John McCarthy, *Formalizing Common Sense: Papers by John McCarthy*

Varol Akman
Bilkent University, Ankara

November 21, 1996

Abstract

This is a review of the above title published by Ablex Publishing Corporation, Norwood, New Jersey, 1990; vi + 256 pages, hardback, ISBN 0–89391–535–1 (Library of Congress: Q335.M38 1989), edited by Vladimir Lifschitz.

Language has never been accessible to me in the way that it was for Sachs. I'm shut off from my own thoughts, trapped in a no-man's-land between feeling and articulation, and no matter how hard I try to express myself, I can rarely come up with more than a confused stammer. Sachs never had any of these difficulties. Words and things matched up for him, whereas for me they are constantly breaking apart, flying off in a hundred different directions. I spend most of my time picking up the pieces and gluing them back together, but Sachs never had to stumble around like that, hunting through garbage dumps and trash bins, wondering if he hadn't fit the wrong pieces next to each other. His uncertainties were of a different order, but no matter how hard life became for him in other ways, words were never his problem."

1 Introduction

These words of Peter Aaron as he sits down to tell the story of his best friend, Benjamin Sachs, in Paul Auster's brilliant novel *Leviathan* [3, page 55], strike me as a good description of what one feels upon reading this important collection of papers by John McCarthy. Since the dawn of theoretical computer science McCarthy has been a fertile source of bright ideas and, like so many other people in the field of artificial intelligence (AI), I have been stimulated by his thinking. However, when I reflect on the primary quality that draws me to his work, I realize that one aspect that stands out among others is the beauty of his prose. What one immediately notices in reading this collection of papers are those very special literary abilities of McCarthy, which are equaled, in my view, only by Marvin Minsky. A major feature of McCarthy's prose is the freshness of his style. His witty and striking examples give great vitality to his presentations which are clearly marked by an absence of confusion and disorder. True, McCarthy has "his uncertainties," but these are really "of a different order"—difficulties that are bred by the tough problems that he invariably grapples with.

Seventeen papers written by McCarthy on the subject of common sense are collected here. Three papers, viz. Chapters 5, 10, and 11 (cf. Appendix A for a list of contents), have previously existed as informal Stanford memos and are published here for the first time. On the other hand, an IJCAI paper [12] and another expository paper [13] are not included. (Nevertheless, their contents are subsumed pretty much by Chapters 6, 8, and 18.) Vladimir Lifschitz, the editor of the volume, made minor corrections and editorial changes to make the collection more accurate and uniform. I might state that he was successful in this endeavor to a great extent. All the remaining bugs I have spotted are obvious typos which are sometimes disturbing (page 2: "casual" instead of "causal") but not fatal.

The book opens with a fine overview (Chapter 1) of McCarthy's research in AI. Lifschitz, the author of the review, offers an admirably succinct account of the development of McCarthy's ideas on common sense from the early days of AI to his current work. Lifschitz's introduction is especially useful in appreciating the dramatically original and permanently influential nature of McCarthy's work. While McCarthy's papers collected in this volume were written over the span of almost three decades, Lifschitz correctly observes that the underlying concern has always been the same: to understand and model the intellectual ability realized by human common sense.

2 The JMC works

We learn from Minsky that McCarthy's general programme to formalize common sense took shape almost 35 years ago [17, page 323]: "In 1959, John McCarthy came to MIT from Dartmouth, and we started the MIT Artificial Intelligence Project. We agreed that the most critical problem was of how minds do common sense reasoning. McCarthy was more concerned with establishing logical and mathematical foundations for reasoning, while I was more involved with theories of how we actually reason using pattern recognition and analogy."

Chapter 2 presents in detail how McCarthy went about to accomplish this goal. The *advice taker* is the proposed program for commonsense reasoning. McCarthy gives the following 'working' definition: "a program has common sense if it automatically deduces for itself a sufficiently wide class of immediate consequences of anything it is told and what it already knows" [page 9]. The *advice taker* system has the following main features:

1. There is a method of representing expressions in the computer. (Motto: "In order for a program to be capable of learning something it must first be capable of being told" [page 11].) Certain of these expressions are declarative sentences in a logical system.

2. There is a deduction routine which when presented a set of premises will deduce a set of 'immediate' conclusions. The intelligence of *advice taker* is not embodied in this deduction routine; rather, it is embodied in the procedures which choose the input to the deduction routine.

