An Instructional Environment for Practicing Argumentation Skills

Vincent Aleven and Kevin D. Ashley
Intelligent Systems Program,
Learning Research and Development Center, and
School of Law
University of Pittsburgh
Pittsburgh, PA 15260
aleven+@pitt.edu, ashley+@pitt.edu

Abstract

CATO is an instructional environment for practicing basic skills of legal research: to use cases in arguments about a problem situation and to test a theory about a legal domain. Using the CATO tools, law students analyze a legal problem, frame queries of CATO's database of legal cases, and judge how relevant the retrieved cases are to their developing argument or theory. CATO aids learning by making explicit an abstract model of the process of argument. It allows students to focus on the high-level argumentation issues, by assisting the student in various ways. By providing an abstract representation of the text of cases, it helps students to reason about the texts and helps guide their critical analysis of the texts. CATO makes available opportunities for practice that are hard to set up with traditional instructional methods.

CATO differs from other instructional environments in the following respects: Few instructional environments focus on argumentation skills. Although there are other instructional environments in which students work with an abstract representation of the task domain, abstracting from text is unusual. CATO demonstrates a contribution that case-based reasoning techniques can make to instructional environments.

1. Introduction

Books, movies, and computers all have the ability to create worlds in which things are a little bit different (or even vastly different) from the way they are in the real world. However, only computers create interactive worlds, worlds in which the audience actively participates and influences the turn of events. In a typical computer game, for instance, the audience faces the task of flying a spacecraft to a far away part of the Universe in order to put down a rebellion (no mean task). Players may become astonishingly skilled at such games, especially if this is regarded as "cool" by one's peers (the competition).

The potential for education has not been lost upon AI researchers. They have created environments in which students practice skills that (unlike the skill of flying a spacecraft) can be applied to solve problems in the real world. While the emphasis is on cognitive skills rather than motor skills (as in the computer game mentioned above) and graphic simulations are not always the norm, these environments have in common with the computer games that they pose a task and provide tools to solve the task, and that the students work in a world that may be slightly simpler than the real world.

These environments are usually modules of larger instructional systems, "intelligent tutoring systems" (ITSs). Such systems take an active role in communicating the domain knowledge to the student, adapting to the individual student's needs as much as possible [Wenger, 1987]. However, even an environment module standing alone can have considerable pedagogical advantages.

First, an environment module, perhaps more so than any other part of the tutoring system, can communicate to the student a conceptualization of the domain. If this conceptualization is sound, it may greatly help the student in constructing a mental model of the domain [Miller, 1988]. "[T]he interface language can rival the generation of active communication steps in pedagogical importance." [Wenger, 1987, p. 317] In particular, the environment module may make explicit properties of the domain that were previously hidden or implicit. [Burton, 1988]. For example, the graphical display of the Geometry tutor explicitly shows that geometry proofs are not linear, as most textbooks suggest, but are tree-structured and can be developed either by going forward from the premises or backward from the goal [Anderson, *et al.*, 1985]. An environment module may also present an abstract view of the task, in order to focus the student on what is important [Burton, 1988]. For example, TASK lets students work with a highly abstract simulation of a fault-diagnosis task [Rouse and Hunt, 1984].

Second, an environment module may assist a student by taking over certain parts of the problem-solving task. This enables the students to concentrate on the high-level structure of the task, without being overwhelmed by details. [Burton, 1988] In Algebraland [Brown, 1985], for example, the computer does the calculations involved in algebraic operations, so that students can concentrate on their applicability conditions.

Finally, an environment module may create opportunities for practice that would not otherwise be available. For instance, Sherlock enables students to practice trouble-shooting skills by simulating a complex electronic test station that is simply not available for educational use [Lajoie and Lesgold, 1992].

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

Text	Factors
In 1980, a restaurant owner named Mason developed a combination of Jack Daniel's whiskey, Triple Sec, sweet and sour mix, and 7-Up to ease a sore throat. He promoted the drink, dubbed "Lynchburg Lemonade" for his restaurant, "Tony Mason's, Huntsville," served it in Mason jars and sold t-shirts. Mason told the recipe only to his bartenders and instructed them not to reveal the recipe to others. The drink was only mixed out of customer's view. Despite its extreme popularity (the drink comprised about one third of the sales of alcoholic drinks), no other establishment had duplicated the drink, but experts claimed it could easily be duplicated.	F6 Security-Measures (p) F15 Unique-Product (p) F16 Info-Reverse-Engineerable (d)
In 1982, Randle, a sales representative of the distillery, visited Mason's restaurant and drank Lynchburg Lemonade. Mason disclosed part of the recipe to Randle in exchange, Mason claimed, for a promise that Mason and his band would be used in a sales promotion. Randle recalled having been under the impression that Mason's recipe was a "secret formula."	F1 Disclosure-In-Negotiations (d) F21 Knew-Info-Confidential (p)
Randle informed his superior of the recipe and the drink's popularity. A year later, the Distillery began using the recipe to promote the drink in a national sales campaign. Mason did not participate in the promotion or receive other compensation.	

