

Evaluating a Learning Environment for Case-Based Argumentation Skills

Vincent Alevan 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 intelligent learning environment designed to help beginning law students learn basic skills of making arguments with cases, through practice in theory-testing and argumentation tasks. CATO models ways in which experts compare and contrast cases, assess the significance of similarities and differences between cases in light of general domain knowledge, and use the same general knowledge to organize multi-case arguments by issues. CATO communicates its model to students by presenting dynamically-generated argumentation examples and reifying (i.e., making visible) argument structure. Also, the CATO Tools reduce some of the distracting complexity of the students' task.

We evaluated CATO in the context of a second-semester legal writing course taught at the University of Pittsburgh School of Law. We found that 7.5 hours of CATO instruction led to a statistically significant improvement in students' basic argumentation skills, comparable to that achieved, in the same amount of time, by an experienced legal writing instructor teaching small groups of students in a more traditional way. On a more advanced memo-writing assignment, meant to explore the "frontier" of the CATO instruction, students taught by the legal writing instructor did better, indicating that more is needed if CATO is to help students to improve their memo-writing skills.

1. Introduction

CATO is an intelligent learning environment, designed to help beginning law students learn basic skills of making arguments with cases. Students learn to use cases to support conclusions about how a problem should be decided, to select the cases that support the strongest possible argument, and to organize their arguments coherently, by the issues that the claim raises. Also, they learn to assess and explain the significance of similarities and differences between cases in light of more general domain knowledge. Since in the legal domain, arguers have a certain latitude in interpreting cases, students also learn to select the interpretations strategically, in a way that advances their argument. These skills of reasoning with cases are among the more important of the lawyering skills that American legal

education seeks to teach. But they are difficult to learn and time-consuming to teach. If computer-based instruction can help students get some extra practice in basic skills and free up instructor's time for more advanced topics, much would be gained.

CATO employs a computational model of case-based argumentation that addresses eight basic argument moves and more elaborate multi-case arguments. The model includes a "Factor Hierarchy," which represents more abstract, but still domain-specific, legal knowledge about the meaning of the factors used to represent cases. CATO uses the Factor Hierarchy for a number of purposes, among them to organize multi-case arguments by issues and to make arguments about the significance of distinctions. To generate the latter, CATO strategically selects alternative interpretations of cases to elaborate a "deeper" (or more abstract) contrast or parallel between cases.

CATO communicates its argumentation model to students through dynamically-generated examples and reification. Upon students' request, CATO presents examples of basic argument moves and issue-based arguments, using cases selected by students. CATO also uses the model to reify (i.e., make visible on the computer screen) argument structure that is not visible in more traditional instruction. Finally, the CATO environment provides a case database and tools for analyzing, retrieving and comparing cases, which make students' tasks more manageable than if a full-text retrieval system were used. This results in an increased focus on the skills targeted in the instruction, making arguments with cases.

We evaluated CATO in the context of a second-semester legal writing course at the University of Pittsburgh School of Law, in a study involving 30 first-year law students. The purpose of the experiment was to find out how the CATO instruction compares with the best current way of teaching the same skills.

We would like to thank Kevin Deasy of the University of Pittsburgh School of Law for giving us the opportunity to evaluate CATO in the context of his legal writing course. He and Steffi Brünghaus of the Intelligent Systems Program provided very valuable assistance with the evaluation study. This research has been supported by the National Science Foundation, West Publishing Company, Digital Equipment Corporation, Tektronix, the National Center for Automated Information Research, and the University of Pittsburgh ECAC Advanced Instructional Technology Program.

Permission to make digital/hard copy of all or part of this work for personal or classroom use is granted without fee provided that the copies are not made or distributed for profit or commercial advantage, the copyright notice, the title of the publication and its date appear, and notice is given that copying is by permission of ACM, Inc. To copy otherwise, to republish, to post on servers or to redistribute to lists, requires prior specific permission and/or fee.

ICAIL-97, Melbourne, Australia © 1997 ACM 0-89791-924-6/97/06.\$3.50

CATO Database - Contains textual summaries (squibs) and factor sets for 147 trade secrets cases. The query language lets students retrieve cases with any boolean combination of factors.

Factor Browser - Provides information about CATO's set of 26 factors for trade secrets law.

Case Analyzer - Lets students compile a list of applicable factors for a case and generates feedback, comparing student's list of factors to that stored in the CATO Database for same case.

Argument Maker - Presents argumentation examples in the context of the student's on-going work. Also, conducts mini-dialog to give students practice in identifying distinctions.

Issue-Based Argument Window - Presents examples of arguments, organized by issues, with multiple cases selected by students.

Squib Reader - Displays squibs of retrieved cases.

Figure 1: The CATO Tools

CATO uses abstract knowledge in a novel way to assess and explain similarities and differences between cases. It has been recognized that models of case-based legal argument must incorporate abstract knowledge related to the features used to represent and compare cases. But other attempts have not focused on reasoning symbolically about the significance of similarities and differences. In CABARET, factors are linked to statutory predicates, but unlike CATO, CABARET does not use this knowledge strategically to elaborate a parallel or contrast between cases [Rissland and Skalak, 1991]. In GREBE, links from criterial case facts to statutory predicates are used for structure-mapping but not explaining the meaning of similarities and differences [Branting, 1991]. In BANKXXX, factors are linked to the legal theories to which they are relevant, yet these links are used mostly to guide search for cases and theories relevant to a problem [Rissland, et al., 1996]. CATO's use of the Factor Hierarchy is a first step toward incorporating teleological elements into case-based reasoning systems [Berman and Hafner, 1993]. Although CATO's Intermediate Legal Concerns and Legal Issues do not directly represent the legal purposes underlying the domain, they are closely related. They appear to be a suitable point to which an explicit representation of the purposes could be linked.

