Introductory AI for Whom? Presenting AI to the Non-Scientist

Rebecca E. Skinner,
Visiting Scholar, Knowledge Systems Laboratory,
Computer Science Department, Stanford University
701 Welch Road, Building C, Palo Alto, CA 94304.
415 725 3855. fax 415 725 5850.
skinner@hpp.stanford.edu

Abstract

Artificial Intelligence, as a set of tools for describing, analyzing and creating artifacts, has much to offer the humanities and social sciences. These ideas may be applied to far more informal human artifacts such as culture, as a means to better understand the non-engineered complex artifacts which surround us. This paper describes "AI for Everyone", a course for advanced undergraduates which is intended to introduce various ideas in knowledge representation and AI to the non-scientist. The author is currently proposing that this course be offered at Stanford University during 1994-1995. The course introduces the concepts of (a) design of complex artifacts and (b) internal representation derived from cognitive science, as seen in the progress of AI research. The class focuses on symbolic representation techniques, with two sessions on reiterations of other kinds of intelligence (Nouvelle AI, Connectionism). Also included are sessions on critiques of symbolic representation and on potential means to apply KR concepts to fields outside computing. No technical background in computing or cognitive sciences is assumed.

Introduction

The problem as posed assumes that Introductory AI as a set of programming techniques should be taught to computer scientists, and only to computer scientists. The author questions this premise. She suggests that instead certain of the analytical approaches developed by AI be introduced to social science and humanities students. The sciences of the complex, of which AI is the pre-eminent representative addressing cognitive complexity, should be taught to far more people. A well-defined subset of existing KR concepts should be incorporated into the base of knowledge which defines a civilized, educated person.

The state of knowledge concerning the sciences of the complex in the educated public and in the fields of humanities and social sciences is very poor (1). AI, cognitive science, and their commercial applications have been presented to the public, including to highly educated non-technical professionals, by the Dreyfus brothers and business journalists respectively. The latter present AI as a simplistic, inhumane, and failed attempt to replace human intelligence with that of machines ("the Frankenstein model"). The field's leaders are cognizant of this unsavory public image of AI as either "impossible or invisible" (see reports on AAAI Board meetings and the AAAI's 1993 round of ballots for Councilors). However, not much is being done about the situation. To those in the thick of AI research, this institutional intellectual passivity may not be either obvious or important. Why should the non-specialist know anything about this field ? AI software implementations are typically highly transparent, so there is relatively little need for the average PC user to understand AI. According to this interpretation, the public's ignorance about AI is irritating but innocuous. The author asserts that this is a grave misreading of the situation. Actually AI has much to offer the humanities and social sciences.

The field concentrates on the development of useful software (and hardware) tools based on two distinct exemplars of intelligence: "AI as empirical inquiry", entailing the reiteration of logical cognizant intellectual processes, and the recreation of intelligence from the bottom up (epitomized by Brooks' as "elephants don't play chess"). However, both 'AI as complexity' and 'AI as simplicity' orientations overlook the very general definition of the term "tool": "an implement or machine used to do work or perform a task". Ideally the pragmatic software tool orientation of the current AI world could be complemented by a wider understanding of very basic concepts of AI, in a non-technical form. In a sense such understanding is part of one of the basic imperatives of AI as empirical inquiry and tool creation, or the creation of a Knowledge Medium (Stefik, 1986). What is a Knowledge Medium if it is not a means for conveying thought ? The latter is of course, done by software and material tools, but it is also done by thinking, and is part of the collective store of wealth available to any given culture. The simple potential usage of knowledge representation as a means to analyzing problems in a more or less formal way offers a means to help people think more clearly. The descriptive and analytical potency of the basic concepts and issues of AI (in symbolic and connectionist forms alike) hold regardless of whose NCAI demo is more impressive. In this light, the exclusive focus on software tools amidst the field's wealth of ideas about knowledge, intelligence, information, social interactions (collaboration, interaction) is curious rather than natural. The introversion of AI among academic disciplines seems incongruent and inappropriate.

