Using a Well-Structured Model to Teach in an Ill-Structured Domain*

Vincent Aleven and Kevin D. Ashley

University of Pittsburgh
Intelligent Systems Program, School of Law,
and Learning Research and Development Center
Pittsburgh, Pennsylvania 15260
aleven+@pitt.edu, ashley+@pitt.edu

Abstract

Our goal is to develop a tutoring system, called CATO, that teaches law students skills of making arguments with cases. CATO's domain model provides a plausible account of legal arguments with cases, but is limited in that it does not represent certain background knowledge. It is important, however, that students learn to apply and integrate this background knowledge when making arguments with cases. Given that modeling this background knowledge is difficult in an ill-structured domain like legal reasoning, it is worth exploring how effectively one can teach with a model that represents argument structure but relatively little background knowledge.

The CATO instructional environment, comprising a case database and retrieval tools, enables students to apply the CATO model to a specific problem. In a formative evaluation study with 17 beginning law students, we compared instruction with the CATO environment, under the guidance of a human tutor, against more traditional classroom instruction not based on the CATO model. We found that human-led instruction with CATO is as good as, but not better than, classroom instruction. However, answers generated by the CATO program received higher grades than the students' answers, suggesting that the model can potentially be employed to teach even more effectively. Examples drawn from protocols show that students were able to use the CATO model flexibly and integrate background knowledge appropriately, at least when guided by a human tutor.

Introduction

In this paper, we report a formative evaluation study of CATO, a computer-based instructional environment for teaching law students skills of making arguments with cases. CATO is based on a computational model of case-based legal argumentation developed in previous research (Ashley, 1990; 1991). Legal experts find that arguments based on the CATO model are reasonable, even if they are not always optimal, and that the model provides a plausible account of legal arguments with cases (Rissland, 1990). On the other hand, the model does not represent all types of case-based arguments; nor does it represent all background knowledge that human reasoners bring to bear. And while the model provides criteria for saying that some arguments are better than others, arguments that do not conform to the model may be reasonable, too, occasionally even better than arguments based on the model. We believe that these characteristics are exactly what one would expect in an ill-structured domain like legal reasoning.

Many authors, however, seem to imply that the fact that certain knowledge is not included in a domain model makes it unfit for teaching. In order to use the MYCIN knowledge base for teaching, much knowledge had to be made explicit (Clancey, 1983). Also, if a domain model is to be used as a "standard" to evaluate student behavior (Wenger, 1987), the model must be capable, at minimum, of generating all reasonable solutions. The domain model of a model-tracing tutor, for example, must generate all solution paths that are worth teaching (Anderson, et al., 1990).

In a domain like legal reasoning, however, one cannot reasonably expect to represent all background knowledge. This domain is ill-structured and indeterminate. Provably correct answers do not exist. The domain involves large amounts of knowledge, legal knowledge as well as knowledge about the regulated domain (e.g., the corporate world), neither of which can be easily circumscribed. In light of these difficulties, one has to ask: What contribution can an ITS make in a domain like the current? What knowledge needs to be represented and what can be left unrepresented, and with what effect on instructional efficacy? And if certain knowledge is not represented in a form that the program can understand, is it useful to make it available through other means?

In this research, we explore the hypothesis that one can provide useful computer-based practice with a model of case-based argumentation that includes mostly argumentation knowledge, but contains relatively little knowledge about the legal domain. *A priori,* this hypothesis seems reasonable. The CATO model enables a program to compute the relevance of cases and to generate reasonable arguments. Also, the model seems to provide a useful conceptual framework for comparing and contrasting cases and making arguments with cases. Nonetheless, there is a danger that students will rely on the model too much and apply it without drawing on their background knowledge. The model is a useful stereotype, but like all stereotypes, it can be overused, taken too seriously, applied too mechanically.

We investigated whether a human tutor could employ the CATO model and instructional environment to teach argumentation skills to beginning law students. In a controlled experiment, we compared human-guided instruction with CATO against classroom instruction that teaches the same material without the use of the CATO model or tools. In this paper, we describe the CATO model of case-based argumentation, present the results of our experiment, and discuss examples in which students used the model intelligently, integrating their background knowledge in their arguments.

* This work is supported by an NSF Presidential Young Investigator Award and grants from the National Center for Automated Information Research, West Publishing Company and Digital Equipment Corporation.