CATO seems to be unique among intelligent instructional systems for the legal domain in its focus on teaching case-based argumentation. In most legal tutoring systems (as in CATO), students practice legal analysis tasks, usually to infer the legal consequences of a given fact situation. But in almost all of these systems, the goal is to teach substantive law rather than argumentation skills. Unlike CATO, most systems rely on a rule-based representation of the legal knowledge [Sherman, 1989; Routen, 1992; Span, 1993] and do not ask students to make arguments on both sides of an issue. One system teaches a domain-independent process of solving legal cases, but the reasoning is primarily rule-based, not case-based [Muntjewerff, 1994]. Teaching methods in these systems differ, too, many of them relying on student modeling.

Session 1. Introduction to CATO and factors to represent cases.

Sessions 2-3. Introduction to WordPerfect, introductory practice with CATO tools, and theory-testing exercises.

Sessions 4-9. Argumentation problem based on *Mason v. Jack Daniel Distillery*.

1. Analyze *Mason* facts, identifying factual strengths and weaknesses related to trade secrets misappropriation claim. Map to applicable factors using the Case Analyzer tool.
2. Practice basic argument moves, comparing and contrasting *Mason* to cases in CATO Database. Compare arguments to CATO's arguments with the same cases, generated by Argument Maker. Study CATO's "recipes" illustrated in argument examples.
3. Search CATO Database for relevant cases that plaintiff can cite in his argument. Evaluate relevance of retrieved cases using comparison of factors, demonstrations of basic argument moves generated by CATO (Argument Maker), and the textual summary (Squib Reader).
4. Identify issues in *Mason*, using issue-based argument generated by CATO as guide (Issue-Based Argument Window). Select the most relevant cases, organize plaintiff's argument by issues, write a short argument, and compare with CATO's issue-based argument.
5. Outline defendant's response and compare to CATO's argument.

Figure 2: Overview of the CATO instruction for the *Mason* argumentation problem

2. Overview of the CATO Model and Instructional Environment

CATO's model of case-based legal argument addresses arguments in which a plaintiff and a defendant use cases to justify conclusions about how a problem should be decided. The model addresses eight basic argument moves, some of which were first used in HYPO [Ashley, 1990]:

1. Analogizing a problem to a past case with favorable outcome,
2. Distinguishing a case with unfavorable outcome,
3. Downplaying the significance of a distinction,
4. Emphasizing the significance of a distinction,
5. Citing a favorable case to emphasize strengths,
6. Citing a favorable case to argue that weaknesses are not fatal,
7. Citing a more on point counterexample to a case cited by an opponent,
8. Citing an as on point counterexample.

These argument moves are building blocks of more elaborate, multi-case arguments organized by issues. The factual strengths and weaknesses of cases are represented using factors, abstractions of facts that tend to make a case stronger for a side [Ashley, 1990]. As mentioned, the Factor Hierarchy represents normative knowledge about the meaning of factors and is used to organize arguments by issues and to select interpretations of distinctions to emphasize or downplay [Aleven and Ashley, 1996].

Case facts

Mason v. Jack Daniel Distillery

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.

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."

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.

Factors

F6 Security-Measures (p)

F15 Unique-Product (p)
F16 Info-Reverse-Engineerable (d)

F1 Disclosure-In-Negotiations (d)

F21 Knew-Info-Confidential (p)

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

Issues

Is plaintiff's information a trade secret?
F101 Info-Trade-Secret

Did defendant acquire plaintiff's information through improper means?
F110 Improper-Means-Conclusion

Did a confidential relationship exist between plaintiff and defendant?
F114 Confidential-Relationship

Figure 3. Applicable factors for *Mason*. Issues that CATO identifies on the basis of the factors.

The CATO environment is designed to help students to learn the skills addressed by the model. Students practice two general kinds of tasks: Theory-testing and argumentation. In the former, students test a general theory about the domain against the cases in CATO's Database. In the latter, they produce a written argument outline, supported by cases selected from the Database. The environment comprises an integrated set of tools, shown in Figure 1, which students can use to find and evaluate cases to cite in arguments and to evaluate theories. The tools also present argumentation examples.

The CATO Database contains, for 147 trade secrets cases, a textual summary of the case opinion ("squib") and a set of applicable factors. Thus, students work in parallel with textual descriptions and factor representations of cases. CATO's query language lets students retrieve cases with any boolean combination of factors. In theory-testing tasks, students use CATO queries to test generalizations about the relative importance of factors. In argumentation tasks, students query the database to get practice in expressing constraints that cases must satisfy to be useful in an argument.

The Argument Maker and Issue-Based Argument Window present, upon students' request, dynamically-generated argumentation examples with cases from the CATO Database. The examples involve the students' current problem and cases that they have retrieved from the Database and consider citing in an argument. Students study the examples and compare their own arguments against CATO's. Also, the examples of basic argument moves help in evaluating whether and how to use retrieved cases in arguments.

During the evaluation study, reported in Section 6, students used CATO to work through argumentation problems in a traditional casebook chapter on trade secrets law. A set of 4 workbooks provided instructions, aimed at making sure that students successfully complete the tasks and study CATO's basic argument moves and issue-based arguments in the process. The instructions are at times rather detailed but also leave room for self-guided activity. Each student used CATO for 9 sessions of 50 minutes and performed the tasks outlined in Figure 2. During the first session, a human instructor (Kevin Ashley) introduced CATO and guided students in an introductory task to learn factors. From then on, students, collaborating in pairs, worked with CATO without the guidance of an instructor. In the second and third sessions, they did additional exercises to become familiar with the CATO tools and did two theory-testing exercises. During the last five or six sessions, students worked on an argumentation problem based on *Mason v. Jack Daniel Distillery*. We note that nothing in CATO is specific to *Mason*. Any interesting trade secrets case could have been used.