3. Each object has a 'property list' which includes the specific things one would like to say about it.

4. The program operates by cyclically applying the deduction routine to a list of premises and by obeying some of the conclusions that have the form of imperative sentences.

The *advice taker* proposal was so much ahead of its time that it generated a heated debate during the 1958 *Teddington Conference on the Mechanization of Thought Processes* in which it was presented. In the discussion of the paper, which constitutes the latter half of Chapter 2, Y. Bar-Hillel and O. G. Selfridge raise certain objections which make the reading of the paper all the more enjoyable, e.g., Bar-Hillel: "Dr. McCarthy's paper belongs in the

Journal of Half-Baked Ideas." McCarthy's famous 1963 memo "Situations, actions, and causal laws" (now superseded by Chapter 3) was in fact an answer to Bar-Hillel's comment that the paper involved some philosophical presuppositions.

Chapter 3 (written jointly with Pat Hayes) studies some of the problems of philosophy which arise in any serious discussion of a program such as the *advice taker*. It is a rich paper which touches on a wide range of issues and cannot be done justice in the space available to this review. The situation calculus is introduced, and the frame problem is stated in stunning clarity. This is a computational epistemology paper. More specifically, we want a program such as the *advice taker* to decide what to do by deducing in a formal language that a certain approach will lead it to its goal. Philosophically speaking, this would amount to formalizing the concepts of causality, ability, intention, and above all knowledge (including self-knowledge). McCarthy and Hayes offer an explanation of *can*, *causes*, and *knows* in terms of an automata-theoretic representation. A crucial distinction, with roots in the *advice taker* work, is made between an *epistemologically adequate* representation and a *heuristically adequate* one. The former is a representation of the world in such a form that the solution of problems follows logically from the facts. The latter is the mechanism that efficiently solves the problem on the basis of available information. Since McCarthy and Hayes believe that the primary epistemological problems are far from being solved, they concentrate on epistemologically adequate representations.[1] The second part of this chapter formalizes the concepts of situation, fluent, future operator, action, strategy, and knowledge. Finally, the fourth part offers a concise review of modal logic, logics of knowledge, tense logic, etc. as reported in the philosophical logic community.

Chapter 6 keeps up the vivacity of Chapter 3. McCarthy, all the more epistemologically minded now, asserts that there are numerous areas of knowledge that we still do not know how to represent in an adequate way. A partial inventory includes propositional attitudes, concurrent events and actions, properties of space and time, causes and abilities, and properties of materials. McCarthy is especially fascinated by the first item in this list. To use a situation-theoretic terminology, he thinks that the belief states people are in are *systematically related* to the types of situations they are in. Accordingly, learning about one would often provide information about the other. In Chapter 8 he presents (albeit informally and with sketchy examples) a theory of concepts and propositions and illustrates its use in representing knowledge and necessity, two well-known modalities. A classical paradox investigated by McCarthy is that of *intensional contexts*—the death warrant of many a philosopher of language. Briefly, how is it that the telephone numbers of Alice and Bob are the same, and yet possess different properties, e.g., Carol may know the former but not the latter? McCarthy explains this by saying that "Alice's telephone number" and "Bob's telephone number" are different *concepts* with the same denotation. For the reader who may wonder what is meant by *concept*, McCarthy playfully remarks: "He [the reader] will have to remain nervous; no final commitment will be made in this paper" [page 121].

Chapter 10 covers further ground by discussing the formalization of two puzzles involving knowledge. These are the well-known *three wise men* puzzle and another little riddle on sums and products of integers that McCarthy calls the puzzle of *Mr. S and Mr. P*. McCarthy's axioms for these puzzles are written in first order logic and are based on a possible worlds

[1] In fact, the only place in this collection where McCarthy may be said to be dealing with heuristic adequacy is Chapter 11. This chapter, scrutinizing the celebrated doctrine of Robert Kowalski regarding programming as controlled deduction (as in the mock equation $algorithm = logic + control$), tries to propose a notion of how control is to be added to an expression of the logic of a specific yet interesting class of programs.

approach à la Kripke. While they are adorned with various innovations (the ability to express joint knowledge of several people, the ability to express change which occurs when someone learns something, etc.) there is a pervasive ad hoc quality to them. I must add that common knowledge and its variants (mutual belief, public information) receive a principled treatment—based on situation theory [7]—in Barwise's "On the model theory of common knowledge" [4, Chapter 9].