Roger Schank has defined AI as an algorithmic approach to various domains: " AI should in principle be a contribution to a great many fields of study... AI is, potentially, the algorithmic study of processes in every field of inquiry. As such, the future should produce AI/anthropologists, AI/ doctors, AI/political scientists and so on. It is just a matter of time until AI becomes part of other fields..." (Schank 1991, p7).

On the contrary, such individuals have not flooded AAAI on their own. It seems most unlikely that they will unless the field actively undertakes to acquire a wider audience in non-CS university circles. Applying existing concepts to new contexts is essential to intellectual progress (consider the importance of analogy in KR and learning), but is sufficiently strenuous that people in AI must work to develop analogies between AI and semantically and syntactically distant fields. Should such an effort be undertaken, AI could make a significant contribution to society as an intellectual infrastructure or "public good" for problem analysis and problem-solving. In a more abstract sense than the concept is usually understood, the rudiments of knowledge representation could be generally introduced in higher education. The proliferation of AI as intellectual edifice could and should complement the proliferation of AI as software tools.

One contribution to this effort could be to introduce courses entitled "AI for Everyone". A version of the prospective course syllabus follows. The author is currently trying to arrange to teach this course in a humanities department at Stanford University. Many people outside AI are very curious about the field, but see it as entirely inaccessible. Yet many basic ideas in AI, such as internal representation, inferencing, heuristic versus algorithmic solutions, search strategies, frames and scripts, schemata, etc., may be explained without reference to FOPC. Both the student and the teacher could benefit from such a course. Figuring out how to explain the basic concepts of the sciences of the complex in general and AI in specific without the assumptions that the student knows FOPC and programming languages would be a fine excercise for AI graduate students. Perhaps teaching such a course could help to remind AI practitioners, most of whose work is effectively engineering, that they are (ostensibly) AI intellectuals as well. The fragmentation of AI as a field into discrete subdisciplines which make very different assumptions about the nature of "intelligence" or ignore the latter Big Question entirely is another salient problem. The challenge of presenting the field in simple natural language (that is, English) to a non-technical audience would be a useful 'thought experiment' as well. It might even yield AI graduate students.

An historical approach, which introduces the development of symbolic representation technology (intellectual technology, that is) in the form of the Physical Symbol System in mid-century, and follows its elaboration over the past several decades, could be a reasonable teaching format. This should incorporate a judicious allowance for non-symbolic information processing models. Such a course could introduce the field in simple canonical form, commencing with a presentation of the concept of the sciences of the artificial as stated by Simon in The Sciences of the Artificial. This should be followed by a basic introduction to knowledge representation concepts, such as declarative versus procedural decomposition of problems, state space representations and operators, heuristic versus algorithmic solutions, frames and objects, and various forms of search. The historical development from syntactic search procedures to heuristic search and complementary knowledge representations should be illustrated by reference to prominent programs. The syllabus currently being developed also includes lessons on computer languages, connectionist and "Nouvelle" AI challenges to the 'Classic' PSS tradition, and the novel developments of the recent years, including VLKB, generic task representations, and meta-tools of the knowledge sharing variety. Even a simple presentation of canonical approaches to symbolic representation and to non-symbolic information processing could provoke a greater degree of introspection in bright people who have been deprived of such tools. This would also help to place AI in the context of the sciences relating to complex artifacts. This development has been extraordinarily fruitful, yet most non-scientists only see the artifacts of computing in the opaque form of PCs and operating software. Computing in general and AI in specific go much deeper than word-processing packages, spreadsheets, and computer graphics, but the outside world won't know it unless people from AI present this knowledge to a wider audience.

The creation of a syllabus for teaching an Introduction to AI could help to illuminate AI's undeserved invisibility, as well as to introduce the field's analytical wealth through culture, rather than solely through software tools. If technology is culture, then AI should be high culture.

Notes

(1) This is not an exaggeration. The writer began studying commercial AI and KBS under the auspices of the economics of technological change. When trying to explain the essentials of AI and reason for her interest, she has been met with comments such as: "Who is Noam Chomsky ?"; "I read in the New Republic that AI has failed completely"; and, most commonly: "Hubert Dreyfus says..." by distinguished tenured professors in various social science fields at a major research university.

