Artificial Intelligence and Learning Environments: Preface

William J. Clancey
*Institute for Research on Learning, 2550 Hanover Street,
Palo Alto, CA 94304, USA*

Elliot Soloway
*Department of EECS, University of Michigan, Ann Arbor,
MI 48109, USA*

The promise of computer-aided instruction (CAI) has always been individualized instruction: providing a learner with an environment that is tailored to his or her learning needs and goals. Although there have been notable successes (e.g., see Larkin et al. [9]), the architecture of CAI systems has been inadequate to provide robust and rich learning environments. Starting in the early 1970s, researchers applied an AI perspective to the problem of creating learning environments. The architecture that evolved during this period suggested that an intelligent CAI system (ICAI) would need: (1) an explicit model of the domain and an expert program that can solve problems in the domain, (2) a model of the student that identifies, at a fine-grained level of detail, what the student understands, and (3) a tutoring model that can provide instruction to remediate misconceptions and/or present new material. Not surprisingly, during this period there was also considerable effort in exploring the psychological questions underlying learning, teaching and understanding. Attempting to summarize and spotlight research in this area, Derek Sleeman and John Seely Brown in 1978 edited a special issue of the *International Journal of Man-Machine Studies* [10]. The title of the book, *Intelligent Tutoring Systems*, gave the field its most commonly used name, ITS.

Rationale for the Collection

Several years ago, we decided that it was again time to assemble a special journal issue that would bring to a general AI audience the advances, insights, and problems of this niche field. ITS research was quite active, with consider-

able financial support from ONR, NSF and other government agencies. Research efforts since the 1970s are particularly interesting because they critically test and improve upon existing AI techniques. For example:

(1) Existing computational models are adapted for use in instructional programs (e.g., Woolf applies discourse analysis in the MENO-TUTOR [13]; Langley et al. [8] apply problem space and discrimination learning to student modeling in ACM).

(2) Explanation and modeling capabilities of existing instructional programs are improved by second-generation knowledge representations (e.g., NEOMYCIN (Clancey and Letsinger [5]), repair theory (VanLehn [12]), STEAMER (Hollan et al. [7]).

However, we found it difficult to gather together a large collection of papers. Excellent work had already been published elsewhere, schedules were over-committed, and publication standards for the journal *Artificial Intelligence* posed relatively narrow and strong constraints.

Research Standards

Our reviewing experience suggested the following characteristics for selecting articles for publication:

(1) The project must make an original and significant contribution to AI, with application to instruction. That is, the research must make a clear statement about the *qualitative modeling of processes* in an instructional setting:

    (a) modeling a *subject matter system* (e.g., a steam-propulsion plant, the neurological system, electronic circuit), *or*

    (b) modeling the *reasoning process* for some task (e.g., diagnosis, control), *or*

    (c) modeling *communication processes* (e.g., Socratic discourse).

(2) The article must make a clear statement of how research goals and assumptions impact on the architecture, and compare other architectures designed for similar goals.

(3) If based on psychological experimentation, the research must at least propose an AI architecture design that the experiments support (regarding the model of the domain, reasoning, or teaching processes, e.g., a knowledge representation and control regime that would explain data about student misconceptions).

(4) If a working program is presented without psychological experimentation, the research must demonstrate generality through multiple problems, domains, and/or student trials.

Overview of the Articles

The articles selected for this special issue are representative of the state-of-the-art:

– John Anderson et al.'s paper provides a crisp characterization of the cognitive theory that informs the design of his tutors in the domains of introductory LISP programming, high-school geometry, and introductory high-school algebra.

– Building on earlier work in plan-based program recognition, Lewis Johnson describes a high-performance system, PROUST, that can identify the nonsyntactic bugs in student-generated programs, for a class of nontrivial introductory programming exercises.

– Drawing on research ranging from qualitative reasoning to cognitive models of learning and teaching, Barbara White and John Frederiksen detail the design of and rationalization for their system, QUEST, which attempts to help students acquire successively more sophisticated qualitative models of basic electric circuits.

Research Trends of the 1980s

We can use these papers as examples to understand how the field has matured in the past decade.

First, the papers demonstrate a wide-ranging exploration of alternatives to tutoring dialogues. Anderson et al. demonstrate that one can build highly detailed and accurate models as students solve a problem, by constraining the possibilities for making mistakes. On the other hand, White and Frederiksen argue that, for physical systems such as electronic circuits, one can build models that are self-evident, revealing for students how circuits work. Students can acquire and debug these models on their own, without the need for a detailed student model or highly tuned tutoring strategies. Finally, Johnson explores a situation where one doesn't have control of the learning environment, and instead accepts whatever the students produce. Students create PASCAL programs using "standard" techniques and submit their 60–100-line programs to PROUST for analysis. The point here is not new to AI—there are no guaranteed approaches, we are still at the stage of investigating domains and contrasting alternative methods.

Second, all three research efforts are directly motivated by psychological concerns. Anderson et al. are absolutely clear on this point: Their goal is to do psychology—developing models of learning and understanding—and their methodology is to implement tutoring systems that allow them to explore their theory. Similarly, White and Frederiksen's interest in qualitative reasoning is not so much to advance that field, per se, but to understand the interplay of

qualitative reasoning, human expertise, and instruction. While PROUST is a performance program, not primarily conceived as a cognitive model, Johnson's research was directly motivated by psychological research on the nature of programming expertise.

Third, the research is all strongly driven by empirical trials.