3. Use of the Factor Hierarchy to Generate Examples of Issue-Based Arguments

Mason, shown in Figure 3, is a dispute between a bar owner and the Jack Daniel Distillery. Mason complained that the Distillery had misappropriated his secret recipe for a mixed drink, reneging on a promise that Mason's band would be used in the Distillery's national sales promotion. Trade secrets law protects owners of commercially valuable information against competition by parties

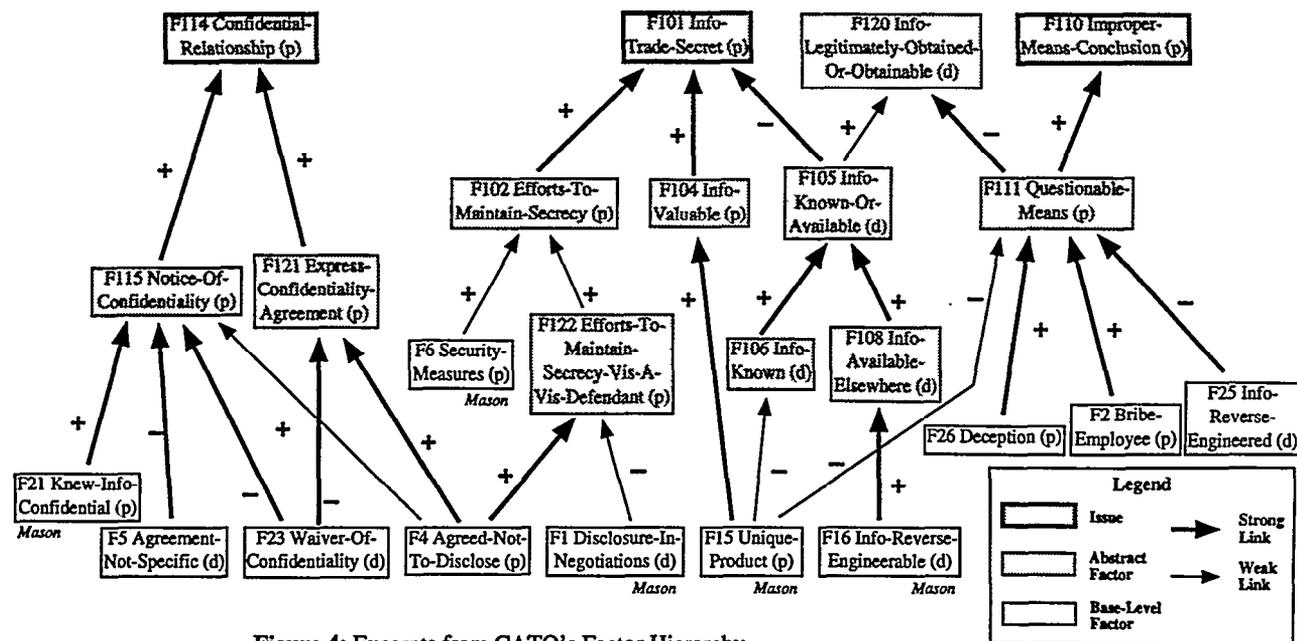


Figure 4: Excerpts from CATO's Factor Hierarchy

who obtained the information through improper means or a breach of faith. Students were asked to evaluate the merits of Mason's claim against the Distillery and outline an argument on behalf of plaintiff, supported by cases selected from CATO's Database.

Students start out the analysis of *Mason* by identifying factual strengths and weaknesses for each side and mapping these to applicable factors (Figure 2, step 1), using CATO's Case Analyzer. This tool lets students browse CATO's set of 26 factors (which initially they do not know) through hierarchical menus and select the factors that they think apply in the given problem. The Case Analyzer provides feedback stating how the list of factors compiled by the student compares to the list for that case stored in CATO's Database. Students are free to follow or ignore the advice, since there is no single correct set of factors for any given case. A student could identify as applicable in *Mason* the five factors shown in the middle of Figure 3. In generating argumentation examples, CATO adopts the students' analysis of the problem.

Mason presents factors that favor plaintiff and factors that favor defendant, but there is no authoritative weighting scheme to resolve the conflict. Instead, lawyers make arguments by analogy to past cases that present a similar combination of factors, arguing that the same result should apply in the current problem. In the CATO instruction, students are taught, by example, to organize their arguments by issues. Using its Factor Hierarchy, CATO identifies issues in a problem and organizes arguments by issues, using cases selected from the CATO Database.

CATO's Factor Hierarchy, excerpted in Figure 4, represents knowledge about what the factors mean. It links the "base-level factors", shown at the bottom, to an intermediate layer of legal concerns and, at the top of the Hierarchy, to legal issues, indicated by thick boxes. The higher-level factors (or "abstract factors") represent nor-

mative knowledge of the domain. Each abstract factor stands for two opposing conclusions, one favoring plaintiff, one favoring defendant. For example, abstract factor F102 represents a conclusion that "Plaintiff took efforts to maintain the secrecy of its information" but also a conclusion that "Plaintiff showed a lack of interest in maintaining the secrecy of its information." The links represent a generic support relation. They are labeled with a "+" or "-" indicating whether they link conclusions that favor the same or opposing sides. The links can be weak (thin arrow) or strong (thick arrow), indicating the level of support that they provide. CATO uses the link strength to evaluate whether a conclusion associated with an abstract factor is "supported" (i.e., worth arguing for) in a given case or whether potential support is "blocked" by opposing factors [Alevan and Ashley, 1996]. This enables it to identify conflicts whose resolution is not controversial and conflicts that should be played out in case-based arguments. The Factor Hierarchy contains 50 links, 26 base-level factors, and 16 abstract factors, 5 of which are legal issues.