Bibliography

D. Bobrow edited volume of Artificial Intelligence. 1991.
Haugeland, John. Artificial Intelligence: The Very Idea. 1982.
Kuhn, Thomas." A Function for Thought Experiments". The Essential Tension: Selected Studies in Scientific Change and Tradition. 1977.
Schank " What is AI anyway ?". Partridge and Wilks, eds., The Foundations of Artificial Intelligence: A Sourcebook. New York: Cambridge University Press. 1991: 3-13.
Stefik, Mark." The Knowledge Medium". AI Magazine 1986.

Proposed Syllabus

AI for Everyone: An Introduction to the Sciences of the Complex for Humanities and Social Science Majors

Brief Description:
The one-semester course for juniors and seniors will provide an historical overview of Artificial Intelligence with a symbolic representation slant, and the Sciences of the Artificial to humanities and social science majors.
Requirements:
Reading, class attendance twice per week, and two examinations.
There are no prerequisites.

Week 1. Introduction
This lecture defines the concept of the sciences of the complex, artificiality and complexity. It then explains how a cluster of disciplines has emerged to examine artifacts, or engineered systems of both physical and ephemeral forms (i.e., software). It suggests how this study in turn provides novel and insightful means by which to consider the human phenomena traditionally studied in the humanities and social sciences. The lecture suggests reasons why art history and comparative literature majors should care about this discipline. Such rationales call upon pragmatism (the increasing presence of computation in daily life in industrial nations); cynicism (technology is power); historical interest (this is one of the century's most significant intellectual developments); and philosophical interest (the methods of the sciences of the complex may be applied outside the usual domains).
Terms to be introduced: artifact, sciences of the complex, technology, representation; artificial-natural, AI, cognitive science, computer, hardware, software.
Reading: (some items over the course of the semester): S and H. Dreyfus, What Computers Still Can't Do; Haugeland, J.Artificial Intelligence: The Very Idea. Johnson-Laird, P. The Computer and the Mind. Selected chapters. Gardner, H. The Mind's New Science: A History of the Cognitive Revolution. McCorduck, P.Machines Who Think. Stillings, N.S. et al. Cognitive Science: An Introduction. Simon, H. The Sciences of the Artificial. (2nd ed.).1980. Turing, A.M." Computing Machinery and Intelligence". (1950).

Week 2. Historical Precedents
Artificial Intelligence and cognitive science emerged from various methodological and substantive theoretical roots during the middle decades of the Twentieth century. These roots include formal logic in philosophy; operations research; and cybernetics, or the study of automata. These necessary underpinnings entailed (respectively) improvements in formal logic (Boole, Russell and Whitehead, Frege, Godel); improvements in the representation of information derived from electronics (Shannon's concept of binary-based representation of switching); and inquiries into the nature of complex intelligent systems (anthropomorphic or otherwise) under the auspices of research into intelligent automata (Wiener, Von Neumann, and Newell Shaw and Simon, and Turing). Given the prevailing regime of Watsonian-Skinnerian behaviorism, the agenda for exploring intelligence through cybernetics was practically sub rosa.
Terms to be introduced: the Cognitive Revolution, automata, Behaviorism, cybernetics, operations research, information representation, internal representation, Von Neumann machine.
Reading: Gardener chapters.
Heims, S. The Cybernetics Group. 1991.
Regis, E. Who Got Einstein's Office ? Eccentricity and Genius at the Institute for Advanced Study.

Week 3. The 1950s: Formation of the AI Research Programs
The formation of agendas for studying intelligent information processing was slow and collaborative. Research directions which now seem natural were by no means so before 1956. (Alan Newell:"Deep scientific ideas are exceedingly simple. Others usually see them as trivial"). Directly inspired by the Cybernetic modelling of information processing, several strands of cognitive science and AI coalesced during this period. These include the Chomskian linguistic agenda, the most directly argumentative with Behaviorism; Hebb's proto-connectionist biological metaphor for modelling information processing; and the representation of thought through computational manipulation of symbolic systems. All three will be sketched, and the latter examined in some detail. Newell and Simon's General Problem Solver experiments at Carnegie-Mellon ushered in the modeling of intellectual processes through weak search. This yielded various atomistic syntactic problem-solving and search methods (the latter issue continued into next week).
Terms to be introduced: human-computer interaction, protocol, biological metaphor, toy problem, neuron, intelligence, information processing, computational metaphor, symbolic representation, weak versus strong methods, deep structure.
Reading: Scheerer, E." Toward a History of Cognitive Science". International Social Science Journal 40. Sections from Gardner, McCorduck, Stillings et al.

