alternative hypotheses, paths, or solutions, relative to the proposed refinements of those alternative hypotheses, paths, or solutions.⁶ These considerations suggest that traditional conceptions should not be taken for granted, especially if we assume that this is what Pearl intends by his observation that “Probability theory is today our primary (if not the only) language for formalizing concepts such as “average” and “likely,” and therefore it is the most natural language for describing those aspects of (heuristic) performance that we seek to improve” (p. 139).

On general theoretical grounds, I think, there are excellent reasons to suppose that (a)-(f) are fundamental problems in AI science and that an extensional probabilistic analysis of this sort simply cannot lead to their effective solutions. In order to understand the traditional approach, however, this book is recommended with the reservations implied above, namely, that the author has omitted basic definitions that might not be familiar to some readers, and that serious difficulties seem to confront the theoretical framework he apparently endorses, where these difficulties are especially severe from an epistemological perspective. The thorough justification of Pearl’s approach as an epistemological framework for AI science with respect to the processing of heuristic information, moreover, is not provided by his analysis. Readers who are interested in these underlying (and quite pressing) theoretical concerns are advised to look elsewhere, therefore, since not only was it not Pearl’s intention to address them here, but it also appears as though his approach could not succeed if that were his goal.

References

- Bierre, Pierre (1985) The professor’s challenge *AI Magazine* 5(2):60-70
- Fetzer, James H (1981) *Scientific knowledge: Causation, explanation, and corroboration* Dordrecht, Holland: D. Reidel Publishing Company
- Fetzer, J H & Donald E Nute (1979) Syntax, semantics, and ontology: A probabilistic causal calculus. *Synthese* 40 (1979):453-495
- Gabbay, Dov M. (1982) Intuitionistic basis for nonmonotonic logic. *Proceedings of the Conference on Automated Deduction, Springer Lecture Notes in Computer Science*, No 6 Berlin: Springer-Verlag

⁶Fetzer (1981), p. 276

- Kuhn, Thomas S (1970) *The structure of scientific revolutions* Second Edition Chicago: University of Chicago Press
- McCarthy, John (1977) *Epistemological problems of artificial intelligence*. *IJCAI-77*: 1038:1044
- McCarthy, John (1980) Circumscription: A form of nonmonotonic reasoning *Artificial Intelligence* 13: 41-72
- Nilsson, Nils J (1983) Artificial intelligence prepares for 2001. *AI Magazine* 4(4):7-14
- Nute, Donald E. (1980) *Topics in conditional logic* Dordrecht, Holland: D. Reidel Publishing Company
- Nute, Donald E (work in progress) A logic and a program for non-monotonic reasoning
- Pentland, A P & M A Fischler (1983) A more rational view of logic, or, up against the wall, logic imperialists! *AI Magazine* 4(4): 15-18.
- Reiter, R (1980) A logic for default reasoning. *Artificial Intelligence* 13 80-81
- Ringle, Martin (1979) Philosophy and artificial intelligence. In M. Ringle, (Ed.), *Philosophical Perspectives in Artificial Intelligence* Atlantic Heights, New Jersey: Humanities press.

Terry L. Rankin
University of Georgia
Athens, Georgia

Artificial Intelligence and Robotics:

Five Overviews. William D. Gevarter. Business/Technology Books, 1984. 564 pp.

Gevarter originally wrote the text for *Artificial Intelligence and Robotics; Five Overviews*, while he was a Research Associate at the National Bureau of Standards, sponsored by the National Aeronautics and Space Administration. Gevarter’s work was published by the National Bureau of Standards as a set of five volumes, and this book, published by Business/Technology Books, is Gevarter’s original work with some minor changes and corrections. Because Gevarter’s original work was very popular, Business/Technology Books (B/T Books) has performed a useful service by binding the individual volumes into one book.

In this work, Gevarter has surveyed the field of AI and has compiled a vast amount of information relating to what is being done in AI, where AI is being done, and who is doing it. The technical depth of the book is between that of a market survey report and an introductory text on the subject. The breadth of topics covered in *Artificial Intelligence and Robotics: Five Overviews* is described by the titles of the individual volumes which are bound together: *Robotics: An Overview (volume 1)*, *Expert Systems: An Overview (volume 2)*, *Computer Vision: An Overview (volume 3)*, *Computer Based Natural Language Processing: An Overview (volume 4)*, *Artificial Intelligence: An Overview (volume 5A-The Core Ingredients)*, *Artificial Intelligence: An Overview (volume 5B-Fundamental Applications Areas)*, *Artificial Intelligence: An Overview (volume 5C-Basic*

Topics).