CATO identifies issues in a problem simply by following links in the Factor Hierarchy upward from the base-level factors of the problem to the issues at the top. Using this method, it identifies three issues in the *Mason* problem, shown on the right in Figure 3. For example, the issue of whether plaintiff's information is a trade secret, corresponding to abstract factor F101 Info-Trade-Secret, is related to four factors in *Mason*.

An argument generated by CATO on behalf of plaintiff in *Mason*, shown in Figure 5, illustrates its method for generating issue-based arguments. The argument about each issue follows a basic rhetorical format, explicated on the right: first to emphasize the strengths related to the issue and then to downplay any weaknesses. CATO emphasizes strengths (favorable factors) by pointing to reasons

Argument for Plaintiff in the *Mason v. Jack Daniel Distillery* problem

Plaintiff should win a claim of trade secrets misappropriation. Plaintiff's information is a trade secret [F101], a confidential relationship existed between plaintiff and defendant [F114], and defendant acquired plaintiff's information through improper means [F110].

Plaintiff's information is a trade secret [F101]

Plaintiff's information is a trade secret [F101]. Restatement 1st of Torts s 757, and Comment b, factors 1-6 (1939). In the case at hand, plaintiff took measures to keep its information secret [F6] and plaintiff's product was unique on the market [F15]. This shows that plaintiff took efforts to maintain the secrecy of its information [F102], the information apparently was not known or available outside plaintiff's business [F105], and plaintiff's information was valuable for plaintiff's business [F104]. In *M. Bryce & Associates, Inc. v. Gladstone*, 107 Wis.2d 241, 319 N.W.2d 907 (Wis.App.1982), which held for plaintiff, there was similar evidence that plaintiff's information is a trade secret [F101]. In *Bryce*, as in the case at hand, plaintiff adopted security measures [F6], and plaintiff won.

[To emphasize strength: Need cases won by plaintiff with F15.]

The fact that plaintiff conveyed its information to defendant in the course of negotiations [F1] does not preclude a conclusion that plaintiff's information is a trade secret [F101], especially where plaintiff took security measures to protect the information [F6]. *Bryce*.

The factual strengths favoring plaintiff warrant the conclusion that plaintiff's information is a trade secret [F101], even though plaintiff's information could be discovered by reverse engineering plaintiff's product [F16].

[To downplay weakness: Need cases where plaintiff won in spite of F16, preferably cases with F6 or F15.]

A confidential relationship existed between plaintiff and defendant [F114]

Defendant acquired plaintiff's information through improper means [F110]

[argument justifying decision for plaintiff on these issues]

To justify a favorable decision on an issue:

Point to strengths related to issue, say why they matter.

Show case(s) in which these strengths led to favorable outcome.

Discuss weaknesses related to issue; point to strengths that compensate.

Show case(s) that had favorable outcome in spite of weaknesses.

Figure 5. Issue-based argument generated by CATO. References to elements of the Factor Hierarchy are marked with square brackets.

why they matter (corresponding to intermediate factors in the Factor Hierarchy) and presenting cases in which the same strengths were present and had favorable outcome. CATO downplays weaknesses by pointing to factors that may compensate and by citing cases that had a favorable outcome in spite of having those same weaknesses. It identifies the compensating factors using the Factor Hierarchy: The compensating factor must share an intermediate factor with the weakness. Generally, these arguments involve multiple cases. Here, CATO was given only one case, *Bryce*, which it uses to emphasize several strengths and to downplay a weakness. Where it does not have cases to cover an argument need, it inserts a note in the argument specifying the desired cases. These notes may prompt students to run queries for the specified cases. CATO can generate this type of argument for any problem and any small set of selected cases to cite. The algorithm is presented in [Aleven and Ashley, 1996].

During the CATO instruction, students also identify issues in *Mason* and organize their argument by issues, after studying two example issue-based arguments generated by CATO with cases suggested by Workbooks (Figure 2, steps 4 and 5). CATO shows arguments of this type without the annotation shown in Figure 5, in its Issue-Based Argument Window. Students compare their own argument outline to a third issue-based argument generated by CATO. We had planned that this would be an ar-

gument citing the same cases as the students had selected (which would have been easy to generate), but time permitted us only to provide a generic handout. The students' argument outlines were approximately one page in length.

4. A Difficult Skill for Students to Learn: Distinguishing

An important goal in the CATO instruction is for students to learn effectively to analogize and distinguish cases, major analytical tools for legal professionals. Also, students must learn to relate the similarities and differences between cases to more abstract knowledge, in ways that help their argument. Distinguishing in particular is difficult for students to learn. CATO helps students learn this skill in various ways (Figure 2, step 2). It presents dynamically-generated examples of arguments analogizing and distinguishing a problem and a case. Also, CATO conducts a "Mini Dialog," in which students get practice in identifying distinctions between cases and receive feedback. Finally, the program generates examples of arguments about the significance of distinctions, drawing on abstract normative knowledge.

In the course of working on their argument about the *Mason* problem, students could retrieve *Bryce* from the CATO Database and evaluate whether it is worth citing in an argument. A comparison of the factors of *Mason* and

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

Bryce (p)
 = F1 Disclosure-In-Negotiations (d)
 * F4 Agreed-Not-To-Disclose (p)
 = F6 Security-Measures (p) = shared factor
 * F18 Identical-Products (p) * distinction
 = F21 Knew-Info-Confidential (p)

Arguments about the significance of distinction F16

⇒ Plaintiff's argument *downplaying* distinction F16 in *Mason*.

In *Mason*, plaintiff's product information could be learned by reverse-engineering [F16]. This was not so in *Bryce*. However, this is not a significant distinction. In *Bryce*, plaintiff disclosed its product information in negotiations with defendant [F1], yet plaintiff won. In both cases, therefore, defendant obtained or could have obtained its information by legitimate means [F120]. But plaintiff may still win.