Week 4. The 1960s: Early Development of Search and Problem-Solving
The PSS research program yielded both significant cognitive science research and powerful KR techniques. Some of these, especially the GPS-based semantically blind search methods, are practically taken for granted at the present, perhaps because of their application to toy problems. The sessions will introduce the concepts of internal representation, intelligence as goal-seeking, and search as movement through a problem space using operators.
Terms to be introduced: semantics-syntax, problem space, state space representation, top-down versus bottom-up, physical symbol system, search, pruning

search spaces, graphs, declarative versus procedural information. heuristic, algorithm, operator, operand.
Reading: TBA

Week 5a. The Tropism to Domains in the Late 1960s

The syntactic approach to blind search in toy spaces shifted pivotally to search of more ambitious domains (areas of expertise) during the late 1960s. This session will describe two of the strategies of this period which moved toward systematic representation of difficult (rather than toy) problems, such as the modeling of scientific reasoning in the Dendral project at Stanford, and the introduction of production systems at CMU in the late 1960s.

Terms to be introduced: domain representation, control structure, generate and test; production systems, control structure;
Reading: Chapter from Newell's Unitary Theories of Cognition; other TBA.

Week 5b. Computer Languages

A language is a protocol for communication, between people, agents, or other systems. Computer languages describe entities and activities, just as natural languages do, but in a far more formal and rigidly defined fashion. Natural language and computer languages however differ in formality, lexical and referential ambiguity, syntactic idiosyncrasies, and nuances of semantics. This lecture will explain the need for formal languages to control engineered artifacts, and the distinctions between low-level and higher-level computational languages.

Terms to be introduced: computer versus natural language; syntax, semantics, lexical ambiguity, protocol; functional versus procedural language, object, encapsulation, global versus local variable; high-level language, platform, environment, application, operating system, binary, Boolean;
Reading: (selected chapters): Brookshear, J.G. Computer Science: An Overview. Baron, N. Computer Languages: A Guide for the Perplexed. 1986.

Week 6. Changing the Scope of Knowledge Representation in the 1970s

Despite the reach of GPS and similar programs, much of its initial technical obstacles had involved very minute manipulation of particles of thought. In the early 1970s, the technology for representing small-scale information manipulation had advanced sufficiently to allow several large-scale concepts of AI representation of intelligence. These include production systems and rules (Newell and Simon, Buchanan and Shortliffe), frames (Minsky), schemata (Bobrow and Collins), and scripts (Schank). These ways of representing intelligence (usually based on cognitive science research on people) were also ways of creating artifacts that had certain forms of knowledge. Thus the concepts embodied both scientific and technological knowledge.

Terms to be introduced: rules, objects, frames, domains, script, expert systems, knowledge base systems; conceptual primitive, blackboard;
Reading: Gardener 1985 and Stillings et al., selections; Selections from Bobrow, Minsky, Newell and Simon, Schank, MYCIN Project.

Week 7. Representing Other Forms of Intelligence

The cognitive metaphor as presented in this course certainly cannot presume to embrace the entirety of human experience and reaction to the environment. Much of the latter includes emotion, physical reaction, and instant reactions to the environment. The latter should be thought of as different sorts of intelligence. "Nouvelle AI" replicates simple, rapid reactions to the physical environment using robotic techniques. The biological metaphor, which tries to reiterate the rudimentary intelligence captured in the neural system of living creatures (not necessarily humans), is a useful if very distinct means to understand intelligence. Still another approach, closer to the Classical symbolic genre, is the examination of intelligence in a modular perspective. There are various components in conscious cogitation, including memory, analogical reasoning, spatial reasoning, etc.