In Chapter 4 we find McCarthy musing on the criteria of an ideological proposition, as opposed to a scientific one. Approximately two decades ago, James Lighthill, a Cambridge University professor with no previous acquaintance with AI, wrote a review of the field at the request the Science Research Council of U.K. The purpose was to help the Council settle requests for support of work in AI. In a run of the mill report, Lighthill painted a distorted picture of AI and it is commonly thought that the development of AI in England suffered considerably as a result. With his usual cultured style, McCarthy dissolves the ideological propositions of Lighthill into chatter. Still, McCarthy is well-aware of the social aspects of AI. A society cannot fund a discipline unless its members have common feelings about the discipline's rate of progress. How can we devise scientific measurements and chart a map of our territory so that we have a guide to the conduct of AI in the decades to come? This question—one that does not belong to the realm of science and yet is necessary to it—gets a painstaking analysis from McCarthy. While the Lighthill report has all the ingredients to be shelved as a cheap shot, it is unfortunately true that AI research suffers from major deficiencies and so far has been only moderately triumphant. McCarthy's verdict is that much work in AI has signs of a childish disorder, the *look ma, no hands* disease: someone writes a program to do something no program has done before and concocts a paper pointing out that it is intelligent, without due regard to the identification of the intellectual mechanism (if any) behind it. His quarterly message as the president of AAAI (Chapter 14) reiterates this point and concludes that there is a need to think intensively about how to make AI 'experiments' that are really informative. This is probably feasible once we have a good problem domain—"the AI analog of what the *drosophila* did for genetics" [page 188]—but McCarthy rejects chess, seemingly a sensible candidate, on the grounds that chess programming is carried on at a level of sport (fast hardware!) rather than science and there is little analysis of the heuristics involved.

Two papers in this collection (Chapters 5 and 12) belong to the interface between commonsense reasoning and natural language semantics. Chapter 5 quotes a newspaper story describing a rather outlandish adventure of a certain Mr. Hug in a freight elevator. Before a program is declared to have digested the story of Mr. Hug, it would have to be able to answer a bunch of questions that cover important aspects of the story. McCarthy makes up such a list in order to get some intuitive idea of what commonsense skills are involved in penetrating the surface composition of the given narrative. He then enumerates the difficult subproblems that need to be resolved in order to understand such stories in a genuine way:

1. Construction of a parser that takes English into what McCarthy calls ANL, an *artificial natural language* to express the assertions of the sentences in a way free from dependence on the grammar of English.

2. Construction of an understander that extracts the *facts* from a text rendered in ANL.[2]

[2]It appears, at least to me, that McCarthy has in mind an information-based approach when he says 'facts.' I would conjecture that he has in mind information, or more specifically the flow of information,

3. Expression of the general information about the world and the construction of a *problem solver* which would allow question-answering by formal reasoning from the facts and the general information.[3]

Chapter 7 includes a highly intricate discussion on ascribing mental qualities to programs (robots) and is the technical paper on which Chapter 13 is based. Interwoven with laborious arguments, it defies a condensed analysis. According to McCarthy, crediting a machine or a computer program with qualities such as beliefs, knowledge, free will, intentions, abilities, or wants must be considered justifiable when such an attribution expresses the same information about the program that it expresses about a person. Such a practice is especially useful when the attribution helps us grasp the structure of the program, its behavior, etc. McCarthy believes that a major bottleneck in defining mental and intentional concepts has been the vulnerability of the methods of definition that have been employed. He introduces two kinds of definition, *definition relative to an approximate theory* and *second order structural definition*, and applies them to defining mental qualities. The first proposal declares that a concept is meaningful only in the theory and cannot be defined with more accuracy than the theory allows. The second is best explained by the second [sic] order predicate $\beta(W, M, B)$. This asserts that the first order predicate B is a good notion of belief for the machine M in the world W. In this context, 'good' means that the beliefs that B ascribes to M agree with our intuitions of what beliefs M would have.

