Intelligent Tutoring, but Not Intelligent Enough

Heinz Mandl and Alan Lesgold (Eds.)
Learning Issues for Intelligent Tutoring Systems
(Springer-Verlag New York);
3-540-96616-1 (Springer-Verlag Berlin). $20.00 paperback

Joseph Psopta, L. Dan Massey, and Sharon A. Mutter (Eds.)
Intelligent Tutoring Systems: Lessons Learned
ISBN 0-8058-0023-9 (hardcover);
0-8058-0192-8 (paperback). $79.95 hardcover; $34.50 paperback

Review by
Patrick Suppes

Heinz Mandl, head of learning research and director of the German Institute for Distance Studies at the University of Tübingen (West Germany) and president of the European Association for Research on Learning and Instruction, is coeditor, with E. DeCorte, S. N. Bennett, and H. F. Friedrich, of Learning and Instruction. Alan Lesgold, professor of psychology and intelligent systems at the University of Pittsburgh (Pennsylvania), is coeditor, with R. Glaser, of Foundations for a Psychology of Education. Joseph Psopta, research psychologist at the Army Research Institute for the Behavioral and Social Sciences (Alexandria, Virginia) and the University of Maryland Institute for Advanced Computer Studies (College Park), is coauthor, with M. J. Farr, of Intelligent Tutoring Systems Encountering Reality. L. Dan Massey, manager at Systems and Software Engineering Department, BBN Systems and Technologies, Inc. (Cambridge, Massachusetts), is author of Probability and Statistics. Sharon A. Mutter is research assistant professor in the Human Performance Laboratory at the Catholic University of America (Washington, DC). Patrick Suppes, Lucie Stern Professor of Philosophy and director of the Institute for Mathematical Studies in the Social Sciences at Stanford University, is recipient of the American Psychological Association's Distinguished Scientific Contribution Award and is a member of the National Academy of Sciences. He is coauthor, with D. H. Krantz, R. D. Luce, and A. Tversky, of Foundations of Measurement, Vols. 2 and 3.

During the past decade or so there has been a substantial research and development effort, not only in this country but in many other parts of the world, to create intelligent tutoring systems. These efforts have as their overall objective the application of current research in cognitive science to computer-
based instruction or training. The hubris of a new movement in education is to be seen throughout the pages of both books under review. Taken together, they give a good sense of current work. The book edited by Mandl and Lesgold contains 13 articles by 19 authors. The longer book, edited by Psotka, Massey, and Mutter, contains 18 articles by 39 authors, as well as introductions to various sections by the three editors. There is much overlap in authorship between the two volumes, and it is fair to say that many, if not all, of the most active researchers in intelligent tutoring systems have written for one of the two volumes.

In their preface, Mandl and Lesgold nicely summarize the four major components that are usually to be found in an intelligent tutoring system. These are the expert knowledge component, the learning modeling component, the tutorial planning component, and the communication component. Both volumes contain descriptions of many interesting developments and a wealth of general remarks about the process of constructing intelligent tutoring systems. It is necessary to restrict any substantiation of this general claim to a few salient examples. In the Mandl and Lesgold volume, for example, VanLane’s article on the theory of impasse-driven learning is notable for its detailed attention to past theorizing, building as it does on the “buggy” studies of Brown and Burton and on his own previous work on repair theory, which attempted to give an analysis of the origin of bugs. VanLane’s article is one of the few in either volume that acknowledges and refers to the extensive earlier research on the learning of elementary mathematics. His references go all the way back to Buswell’s work in the 1920s. His new idea concerns what happens when the child who is learning elementary arithmetic encounters an impasse.

One of the few articles to present any systematic data is that by Fisher and Mandl (in Mandl and Lesgold) on the improvement of the acquisition of knowledge. Unusual for these two volumes, an experiment with well-defined variables was conducted, and statistical analyses of the data are presented. In addition, the article is good on prior ideas of learning that go back to the classic work of Thorndike in the early part of the century. Some systematic data are also to be found (in Psotka et al.) in the article by Tenney and Kirkman on the development of trouble-shooting expertise in radar mechanics. In both cases where data are presented, no sophisticated comparison with theoretical quantities derived from an explicit model is given.

Wilkins, Clancey, and Buchanan have a useful and detailed chapter (in Psotka et al.) on using and evaluating differential modeling. The function of differential modeling is to identify the differences between the student problem solver and the computer-based expert system. This is one conceptual approach to the problem of identifying and correcting the student’s approach to problem solving.

The article by Scanlon and O’Shea (in Mandl and Lesgold) on learning elementary physics at the high school or university level reports interesting qualitative data from a long-term project. A small number of students were studied in detail, especially by the use of verbal protocols obtained from the students. Scanlon and O’Shea have some significant conclusions that run counter to much of the conventional wisdom in cognitive science as applied to instruction. Perhaps the most important conclusion is that the use of multiple representations in problem solving can be more a hindrance than a help to the student. They found that students who were not doing well moved back and forth between graphic representations and representations in terms of equations, without understanding either representation adequately. The main point is that students who have a bad understanding of both representations will move from one to another seeking a solution and will reach incorrect solutions in both approaches. The key observation is that students tend to leave one representation without thoroughly analyzing how it may be used and move on to another representation too quickly.