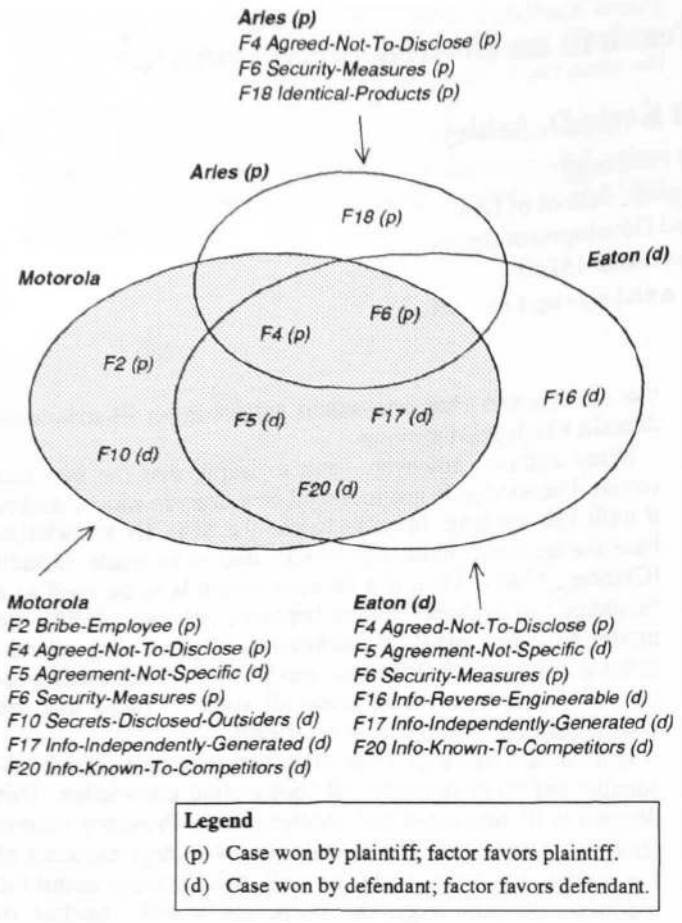

Figure 1: Comparing and contrasting cases in terms of factors.

Case-Based Reasoning in the Legal Domain

While one may think of the law as a system of rules, attorneys very frequently reason with cases: in order to evaluate how a problem should be decided, they compare and contrast it to past cases. A legal problem is a "fact situation," described in narrative text, about which a legal dispute has arisen. The issue to be decided is whether the plaintiff is entitled to some form of legal relief against the defendant. Attorneys employ cases for two reasons: Where the legislature has not provided a detailed statute, an attorney must look to court decisions to determine what the law is and how it has been applied. In statutory domains, attorneys use cases to interpret open-textured terms for which the statute provides no definition. Skills of reasoning with cases are important in other domains as well, including practical ethics, business, and political science.

The CATO model covers arguments analogizing a problem to past cases and arguments in which past cases are used to emphasize strengths, downplay weaknesses, and cover the opponent's bases (i.e., the weaknesses that the problem presents). It also covers responses to such arguments by distinguishing or citing counterexamples. The CATO model provides an abstract argument plan to guide the organization of an overall argument supported by multiple cases, and a set of relevance criteria for selecting the cases (from any given set of candidates) that will make the most effective argument (Aleven and Ashley, 1993; Ashley and Aleven, 1992).

In the CATO model, cases are represented as sets of factors, abstractions of facts that tend to influence the outcome of a legal claim (Ashley, 1990). Figure 1 illustrates how an arguer can make arguments about the *Motorola* problem by comparing and contrasting it to past cases in terms of factors. *Motorola* involved a claim for trade secrets misappropriation; the plaintiff complains that the defendant (often a corporate competitor) has used plaintiff's trade secret (often, valuable product development information) to gain an unfair competitive advantage. Let us assume that an arguer has read the narrative description of the *Motorola* facts and decided that the following factors apply: On the one hand, it helps the plaintiff that it secured a nondisclosure agreement from the defendant (F4), that it took measures to keep its information secret (F6), and that defendant tried to lure away plaintiff employees by offering a bribe (F2). On the other hand, in defendant's favor, the nondisclosure agreement was not specific as to what plaintiff regarded as its secret (F5), the plaintiff disclosed its alleged secret to outsiders (F10), the information was known to competitors (F20), and the defendant developed its information through its own efforts (F17).

