Preface: Improving Instruction of Introductory AI

Marti A. Hearst
Xerox PARC
3333 Coyote Hill Rd.
Palo Alto, CA 94304
hearst@parc.xerox.com

Introduction

This symposium was motived by a desire to address the oft-voiced complaint that introductory Artificial Intelligence is a notoriously difficult course to teach well. The response to the initial announcement was overwhelmingly enthusiastic. From the comments we received, it became clear that a large number of instructors are struggling with this course and feel that the time is ripe for such a symposium.

The papers collected here present constructive suggestions for many of the problems posed by the call for participation. A central part of the call requested that the authors identify underlying themes that can be used to help structure the material, even if they do not precisely reflect all of the current viewpoints of the field. An unforeseen consequence is that these working notes reveal not only trends in teaching introductory AI, but also trends in how its members currently perceive the field. This is discussed in more detail in the Summary section below.

These notes only supplement the activity that will take place at the symposium itself. The first and most important goal of the symposium is to provide a forum in which colleagues from different universities and other institutions can meet to discuss issues relating to instruction of introductory AI specifically. Many participants have expressed a desire for the opportunity to discuss ideas, strategies, and problems with more experienced faculty members from other universities. This kind of interaction will be encouraged during small group discussions and discussion time between scheduled presentations and panels.

Synopsis of the Working Notes

This section summarizes the papers and their corresponding sessions, in order of appearance.

Strategies for Unifying Themes

The first two sessions describe overall strategies for teaching AI. The authors of these papers were asked to address some subset of the following questions:

- What underlying themes can be used to help structure the material and avoid the "smorgasbord" problem?
- What is the role of cognitive motivation, if any? Should there be an emphasis on simulating intelligence, modeling the mind, etc.?
- What is the role of formalism?
- Should commercially feasible aspects be addressed, and if so, how?
- How should we address the question, is it still AI if it can be done?
- What role should advanced areas (e.g., vision, robotics) play?
- What kind of hooks should be left for more in-depth courses (e.g., graduate ML, NLP, vision, KR, connectionism, etc.)?
- What is the role of historical developments? Should this be a structuring theme?
- What goals do we have for this course? What should our students walk away with? How does our course meet these goals?
- What should be the interactions with other subdisciplines, e.g.: cognitive systems, intelligent systems?
- What kind of educational program is the curriculum a part of? One semester or a quarter overview? A two or three semester series centered around a theme (e.g., KR)?
- How much overlap should undergraduate versus graduate curricula have? Should a graduate course be remedial introduction for those who never took undergraduate AI? What are the positive and negative results?
- How can we acknowledge and take advantage of the diversity of student backgrounds, perspectives, and learning styles?

The most important question is what overall theme serves to tie the information into a coherent whole, as opposed to a disjoint "smorgasbord". The first two sessions, and their corresponding papers, are supposed

to address this issue in the context of presenting an overall syllabus for an introductory course.

The first set of papers describes approaches organized around traditional AI concepts. Korf emphasizes search, knowledge representation, and specialized applications, Epstein & Teller focus on problem solving, and Ginsberg focuses on a neat, or formal, approach. Epstein & Teller bring up a point, echoed by Zilberstein, Russell & Norvig, and others, that it is important for students in an introductory AI course to develop at least one project or program that "does something smart," since this gets students excited about the field.

The next three papers, by Russell & Norvig, Nilsson, and Zilberstein, describe a new strategy: the *agent-centric* approach to teaching AI. We think this new viewpoint is exciting, and promising enough to merit an entire session. In these papers, the unifying theme of intelligent agents casts the problem of AI as that of describing and building agents that receive percepts from the environment and perform actions. Starting with a simple definition of the agent's internal states and behaviors and working towards more complex ones allows for the gradual introduction of the relevant AI concepts. In Nilsson's framework, the simplest agents have no state and are simply functions that react to their inputs producing output. In the next step the agents are given bits of state (making them Turing-equivalent if given a infinite number of bits). This is followed by a more complex, iconic form of representation. Eventually this progression leads to the incorporation of logic and planning, and language arises from the need for communication among agents. Other AI subject areas, such as vision and robotics, can also be cleanly incorporated into this model, and learning naturally permeates every level of development (as opposed to being a unit added towards the end of the course). Additionally, as Russell & Norvig point out, the agent theme forces the instructor to consider topics that are often ignored or assumed away by other treatments of the subject matter. One example is the need to show how goal-based agents acquire their goals. The agent theme also brings to light the emerging importance of the distinction between AI as the study of thought processes and as the study of action and behavior.