Legend
(p): factor favors plaintiff
(d): factor favors defendant

Figure 1: Description of the facts of the *Mason* problem; Applicable factors that the students identified.

However, to our knowledge few systems have demonstrated these advantages in the context of teaching argumentation skills. We are developing CATO, an ITS that teaches law students basic skills of legal research: To use cases to make arguments about a problem situation and to test legal theories against cases. So far, we have implemented an environment module, based on case-based reasoning techniques. It lets students represent a problem situation, retrieve past cases that are relevant to the analysis, and judge the cases' relevance to the developing argument or to the theory being tested. To assist students with text interpretation, CATO provides (pre-stored) abstract representations of the texts of the legal cases.

In this paper, we show examples of students using CATO, guided by a human tutor, to illustrate that CATO achieves the advantages of environment modules mentioned above. We discuss CATO's strengths and limitations and compare it to other instructional environments.

2. CATO as Environment for Teaching Legal Research and Argumentation

The CATO environment provides students with tools to do legal research, tools that differ from those that are normally used. We explain why after we describe the tools in the context of (actual) examples of students using CATO.

Our goal is to teach first-year law students basic skills of making arguments with cases and using cases to test theories. The legal domain has a long-standing practice of analyzing problems by comparing and contrasting them to relevant past cases. Past cases can be cited as primary authorities in arguments. Therefore, students have to learn various argument strategies involving cases. Also, they have to learn to use full-text retrieval systems such as Lexis and Westlaw to find cases that support their arguments in automated case law databases.

Likewise, formulating and testing legal rules is an important activity that students have to learn. "[O]ut of the matching of a number of related cases it is your [the student's] job to formulate a rule that covers them all in harmony, if that can be done, and to test your formulation against possible variants on the facts. Finally, to test it, if there is time, against what writers on the subject have to say, and against other cases." [Llewellyn, 1930, p. 52]

Working with CATO, the student's task is to analyze a problem (a set of facts that led to a legal dispute) and then to outline arguments on behalf of the plaintiff and defendant that each should win the legal dispute. The argument outline should list the most relevant cases that each party can cite. The cases will be selected from the CATO database, which currently contains 45 legal cases. CATO provides tools to retrieve cases and to present information about the cases that helps students judge their relevance. So far, we have run the program with 9 first-year law students.

In trade secrets cases, the plaintiff and defendant are often corporate competitors. The plaintiff complains that the defendant has gained access to information plaintiff deemed confidential (often, technical knowledge developed at considerable expense) and has used that information to gain an unfair competitive advantage, for instance, by developing and marketing a product to compete with plaintiff's. To win the case, the plaintiff must show that it actively tried to protect its secret and that the defendant acquired the information by improper means or in breach of a confidential relationship.

In our first example, two first-year law students (guided by a human tutor) use CATO to analyze a problem situation based on *Mason v. Jack Daniel Distillery*. Figure 1 shows a summary of the facts of *Mason*. After reading this description, the students identify facts that tend to strengthen each party's position in the dispute. This requires some knowledge of trade secrets law and of the world of corporate competition, where disputes about trade secrets arise. CATO provides a vocabulary to state the result of this analysis, namely, a set of 21 *factors* for trade secrets law. Factors are abstractions of facts that tend to strengthen or weaken a party's position on a legal claim. The factors that the students found were present in the *Mason* problem are shown in Figure 1. It should be mentioned that there is no single right set of factors for any case or problem; there is room for interpretation.

The next step is to develop arguments for the plaintiff and defendant in *Mason*. Both parties need to convince the court that the favorable factors present in the problem outweigh the unfavorable ones. An important argument strategy, taught from day one in law school, is to draw an analogy to a past case that presented a similar fact pattern as the problem situation and had the outcome that is desired in the current problem. (We call this strategy: Citing a representative example.) Therefore, the students need to find cases in the CATO database that they can cite as representative examples.

Like the problem situation, the cases in CATO's database are represented as lists of factors. They have been analyzed by the system developers, to identify the factors that apply. CATO offers a query language that allows one to retrieve cases with any boolean combination of factors.