The Venn diagram indicates how *Motorola*'s factors compare to those of two other cases: *Aries*, a case won by the plaintiff, and *Eaton*, a case won by the defendant. The plaintiff in *Motorola* can cite *Aries* to support an argument that it should win. *Aries* shares two factors with *Motorola*: F4 Agreed-Not-To-Disclose and F6 Security-Measures, both of which favor the plaintiff. Plaintiff can analogize *Motorola* to *Aries*, citing the shared factors as relevant similarities and arguing, in effect, that these factors warrant a favorable decision in *Motorola* (as they did in *Aries*).

The defendant, however, can respond by distinguishing *Aries*, pointing out the relevant differences, that is, the factors that *Aries* and *Motorola* do not share and that push to an opposite result in each case. For example, in *Aries*, the defendant developed a product that was identical to plaintiff's (F18). This was not so in *Motorola*. Also, *Motorola* has four pro-defendant factors that do not apply in *Aries*, as is clear from the Venn Diagram (F5, F10, F17, F20). The presence of these opposing factors, defendant argues, warrant opposite outcomes in the two cases. The defendant can make an even more devastating response to plaintiff's argument citing *Aries*: It can cite the *Eaton* case, which was won by the defendant, as a counterexample. Notice that *Eaton* shares a more inclusive set of factors with the *Motorola* problem than does *Aries*: It has all factors that *Aries* shares with *Motorola*, but shares additional factors with *Motorola* as well (namely, F5, F17, and F20). When these additional shared factors are taken into account, defendant argues, a decision in favor of the defendant is warranted, as it was in *Eaton*.

The instructional goal in CATO is for students to learn to make the types of arguments covered by the CATO model and to make better use of commercially available case law databases to find cases to cite in such arguments. Currently, the CATO environment provides tools and resources that help students apply the CATO model to make arguments about a problem; it does not do any active tutoring. It provides a database containing, for each of 45 trade secrets

cases, a list of factors and a textual summary of the case (a "squib"). CATO also provides tools for retrieving, displaying and comparing cases in terms of factors. The CATO query language enables students to retrieve cases from CATO's database with any boolean combination of factors. CATO can list the factors for a case retrieved from the database, and can display a comparison of the factors of two cases, marking the shared factors and the distinctions. CATO is based on a knowledge base implemented in Loom (Ashley and Aleven, 1994). As is described in the next section, students (guided by a human tutor) have used these tools to develop case-based arguments about legal problems.

While the CATO model supports useful argument-making, it is a simplification. Attorneys do not use a fixed vocabulary of factors to represent and compare cases; they are free to invent their own terms to characterize the legally relevant facts. They draw on various types of legal knowledge to reason about why certain facts matter and employ knowledge about the corporate world to interpret the narrative descriptions of cases. The CATO model does not include background knowledge related to what factors mean and why they matter, at least not in a form that CATO can handle.

It is important that students learn to apply and integrate this background knowledge when they make arguments with cases. We make some of it available through a preparatory lecture, by providing textual summaries of cases, through the names of factors, and via a human tutor. (It is our goal, however, that students will use CATO without a human tutor.) It is our impression that it is not very difficult for students to understand what an individual factor means or why it matters. The problem is invoking this knowledge at the right time while making arguments with cases. The question, then, is whether the CATO conceptual framework prompts students to integrate their background knowledge appropriately.

An Experiment with the CATO system

We conducted a formative evaluation study to investigate whether beginning law students learn useful argumentation skills as a result of practice with the CATO system, under the guidance of a human tutor (Kevin Ashley). We compared human-led CATO instruction against classroom instruction that teaches the same material, but without using the CATO model or system. The design of this experiment is somewhat unusual, compared to other evaluation studies of ITSs (Shute and Regian, 1993), in that a human tutor is involved in the CATO instruction. Obviously, this means that any observed improvement in students' learning cannot be attributed to CATO alone. However, given that only the CATO model but not the CATO tools had before been used for tutoring[1], we wanted to see if a human tutor, "constrained" to using these tools, could succeed. Also, we wanted to generate protocols of teaching with these tools.

The subjects in the experiment were first-semester law students of the University of Pittsburgh School of Law, divided into an experimental group comprising 7 students, and a control group of 10 students. At the start of the experiment, all students received preparatory instruction about trade secrets law and about the use of the Westlaw legal information retrieval system. All students then took a written, take-home pre-test exam involving argument-making and case-finding questions, designed to test students' skills in making arguments with cases and framing queries to find relevant cases using Westlaw, a commercially available case law database.