⇐ Defendant's argument *emphasizing* distinction F16 in *Mason*.

In *Mason*, plaintiff's product information could be learned by reverse-engineering [F16]. This was not so in *Bryce*. This distinction is highly significant. It shows that in *Mason*, plaintiff's information was available from sources outside plaintiff's business [F108].

This was not so in *Bryce*.

Arguments about the significance of distinction F4

⇒ Plaintiff's argument *downplaying* distinction F4 in *Bryce*.

In *Bryce*, defendant entered into a nondisclosure agreement with plaintiff [F4]. This was not so in *Mason*. However, this difference is not significant.

First, in *Mason*, plaintiff took security measures to protect the information [F6]. Therefore, in both cases, plaintiff took efforts to maintain the secrecy of its information [F102].

Second, in *Mason*, defendant knew that plaintiff's information was confidential [F21]. Therefore, in both cases, defendant was on notice that using or disclosing the information would be a breach of confidentiality [F115].

⇐ Defendant's argument *emphasizing* distinction F4 in *Bryce*.

In *Bryce*, defendant entered into a nondisclosure agreement with plaintiff [F4]. This was not so in *Mason*. This is a marked distinction. It shows that in *Bryce*, there was an express agreement to keep the information confidential [F121].

This was not so in *Mason*.

Figure 6: Argument exchanges generated by CATO about the significance of two distinctions

Bryce indicates that *Bryce* is not a perfect match with *Mason* (Figure 6, at the top): There are similarities, which plaintiff could seize on to analogize *Mason* to *Bryce*, but there are distinctions as well, which defendant could exploit to distinguish *Bryce*. The Argument Maker tool generates examples these argument moves and others, using cases retrieved from CATO's Database. Specifically, the Argument Maker generates examples of the first five argument moves listed in Section 2. Students learn initially about these argument moves by reading a general description in the Workbook. The Workbook instructions ask them to write their own arguments, study CATO examples, and compare their own arguments to CATO's.

The Argument Maker conducts a Mini Dialog, illustrated in Figure 7, intended to help students learn to identify distinctions between cases. In particular, it is important that students learn to differentiate between mere differences and real distinctions. (Unshared factors are distinctions only if they make the cited case stronger for the side that cited it.) Attempts at distinguishing that are not based on relevant differences are at best ineffective and at worst help the opponent. The Mini Dialog tries to make students aware of this by showing the argument an opponent might make in response to an ineffective attempt at distinguishing. In the top pane of the Argument Maker window, CATO displays the factors of a problem and case (without the "=" and "*" marks) and asks students to click on the factors that they think are distinctions (Figure 7, annotation 1). CATO then generates an argu-

ment distinguishing the case based on the factors selected by the student. If the student selected genuine distinctions, CATO acknowledges that the argument is effective (Figure 7, annotations 2 and 6). If the student selected differences that are not distinctions, CATO still presents an argument based on those differences, even though the argument is "misaligned" (Figure 7, annotations 3 and 4). It then shows how an opponent may take advantage of the misaligned argument by an astute response (Figure 7, annotation 5). Students can invoke this Mini Dialog for any pair of cases.

In order to help students to assess and explain the significance of similarities and differences in light of more general legal knowledge about the domain, CATO produces arguments about the significance of distinctions between cases. In these arguments, CATO draws a deeper contrast or parallel between the cases, strategically characterizes the significance of a distinction in terms of more abstract legal knowledge. CATO elaborates this contrast or parallel, marshaling appropriate base-level factors as evidence. The abstract characterizations cited in these arguments correspond to abstract factors in the Factor Hierarchy. This is illustrated in Figure 6, which shows CATO's arguments about two distinctions between *Mason* and *Bryce*, downplaying the distinctions on behalf of plaintiff, the party interested in establishing that *Mason* is similar to *Bryce*, and emphasizing the distinction on behalf of defendant, interested in establishing that the cases are dissimilar.

Interestingly, one characterization of a distinction's meaning may be used to draw a parallel between two cases, while a different sense of why the distinction matters may lead to a contrast. For example, two interpretations of the significance of distinction F4 (in *Bryce*, defendant had entered into a nondisclosure agreement; this was not so in *Mason*), are also warranted of the facts of *Mason*, although by different factors (Figure 6, right column). These interpretations are therefore used to downplay the significance of the distinction: In *Mason*, as in *Bryce*, there was evidence that plaintiff took efforts to maintain the secrecy of the information (abstract factor F102 Efforts-To-Maintain-Secrecy) and that defendant was on notice that plaintiff would consider use of its information a breach of confidence (abstract factor F115 Notice-Of-Confidentiality). In *Mason*, however, there are no factors that support a conclusion that there was an express agreement to keep the information confidential (F121 Express-Confidentiality-Agreement), a third sense in which factor F4 matters (see Figure 4). This abstract factor therefore becomes the focus of an argument emphasizing the F4 distinction. Similarly, the arguments downplaying and emphasizing distinction F16 focus on different interpretations of that factor's significance.

The crucial step in generating these arguments is strategically selecting the "focal abstractions," that is, the interpretations of cases on which the arguments focus. In general, context sensitivity in case comparisons means knowing which similarities and differences are salient in different circumstances: Which should an arguer focus upon and which should it ignore. Here, it means knowing which abstract interpretations of cases to are salient. CATO's arguments about the significance of distinctions are sensitive to the arguer's viewpoint, as illustrated here, and are also sensitive to the two cases being compared [Aleven and Ashley, 1996].