The queries that the students tried to find cases with similar sets of factors as the *Mason* problem are shown in Figure 2. (The queries were typed by a human operator, since we did not want students to have to learn the query language.) With their first query, the students find out that there are no cases in the CATO database that have all factors that are present in the *Mason* problem. However, even cases that share some, but not all, factors can be useful to cite in an argument. The students therefore relax the query constraints to ask for cases that have one or more factors in common with the problem. Unfortunately, this second query returns too many cases. It would be too much work to inspect them all. With the third query, the students focus on what they perceive are the most important factors in the problem and finally find the right level of specificity: The query returns a manageable number of cases. (It should be mentioned, however, that this last query does not cover all ways in which a case may be relevant. A more systematic approach is needed, and indeed the students were more systematic later on.)

To decide which of the retrieved cases (if any) to cite in an argument, the students need to compare them to the problem situation. CATO provides various tools that present useful information, but leaves the final judgement as to which cases are worth citing up to the students. CATO has a tool for listing the applicable factors of a retrieved case. It can also display a comparison of the factors of a retrieved case and the problem. In Figure 2, for example, CATO shows a comparison of the *Boeing* case and the *Mason* problem. CATO marks the factors that the two cases have in common and also marks the distinctions (those unshared factors that push toward an opposite result in the two cases). In general, the more shared factors a case has, and the fewer distinctions, the better the argument citing it as a representative example. Finally, for any case in its database CATO can display a "squib," a summary of the case, which includes a brief description of the facts (like the description shown in Figure 1).

After reviewing the comparison of the *Boeing* and *Mason* factors and reading the *Boeing* squib, the students decide that *Boeing* is not ideal. They go on to check the other retrieved cases and find better cases, but this is not shown in Figure 2.

```
1> (list-cases f6 f15 f21 f1 f16)

List all cases
  with factors
    F6 Security-Measures (p)
    F15 Unique-Product (p)
    F21 Knew-Info-Confidential (p)
    F1 Disclosure-In-Negotiations (d)
    F16 Info-Reverse-Engineerable (d).

None.
```

```
2> (list-cases (:or f6 f15 f21 f1 f16))

List all cases
  with one or more of factors
    F6 Security-Measures (p)
    F15 Unique-Product (p)
    F21 Knew-Info-Confidential (p)
    F1 Disclosure-In-Negotiations (d)
    F16 Info-Reverse-Engineerable (d).

22 cases won by plaintiff, 10 cases won by defendant.
```

```
3> (list-cases f1 f6)

List all cases
  with factors
    F1 Disclosure-In-Negotiations (d)
    F6 Security-Measures (p).

Cases won by plaintiff:
Boeing (p)
Bryce (p)
Digital Development (p)

3 cases won by plaintiff.
```

```
4> Comparison of factors

Mason
  - F1 Disclosure-In-Negotiations (d)
  - F6 Security-Measures (p)
    F15 Unique-Product (p)
  * F16 Info-Reverse-Engineerable (d)
  - F21 Knew-Info-Confidential (p)

Boeing (p)
  - F1 Disclosure-In-Negotiations (d)
  * F4 Agreed-Not-To-Disclose (p)
  - F6 Security-Measures (p)
    F10 Secrets-Disclosed-Outsiders (d)
  * F12 Outsider-Disclosures-Restricted (p)
  * F14 Restricted-Materials-Used (p)
  - F21 Knew-Info-Confidential (p)

- shared factor
* distinction
```

Figure 2: Transcript of students using the CATO tools to find relevant cases

In this example, students, using the CATO tools, were able to find useful cases with only a moderate amount of trial and error. We should mention that there is no single