Innovative Approaches and Controversial Stances

The next three papers, corresponding to the next session, present innovative ways to view the task of teaching introductory AI. Batali explores the role of traditional topics associated with AI in the context of a complete cognitive science program. Batali points out that the AI approach to cognitive modeling is no longer the exclusive approach, and that other paradigms such as dynamic systems theory and neurologically-motivated models are increasingly important, in part because they can be more readily aligned to what is known (incompletely) about the brain and human behavior. Nevertheless, there are conceptual paradigms that are associated with AI that are important for the cognitive scientist to know about, including symbolic programming, search, structured knowledge representation, declarative knowledge representation, and constraint propagation.

Martin introduces the notion of "AIgorithms," emphasizing the need to fully integrate AI methods with that of core computer science, and eradicate the "AI Ghetto". Martin focuses on using AI problems to motivate the material in computer science theory courses, and suggests pointing out the parallels between AI approaches and core computer science. Spector also advocates the use of AI as a structure around which to organize larger computer science curricula, but suggests a more radical stance in which AI instruction is distributed in a broad, interdisciplinary curriculum taught through many courses. (Russell & Norvig also emphasize the need to link AI approaches to the related areas in computer science, e.g., linking the students' knowledge of BNF formalims to that of specifying ATN grammars. Klassner also emphasizes this need for integration.)

The next session is a panel entitled "What should a graduate of AI-101 be expected to know?" Hirsh frames the problem and will lead the discussion of the panel. Hayes & Ford have written a position paper stating, in part, that AI is one of the great ideas of the century, and urging the instruction of the foundational ideas of the field in order to enable students to be able to respond to the ideas and arguments of philosophers and others. They also advocate the importance of weeding out what is historical from what is important, warning against the "stamped-earth-foundation" theory of the subject. Skinner continues in the controversial framework, suggesting a curriculum oriented toward non-scientists and centered around non-engineered complex artifacts.

The Role of Programming

The next two sessions focus on issues related to programming in introductory AI courses. In the call for participation, contributors were asked to consider the following questions: Is programming useful or a time waster? Is it better to use existing tools, or is a compromise – modifying existing programs – better? From the papers received, the overall consensus seems to be that programming is very important, and the papers present several suggestions for how best to incorporate programming into the curriculum.

The papers by Eiselt, E. Walker, Goel, and Kumar & Wyatt address the general programming questions. Both Eiselt and Kumar & Wyatt suggest a required companion course that presents a combination of functional (non-imperative) programming, software design, and basic knowledge representation and data struc-

tures. E. Walker describes the use of computers during lecture and advocates the generous use of demonstrations of existing systems. Walker emphasizes a balance between existing code and code written by students, and supplies the students with code that can be easily extended without unduly restricting their designs. Walker also suggests allowing students to work in groups, but having each report on what parts they actually programmed. This allows weak programmers to learn from the design phase but gives proper credit to those students who invest a large amount of time in the programming phase. Goel suggests a design-centric approach to teaching AI, giving the students interrelated projects that demonstrate the competence and performance differences of different AI concepts and methods.

The companion session will include presentations of repositories of tools to accompany AI courses, as well as two modern, innovative instructional systems. Vastola & E. Walker present a flexible graphical tool for teaching reasoning with uncertainty based on Dempster-Shafer theory. M. Walker & Jordan describe Design-World, a framework for studying communication among agents, and a testbed for experiments on the processing limitations of such communicative actions. Papalaskari *et. al* will describe Flair (Flexible Learning with an Artificial Intelligence Repository), a repository of educational material and a highly visual computing environment for use in laboratories associated with the introductory undergraduate Artificial Intelligence (AI) course. Luger & Stubblefield describe a repository of basic programs for introductory and advanced AI courses.

Small College Issues, Non-American Approaches, Other Issues

The next two time periods will consist of "break-out" sessions, in which the participants choose to join one of a set of discussions about particular issues. In the first session, the groups meet and discuss the issue, and in the second, the results of the discussions are presented to all symposium participants.