In the CATO instruction, students studied arguments about the significance of distinctions in the context of the *Mason* argumentation problem (Figure 2, step 2), using CATO's Argument Maker tool. The Workbooks instructed students to try to write one or two of these arguments for themselves and compare them against CATO's. Also, students study these arguments when evaluating the relevance of potentially interesting cases to cite (Figure 2, step 3).

5. Making Students' Research Processes More Manageable

CATO's Database of case squibs provides a convenient on-line source of cases for students to cite in arguments or to test theories against. The CATO Database is more useful for instructional purposes, than, say, a book of hard-copy squibs or a traditional full-text retrieval system, because the cases are represented and indexed by the factors that guide their use in arguments. This enables CATO to demonstrate argumentation examples with a wider range of cases and helps to make the legal research processes that students practice more manageable. CATO queries help students to zero in on relevant cases in theory-testing

and argumentation tasks. Also, the factor representations of cases help students get an initial impression of what the cases mean and whether they are relevant to a problem, an impression we encourage them to confirm by reading the squib. Also, the factor representations of cases make it easier for students to interpret CATO's examples and hence, to study more of them. Examples and Mini Dialogs are useful, but if students need to spend much time first to digest the cases that are involved, one could illustrate but few.

CATO's queries are important in theory-testing problems (Figure 2), where students are given a theory about trade secrets law, describing a class of fact scenarios in general terms, and are asked to test it against the cases in the database. The workbook instructions outline a 4-step process, in which students frame a CATO query for cases relevant to the theory and consider whether the retrieved cases confirm or disconfirm the theory. Students inspect the squib or factors of any disconfirming cases to see if these cases invalidate the theory or were decided on grounds not related to the theory. Finally, students modify the theory as needed.

In the context of argumentation problems, students use CATO queries to search for cases (Figure 2, step 3). This is a way of getting students to express the constraints that cases must satisfy in order to be useful in an argument. The Workbook mentions two general research and argumentation strategies: To cite cases that are highly similar to the problem and to cite cases that can be used to downplay weaknesses. In the CATO environment, it is straightforward to run queries to find the cases that are needed to "implement" such strategies. For example, students arguing on behalf of plaintiff in *Mason* will be very interested to know whether there are any cases in which plaintiff won despite the presence of the weaknesses in plaintiff's position, pro-defendant factors F1 Disclosure-In-Negotiations and F16 Info-Reverse-Engineerable. Using CATO's query language, they can find out simply by running a query for cases that have these factors. This is useful practice, especially since students may be inclined to overlook this strategy: They tend to deny weaknesses rather than to argue that they can win in spite of the weaknesses.

This kind of exercise would be far more difficult to accomplish without a database where cases are indexed by the factors used to express theories or argumentation strategies. One could imagine the difficulty of doing this kind of exercise with a full-text retrieval system. The exercise would bog down in finding out what the retrieved cases mean and to what extent they are relevant. Alternatively, an instructor could select a small number of cases in advance. However, we believe there is value in letting students deal with larger numbers of cases and in giving them the freedom and means to explore avenues of analysis not anticipated by the instructor. For example, this forces students to think about what cases they need and what argumentation strategies to pursue, which is likely to help them learn. As a result, there is an increased focus on the parts of the task that we target in the CATO instruction, comparing and contrasting cases.

CATO: Assume you are arguing on behalf of the defendant in *Mason*. Suppose plaintiff cites the *Bryce* case as follows:

Where plaintiff adopted security measures [F6] and defendant knew that plaintiff's information was confidential [F21], even though plaintiff disclosed its product information in negotiations with defendant [F1], plaintiff should win a claim of trade secrets misappropriation, as in *M. Bryce & Associates, Inc. v. Gladstone*, 107 Wis.2d 241, 319 N.W.2d 907 (Wis.App.1982).

How would you respond on behalf of the defendant?

Please type your response in a WordPerfect window, then click on "Continue".

Student: continue

CATO: Assuming that in your response for defendant you distinguished *Bryce*, please "inform" CATO of your answer, as follows: In the top part of this window, click on the factors that you mentioned in your response as the relevant differences of *Mason* and *Bryce*.

If by mistake you clicked on the wrong factor, you can "turn off" that factor by clicking on it again.

Finally, click on "Continue".

Student: f16

CATO: Yes, defendant can distinguish *Bryce*, on the basis of the factor that you selected, as follows:

M. Bryce & Associates, Inc. v. Gladstone is distinguishable, because in *Mason*, plaintiff's product information could be learned by reverse-engineering [F16]. This was not so in *Bryce*.

This factor is a reason that *Mason* is stronger for defendant than is *Bryce*, and therefore, that *Bryce* should not be followed.

Is there any other way in which defendant can distinguish *Bryce*? Please click on other distinguishing factors you may find. Then click on "Continue".

Student: f15

CATO: Suppose defendant made the following argument:

M. Bryce & Associates, Inc. v. Gladstone is distinguishable, because in *Mason*, plaintiff was the only manufacturer making the product [F15]. This was not so in *Bryce*.

Would this really help defendant?

When you have answered this question for yourself, please click on "Continue".

Student: continue

CATO: Plaintiff could respond to defendant's argument, as follows:

Your Honor, as defendant's counsel correctly pointed out, in the current problem, plaintiff was the only manufacturer making the product [F15]. *Bryce*, which was won by plaintiff, does not have this strength on plaintiff's part but is otherwise similar to the problem at hand. Where this additional strength is present, a decision in favor of plaintiff is warranted even more strongly than it was in *Bryce*.

In short, even though factor F15 is a difference, defendant cannot use it to distinguish *Bryce*. Defendant needs to focus on other differences, namely, on those differences that make *Mason* look stronger for defendant than *Bryce*.

How could defendant distinguish *Bryce*, other than by pointing to factor F16? Please click on other distinctions, then click on "Continue".