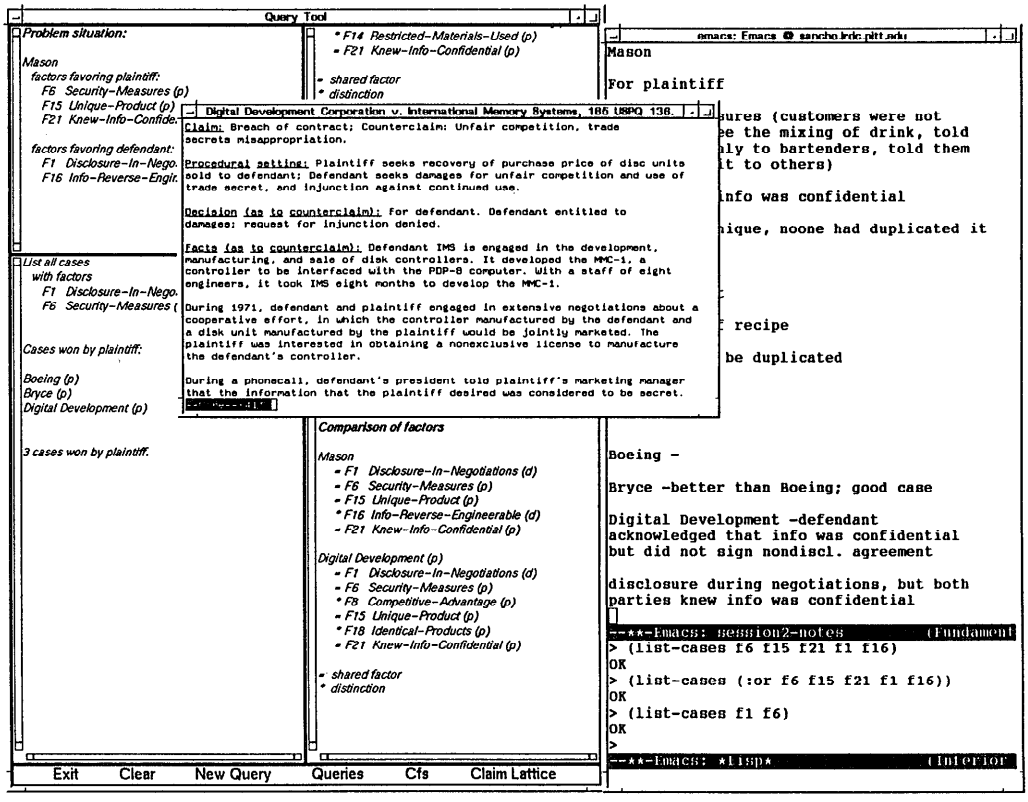

Figure 3: The CATO interface. The factors that the student identified for the problem situation *(Mason)* are displayed in the top left window. There are windows for entering queries (bottom right), displaying the result of the last query (bottom left), and displaying a case's factors or a comparison with the factors of the problem (middle window). Queries are expressed in a Lisp-like syntax. CATO displays squibs (short summaries of cases) in separate pop-up windows. Students can also request to see a history of their queries (middle window) and can make notes about the cases they have seen so far in the top right window.

right solution to this task, as is very characteristic of the legal domain. It is important, however, that students consider all pertinent information in evaluating the relevance of cases and do not miss important lines of cases.

The CATO interface is shown in Figure 3. CATO has a knowledge base which contains explicit definitions for many of the concepts in its model of argument [Ashley and Aleven, 1992, 1993]. This knowledge base is implemented in the knowledge representation system Loom [MacGregor, 1991]. CATO queries are executed by translating them into Loom's query language. The CATO interface is implemented in CLIM and currently also uses an Emacs window.

3. Examples of Instructional Opportunities Students Achieved with CATO

In this section, we show two more actual examples of law students using CATO, to illustrate that CATO enables the students to formulate, implement and test hypotheses and, in this way, obtain practice in legal argumentation and research that would be difficult to achieve otherwise.

In the example of Figure 4, a law student uses CATO to test a hypothesis that two weaknesses in his side's position are not fatal, and to formulate a perceptive argument to establish that assertion. The student had analyzed the *Mason* problem situation and identified two factors that weakened the plaintiff: the plaintiff had disclosed its alleged confidential product information (the beverage recipe) in negotiations with the defendant (factor F1, Disclosure-In-Negotiations) and the plaintiff's recipe could be readily reverse engineered (factor F16, Info-Reverse-Engineerable), both of which factors helped the defendant.

The student, arguing for the plaintiff, needed to counteract the weaknesses represented by factors F16 and F1, in other words, to "cover the defendant's bases". Figure 4 shows three successive queries that the student actually submitted to CATO and CATO's responses. The first two queries show that the student had formulated and was testing the hypothesis that there can be a trade secret misappropriation despite the fact that plaintiff disclosed the information to the defendant in negotiations or that experts could discover the recipe by analysis. The student

> (list-cases f16)

List all cases with factor
F16 Info-Reverse-Engineerable (d).

Case won by plaintiff: *
American Precision (p)

> (list-cases f1)

List all cases with factor
F1 Disclosure-In-Negotiations (d).

Cases won by plaintiff: *
Boeing (p)
Bryce (p)
Digital Development (p)
Space Aero (p)

> (list-cases f1 :not-any f4)

List all cases with factor
F1 Disclosure-In-Negotiations (d)
but without factor
F4 Agreed-Not-To-Disclose (p).

Cases won by plaintiff: *
Digital Development (p)
Space Aero (p)

* Cases won by defendant omitted.

Figure 4: Student's queries for cases that counteract unfavorable factors ("cover the opponent's bases") and avoid a possible distinction.

sought and found pro-plaintiff cases to counteract factors F16 and F1. In these cases, plaintiffs won despite the presence of one or other of those pro-defendant factors.