Chapters 9 and 16 describe in replete detail the idea of circumscription, a renowned invention of McCarthy. So much ink has been spilled (and rightly so!) discussing these landmark contributions of McCarthy that my review of them will be necessarily brief and superficial; cf. [16] for a recent appraisal by McCarthy himself. The property common to all versions of circumscription is the existence of a second-order axiom conjoined to a first-order theory. McCarthy remarks that circumscription is not a nonmonotonic logic but is instead a nonmonotonic enlargement of first-order logic using this axiom. He sees the most straightforward way of using circumscription in a heuristic reasoning program that represents knowledge and belief by sentences of logic. Assuming that the program needs to jump to the conclusion that the objects it can determine to have certain properties are the only objects that do, it should use circumscription. In a nutshell, the objects that can be shown to have a certain property π by reasoning from certain facts Φ are all the objects that satisfy π. Chapter 16 introduces a more powerful form of circumscription which is called *formula* circumscription to distinguish it from the *domain* circumscription and *predicate* circumscription of Chapter 9. McCarthy also offers in this chapter an extended use of circumscription with a proof checker. Employing an interactive theorem prover for the theory of types, he determines what kinds of reasoning are required to use circumscription effectively.

The fragility of contemporary expert systems is examined in Chapter 15. I believe that many people simply discounted the quantifier "some" in the title of this paper and

as popularized by situation theory [4]. Thus, "inference is an activity whereby certain facts (items of information) about the world are used in order to extract additional information (information that is in some sense explicit in those facts)" [7, page 10].

[3] The problem solver can be regarded as a relationship between the content of the story and the background knowledge needed for its understanding. I would like to mention here recent important work by W. Zadrozny and K. Jensen (IBM T. J. Watson Research Center, Yorktown Heights, NY). In [21] they diligently show how to use such background knowledge. Their proposal exploits differences in plausibility of the meanings of words and phrases, and takes advantage of the connections between those meanings. As far as I know, theirs is the first proposal in the precise spirit of ANL.

commonly thought that McCarthy is boldly asking to arm each and every expert system with common sense. That this is not so can best be seen by noticing McCarthy's concern with MYCIN throughout the paper. MYCIN, an experimental program developed by R. Davis, B. Buchanan, and E. Shortliffe for advising physicians on treating bacterial infections, does reasonably well without common sense, provided the doctor using it has common sense and does not step over the program's 'curbs.' Surely, if in the future a program descending from MYCIN, say OURSIN, is put into production use (by selling OURSIN disks to doctors), then one would like it to be more robust vis-à-vis its limitations because after all human life is at stake. McCarthy devises natural scenarios which would cause MYCIN bite the dust simply because it is missing some measure of common sense which is required for usefulness in this domain.

In Chapter 17 McCarthy returns to a problem which he was not able to compose in a satisfactory written form at the time (1971) he first lectured about it: "generality in AI." He maintains that AI programs suffer from a lack of generality and cites two symptoms: (i) a small addition to the idea of a program frequently implies a complete rewrite, and (ii) no one knows how to build a commonsense knowledge base that could be used by any program that needed it. While the ideas McCarthy propounds for achieving generality are formulated in an excessively terse manner, one can still ferret out their essence. For example, a one-page discussion of contexts (cf. Section 3) moves like a meteor across our contemplative horizon and gives rise to multitudinous questions. However, McCarthy is well-aware of this quality of the article and gently comforts the reader: "All this is unpleasantly vague, but it's a lot more than could be said in 1971" [page 235].

Chapter 18, construed as a position paper about the relation between AI and mathematical logic, is a more compact version of [13]. Here, McCarthy hints at a subject he likes to call *metaepistemology*. Metaepistemology is analogous to metamathematics which considers the mathematical properties of mathematical theories as objects. For instance, model theory, a branch of metamathematics, is concerned with the relation between the linguistic expressions of mathematics and the mathematical structures they describe. Metaepistemology, on the other hand, "considers the relation between the world, languages for making assertions about the world, notions of what assertions are considered meaningful, what are accepted as rules of evidence, and what a knowledge seeker can discover about the world" [13, page 186]. With the help of metaepistemology one can study the relation between an (intelligent) agent's rules for accepting evidence and a world in which it is embedded. Metaepistemology is a purely mathematical theory in the sense that it is supposed to contain mathematical theorems about whether certain approaches will discover certain facts about the world.

3 What next?

More recent work of McCarthy, not included in this volume, span two equally important areas whose application to AI is original with him: "contexts" and "speech acts for programming languages." Not surprisingly, the roots of these proposals can be found in some of the papers (Chapters 7, 12, 13, 17, and 18) in the Lifschitz volume.

