Artificial Intelligence[1]

1 Introduction

There have been many definitions of Artificial Intelligence (AI) offered over the years. Perhaps a good one is:

> Artificial Intelligence is a field of science and engineering concerned with the computational understanding of what is commonly called intelligent behavior, and with the creation of artifacts that exhibit such behavior.

This may be examined more closely by considering the field from the points of view of three goals AI researchers have, which might be called Computational Psychology, Computational Philosophy, and Advanced Computer Science.

1.1 Computational Psychology

The goal of Computational Psychology is to understand human intelligent behavior by creating computer programs that behave in the same way people do. For this goal it is important that the algorithm expressed by the program be the same algorithm that people actually use, and that the data structures used by the program be the same data structures used by the human mind. The

[1] This is a "preprint" version of an article that appears in S. C. Shapiro, Ed. *Encyclopedia of Artificial Intelligence, Second Edition.* John Wiley & Sons, Inc., New York, 1992, 54–57.

program should do quickly what people do quickly, should do more slowly what people have difficulty doing, and should even tend to make mistakes where people tend to make mistakes. If the program were put into the same experimental situations that human subjects are subjected to, the program's results should be within the range of human variability.

1.2 Computational Philosophy

The goal of Computational Philosophy is to form a computational understanding of human-level intelligent behavior, without being restricted to the algorithms and data structures that the human mind actually does (or conceivably might) use. By "computational understanding" is meant a model that is expressed as a procedure that is at least implementable (if not actually implemented) on a computer. By "human-level intelligent behavior" is meant behavior that, when engaged in by people is commonly taken as being part of human intelligent cognitive behavior. It is acceptable, though not required, if the implemented model perform some tasks better than any people would. Bearing in mind Church's Thesis (qv), this goal might be reworded as asking the question, Is intelligence a computable function?

In the AI areas of Vision and Robotics Computational Philosophy is replaced by Computational Natural Philosophy (Science). For example, computer vision researchers are interested in the Computational Optics question of how can the information contained in light waves reflected from an object be used to

reconstruct the object. Notice that this is a different question from the Computational Psychology question of how the human visual system uses light waves falling on the retina to identify objects in the world.

1.3 Advanced Computer Science

The goal of Advanced Computer Science is to push outwards the frontier of what we know how to program on computers, especially in the direction of tasks that, although we don't know how to program them, people can perform them. This goal led to one of the oldest definitions of AI: The attempt to program computers to do what until recently only people could do. Although this gets across the idea of pushing out the frontier, it is also perpetually self-defeating in that as soon as a task is conquered, it no longer falls within the domain of AI—AI is left only with its failures as its successes become other areas of Computer Science. The most famous example of this is the area of symbolic calculus. When James Slagle wrote the SAINT program (qv), it was the first program in history that could solve symbolic integration problems at the level of freshman calculus students, and was considered an AI project. Now that there are multiple systems on the market that can do much more than what SAINT did, these systems are not considered to be the results of AI research.

1.4 Heuristic Programming

Computational Psychology, Computational Philosophy, and Advanced Computer Science are subareas of AI divided by their goals. Any given AI researcher probably wanders among two or all three of these areas throughout his or her career, and may even have a mixture of these goals at the same time.

Another way of distinguishing AI as a field is by noting the AI researcher's interest in heuristics rather than in algorithms. Here I am taking a wide interpretation of a *heuristic* as any problem solving procedure that fails to be an algorithm, or that has not been shown to be an algorithm, for any reason. An interesting view of the tasks that AI researchers consider to be their own may be gained by considering those ways in which a procedure may fail to be an algorithm.

The common definition of an algorithm for a general problem P is: an unambiguous procedure that, for every particular instance of P, terminates and produces the correct answer. The most common reasons that a heuristic H fails to be an algorithm are: it doesn't terminate for some instances of P; it has not been proved correct for all instances of P because of some problem with H; or it has not been proved correct for all instances of P because P is not well-defined. Common examples of AI heuristic programs that don't terminate for all instances of the problem they have been designed for are search and theorem proving programs. Any search procedure will run forever if given an infinite search space that contains no solution state. Gödel's Incompleteness

Theorem states that there are formal theories that contain true but unprovable propositions. In actual practice, AI programs for these problems stop after some prespecified time, space, or work bound has been reached. They can then only report that they were unable to find a solution—in any given case, a little more work might have produced an answer. An example of an AI heuristic that has not been proved correct is any static evaluation function used in a computer chess program. The static evaluation function returns an estimate of the value of some state of the board. To be correct, it would return $+\infty$ if the state were a sure win for the side to move, $-\infty$ if it were a sure win for the opponent, and 0 if it were a forced stalemate. Moreover, for any state it is theoretically possible to find the correct answer algorithmically by doing a full minimax search of the game tree rooted in the state being examined. However, such a full search is infeasable for almost all states because of the size of the game tree. Static evaluation functions are still useful, even without being proved correct. An example of an AI heuristic program that has not been proved correct because the problem for which it has been designed is not well-defined is any natural language understanding program or natural language interface. Since no one has any well-defined criteria for whether a person understands a given language, there cannot be any well-defined criteria for programs either.