Then the student did something very interesting in the third query. He screened the cases returned by the second query to select the most convincing evidence for his hypothesis. He sought all the cases with factor F1 Disclosure-In-Negotiations, but without factor F4 Agreed-Not-To-Disclose. That is, he sought cases where the plaintiff disclosed the secrets in negotiations, but where, as in the *Mason* problem, the plaintiff had failed to take the precaution of securing the defendant's agreement not to disclose the secrets. He found two cases, *Digital Development* and *Space Aero* where the plaintiff won despite this lack of care. These are especially good cases to cite because they show that mere knowledge of the confidentiality is enough to bind the defendant even when the defendant did not promise not to use or disclose the information. In other words, with his third query the student filtered out cases that could be distinguished in a way that would be damaging to his position since the distinction would call attention to a troublesome lack of caution on his client's part. The student's queries are good evidence to a teacher that the student not only understands how to make an argument but has a deeper understanding of the problem and the legal domain.

In the example of Figure 5, a first-year law student formulates a more general hypothesis about the legal domain, retrieves cases to test his theory, and revises the theory when it turns out that it is inconsistent with some of the cases that were retrieved.

At the top of Figure 5, having used CATO to make a number of arguments involving a variety of cases, and without prompting from the human tutor, the student announced what amounted to a theory about the domain: he predicted that where secrets had been disclosed to outsiders, plaintiffs would lose no matter what security measures they had taken because "once the information is not secret, what's the point of security measures?" From a pedagogical viewpoint, it is important to inculcate this kind of behavior in law students, that is, formulating a hypothesis about the importance of features in a problem, based upon a rationale concerning the point or purpose of a law. As in other fields, it is even more important that students learn how to test and revise such hypotheses [Lakatos, 1976; Rissland, 1983; Collins and Stevens, 1982]. Seizing the pedagogical opportunity, the tutor challenged the student to test the hypothesis. In response, the student formulated a query to retrieve all cases that were covered by his hypothesis, to see whether the outcome of these cases was, indeed, as he predicted.

To his evident surprise, the student found that some of the retrieved cases were inconsistent with his theory: plaintiffs who had taken security measures won despite disclosures to outsiders. This was puzzling. The student allowed that he would "read the cases and see what's going on". To that end, at the tutor's suggestion, the student inspected one of the retrieved cases and realized that his theory could be saved. Recognizing that his rule was overly general, he revised it accordingly: where the secrets had been disclosed to outsiders, and those disclosures were not limited by restrictions, then plaintiffs would lose despite having taken security measures. The student could have tested that query, as well. However, there was little time left, so the tutor decided to move on to other things.

This was a valuable lesson. In response to a student's spontaneous comment, using CATO's tools, the tutor could empower a student to work through, in a concrete way, a complex, abstract process of formulating, testing, and revising a legal hypothesis. Recognizing the need to engage in this process and learning how to do it are at the core of legal education and legal scholarship; much of what legal scholars do is to formulate, test and revise such theories in the light of existing precedents and anticipated hypothetical problems. But it is hard to identify and communicate such an abstract, complex process to students, especially where a reasoner needs to invent the terminology for formulating the theory, find, read and interpret the relevant cases and hypotheticals, or revise the theory in light of the cases and the purposes of the law.

This is where CATO helps; CATO suppresses some of the distracting complexity so that students can identify and complete a few cycles of the process, but CATO does not simplify the process so much as to make the process seem uninteresting and pointless to the student. Firstly, CATO provided a convenient language in which the student could express his theory and his query to find cases to test the theory.

Secondly, CATO presented just the right information about the retrieved cases to make clear their import to the theory being tested. Retrieved cases are relevant to testing the hypothesis because they are guaranteed to have the specified combination of factors. Since CATO shows the retrieved cases' outcomes, it was readily apparent

Student ... when those two contradict, the secrets being disclosed would cancel out the ... they would make it irrelevant that there were security measures, 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 f6 and f10 or f20. .... According to my theory, all those cases should go for the defendant.

```
> (list-cases f6 (:or f10 f20))

List all cases
  with factor
    F6 Security-Measures (p)
  and with one or more of factors
    F10 Secrets-Disclosed-Outsiders (d)
    F20 Info-Known-To-Competitors (d).
```

Cases won by plaintiff:
Boeing (p)
Data General (p)
Drill Parts (p)
FMC (p)

Case won by defendant:
Eaton (d)

Tutor Now *Eaton* went for the defendant ... How about those others?

Student That boggles my mind!

Tutor How would you resolve this?

Student I would read the cases and see what's going on.