Each student in the experimental group participated in two two-hour sessions of instruction with CATO. During these sessions, students worked in pairs, guided by a human tutor. At the start of their first session, the tutor introduced students to CATO's case representation scheme and query language, showing sample factors, cases, and queries. The rest of the time, the students used the CATO tools to analyze two trade secrets problems and outline an argument with cases selected from CATO's database. After reading the narrative description of the problem facts, students selected CATO factors to represent the factual strengths and weaknesses present in the problem. They then used the CATO query language to implement general research strategies aimed at finding cases to cite in an argument (i.e., cases that share interesting sets of factors with the problem). All students were able, with a small amount of trial and error, to retrieve relevant cases from the CATO database. (As noted below, the human tutor prompted the students in various ways to use the CATO tools. Vincent Aleven handled all student interaction with the CATO program, for example, typing the queries that the students dictated.) The students also used the CATO tools to make initial judgments of what the retrieved cases mean and whether they are relevant to the problem. They inspected the applicable factors of the retrieved cases to get a first impression of the legally relevant facts of each case. They judged the relevance of the retrieved cases by comparing and contrasting them to the problem, in terms of factors. When a case seemed relevant, based on its factors, the students read the squib (i.e., narrative summary of the case), in order to verify their initial relevance judgment and decide whether to cite the case in an argument. All students were able to outline a convincing argument.

The human tutor's main role was to guide the students during this process, that is, to get the students to use the CATO tools and to help them interpret what they had found with the tools. He exerted a fair amount of control. He often suggested things that the students could do next. He confirmed that students had made good arguments or made accurate evaluations of a case's relevance. He summarized arguments that students made in ways that made it clear how they relate to the CATO model. He also helped students keep track of the overall state of their arguments under development and sometimes helped them see how different cases fit into an overall argument.

However, he did not force students to follow any specific research strategy. He did not suggest any specific queries to try, nor did he insist that students agree with his own evaluations of the relevance of cases, or make any arguments he may have had in mind. Also, he did not demonstrate the overall process, nor did he engage in argument exchanges with the students. He also did not engage in long discourses

[1] In a previous study, a human tutor used a lesson plan and computer-generated examples based on the model (Ashley and Aleven, 1992). The students did not use the CATO tools.

	Pre-test		Post-test	
	argument-making	argument-making and case-finding	argument-making	argument-making and case-finding
Experimental Group Average	52	45	41	42
Control Group Average	56	50	46	45
CATO score	68*		78*	

Table 1. Percentage of maximum score on pre-test and post-test. * denotes highest score.

about legal argumentation or trade secrets law. In short, the CATO instruction focused students on the importance of comparing and contrasting cases and gave them practice in a process of making arguments with cases.

The control group students were taught the same material in a classroom setting, which did not involve the CATO model or tools. Instead, students used the Westlaw database to retrieve relevant cases[2]. The control group instruction thus comprised a more traditional way of teaching the same material, using regular tools, but was not part of the regular legal curriculum. We organized two two-hour sessions with the control group students, led by legal methods instructors of the University of Pittsburgh School of Law. Each session started with a lecture about legal research and argumentation, focusing on the use of cases in arguments. The students then analyzed the same two trade secrets problems as the experimental group. Each student used the Westlaw legal information retrieval system to find cases that are relevant to the problem. In a group discussion led by the instructor, the students considered how the cases they retrieved could be used in an argument. During the second session, students engaged in an oral argument exchange, citing the cases they had retrieved.

All students then took a written, take-home post-test exam containing the same types of questions as the pre-test. The pre-test and post-test answers were graded by legal methods instructors (not those who conducted the control group sessions), who were not familiar with the CATO model. The grading criteria were compatible with the CATO model, but were not phrased specifically in terms of the CATO model. For the argument-making questions, we included in the materials to be graded a set of answers generated by CATO, without informing the graders.