2 Early History

Although the dream of creating intelligent artifacts has existed for many centuries, the field of Artificial Intelligence is considered to have had its birth at a conference held at Dartmouth College in the summer of 1956. The conference was organized by Marvin Minsky and John McCarthy, and McCarthy coined the name "Artificial Intelligence" for the proposal to obtain funding for the conference. Among the attendees were Herbert Simon and Allen Newell who had already implemented the Logic Theorist program at the Rand Corporation. These four people are considered the Fathers of AI. Minsky and McCarthy founded the AI Laboratory at M.I.T.; Simon and Newell founded the AI laboratory at Carnegie-Mellon University; and, McCarthy later moved from M.I.T. to Stanford University, where he founded the AI laboratory there. These three universities, along with Edinburgh University, whose Department of Machine Intelligence was founded by Donald Michie, have remained the premier research universities in the field. The name, Artificial Intelligence, remained controversial for some years, even among people doing research in the area, but it eventually was accepted.

The first AI text was *Computers And Thought,* edited by Edward Feigenbaum and Julian Feldman, and published by McGraw-Hill in 1963. *Computers And Thought* is a collection of 21 papers, some of them short versions of Ph.D. dissertations, by early AI researchers. Most of the papers in this collection are still considered classics of AI, but of particular note is a reprint of Alan M.

Turing's 1950 paper in which the Turing Test was introduced.

Regular AI conferences began in the mid to late 1960's. The Machine Intelligence Workshops series began in 1965 in Edinburgh. A conference at Case Western University in Spring, 1968 drew many of the U.S. AI researchers of the time, and the first biennial International Joint Conference on Artificial Intelligence was held in Washington, D. C. in May, 1969. *Artificial Intelligence,* still the premier journal of AI research began publishing in 1970.

For a more complete history of AI, see McCorduck 1979 and the article LITERATURE, AI in this Encyclopedia.

3 Neighboring Disciplines

Artificial Intelligence is generally considered to be a subfield of Computer Science, though there are some Computer Scientists who have only recently and grudgingly accepted this view. There are several disciplines outside Computer Science, however, that strongly impact AI and which AI strongly impacts.

Cognitive Psychology (qv) is the subfield of Psychology that uses experimental methods to study human cognitive behavior. The goal of AI called Computational Psychology above is obviously closely related to Cognitive Psychology, differing mainly in the use of computational models rather than experments on human subjects. However, most AI researchers pay some attention to the results of Cognitive Psychology, and Cognitive Psychologists tend to pay atten-

tion to AI as suggesting possible cognitive procedures that they might look for in humans.

Cognitive Science (qv) is an interdisciplinary field that studies human cognitive behavior under the hypothesis that cognition is (or can usefully be modelled as) computation. AI and Cognitive Science overlap in that there are researchers in each field that would not consider themselves to be in the other. AI researchers whose primary goal is what was called Advanced Computer Science above generally do not consider themselves to be doing Cognitive Science. Cognitive Science contains not only AI researchers, but also Cognitive Psychologists, Linguists, Philosophers, Anthropologists, and others, each using the methodoly of his or her own discipline on a common problem—that of understanding human cognitive behavior.

Computational Linguistics (qv) researchers use computers, or at least the computational paradigm, to study and/or to process human languages. Like Cognitive Science, Computational Linguistics overlaps AI. It includes those areas of AI called Natural Language Understanding (qv), Natural Language Generation (qv), Speech Understanding (qv), and Machine Translation (qv), but also non-AI areas such as the use of statistical methods to find index keywords useful for retrieving a document.