Student: f4 f18

CATO: Yes, defendant can distinguish *Bryce*, on the basis of the factors that you selected, as follows:

M. Bryce & Associates, Inc. v. Gladstone is distinguishable, because in *Bryce*, defendant's product was identical to plaintiff's [F18] and defendant entered into a nondisclosure agreement with plaintiff [F4]. This was not so in *Mason*.

These factors are reasons that *Bryce* is stronger for plaintiff than is *Mason*, and therefore, that *Bryce* should not be followed.

Very good!

1. CATO presents argument analogizing *Mason* to *Bryce*, invites student to respond on behalf of defendant.

2. CATO confirms that student selected distinction. Shows argument based on that distinction, States general principle.

3. Student selects unshared factor that is not a distinction.

4. CATO presents misaligned argument based on that supposed distinction.

5. To explain why the argument is misaligned, CATO presents response showing how misaligned argument sets up damaging response by opponent.

6. Students selects remaining distinctions, CATO confirms that they are distinctions.

Figure 7: Unedited transcript of a mini-dialog with CATO. Annotations are in boxes.

6. Empirical Evaluation of CATO

We conducted an evaluation study to assess the effectiveness of the CATO instruction in comparison to the best known way of teaching the same skills by traditional methods¹. In fact, the control group instruction in this experiment represents a more ideal instructional situation than found in most American law school curricula. The study was carried out in the context of a second-semester legal writing course at the University of Pittsburgh School of Law. The subjects were first-year law students, recruited from those taking the course, and were assigned randomly to an experimental group of 16 and a control group of 14.

During a three-week period, instead of attending their regular class meetings, the students in the experimental group went to a CATO computer lab in the Law School set up specifically for this experiment. Each student worked with CATO for a total of nine 50-minute sessions. After an introductory classroom session, students collaborating in pairs used the CATO program to work through theory-testing and argumentation problems, as described in the previous sections.

The control group instruction was designed to cover, in the same amount of time, the same skills of making arguments with cases. It was based on the same casebook, but was taught in a more traditional way, by an instructor who was not familiar with the CATO model. The control group was divided into 3 groups of, on the average, 8 students. (These groups also included the students who elected not to participate in the experiment.) During four classroom sessions, the course instructor used a Socratic method to present a framework of inquiry for trade secrets law, synthesized from the casebook cases. During two moot court sessions, students made oral arguments. The instructor, combining the roles of judge and teacher, very actively moderated the argument exchange and kept students on their toes by asking questions and putting some pressure on their arguments. The students prepared for these sessions outside of class, spending at least 75 minutes each time.

To assess the improvement in students' argumentation skills, we administered a Basic Argument Skills pre-test and post-test, in-class exams which included a problem and three cases and focused on basic argument moves and relatively uncomplicated argumentation problems. Special care was taken to make the pre-test and post-test equally difficult. Both tests comprised the same questions, except that different cases were used. We also administered a second post-test, a more complex and advanced legal writing assignment, involving a problem and six cases, in which students completed a six-page memo at home in one week. The tests and grading criteria were designed by us, in consultation with the legal writing instructor. All

¹ In the current study, students used CATO without guidance of a human tutor. The study therefore represents a stricter test than a previous study [Aleven and Ashley, 1995], in which we tested an early version of CATO that did not have a Factor Hierarchy and did not generate argumentation examples and students used CATO under the guidance of a human tutor.

	Basic Argument Skills				Memo Writing		
	Pre-Test		Post-Test		Prev.	Post-Test	
Exp. Gr. Avg.	60	C-	70	C+	63	70	B-
Control Gr. Avg.	55	D	68	C	63	79	B+
CATO answers	81	B+	87	A-		62	C

Table 1. Results of the Basic Argument Skills Tests and Legal Memo Writing Assignment

tests were graded in a blind test by the legal writing instructor. Without informing the grader, we included in the materials to be graded, answers generated by the CATO program, formatted so as to disguise the fact that they had been generated by CATO. The answers contained issue-based arguments, similar to the one illustrated in Figure 5.

The results, presented in Table 1, indicate that on the Basic Argument Skills Tests, both groups' scores improved from pre-test to post-test. (The letter grades were computed by converting the numerical grades.) In each group, the improvement was statistically significant, as we determined using a t-test ($p < .05$). There was no significant difference between the groups in pre-test, post-test, or gain scores (t-test, $p > .05$). On the Memo-Writing Assignment, however, the control group scores were higher and the difference was statistically significant (t-test, $p < .05$). The difference did not exist on a memo-writing assignment of the previous semester. While CATO's answers ranked among the best on the Basic Argument Skills tests, CATO's memo did not fare so well.

7. Discussion and Conclusion

The results on the Basic Argument Skills Tests indicate that the CATO instruction leads to statistically significant improvement of students' basic skills of making arguments with cases², comparable to that achieved by an experienced legal writing instructor teaching small groups of students. This result should be interpreted in light of the fact that the control group instruction was a very high standard to compare against. First, the subjects were enrolled in a program that provides extra-intensive legal writing instruction. Students are selected for this program, by the law school's admissions committee, on the basis of various indicators of law school aptitude suggesting a need for special instruction. Second, the control group instruction involved very small groups: 8 students on average. Finally, the control group instructor is the director of the legal writing program at the University of Pittsburgh School of Law as well as the special program described above. Experienced in teaching such students, he is a

² The fact that CATO's grade improved from pre-test to post-test means most likely that CATO's arguments came out better for the post-test problems. It does not necessarily follow that the post-test was easier for students than was the pre-test. We had taken special care to make both tests equally difficult. That students' skills improved from pre-test to post-test is evidenced not only by their grades but also by statements made by the legal writing instructor/grader in a posthoc interview.

dedicated teacher who enjoys an excellent rapport with his students, as evidenced for example by the students' enthusiasm for the oral argument sessions³.