The last three volumes of Gevarter's book (volume 5A, 5B, and 5C) contain a compendium of volumes 2, 3, and 4 (volume 1, on robotics, was not included in the compendium) plus some new material on the basics of AI. The new material includes a glossary of AI, a list of sources for AI information, and introductory AI topics, for example, search methods, knowledge representation, and computational logic.

Artificial Intelligence and Robotics: Five Overviews contains only descriptive material. The book does not contain examples or techniques for the solution of AI problems. Instead it contains the description of the techniques used by AI practitioners and researchers to solve AI problems. The important point is that this book will not help a reader who wants the "how" of AI. As an overview, it describes the "what," "where," and "who." *Artificial Intelligence and Robotics: Five Overviews* is in a form that is sometimes referred to as executive summary, and a book of this type can be useful for students, reference librarians, practicing AI researchers and engineers, and teachers. For example, in the past I have used Gevarter's NBS documents on which this book is based, for planning college courses in AI.

Rather than review the details of all the volumes, only one volume will be examined in detail. Volume 2, on Expert Systems, is the one chosen for the detailed discussion. The approach is to examine the number of pages (double spaced type was used) that are devoted to each of the sections within the volume on Expert systems. The volume on Expert Systems is 64 pages long with the last two pages used for references, which leaves 62 pages for the topic itself. Of the 62 pages covering expert systems, there are 34 pages devoted either to lists, tables or charts, and five blank pages. This leaves only 23 pages of double spaced prose to explain the topic of expert systems. The page count is distributed in the following manner. In the first 15 pages (page 16 is blank and not numbered, as are pages 38, 48, 56, and 58), the purpose and architecture of expert systems is introduced. From page 17 to page 37 there are tables that give the characteristics of widely used expert systems. These tables, one table per expert system, are packed with information on each expert system being explained. Following the tables, there are several pages of descriptive material which categorize expert systems by function. The remaining pages, from 47 to 62, list the places where expert systems are being worked on, where the funding is coming from, and list predictions for future expert systems. The other four volumes, Volumes 1, 3, 4, 5 (A through C), are constructed similarly to the volume on Expert Systems.

It is difficult to locate specific information in Gevarter's book, since it does not have a general index or general table of contents. Each of the five volumes con-

tains its own table of contents, a list of tables, and a list of figures. Since each volume is separately paged, these tables of contents, lists of tables, and lists of figures, only apply to the volume in which they are located. This arrangement is inefficient as well as difficult. Readers without a knowledge of AI will have a difficult time locating specific information.

In 1984 two books by Gevarter were copyrighted, the one being reviewed and one entitled *Artificial Intelligence, Expert Systems, Computer Vision, and Natural Language Processing*, published by Noyes Publications. Both books were based on the same previously mentioned NBS reports. The Noyes book doesn't contain the volume on Robotics, but otherwise most of the tables, figures, and prose are the same in both books. I could not find a reference in either book to the other one, although both books were copyrighted in the same year.

I would recommend *Artificial Intelligence and Robotics: Five Overviews*, if you can afford the price, and you want or need an overview of AI. I personally obtained Gevarter's original work from the National Bureau of Standards, and have found it useful on a number of occasions. Because much of the material in Gevarter's book is related to data such as AI funding, robotic production, or investigators working in AI for the years in which the material was written, the book is dated. From a historical perspective, this may be an asset. One of the most important future uses *Artificial Intelligence and Robotics: Five Overviews* may be that of providing a historical description of the primitive condition of AI at the beginning of the 1980s.

Vernon Williams
Ellicott City, Maryland

Books Received

The following books (see Books Received, next page) were received at the AAAI office since the publication of the last issue of the magazine. We encourage readers to volunteer as book reviewers. If you wish to write a review of any of these books, please contact our Book Review Editor:

Dr. Michael Fehling
Teknowledge, Inc.
525 University Avenue
Palo Alto, California 94301
Telephone: (415) 327-6600
(Arpanet: FEHLING@SUMEX)