4 AI-Complete Tasks

There are many subtopics in the field of AI, as one can see by contemplating the individual articles in this Encyclopedia. These subtopics vary from the consideration of a very particular, technical problem, to broad areas of research. Several of these broad areas can be considered *AI-complete,* in the sense that solving the problem of the area is equivalent to solving the entire AI problem—producing a generally intelligent computer program. A researcher in one of these areas may see himself or herself as attacking the entire AI problem from a particular direction. These areas are also ways of organizing the articles of this Encyclopedia. The following sections discuss some of the AI-complete areas covered by this Encyclopedia, and point to some of the articles relevant to those areas. Not every article in the Encyclopedia is included in the following lists.

4.1 Natural Language

The AI subarea of Natural Language is essentially the overlap of AI anc Computational Linguistics (see above). The goal of the area is to form a computational understanding of how people learn and use their native languages, and to produce a computer program that can use a human language at the same level of competence as a native human speaker. Virtually all human knowledge has been (or could be) encoded in human languages (consider this, and other encyclopedias, textbooks, *etc.*) Moreover, research in Natural Language Understanding has shown that encyclopedic knowledge is required to under-

stand natural language. Therefore, a complete Natural Language using system will also be a complete Intelligent system. The articles in this Encyclopedia relevant to Natural Language include: ARGUMENT COMPREHENSION; COMPUTATIONAL LINGUISTICS; CONVERSATIONAL IMPLICATURE; DEEP STRUCTURE; DICTIONARY; DISCOURSE UNDERSTANDING; ELLIPSIS; GRAMMAR, AUGMENTED TRANSITION NETWORK; GRAMMAR, CASE; GRAMMAR, GENERALIZED PHRASE STRUCTURE; GRAMMAR, PHRASE STRUCTURE; GRAMMAR, SEMANTIC; HERMENEUTICS; LEXICAL DECOMPOSITION; LEXICAL SEMANTICS; MACHINE TRANSLATION; MONTAGUE GRAMMAR; MORPHOLOGY; NATURAL LANGUAGE GENERATION; NATURAL LANGUAGE UNDERSTANDING; PARSING; PARSING, WORD EXPERT; PREFERENCE SEMANTICS; PRESUPPOSITION; QUESTION ANSWERING; SPEECH RECOGNITION; SPEECH SYNTHESIS; SPEECHUNDERSTANDING; STORY ANALYSIS; SYSTEMIC GRAMMAR; TEXT SUMMARIZATION; THESAURUS.

4.2 Problem Solving and Search

Problem Solving is the area of AI that is concerned with finding or constructing the solution to a problem. That sounds like a very general area, and it is. The distinctive characteristic of the area is probably its approach of seeing tasks as problems to be solved, and of seeing problems as spaces of potential solutions that must be searched to find the true one, or the best one. Thus

the AI area of Search is very much connected to Problem Solving. Since any area investigated by AI researchers may be seen as consisting of problems to be solved, all of AI may be seen as involving Problem Solving and Search. The articles in the Encyclopedia that are most directly about Problem Solving and Search include: A* ALGORITHM; AND/OR GRAPHS;; BACKTRACKING; BRANCHING FACTOR; DISTRIBUTED PROBLEM SOLVING; HEURISTICS; MEANS-ENDS ANALYSIS; MINIMAX PROCEDURE; PROBLEM REDUCTION; PROBLEM SOLVING; SEARCH; SEARCH, BEAN; SEARCH, BEST-FIRST; SEARCH, BI-DIRECTIONAL; SEARCH, BRANCH-AND-BOUND; SEARCH, DEPTH-FIRST; SIMULATED ANNEALING.

4.3 Knowledge Representation and Reasoning

Knowledge Representation is the area of AI concerned with the formal symbolic languages used to represent the knowledge (data) used by intelligent systems, and the data structures used to implement those formal languages. However, one cannot study static representation formalisms and know anything about how useful they are. Instead, one must study how they are helpful for their intended use. In most cases, the intended use is to use explicitly stored knowledge to produce additional explicit knowledge. This is what reasoning is. Together, Knowledge Representation and Reasoning can be seen to be both necessary and sufficient for producing general intelligence—it is another AI-complete area. Although they are bound up with each other, Knowledge Representation and

Reasoning can be teased apart according to whether the particular study is more about the representation language/data structure, or about the active process of drawing conclusions.

The articles in this Encyclopedia that are most concerned with Knowledge Representation include: BELIEF RPERESENTATION SYSTEMS; CONCEPTUAL DEPENDENCY; DYNAMIC MEMORY; EPISODIC MEMORY; FRAMES; KNOWLEDGE REPRESENTATION; LOGIC; LOGIC, CONDITIONAL; LOGIC, HIGHER ORDER; LOGIC, MODAL; KOGIC, ORDER SORTED; LOGIC, PREDICATE; LOGIC, PROPOSITIONAL; MEMORY ORGANIZATION PACKETS; MENTAL MODELS; SEMANTIC NETWORKS.