Omission of comment on the many other articles does not mean they are not of interest, but it is not possible to review them individually even in a cursory fashion. Consequently, I turn to some general comments and impressions.

First, there is a tendency throughout the two volumes for the tone to be too self-congratulatory on what has been accomplished, for the results, as yet, are rather limited. In their preface, for example, Mandl and Lesgold state that previously “the only theory available to guide instructional development was behavior theory, which poorly matched the cognitive goals of education” (p. v). However, almost all the current articles are far removed from complex learning, and the cognitive models proposed are simple and schematic in character. Moreover, the subject matter that is treated is ordinarily quite elementary. The several articles on mathematics deal almost entirely with the learning of very simple algorithmic parts (e.g., the four rational operations of arithmetic or similar computations in algebra). The much richer material of intuitive geometry or complex word problems at the level of calculus or beyond is not considered. Looked at from a long historical perspective, what is discussed in these volumes in terms of elementary mathematics is not an enormous intellectual advance on the work of Thorndike early in this century on arithmetic and algebra. In fact, Thorndike’s work was probably more systematically oriented toward the analysis of data in these matters than are any of the articles in these two volumes. Another area in which there is an enormous gap between the schematic models proposed and the rich data available is in the use of verbal protocols, which are themselves not analyzed with any grammatical or semantic sophistication.

Second, there are many references to learning research throughout both volumes, but the references are almost all soft and qualitative in character. Both volumes are lacking in references to data from other experiments or to quantitative theory. In this connection, perhaps the most surprising absence—and one that is characteristic of what often happens in the interface between the sciences and education—is that there is almost no mention of newer developments in learning theory (e.g., the current extensive work on neural nets and connectionism). In other words, there is no use of the newest theory of learning, which itself has many relations to older behavioralist theories.

Third, there is almost no study of the validity of representing a given domain of knowledge in some particular way in spite of the great emphasis on representation of knowledge. In the same vein, there are very few appeals to any fundamental research on representation (e.g., the extensive psychological research on mental representations by Roger Shepard and his collaborators). Surprisingly, there is no reference to Shepard’s work in the index of either volume. As in the case of the recent work on learning just mentioned, this is an old tradition in education that needs correction: There is no attempt to stay abreast of the best current relevant research in psychology.

Fourth, there is a general lack of the critical assessment that is characteristic of good scientific work. For example,
there is no discussion of the fundamental computation difficulties of providing deep assistance to students with an intelligent tutoring system if the domain of problems is complicated. It is very likely that it would not be hard to construct examples in mathematics or science where to pursue the line of reasoning that a student would use would require unreasonably lengthy computations (in technical jargon, nonfeasible computations) to tailor advice precisely to the student's activities. What is disturbing is the absence of detailed theoretical discussions of any problems of this character. Govindaraj, for instance (in Psotka et al.), does discuss intelligent tutoring systems for the diagnosis of failure in complex engineering systems, but he does not provide anything but a very qualitative analysis of the difficulties.

Fifth, in the same vein there are some excellent discussions of troubleshooting and of the problems of reasoning with incomplete knowledge. An example is the article by Massey, DeBruin, and Roberts (in Psotka et al.), but surprisingly enough, here and elsewhere, there is no mention or development of the use of probabilistic methods of analysis, in spite of their long history in other parts of educational research and of their extensive use in contemporary engineering and physics research.

Finally, there is the general spirit that more is better, but there are no comparative data at all to show that the use of an intelligent tutoring system will actually produce better learning results than a much simpler system. Perhaps the most interesting remark that I noticed in this context in the two volumes was the one by Scanlon and O'Shea (p. 262, Mandl and Lesgold) contrasting the difference in approaches to teaching problem solving on the part of physicists and mathematicians in the Open University in Great Britain. The physicists, according to this account, rely on very large amounts of practice with feedback on accuracy of solutions; in contrast, the mathematicians try to teach problem-solving strategies explicitly. There is no attempt to make a claim as to which method is better, but it is the kind of study that is very much needed. There are some anecdotal data that one of the problems of intelligent tutoring systems can be the following: If the tutoring system is too slow, then spending time on diagnosis and explanation is less efficient than having the student work additional exercises, using relatively simple correction procedures when errors are made.

Almost all of us believe that there is certainly something to tutorial methods, but what the two volumes do not really accomplish is the sophisticated presentation of data showing some of the particular strategies that are effective and some that are not effective as tutorial approaches. I have been critical in the preceding remarks, but I want to conclude by stating that my criticisms have been driven by the conviction that the development of effective intelligent tutoring systems is as important as any instructional topic in education. I also emphasize that much interesting and suggestive work is reported in these two volumes, even if the scientific level of most of the reports is disappointing.