The results, shown in Table 1, can be summarized as follows: (1) Overall, the control group performed slightly better on the pre-test and post-test than the experimental group, but the difference was not statistically significant. (2) The answers generated by CATO were graded higher than any student's answers. (The fact that the post-test scores are lower than the pre-test scores does not convey any useful information, because the tests were graded by different graders and were not equally difficult.) We interpret the first finding as saying that instruction with CATO (under the guidance of a human tutor) was as good as, but not better than, classroom instruction. The second finding suggests that the students did not learn everything there is to learn about the CATO model and therefore, that the CATO model can potentially be employed more effectively to teach students than it was in the experiment. On the other hand, it may simply be that CATO's answers were easier to understand because they were better organized or even printed.

Examples of Students' Using the CATO Model and Tools

In this section, we show examples that illustrate that students (at least when guided by a human tutor) were able to use the CATO model and tools as a useful conceptual framework. At the same time, the examples illustrate that students were able to integrate knowledge about factors that is not represented explicitly in the CATO model.

The students in the first example use their knowledge about the meaning of factors to improve on CATO's argument based on factors. They considered how they could use *Eaton*, a case won by the defendant, in an argument on behalf of the defendant in *Motorola*. On the one hand, *Eaton* is very similar to *Motorola* (their factors overlap considerably, as the Venn diagram in Figure 1 indicates) and therefore a good case to cite for the defendant. On the other hand, plaintiff can respond by distinguishing *Eaton*, drawing attention to the distinguishing factors F16 Info-Reverse-Engineerable in *Eaton* and F2 Bribe-Employee in *Motorola*. In other words, CATO would have made the following argument: *Eaton* is distinguishable, because in *Eaton*, plaintiff's information could be discovered through reverse engineering, that is, analyzing plaintiff's product (F16), while this was not so in *Motorola*. Also, in *Motorola*, the defendant offered plaintiff's employees a salary increase if they would switch employment (F2). This was not so in *Eaton*[3].

However, the students made convincing arguments that these seeming distinctions were of no account, reasoning

[2] The Westlaw database contains the opinion of virtually every case ever published. The opinions are the official documents produced by the courts; they are many times longer than the squibs used in the CATO instruction. Westlaw's full-text retrieval system enables users to retrieve documents based on the words that appear in them.

[3] Currently, CATO does not display Venn diagrams, but displays a comparison of the factors of two cases in list format.

about the meaning of the factors, using their knowledge about the corporate world, and taking into account the specific facts behind the factors (which they read in the squib). They claimed that the presence of factor F16 Info-Reverse-Engineerable in *Eaton* did not make much difference in light of the other factors and therefore should not be regarded as a distinction. They reasoned that where information is known in an industry (F20), as in *Eaton* and *Motorola*, it is immaterial whether that information is also reverse-engineerable (F16), because this hardly makes it any more widely available than it is already.

They also argued that the facts related to the F2 Bribe-Employee factor in *Motorola* were not so extreme that they should be regarded as a strength on the part of the plaintiff or a way of distinguishing *Motorola* from *Eaton*. The *Motorola* squib said that plaintiff's employees "received substantial salaries and bonuses" when they switched employment. The students reasoned that this should not be interpreted as an attempt to bribe them to bring trade secrets: "I think it would be fairly easy for the defendant to say ... that in a market-oriented situation bonuses and salaries were being made to encourage ... [plaintiff's former employees] to come over to the company. Not necessarily for any devious means but just because they were skilled workers ..."

In sum, the students were able to improve on the CATO conceptual framework, applying knowledge not represented in the CATO model. CATO framed the argument; the students went beyond the model in evaluating, augmenting, and critiquing the argument. That is good behavior to encourage. Also, just making an argument naturally encourages a response.

In the second example, a student use factors to formulate a general theory about trade secrets law and to test it against cases. Formulating and testing theories is an important aspect of legal scholarship, closely related to argumentation, so this was a valuable exercise. Reflecting on the meaning of the factors in the *Motorola* problem (see Figure 1), he noticed that certain factors were at odds, and went on to predict how courts would resolve this conflict:

> **Student** Actually, aren't the *Motorola* factors contradictory in saying that there were security measures [F6] but yet the secrets were disclosed [F10] and the information was known to competitors [F20] ?
>
> **Tutor** Do you think that ever happens, that you have fact situations with contradictions of that nature?
>
> **Student** Yeah. I guess, I would think that though when those two contradict, the secrets being disclosed [F10] would cancel out the ... they would make it irrelevant that there were security measures [F6], because the only point for security measures is to keep the information secret and once the information is not secret, what's the point of security measures?
>
> **Tutor** How would you test that?
>
> **Student** As far as, if that's true, you would do a search for [cases with] F6 and F10 or F20. ... according to my theory, all those cases should go for the defendant.