The articles in this Encyclopedia that are most concerned with Reasoning include: ABDUCTION; BAYSIAN INFERENCE METHODS; BELIEF REVISION; CIRCUMSCRIPTION; DEMPSTER-SHAFER; EQUALITY INFERENCING; FUZZY LOGIC: OVERVIEW; MEMORY-BASED REASONING; META-KNOWLEDGE, -RULES,-REASONING; QUALITATIVE PHYSICS; REASONING, CASE-BASED; REASONING, CAUSAL; REASONING, COMMONSENSE; REASONING, DEFAULT; REASONING, NONMONOTONIC; REASONING, PLAUSIBLE, REASONING, SPATIAL; REASONING, TEMPORAL; RESOLUTION; RESOLUTION, GRAPH-BASED; RESOLUTION, THEORY; THEOREM PROVING; TRUTH MAINTENANCE; UNIFICATION.

4.4 Learning

Learning is often cited as the criterial characteristic of intelligence, and it has always seemed like the easy way to producing intelligent systems: Why build an intelligent system when we could just build a learning system and send it to school? The articles in this Encyclopedia that are most concerned with learning include: CONCEPT LEARNING; INDUCTIVE INFERENCE; LEARNING, MACHINE; LEARNING THEORY.

4.5 Vision

Vision, or Image Understanding, has to do with interpreting visual images that fall on the human retina or the camera lens. The actual scene being looked at could be 2-dimensional, such as a printed page of text, or 3-dimensional, such as the world about us. If we take "interpreting" broadly enough, it is clear that general intelligence may be needed to do the interpretation, and that correct interpretation implies general intelligence, so this is another AI-complete area. The articles in this Encyclopedia that are most concerned with Vision include: CHARACTER RECOGNITION; COLOR VISION; EARLY VISION; EDGE & LOCAL FEATURE DETECTION; HOUGH TRANSFORM; IMAGE MODELS; IMAGE PROPERTIES; IMAGE UNDERSTANDING; OBJECT RECOGNITION; STEREO VISION; VISUAL PERCEPTION.

4.6 Robotics

The area of robotics is concerned with artificats that can move about in the actual physical world, and/or that can manipulate other objects in the world. The articles in this Encyclopedia that are most concerned with Robotics include: MOTION ANALYSIS; PATH PLANNING & OBSTACLE AVOIDANCE; RANGE DATA ANALYSIS; ROBOT CONTROL SYSTEMS; ROBOT HANDS & END EFFECTORS; ROBOT MANIPULATORS; ROBOTICS; ROBOTS, LEGGED; ROBOTS, MOBILE; SENSORS & SENSOR FUSION; TELEOPERATORS.

5 Applications

Throughout its existence as a field, AI research has produced spin-offs into other areas of Computer Science. Lately, however, programming techniques developed by AI researchers have found application to many programming problems. This has largely come about through the subarea of AI known as Expert Systems. Whether or not any particular program is intelligent, or is an expert is largely irrelevant pragmatically. From the point of view of the field as a whole, probably the best thing about this development is that after many years of being criticized as following an impossible dream by inappropriate and inadequate means, AI has been recognized by the general public as having applications to everyday problems.

The articles in this Encyclopedia that are most relevant to Expert Systems as

a subarea of AI include: BLACKBOARD SYSTEMS; CERTAINTY FACTORS; EXPERT SYSTEMS; RULE-BASED SYSTEMS.

The articles in this Encyclopedia that discuss past, current, and potential applications of AI include:ANIMATION; ART, AI IN; BUSINESS, AI IN; CHEMISTRY, AI IN; COMPUTER INTEGRATED MANUFACTURING; EDUCATION, AI IN; ENGINEERING, AI IN; LAW, AI IN; MANUFACTURING, AI IN; MEDICINE, AI IN; MILITARY, AI IN THE; MUSIC, AI IN; PROGRAMMING ASSISTANTS.

References

[1] E. A. Feigenbaum and J. Feldman. *Computers and Thought.* McGraw-Hill Book Company, New York, 1963.

[2] P. McCorduck. *Machines Who Think.* W. H. Freeman and Company, San Francisco, 1979.

[3] A. M. Turing. Computing machinery and intelligence. *Mind*, 59:433–460, October 1950.

Stuart C. Shapiro

SUNY at Buffalo