McCarthy believes that in order to make logical reasoning useful, one has to formalize the notion of "context of use" of logical sentences. Contexts are abstract objects. Their origin can be traced back to his Turing Award lecture (Chapter 17) where he first proposed them. According to *The New Merriam-Webster Pocket Dictionary*, *context* is "the part of a discourse surrounding a word or group of words that helps to explain the meaning of the word or word group." It is also "the circumstances surrounding an act or event." McCarthy's interest in contexts is not essentially from a linguistic viewpoint as the first *Merriam-Webster* definition implies; rather, it has a more general character as the second definition suggests. Again, McCarthy does not offer a definition, but gives some examples. The basic relation is $ist(c, p)$ which asserts that p (a proposition) is true in c (a context). This itself is asserted in an outer context c_o, i.e., we have something like $ist(c_o, ist(c, p))$. There are useful relations among contexts. For example $specialize\text{-}time(t, c)$ is a context related to c in which the time is specialized to have the value t. Similarly, if p is a proposition, then $assuming(p, c)$ is another context like c in which p is assumed. There is a relation $specializes$ between two contexts: $specializes(c_1, c_2)$ asserts that c_2 involves no more assumptions than c_1 and every proposition meaningful in c_1 is translatable into one meaningful in c_2. A logical system using contexts might afford a sequence of reasoning comparable to natural deduction. This would be realized by the operations of *entering* and *leaving* a context but McCarthy warns the reader that it is not correct to regard contexts as equivalent to sets of assumptions.[4]

I have been inspired by McCarthy's views to write up a short piece on contexts [2], but Guha's doctoral dissertation [9] should probably be regarded as the principal outcome of McCarthy's proposal, as McCarthy himself also indicates in [15]. Guha's work includes formal development of contexts as well as their application in Cyc—a huge multi-domain commonsense knowledge base [10]—largely in the form of what Guha calls *micro-theories* (descriptions of different aspects of the world). Finally, Shoham approaches the task of formalizing context from the perspective of modal logic [19].

A more provocative recent work of McCarthy is Elephant—a proposal for a programming language based on speech acts—which will probably be as fascinating, insightful, and seminal as LISP. Elephant is not yet articulated, to my best knowledge, in the form of a finished manuscript, but is more like thinking-out-loud to lay the foundations of a future programming language. Conceptually, it was McCarthy's study of the conditions under which computer programs may be ascribed beliefs, intentions, and other mental qualities (Chapter 7) that paved the way to Elephant.

The pragmatic origins of Elephant should be espied in the short proposal McCarthy made for a *common business communication language* (CBCL) (Chapter 12). The need for CBCL was suggested to McCarthy when he reflected on a world of future in which companies, well-equipped with online computer systems, wish to eliminate the clerks and have their computers 'speak' directly to each other. Presumably, a design for CBCL should be inspired by the low level protocols used for simple business communications like person-to-person messages. This shows that developing an expressive CBCL is difficult.

Elephant programs are expected to communicate with people (e.g., to process transactions) or other programs (e.g., to interchange electronic data) in sentences of the Elephant input/output language. They perform requests, permissions, promises, etc. which are traditionally called *speech acts* by philosophers [18]. The idea is that certain sentences do not have only a declarative aspect but are essentially actions. More specifically, speaking

[4] More recently, the situation semantics of Barwise and Perry came to be known as the most meticulous study of context in semantics [5]. In situation semantics, situations—limited portions of the reality—are assigned primary philosophical importance. The key insight is that inference is a 'situated' activity, i.e., it is carried out by agents situated in an environment which can be exploited in various ways [4]. There might be some affinity between McCarthy's contexts and contexts as studied in semantics and pragmatics; it should especially be worthwhile to examine the applicability of situation theory in this regard.

a language is performing speech acts—acts such as making statements, giving commands, asking questions, making promises, and so on. While it is true that these acts are in general performed in accordance with certain rules for the use of linguistic elements, the 'unit' of linguistic communication does not seem to be the word (or sentence) but rather the production of the word (or sentence) in the performance of the speech act. This outlook, first formulated by J. L. Austin, was later extended by Searle [18] so that the study of meaning of sentences is now considered equivalent to the study of speech acts. In other words, every meaningful sentence can be used to perform a particular speech act, and conversely, every possible speech act can be given an exact formulation in a sentence. Searle refines this as the "Principle of Expressibility" (*whatever can be meant can be said*) which enables one to identify rules for performing speech acts with rules for uttering certain words or sentences. It is on this consequence the feasibility of McCarthy's proposal resides. Slightly reformulating Searle [18, page 20], we may state that for any meaning M and any speaker S whenever S means (intends to convey, wishes to communicate in an utterance, etc.) M then it is possible that there is some expression E which is a precise expression of M.