When the student executed the query, he found—much to his surprise—that not all retrieved cases were consistent with his prediction. By looking at the factors of one of the retrieved cases, he was able to refine the theory, as is described in (Aleven and Ashley, 1994). This example illustrates that students can do interesting things with factors by applying knowledge that is not represented in CATO. CATO could not have produced the same behavior. Nonetheless, it evoked the behavior: The case representation in terms of factors prompted the student to make a prediction. The rationale for his theory was based on the meaning of the factors, which is not represented in the CATO model. But CATO enabled him to test his theory against a particular database of cases; this would have have been difficult to achieve with a full-text retrieval system (Aleven and Ashley, 1994).

We hope that the examples provide a sense of what the background knowledge is and why it is important. At one level, it is difficult to see how an exercise like the current would make sense at all to students if they did not have at least a superficial understanding of what the factors mean and why they matter. For example, an argument by analogy is compelling only if the similarities between the analogs are relevant to the conclusion. Otherwise, the response will simply be: "So what if the plaintiffs in both cases had blue eyes? What does that have to do with anything?" In other words, to understand analogies in terms of factors, one must understand why the factors matter. The same can be said about other types of arguments, for example, distinguishing. This is why we were concerned that CATO does not represent the background knowledge about what the factors mean, and why we were pleased to find that students found it quite natural to analogize and distinguish cases in terms of factors and were able to use background knowledge about factors to make arguments that are not captured in the CATO model. In each protocol, we found between 4 and 10 instances of students' applying knowledge that is not represented in CATO. Although not all examples are as clear and striking as the ones presented here, they indicate that students are able to integrate their background knowledge within the CATO conceptual framework, at least when guided by a human tutor.

Discussion and Conclusions

In our research project, we explore the hypothesis that an ITS can teach skills of making arguments with cases, using a domain model that contains concepts of reasoning with cases, but relatively little of the background knowledge on which human reasoners draw. The background knowledge that we do represent is important, namely, the factors that influence the outcome of the legal cases. However, we do not represent knowledge related to what the factors mean or why they matter, at least not in a form with which the program can reason. Given that it is difficult to represent the background knowledge, it is worth investigating how far one can carry tutoring in the absence of much explicitly represented background knowledge.

CATO is similar to Belvedere, an instructional environment for teaching scientific argumentation (Cavalli-Sforza and Suthers, 1994). Both are examples of programs that make explicit a structure for argumentation and employ it to teach certain skills of analysis and argumentation. The pro-

grams can reason with the argumentation structure to varying degrees. CATO can evaluate the relevance of cases and make arguments; Belvedere can interpret argument structure to suggest where an argument should be extended. They do not, however, have very complete representations of the background knowledge upon which human practitioners draw in analyzing problems or making arguments. Nor can these programs interpret most free form arguments generated by humans. The question for them is, how successfully can they employ the argumentation structure to teach the analytical and argumentation skills despite the lack of background knowledge or understanding of students' arguments.

CATO employs cases in constructing arguments. Though also case-based, case-based teaching programs developed at Northwestern University (Kass, et al., 1993; Edelson, 1992), do not purport to teach analytical and argumentation skills. They do not employ a structure of argumentation. Instead, their aim is to teach a set of cases, presenting them in a context where they are most likely to be of interest to the student. But like CATO, they represent fairly little of the background information humans employ in interpreting the cases.

We conducted a formative evaluation study to investigate whether a human tutor can employ the CATO model and tools to teach argumentation skills. The tutor guided students in analyzing legal problems and developing an argument outline. Students used CATO's vocabulary of factors to represent the strengths and weaknesses present in a problem, used the CATO query language to retrieve cases that share interesting factors with the problem, and used CATO's tools for comparing cases in terms of factors to interpret what the retrieved cases mean and whether they are relevant. The human tutor exerted a fair amount of control, but his role was constrained to getting the students to use the CATO tools and helping them to interpret what they had found with the tools. We made some background knowledge available through preparatory instruction in classroom format, through textual representations of the cases, and via the human tutor.