The legal memo-writing assignment was a more advanced and complex task than the Basic Argument Skills tests or the the argumentation tasks students carried out with CATO. It involved more cases and posed additional problems of selecting the best argument and composition. It was intended to explore the frontier of the CATO instruction. We interpret the results on the memo-writing assignment as saying that while both the experimental and control group learned basic skills of making arguments with cases, the control group was better able to integrate these skills in a more complex context.

We attribute this to the fact that the control group instruction taught the argumentation skills in a more integrated way than the CATO instruction. The legal writing instructor who taught the control group did not focus on *methods* of argumentation or analysis, but instead taught argumentation in a "holistic" way: He discussed cases and let students make oral arguments with cases. The CATO instruction, on the other hand, initially focused on argument components and addressed integration only in the later stages, when students organized arguments by issues and selected the most relevant cases to cite in their arguments. Unfortunately, most students ran out of time while working on these topics. Also, the CATO model addresses integration only to a limited extent. We made available certain background knowledge in the casebook, the workbooks, and the case squibs, but merely making it available, without demonstrating its use, or teaching students how to use it, evidently is not sufficient in order for students to integrate it effectively in their arguments. Finally, it did not help that CATO's issue-based arguments did not conform to the organization and format that the legal writing instructor wanted to see, as was clear from the comments that he wrote on CATO's memo and the grade this memo received. Yet the limitations revealed by the memo-writing assignment should not overshadow the positive results on the Basic Argument Skills tests.

The evaluation indicates that the CATO instruction is effective in teaching basic skills of making arguments with cases. This result confirms the pedagogical utility of the CATO model and Factor Hierarchy and confirms that the instructional environment, in conjunction with the Workbooks, communicates the model effectively. The experiment helped us identify areas in which CATO must be improved, if it is to prepare students for a memo-writing task. First and foremost, the issue-based arguments must conform to the format and structure that the legal writing

instructor teaches. Also, since students find it difficult to identify issues and use them to organize arguments, it is important to place issues and the Factor Hierarchy more in the foreground. To improve CATO's arguments, we must address the critiques of the legal writing instructor, for example, that some arguments are conclusory, that a more detailed discussion of case law is needed, and that it is desirable to present a synthesis of cases. This may be achieved in part by improving the argument generation routines and templates, but will also require extensions to the CATO model. It may be sufficient (and feasible) to use memo outlines that are not completely filled in. Finally, it is interesting to investigate whether changes in the program's teaching strategies could improve its effectiveness. For example, CATO could use the Factor Hierarchy to engage students in a discussion of how in individual cases the factual strengths and weaknesses relate to issues. Also, it could use its model to generate counterarguments and perhaps to evaluate student behavior. We are particularly interested in exploring how CATO could be integrated more with classroom instruction. For example, a human instructor could demonstrate CATO in the classroom, using an LCD panel to project its output onto a big screen. Students could then use the program in a lab, working on additional problems to solidify their understanding.

References

- Aleven, V., and K. D. Ashley, 1996. How Different is Different? Arguing about the Significance of Similarities and Differences. In *Advances in Case-Based Reasoning: Proceedings EWCBR-96*, edited by I. Smith and B. Faltings, 1-15. Berlin: Springer.
- Aleven, V., and K. D. Ashley. Doing Things with Factors. In *Proceedings of the Fifth International Conference on Artificial Intelligence and Law*, 31-41. New York: ACM, 1995.
- Ashley, K. D., 1990. *Modeling Legal Argument: Reasoning with Cases and Hypotheticals*. Cambridge, MA: MIT Press.
- Berman, D. H., and C. D. Hafner, 1993. Representing Teleological Structure in Case-Based Legal Reasoning: The Missing Link. In *Proceedings of the Fourth International Conference on Artificial Intelligence and Law*, 50-59. New York: ACM.
- Branting, L. K., 1991. Building Explanations from Rules and Structured Cases. *Int. J. Man-Machine Studies* 34: 797-838.
- Muntjewerff, A. J., 1994. Towards Automated Training of Legal Problem Solving. In *Proceedings of Ed-Media 94: World Conference on Educational Multimedia and Hypermedia*, edited by T. Ottman and I. Tomek, 410-415.
- Rissland, E. L., and D. B. Skalak, 1991. CABARET: Rule Interpretation in a Hybrid Architecture. *Int. J. Man-Machine Studies* 34: 839-887.
- Rissland, E. L., Skalak, D. B., and M. T. Friedman, 1996. BankXX: Supporting Legal Arguments through Case-Based Search. *Artificial Intelligence and Law*, 4: 1-71.
- Routen, T., 1992. Reusing Formalisations of Legislation in a Tutoring System. *Artificial Intelligence Review*, 6: 145-159.
- Sherman, D. M., 1989. Expert Systems and ICAI in Tax Law: Killing Two Birds with One AI Stone. In *Proceedings of the Second International Conference on Artificial Intelligence and Law*, 74-80. New York: ACM.
- Span, G., 1993. LITES, an Intelligent Tutoring System for Legal Problem-Solving in the Domain of Dutch Civil Law. In *Proceedings of the Fourth International Conference on Artificial Intelligence and Law*, 76-81. New York: ACM.
- Wenger, E., 1987. *Artificial Intelligence and Tutoring Systems: Computational and Cognitive Approaches to the Communication of Knowledge*. Los Altos, CA: Morgan Kaufmann.

³ It would have been informative to include in the experiment a second control group of students who received no instruction in the targeted skills or engaged only in self-guided study. However, this was not feasible for various reasons, not the least of which is that one cannot have students enrolled in a course do nothing or be without an instructor's guidance for three weeks. Also, given the choice, a comparison against a high standard is more interesting.