In speech act theory, a distinction is made between *illocutionary acts*, such as warning someone of a snake, and *perlocutionary acts*, such as scaring or alarming him by doing so. Procedures for human execution are frequently specified as perlocutionary acts, e.g., "Get him to prepare breakfast." On the other hand, illocutionary acts are analogous to behavioral specifications relating inputs and outputs. McCarthy finds it worthwhile to formulate both illocutionary and perlocutionary specifications for the same program and to relate them. Assuming that one has available an axiomatic scientific theory of the domain under consideration, one might try to prove that a program meeting certain illocutionary specifications does also fulfill certain perlocutionary specifications. I have argued at length elsewhere [1] that justifying such axiomatic theories is not easy, especially when a commonsense model of the physical world is in question. In order not to be misunderstood, I must note that McCarthy appreciates the difficulty of this problem. In an enlightening discussion [13, page 163] he states: "From the problem solving or goal-achieving point of view, the common-sense world is characterized by a different *informatic situation* than that *within* any formal scientific theory. In the typical common-sense informatic situation, the reasoner doesn't know what facts are relevant to solving his problem. Unanticipated obstacles may arise that involve using parts of his knowledge not previously thought to be relevant."

4 Conclusion

A long time ago, Voltaire wrote in his *Philosophical Dictionary* that *sens commun* (common sense) means "good sense, crude reason, the beginnings of reason, the first notion of ordinary things, a state midway between stupidity and intelligence" [20, page 377]. Today, thanks to McCarthy, we are able to form a much more refined view of this popular phrase. Unlike Forguson's very readable but nontechnical monograph [8], the theory of common sense, as developed by McCarthy and cast into this volume by Lifschitz, ofttimes includes beautiful mathematical constructions and difficult results. Researchers interested in building commonsense theories (or advice takers) should read this book (and other works such as [6], [10], and [17]) in order to comprehend the fascinating world of commonsense reasoning.

Many of the papers in this volume are worth reading many times because each time one finds something that one failed to appreciate in a prior reading. Here is a suggested reading strategy, adapted from Matt Ginsberg's introduction to his *Readings in Nonmonotonic*

Reasoning (Morgan Kaufmann Publishers, Inc., Los Altos, CA, 1987): (1) Put the book aside for at least a month, and think about commonsense reasoning before you are swayed by what McCarthy has done. (2) Read the book from cover to cover. (3) Repeat the first two steps, as many times as you like. McCarthy handles with ease a group of problems that have been at the focus of AI debate: knowledge and action, epistemological adequacy, nonmonotonicity, contexts, ... What is more, he manipulates them all inside the unified theoretical framework of mathematical logic. It may be envisioned that in the years to come researchers comprehending and emulating the subtlety, analytical powers, and technical grasp of McCarthy will extend this body of work to an even greater extent.

5 Acknowledgments

I thank Mark Stefik and Stephen Smoliar for their positive response when I first approached them with an offer to write this review. Their thoughtful advice on content and organization has proved vital throughout the development of the manuscript. Other sources I found useful in this regard have been [6] and [11], although there is no overlap between my review and the latter.

My research is supported in part by two grants: TBAG–992, awarded by the Scientific and Technical Research Council of Turkey, and TU-LANGUAGE, awarded by the NATO Science for Stability Programme III.

6 Appendix A. Table of Contents of *Formalizing Common Sense*

1. Understanding Common Sense: John McCarthy's Research in Artificial Intelligence, *by Vladimir Lifschitz*
2. Programs with Common Sense
3. Some Philosophical Problems from the Standpoint of Artificial Intelligence
4. Review of "Artificial Intelligence: A General Survey"
5. An Example for Natural Language Understanding and the AI Problems It Raises
6. Epistemological Problems of Artificial Intelligence
7. Ascribing Mental Qualities to Machines
8. First Order Theories of Individual Concepts and Propositions
9. Circumscription—A Form of Nonmonotonic Reasoning
10. Formalization of Two Puzzles Involving Knowledge
11. Coloring Maps and the Kowalski Doctrine
12. The Common Business Communication Language
13. The Little Thoughts of Thinking Machines