Tutor Let's just take a look at *Data General*, that's a good one.

```
Data General (p)
  F6 Security-Measures (p)
  F10 Secrets-Disclosed-Outsiders (d)
  F12 Outsider-Disclosures-Restricted (p)
  F14 Restricted-Materials-Used (p)
  F18 Identical-Products (p)
```

Student Outsider-Disclosures-Restricted ... Restricted-Materials-Used ...

Tutor Can you save your theory?

Student Well, actually, somewhat in that secrets disclosed to outsiders means that it might be a limited disclosure uh within your factor.

Tutor That is captured by the addition of another factor in this *Data General* situation: outsider disclosures are restricted and also that the defendant used these restricted materials.

Student So that would be more like ... limited disclosure would be distinguished from general disclosure, or, competitors knowing ...

Tutor So did Data General discredit your theory?

Student Not in a wide theoretical ... only on the basis of these factors but not in a theoretical sense

Figure 5: Using CATO to test a theory about the domain.

which of the cases were consistent and which were inconsistent with the theory.

Thirdly, CATO supports inspecting the cases to see if counterexamples really invalidate the theory or whether the theory can be revised. By displaying the factors of even one of the retrieved cases, CATO provided enough information for the student to modify the theory, a process that could have been repeated at least one more time. The student can also inspect the case squib to confirm whether the sets of factors assigned to the case in CATO's case base really are accurate or whether, on a closer reading, they might be reinterpreted.

It is hard to set up this kind of learning opportunity without CATO. Students frequently espouse theories in law school classes, but getting students to test and modify them is hard to arrange. The class typically reads only two or three cases, selected by the case book writer to illustrate the writer's theory, but not necessarily those invented by students. In class, it is hard and time consuming to get a consensus even about the facts of those cases all have read. With CATO one can test the hypothesis over a relatively large database of cases, point physically to an abstract summary of the case for purposes of discussion, modify queries to reflect revised theories, and cycle through the process repeatedly. Conceivably, with an overhead projecting screen, this could even be done in front of an entire class. Trying to conduct such a classroom exercise with full-text legal retrieval tools would founder with the complexity of interpreting what the retrieved cases say. Without reading a large amount of text, one cannot even determine who won a case retrieved from a full-text retrieval system.

4. Advantages and Limitations of the CATO Environment

CATO realizes various advantages of instructional environments. First, CATO aids learning by communicating a conceptualization of the domain, a view of the process of legal argumentation and research that is useful for a first-year law student to learn. (See [Ashley, 1991] and [Aleven and Ashley, 1993].) As illustrated by the examples of students using CATO, cases are used, in various argument moves, to justify assertions about a problem, or to test a proposed theory about a legal domain. The issues that a problem presents, or were resolved in a past case, are represented as a list of factors. Queries for relevant cases are also expressed in terms of factors.

CATO's model is not normally applied in legal research and is not currently taught in law schools. Although factors seem to be a fairly intuitive concept (e.g., some statutes list factors, one of our students commented: "I think factors came out because ... that was just inevitable."), it is not common in the legal field to explicitly represent, retrieve, and compare cases in terms of factors. Therefore, like the Geometry tutor, CATO makes explicit properties of its domain that are normally hidden or implicit.

CATO, like AlgebraLand [Brown, 1985], does some of the work for the students, so that they can focus on high-level aspects of the task domain. As the examples illustrate, the CATO tools keep track of rather a lot of useful information and present it on request. Also, the query language and tool for comparing cases make it relatively easy for students to find the cases in CATO's database that they need to support their argument or test their the-

ory. CATO also assists students by providing abstract representations of the text of the cases (*i.e.*, lists of factors). This saves them the work of reading the cases and developing an interpretation. Squibs also greatly reduce the amount of text that students need to read. As a result, students can concentrate on the high-level aspects of the process (comparing cases, expressing an argument need in CATO's query language, deciding which cases to cite in an argument) without being distracted by details. Also, the CATO tools make it easier for a tutor to guide students through a few cycles of the process of developing an argument or testing a theory; this would be difficult to realize with traditional instructional methods.

One may object that CATO leaves out a fundamentally important component of legal research: reading and interpreting the opinions of cases. We admit that CATO gets much leverage out of reducing the amount of text that needs to be read. However, we believe that exercise with CATO will actually help students learn to read cases. Having worked with CATO, they may read cases with a better understanding of what it is they are looking for. (Indeed, CATO queries are good models of the questions to have in mind when reading cases.) Also, during our CATO sessions, we pointed out to students that reading the cases is still important and we encouraged them to read the squibs of cases they deemed interesting based on the comparison of their factors. Our experience is that students did regularly read the squibs.