Strikingly, these themes are not dominant in other areas of AI today. A great deal of AI research emphasizes building autonomous agents, rather than viewing AI as just a modeling tool, exemplified by QUEST and PROUST, and implicit in Anderson's de-emphasis of complex teaching dialogues. Second, the questions that AI researchers ask (e.g., concerning the formal properties of a representation and reasoning strategy) need not have psychological relevance. Finally, although empirical trials are important in knowledge engineering and natural language research in general, there is probably no other subarea in which evaluation is so important and yet so difficult as instruction. Here we inherit all the legendary educational problems of evaluating teaching methods and testing what students have learned.

Cognitively oriented, empirical research focusing on AI programming as a modeling methodology can provide insight into a variety of AI problems: Anderson et al.'s claims about the representational and process properties of PUPS (PenUltimate Production System) architecture, White and Frederiksen's perspective on qualitative reasoning, and Johnson's observations on adequacies of various diagnosis strategies, are all pertinent to AI research and deserve attention.

New Perspectives

A number of new perspectives and techniques are shaping the field today (Pea and Soloway [10]):

(a) viewing student modeling as machine learning (and, hence, knowledge acquisition) research;

(b) applying qualitative simulation to modeling complex processes in engineering (e.g., manufacturing process control);

(c) using graphics for explanation;

(d) relating instructional dialogues to natural language discourse research;

(e) determining the extent to which student modeling is possible or necessary;

(f) designing shells for use in multiple domains by teachers;

(g) defining sequences of activities that drive learning by exploration (e.g., through problem-solving failures);

(h) reconceiving epistemological assumptions about the nature of expertise and the differences between novices and experts (Brown, Collins and Duguid [1]);

(i) tackling less formal domains (e.g., history), emphasizing case-based reasoning (Farrell [6]).

Looking towards the future, we believe several shifts in practice need to take place in order for the field to progress:

– More expert systems need to be used for teaching, e.g., there appears to be a preponderance of math, science, and programming applications without balanced concern for engineering, medicine, and social sciences (see [4] for discussion).

– Systems need to be evaluated by applying them in different domains, to determine the generality of the methods and relate them to domain properties. Almost no one studies when programs break and why. Yet, it is precisely failure that helps us understand why programs work, what leverage we are getting from our techniques (see, for example, [3, Chapter 7] or [2, Fig. 6]). Following this advice requires expensive, time-consuming effort, and possibly larger groups working together.

– And most importantly, the small number of knowledgeable AI researchers working on these problems most significantly limits our progress. More support is needed for student researchers. A good goal, albeit controversial, is to provide convincing demonstrations so that engineers, physicians, and other practitioners take for granted that computers are preferable to books for instruction. A secondary goal would be to merge expert systems, natural language, machine learning, knowledge representation, and qualitative simulation in a way that demonstrates synergy in an integrated system.

With focus from a better understanding of the design of practical learning environments and computer power to deliver applications inexpensively, there is good reason to be optimistic that this field will continue to receive the support and research attention that enabled its growth in the past decade.

REFERENCES

1. Brown, J.S., Collins, A. and Duguid, P., Situated cognition and the culture of learning, *Educ. Researcher* **18** (1) (1989) 32–42; also: IRL Rept. IRL88-0008, Institute for Research on Learning, Palo Alto, CA (1988).
2. Clancey, W.J., From Guidon to Neomycin and Heracles in twenty short lessons, ONR Final Report 1979–1985, *AI Mag.* **7** (3) (1986) 40–60.
3. Clancey, W.J., *Knowledge-Based Tutoring* (MIT Press, Cambridge, MA, 1987).
4. Clancey, W.J., Functional principles and situated problem solving, *Behav. Brain Sci.* **10** (3) (1987).
5. Clancey, W.J. and Letsinger, R., NEOMYCIN: Reconfiguring a rule-based expert system for application to teaching, in: W.J. Clancey and E.H. Shortliffe (Eds.), *Readings in Medical Artificial Intelligence: The First Decade* (Addison-Wesley, Reading, MA, 1981) 361–381.
6. Farrell, R., Facilitating self-education by questioning assumptive reasoning, in: *Proceedings AAAI-88*, St. Paul, MN (1988) 2–6.
7. Hollan, J.D., Hutchins, E.L. and Weitzman, L. STEAMER: An interactive inspectable simulation-based training system, *AI Mag.* **5** (2) (1984) 15–27.
8. Langley, P., Ohlsson, S. and Sage, S., A machine learning approach to student modeling, Tech. Rept. CMU-RI-TR-84-7, Robotics Institute, Carnegie-Mellon University, Pittsburgh, PA (1984).

9. Larkin, J. and Chabay, R.R., *Computer Assisted Instruction and Intelligent Tutoring Systems*: *Shared Issues and Complementary Approaches* (Erlbaum, Hillsdale, NJ, 1990).
10. Pea, R. and Soloway, E., Mechanisms for facilitating a vital and dynamic education system: Fundamental roles for education science and technology, Final Report for the Office of Technology Assessment, U.S. Congress (1987).
11. Sleeman, D. and Brown, J.S. (Eds.), *Intelligent Tutoring Systems* (Academic Press, New York, 1982).
12. VanLehn, K., Felicity conditions for human skill acquisition: Validating an AI-based theory, Interim Rept. CIS-21, Cognitive Instructional Science Group, Xerox PARC, Palo Alto, CA (1983); also: Ph.D. Thesis, MIT, Cambridge, MA.
13. Woolf, B. and McDonald, D., Context-dependent transitions in tutoring discourse, in: *Proceedings AAAI-84*, Austin, TX (1984) 355–361.