14. AI Needs More Emphasis on Basic Research
15. Some Expert Systems Need Common Sense
16. Applications of Circumscription to Formalizing Common Sense Knowledge
17. Generality in Artificial Intelligence
18. Mathematical Knowledge in Artificial Intelligence

7 References

1. V. Akman and P. J. W. ten Hagen, The power of physical representations, *AI Magazine* **10** (3) (1989) 49–65.
2. V. Akman and E. Tın, What is in a context?, in: E. Arıkan, ed., *Proceedings of the 1990 Bilkent International Conference on New Trends in Communication, Control, and Signal Processing, Volume 2* (Elsevier, Amsterdam, 1990) 1670–1676.
3. P. Auster, *Leviathan* (Viking Penguin, New York, 1992).
4. J. Barwise, *The Situation in Logic (CSLI Lecture Notes Number 17)* (Center for the Study of Language and Information, Stanford, CA, 1989).
5. J. Barwise and J. Perry, *Situations and Attitudes* (The MIT Press, Cambridge, MA, 1983, fourth printing 1986).
6. E. Davis, *Representations of Commonsense Knowledge* (Morgan Kaufmann Publishers, Inc., San Mateo, CA, 1990). (Also see *Artificial Intelligence* **61** (1) (1993) for several reviews of this book and a response by Davis.)
7. K. Devlin, *Logic and Information* (Cambridge University Press, New York, 1991, reprinted 1992).
8. L. Forguson, *Common Sense* (Routledge, London, 1989).
9. R. V. Guha, Contexts: A formalization and some applications, MCC Technical Report Number ACT–CYC–423–91 (Microelectronics and Computer Technology Corporation, Austin, TX, 1991).
10. D. B. Lenat and R. V. Guha, *Building Large Knowledge-Based Systems: Representation and Inference in the Cyc Project* (Addison-Wesley Publishing Company, Inc., Reading, MA, 1989). (Also see *Artificial Intelligence* **61** (1) (1993) for several reviews of this book and a response by Lenat and Guha.)
11. V. W. Marek, Review of *Formalizing Common Sense: Papers by John McCarthy*, *SIGART Bulletin* **4** (1993) 12–13.
12. J. McCarthy, M. Sato, S. Igarashi, and T. Hayashi, On the model theory of knowledge, in: *Proceedings of the Fifth International Joint Conference on Artificial Intelligence (IJCAI–77)* (The MIT Press, Cambridge, MA, 1977).

13. J. McCarthy, Artificial intelligence, logic, and formalizing common sense, in: R. H. Thomason, ed., *Philosophical Logic and Artificial Intelligence* (Kluwer Academic Publishers, Dordrecht, Netherlands, 1989) 161–190. (Also available via McCarthy's web site http://www-formal.stanford.edu/jmc/home.html)

14. J. McCarthy, Elephant 2000: A programming language based on speech acts, Manuscript (29 pp.), Computer Science Department, Stanford University, Stanford, CA, 1994. (Available via McCarthy's web site http://www-formal.stanford.edu/jmc/home.html)

15. J. McCarthy, Notes on formalizing context, in: *Proceedings of the Thirteenth International Joint Conference on Artificial Intelligence (IJCAI–93)* (Morgan Kaufmann, Mountain View, CA, 1993). (Also available via McCarthy's web site http://www-formal.stanford.edu/jmc/home.html)

16. J. McCarthy, History of circumscription, *Artificial Intelligence* **59** (1993) 23–26.

17. M. Minsky, *The Society of Mind* (Simon and Schuster, New York, 1986). (Also see *Artificial Intelligence* **48** (3) (1991) for several reviews of this book and a response by Minsky.)

18. J. R. Searle, *Speech Acts: An Essay in the Philosophy of Language* (Cambridge University Press, New York, 1969, reprinted 1990).

19. Y. Shoham, Varieties of context, in: V. Lifschitz, ed., *Artificial Intelligence and Mathematical Theory of Computation: Papers in Honor of John McCarthy* (Academic Press, San Diego and London, 1991) 393–408.

20. Voltaire, *Philosophical Dictionary* (Penguin Books Ltd., Harmondsworth, Middlesex, England, 1972, reprinted 1986).

21. W. Zadrozny and K. Jensen, Semantics of paragraphs, *Computational Linguistics* **17** (2) (1991) 171–210.