The following session and its corresponding three papers describe issues confronting instruction at smaller universities and teaching colleges. Danyluk revisits a topic touched on earlier; that is, the importance of folding AI into core computer science instruction, but in this case to attract more students to the field of computer science, mentioning a prior success with a graphics course. Klassner emphasizes the need to present the field from a practical viewpoint for students who are not intending to pursue a PhD in the field, and suggests a curriculum with the dual goals of making the human interface to computers anthropocentric and making a strong link to other computer science courses. Hodgson points out the importance of teaching students how to tackle poorly specified problems, and suggests the use of a large, well-structured system that can have practical applications in the working world.

The final two papers correspond to one of the "break-out" group topics, describe the state of AI research in two non-American countries. Aiello & De Rosis describe AI curricula in Italy, and Gómez & Juristo discuss the state of AI in Spain.

Summary

An exciting, unexpected aspect of this symposium is the emergence of new attitudes towards the field of AI as reflected in new strategies for teaching its introductory course. Several themes arise in the contents of these working notes, and are summarized here.

Below are listed some of the strategies presented for fixing the "smorgasbord" problem:

- agent-centric
- search-centric
- focus on formal methods
- focus on problem solving
- focus around a design science
- focus on design of complex artifacts
- in the spirit of a liberal arts education

Most striking is the development of the agent-centric theme. This theme allows for the incorporation of some of the more recent research (e.g., simple agents, models of inter-agent communication and interaction, etc.) as well as providing a natural framework for folding in ideas that build upon one another. This works because agents can be described in terms of progressing from no state to simple state to iconic representations to more complex representations (including representation of self) to goals to planning to communication, and so on. Issues that are not central to older texts but that occupy important positions in research today (e.g., learning and uncertainty and the role of behavior in cognition) can be discussed throughout this progression. Thus this new trend in instruction seems to be reflecting trends in research and in attitudes towards the field as a whole.

Another strongly sounded theme is the need to distinguish the old from the important. It is now a clearly recognized problem that many AI courses have taught methods and topics only because they were done long ago. Several partipants have voiced the need to change this, and have suggested constructive ways to do so.

Another strong theme is the need to integrate AI with core computer science, both to increase the use of AI methods and to enhance interest in computer science. A subtheme is the need to point out when the two overlap, as in compiler theory and NLP parsing. A related subtheme, pointed out by Hayes & Ford, is the need to integrate AI methods internally as well. They use the example that resolution really involves matching and search, for example, and that TMS and theorem proving are related but are usually taught from different disciplines.

Another theme, focussed around programming issues, is the need to have a project-based orientation and an emphasis on having students solve open-ended, non-toy problems. Several participants advocate the need for a practical approach, while others lay more importance on an understanding of the underlying deep questions.

Finally, the symposium has succeeded in advance in at least one of its goals; that is in bringing together people from many different kinds of affiliations. The following institutions are among those represented by invited attendees:

Beckman Institute, Urbana, IL
Bryn Mawr College
Cornell University
Georgia Institute of Technology
Hampshire College, Amherst, MA
Harvard University
Hunter College, The City University of New York
Information Sciences Institute, Marina del Rey
Knowledge Systems Laboratory
Massachusetts Institute of Technology
McGill University, Montreal
Mitsubishi Electric Research Laboratories
Morgan Kaufmann Publishers
Paris-Sud University, France
Purdue University, West Lafayette, IN
Rensselaer Polytechnic Institute, Troy, NY
Rutgers University
St. Joseph's University, Philadelphia, PA
Stanford University
Sun Microsystems Laboratories, Chelmsford, MA
Temple University, Philadelphia, PA
Universita' di Roma La Sapienza, Italy
University of California, San Diego
University of California, Berkeley
University of California, Los Angeles
University of Colorado, Boulder
University of Florida, Gainesville
University of Geneva, Switzerland
University of Massachusetts, Amherst
University of Michigan, North Campus
University of New Mexico, Albequerque
University of Northern Iowa, Cedar Falls
University of Oregon
University of Pennsylvania, Philadelphia
University of Waterloo, Ontario, Canada
University of West Florida
Villanova University, Villanova, PA
West Chester University, West Chester, PA
Williams College, Williamstown, MA
Xerox PARC

Acknowledgements

We would like to acknowledge the generous support of NSF, which has provided us with a grant to support the travel of many of the workshop's participants. This grant is being administered as the Workshop supplement to Lynn Stein's NYI grant number 9357761-IRI. Oscar Garcia of NSF, Lynn Stein of MIT, and Bonnie Webber of the University of Pennsylvania were instrumental in making this grant come about. AAAI has also generously supplied extra funds in support of this symposium.