While exercises with CATO involve important legal research skills, they do not, of course, cover all of legal research. There are additional criteria for selecting cases to cite in arguments, for example, the date when a case was decided, the court's pedigree and jurisdiction, whether the case was overturned by subsequent decisions, and the case's procedural setting. Moreover, in addition to cases, legal arguments may cite statutory rules and may refer to public policy or legislative intent.

Another limitation is that CATO does not reason about, or represent, the reason why certain facts have theoretical legal significance and others do not. We recognize, it is very important not to discourage such reasoning on the part of students. We would like students to reason in terms of the theory of the legal domain. and to make predictions based on their theories. As the second and third examples illustrate, CATO provides support for formulating and testing theories.

Currently, CATO provides useful tools to engage law students in exercises in legal research, but the guidance of a human tutor is still needed. To develop an ITS that students can use without the intervention of a human tutor, CATO's pedagogical capabilities need to be extended.

While the examples illustrate that law students have used CATO in interesting exercises in legal research, they do not show that CATO improves students' learning. To test this, we are conducting a controlled experiment with 17 first-year law students to evaluate how effective practice with CATO (under the guidance of a human tutor) is compared to classroom instruction that teaches the same material but does not use CATO.

While the advantages that are described above have been demonstrated in other instructional environments, CATO differs from previous systems in various ways. Argumentation is an unusual domain for instructional environments, perhaps because it is so closely tied to text interpretation, which is notoriously hard to implement on a computer. In one project, students are provided with tools to represent (in a graphical notation) their arguments about well-known scientific problems [Cavalli-Sforza, *et al.*, 1993]. Also, EUCLID is an environment that offers (graphical and other) tools for constructing, comprehending, and assessing arguments but is not primarily intended for educational use [Smolensky, *et al.*, 1987].

One of the reasons that CATO is able to make some headway is that it presents an abstract representation of the texts of legal cases. While other instructional environments also suppress some of the complexity by providing an abstract view of the domain, for example, TASK [Rouse and Hunt, 1984], CATO is one of the few where students work with abstract representations of text. This is facilitated by the fact that we are dealing with a highly uniform set of texts.

Case-based teaching systems developed at Northwestern, like CATO, support indexing and retrieval of cases for pedagogical purposes [Schank, 1990; Ferguson, *et al.*, 1992]. These systems are based on the hypothesis that "good teaching is good story telling." The instructional goal is "to teach the cases." Cases are retrieved when students need or can benefit from the information that they contain. The instructional goal in CATO is not to teach the cases; it is to teach a process of reasoning with cases. Working with CATO, students justify assertions about a problem by comparing and contrasting it to past cases. This does not happen in the systems developed at Northwestern. In CATO, case retrieval is done mainly to give students practice in accurately expressing what kind of cases they need to make a good argument.

Techniques similar to those used in CATO have been used in the domain of software engineering, to support the reuse of software components. For example, Prieto-Díaz developed a system, comprising (among other things) an indexed library of computer programs and a query language, that software engineers can use to find a program that fits the constraints of the current problem closely enough so that it can be easily adapted [Prieto-Díaz, 1987]. While the case comparisons supported by Prieto-Díaz's system are certainly helpful to find useful past cases, they do not seem to be of much help in the ensuing step of adapting a past case (i.e., computer program) to fit the constraints of the current problem. This would tend to make practice with case retrieval and comparison of limited utility for those wanting to learn the skills involved in software reuse. By contrast, CATO supports drawing inferences about the problem by comparing it to past cases. Case comparisons are the building blocks of legal arguments [Ashley and Aleven, 1992]. It is for this reason that case comparison and retrieval is a very central activity in the CATO tutorial.

5. Conclusion

We have shown three examples that illustrate that law students have used the CATO environment to do useful and interesting exercises in legal research, under the guidance of a human tutor. Generally speaking, the students were positive about their experiences with CATO, as is evident from the following quotes: "CATO makes it simple" and "It kind of does what you do when you are researching a case."

CATO is based on the hypothesis that if one provides students with an environment where they have a meaningful task (retrieving cases to make arguments) and a retrieval language that reflects important aspects of the model of argument, that one can teach them a process of argumentation, including basic argumentation and research skills, such as formulating, testing, and revision of hypotheses and critical analysis of texts. This appears to be a new technique.

CATO realizes three advantages that have been demonstrated in previous instructional environments. CATO communicates a conceptualization of the domain, namely, an abstract view of a process of legal argument and research that is useful for first-year law students to learn. The process is based on an AI model of reasoning with cases and is not currently applied in legal education or practice. By providing abstract representations of the text of legal cases and by providing tools that allow easy access to relevant cases, CATO removes distracting complexity and allows the students to focus on the high-level issues involved in argumentation and theory-testing. For the same reasons, CATO enables opportunities for practice that would be difficult to arrange otherwise.