We found that the CATO instruction was as effective as more traditional classroom instruction. Arguments generated by CATO were scored higher than students' arguments, suggesting that the model can be used to teach even more effectively. Examples drawn from the protocols of CATO instruction indicate that, equipped with some background knowledge, students applied the model flexibly, as a useful conceptual framework, and were able to integrate the background knowledge into their arguments. Admittedly, we cannot rule out the possibility that it was mostly the human tutor's skills and background knowledge that enabled students to do so. However, as the examples illustrate, students often applied background knowledge on their own initiative.

The challenge for the future is to extend CATO so that students can use it without needing the continuous guidance of a human tutor. CATO will have a process model and will use this to guide students in constructing arguments. It will generate (written) arguments and explanations of the relevance of cases (Ashley and Aleven, 1994), so that students can compare their own solutions against those based on the CATO model. Also, CATO will use small, carefully selected collections of cases to illustrate argumentation issues and provide focused exercises (Ashley and Aleven, 1992).

With these extensions in place, CATO will provide some of the control that the human tutor provided, for example by suggesting things to do next. However, it will not be able to summarize students' arguments or relate them to the CATO model, as the human tutor did. Instead, it will present the CATO solutions for students to compare against their own. Also, CATO will not be able to demonstrate the use of the background knowledge or detect situations where students use it unproductively. Future evaluation studies will bear out whether such a system will help students learn to use the CATO conceptual framework effectively and integrate their background knowledge to make good arguments.

References

Aleven, V. and Ashley, K. D. (1994). An Instructional Environment for Practicing Argumentation Skills. In *AAAI-94: Proceedings of the Twelfth National Conference on Artificial Intelligence,* 485–492. Menlo Park, CA: AAAI Press.

Aleven, V. and Ashley, K. D. (1993). What Law Students Need to Know to WIN. In *Proceedings of the Fourth International Conference on Artificial Intelligence and Law,* 152–161. New York: ACM.

Anderson, J. R.; Boyle, C. F.; Corbett, A. T.; and Lewis, M. L. (1990). Cognitive Modeling and Intelligent Tutoring. *Artificial Intelligence,* 42: 7–49.

Ashley, K. D. (1990). *Modeling Legal Argument: Reasoning with Cases and Hypotheticals.* Cambridge, MA: MIT Press.

Ashley, K. D. (1991). Reasoning with Cases and Hypotheticals in HYPO. *International Journal of Man-Machine Studies,* 34, no. 6: 753–796.

Ashley, K. D., and Aleven, V. (1994). A Logical Representation for Relevance Criteria. In *Topics in Case-Based Reasoning; Selected Papers from the First European Workshop, EWCBR-93,* edited by S. Wess, K. D. Althoff, and M. M. Richter, 338–352. Lecture Notes in Artificial Intelligence, 837. Berlin: Springer-Verlag.

Ashley, K. D. and Aleven, V. (1992). Generating Dialectical Examples Automatically. In *AAAI-92: Proceedings of the Tenth National Conference on Artificial Intelligence,* 654–660. Menlo Park, CA: AAAI Press.

Cavalli-Sforza, V. and Suthers, D. D. (1994). Belvedere: An Environment for Practicing Scientific Argumentation. In *Working Notes, AAAI Workshop on Computational Dialectics,* edited by R. P. Loui and T. Gordon, 39–44. Sponsored by AAAI, 1994.

Clancey, W. J. (1983). The epistemology of a rule-based expert system: a framework for explanation. *Artificial Intelligence,* 20, no. 3: 215–251.

Edelson, D. C. (1992). When Should A Cheetah Remind You of a Bat? Reminding in Case-Based Teaching. In *AAAI-92: Proceedings of the Tenth National Conference on Artificial Intelligence,* 667–672. Menlo Park, CA: AAAI Press.

Kass, A.; Burke, R.: Blevis, E.; and Williamson, M. (1993). Constructing Learning Environments for Complex Social Skills. *The Journal of the Learning Sciences,* 3, no 4: 387–427.

Rissland, E. L. (1990). Artificial Intelligence and Law: Stepping Stones to a Model of Legal Reasoning. *Yale Law Journal* 99: 1957–1982.

Shute, V. J., and Regian, J. W. (1993). Principles for Evaluating Intelligent Tutoring Systems. *Journal of Artificial Intelligence and Education,* 4, no. 2/3, 245–271.

Wenger, E. (1987). *Artificial Intelligence and Tutoring Systems: Computational and Cognitive Approaches to the Communication of Knowledge.* Los Altos, CA: Morgan Kaufmann.