Is it true, as Marvin Minsky is purported to have said, that "Consciousness is overrated" ? If consciousness is many things, how can each of these be translated into useful artifacts ?

Terms to be introduced: massive parallelism 'wetware', 'meat machines', emergent states, perceptrons, Connectionism, robot, knowbot, neural network, modularity of mind, faculties theory;
Reading: M. Minsky and S. Papert, Perceptrons. H. Moravec, Mind Children; D. Hillis article in Wired October 1993. Minsky, The Society of Mind. Daedalus 1988 Special Issue on AI article on Connectionism.

Week 8, a. Critiques of AI, KR and Technology in General:
Week 8, b. AI Views of Emotions, Music, and Creativity:

a) Critiques of AI and cognitive science have been virulent, although most practitioners have paid no notice. The arguments include a) ad hominen critiques of the social skills of hackers (Turkle, Weizenbaum); b) assertions of a simplistically mechanistic worldview entailed in symbolic abstractions (Merchant); c) a more open phenomenological attack on the 'atomism' ingrained in the PSS model (Searle and Dreyfus brothers). However, harsh macho functionalism and 'atomism' are out of style in AI itself, suggesting for a start that the critiques are dated. Given the prominence of distributed and situated knowledge and heterogenous representation among the 'new sensitive AI researchers' of both the Nouvelle AI and the KR communities, are these critiques still valid ?
b) Recent work in AI instead suggests the potential contribution of KR to art, entertainment, and developmental and clinical psychology.

Terms to be introduced: Luddites; phenomenology versus functionalism; intentionality, fringe consciousness;
Reading (a): Selections: D. Harraway; C. Merchant, The Death of Nature; S. Turkle, The Second Self; J. Searle; J. Weizenbaum, Computer Power and Human Reason.

Reading (b): Articles on AI and Music, AI Magazine 1991; work by J. Bates; D. Gelernter. The Muse in the Machine: Computerizing the Poetry of Human Thought. 1994. Schank, R." Creativity as a Mechanical Process". In R. Sternberg, ed. The Nature of Creativity. 1988: 220-238.

Week 9. Current Work in KR:
a. Current Work in Artificial Intelligence:

Criticism of the first generation of KB systems on the grounds of their 'brittle and isolated' quality have been taken very seriously in the KR community. Current research in this community emphasizes the creation of shared knowledge representations. Ongoing research in this effort focuses on the technical and representational challenges of building intermediate "interlingua", a kind of Esperanto between disparate existing computer languages and knowledge representations in those languages. This research challenge, intensified by the end of the Cold War, forces the PSS/ KR community to provide practical software in a shorter time frame than it is used to. The concept of an "intelligent agent" (undefined) is replacing the knowledge base system as the predominant artifact and vehicle for AI platforms. These are less directly anthropomorphic and possibly complementary to more numerous if less lofty applications than KBS seems to have been.

Terms to be introduced: CYC; very large knowledge base; device modeling; agents; knowledge sharing; Knowledge Medium;

b. In-class field trip to an ÀI Lab, TBA: How Things Work project at the Heuristic Programming Project ? Tools for Medical Informatics KR at KSL- CAMIS ?
Reading TBA.

Week 10. Potential Contributions of AI and Cognitive Science to other fields;

The sciences of the complex consist of ways to understand human intelligence and means by which to create new artifacts. This has not been used to analyze the artifacts- physical or intellectual- of the past. Apart from their intrinsic interest, these concepts could be used in the humanities and social sciences in various ways. The descriptive terminology concerning various forms of internal representations and cognitive emulation could be used for the discipline of history. This would pertain specifically to intellectual history (of given cultures, eras, etc.), as well as to the discrete studies of women and different ethnic and economic groups which are now prevalent. The contributions of these fields to economics is currently being explored, and cognitive anthropology is an established field of study. The breadth of domain, extensibility of existing concepts, and characteristics of search or problem-solving which demarcate a given cultural artifact may be described by KR's interdisciplinary terminology in a way not now possible.

Week 11. Final Examination

Acknowledgements: Thanks to Lee Brownston of the Knowledge Systems Lab at Stanford for insightful comments.