CATO is one of the very few instructional environments for argumentation skills. Few instructional environments present students with an abstracted representation of text, as a way of focusing them on what is important about the text and getting them to reason with texts. Finally, although other instructional environments are based on case-based reasoning techniques, none of these engages students in a process of retrieving, comparing, and contrasting cases.

References

Aleven, V., and K. D. Ashley, 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., C. F. Boyle, and G. Yost, 1985. The Geometry Tutor. In *IJCAI-85: Proceedings of the International Joint Conference on Artificial Intelligence*, 1–7. Los Altos, CA: Morgan Kaufmann.

Ashley, K. D., 1991. *Modeling Legal Argument: Reasoning with Cases and Hypotheticals.* Cambridge, MA: MIT Press.

Ashley, K. D., and V. Aleven, 1993. Using Logic to Reason with Cases. In *First European Workshop on Case-Based Reasoning (EWCBR)*, edited by M. M. Richter, S. Wess, K. D. Althoff, and F. Maurer, 373-378. SEKI Report SR-93-12 (SFB 314), Dept. of Computer Science, University of Kaiserslautern.

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

Brown, J. S., 1985. Idea Amplifiers: New Kinds of Electronic Learning Environments. *Educational Horizons*, 63: 108–112.

Burton, R. R., 1988. The Environment Module of Intelligent Tutoring Systems. In *Foundations of Intelligent Tutoring Systems*, edited by M. C. Polson and J. J. Richardson, 109–142. Hillsdale, NJ: Lawrence Erlbaum.

Cavalli-Sforza, V., J. D. Moore, and D. D. Suthers, 1993. Helping Students Articulate and Criticize Scientific Explanations, In *Proceedings of the World Conference on AI and Education (AIED-93)*, 113-120. Edinburgh, Scotland, August 23-27.

Collins, A., and A. L. Stevens, 1982. Goals and Strategies of Inquiry Teachers. In *Advances in Instructional Psychology*, edited by R. Glaser, 65–119. Hillsdale, NJ: Lawrence Erlbaum.

Ferguson, W., R. Bareiss, L. Birnbaum, and R. Osgood, 1992. ASK Systems: An Approach to the Realization of Story-Based Teachers. *The Journal of the Learning Sciences*, 2(1): 95–134.

Lajoie, S. P., and A. Lesgold, 1992. Apprenticeship Training in the Workplace: Computer-Coached Practice Environment as a New Form of Apprenticeship. In *Intelligent Instruction by Computer: Theory and Practice*, edited by M. J. Farr and J. Psotka, 15–36. Washington DC: Taylor and Francis.

Lakatos, I., 1976. *Proofs and Refutations.* London: Cambridge University Press.

Llewellyn, K. N., 1930. *The Bramble Bush: On our Law and Its Study.* Dobbs Ferry, New York: Oceana Publications, 1960 edition.

MacGregor, R., 1991 The Evolving Technology of Classification-Based Knowledge Representation Systems. In *Principles of Semantic Networks: Explorations in the Representation of Knowledge*, edited by J. F. Sowa, 385–400. San Mateo, CA: Morgan Kaufmann.

Miller, J. R., 1988. The Role of Human-Computer Interaction in Intelligent Tutoring Systems. In *Foundations of Intelligent Tutoring Systems*, edited by M. C. Polson and J. J. Richardson, 143–189. Hillsdale, NJ: Lawrence Erlbaum.

Prieto-Díaz, R., 1987. Classifying Software for Reusability. *IEEE Software*, 4 (1):6-16.

Rouse, W. B., and R. M. Hunt, 1984. Human Problem Solving in Fault Diagnosis Tasks. In *Advances in Man-Machine Systems Research, Vol. 1*, edited by W. B. Rouse. Greenwich, CT: JAI Press.

Schank, R. C., 1990. Case-Based Teaching: Four Experiences in Educational Software Design. *Interactive Learning Environments* 1(4): 231-253.

Rissland, E. L., 1983. The Ubiquitous Dialectic. Tech. Rep. 83-15. Dept. of Computer and Information Science, University of Massachusetts, Amherst, MA.

Smolensky, P., B. Fox, R. King, and C. Lewis, 1987. Computer-Aided Reasoned Discourse, or, How to Argue with a Computer. Tech. Rep. CU-CS-358-87, Depts. of Computer Science and Linguistics, and Institute of Cognitive Science, University of Colorado, Boulder, CO.

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