2000

Designing Electronic Casebooks that Talk Back: The Cato Program

Kevin D. Ashley
University of Pittsburgh School of Law, ashley@pitt.edu

ARTICLES

DESIGNING ELECTRONIC CASEBOOKS THAT TALK BACK: THE CATO PROGRAM

Kevin D. Ashley*

ABSTRACT: Electronic casebooks offer important benefits of flexibility in control of presentation, connectivity, and interactivity. These additional degrees of freedom, however, also threaten to overwhelm students. If casebook authors and instructors are to achieve their pedagogical goals, they will need new methods for guiding students. This paper presents three such methods developed in an intelligent tutoring environment for engaging students in legal role-playing, making abstract concepts explicit and manipulable, and supporting pedagogical dialogues. This environment is built around a program known as CATO, which employs artificial intelligence techniques to teach first-year law students how to make basic legal arguments with cases. Ongoing improvements in CATO point the way for electronic casebooks to engage students in realistic analogical legal arguments. By reorganizing the electronic casebook's explicit information about cases and implicit knowledge of argumentation along the lines of CATO's knowledge sources, it is possible to orchestrate a real dialogue between a book and its reader.

CITATION: Kevin D. Ashley, Designing Electronic Casebooks that Talk Back: The CATO Program, 40 Jurimetrics J. 275–319 (2000).

---

*Kevin D. Ashley is Professor of Law and Intelligent Systems and Senior Scientist, Learning Research and Development Center, University of Pittsburgh, Pittsburgh, Pennsylvania, ashley+@pitt.edu. This research was supported by grants from 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. Vincent Aleven designed and built the CATO program as his Ph.D. dissertation project in the University of Pittsburgh Graduate Program in Intelligent Systems (ISP). His thoughtful comments on a draft of this article have been invaluable. Very helpful assistance also was provided by Stefanie Brüninghaus, Ravi Desai, and Sean Bamford.

Some of the most interesting aspects of a legal casebook are the author's intentions regarding the selection and grouping of the cases included, the manner in which the cases are to be compared, and the lessons to be drawn from those comparisons. According to Karl Llewellyn, "the foundation of the case system" consists of:

> a series of opinions which in some manner are related. They may or may not be exactly alike in their outcome. They are always supposedly somewhat similar on their legally relevant facts. Indeed, it is the aspects in which their facts are similar which give you your first guidance as to what classes of fact will be found legally relevant, that is, will be found to operate alike, or to operate at all, upon the court. On the other hand, the states of fact are rarely, if ever, quite alike. And one of the most striking problems before you is: when you find two cases side by side which show a difference in result, then to determine what difference in their facts, or what difference in their procedural set-up, has produced that difference in result. Those are the two problems which must be in your mind as you examine the language of the opinions. I repeat them. First, what are the significant categories of facts, and what is their significance to the court? Second, what differences in facts or in procedural set-up produce differences in the court's action when the situations are otherwise alike?[1]

If law students perform the "case system game" as Llewellyn calls it, if they proceed by "a rough application of the logical method of comparison and difference," the author's intended lessons will be revealed.[2] Presumably, students learn the game by preparing the cases and participating in the Socratic oral discussions that typify the American legal classroom experience. Yet, the process of comparison and difference is rarely explained systematically and often remains mysterious. Some students seem never to learn how to play the case system game. Some learn only belatedly because they lack certain background skills upon entering law school. And, there are indications that fewer students will have the chance to participate in Socratic classroom discussions. As distance learning shifts the law school venue out of the classroom, the question arises whether students can learn the game via email and an on-line forum.

At the same time, the advent of electronic *casebooks* promises unprecedented flexibility in presentation, connectivity, and interactivity. However, these additional degrees of freedom may overwhelm students' attempts to compare and contrast cases. If casebooks are to achieve their pedagogical goals, we will need new methods for guiding students.

This article describes research to demystify the method of comparison and difference. It aims to enable electronic casebooks to generate examples of arguments comparing cases and to engage students directly in legal argument dialogues where the student stakes out a position and the casebook argues back. The research centers on the CATO program, an intelligent tutoring environment

---

1. KARL LLEWELLYN, THE BRAMBLE BUSH: ON OUR LAW AND ITS STUDY 49 (Oceana Publications 1986) (1960).

2. *See id.*

for teaching law students to reason with cases.[3] An "intelligent tutoring environment" is a computerized instructional environment that employs artificial intelligence (AI) techniques.[4] A subfield of computer science, AI endeavors to create computer programs whose behavior, if performed by a human being, would be regarded as intelligent.[5] Researchers in the more specialized subfield of AI and Law apply AI techniques to model legal reasoning.[6] CATO employs an AI model of legal reasoning, specifically, of argument supported by cases. Designing the web-based version of CATO has raised important questions about which pedagogical goals should underlie the use of electronic casebooks in legal education. What new opportunities do electronic casebooks provide for helping students learn and sharpen intellectual legal skills? How best can such opportunities be implemented?

Experience with CATO indicates that there are three related areas for new contributions from electronic casebooks in legal pedagogy:

> (a) Constructing role-playing environments. Electronic casebooks can open new avenues for engaging students in exercises in which they play the role of legal professionals in realistic contexts. This is one of the most effective of pedagogical techniques.
> (b) Making abstract concepts explicit and manipulable in a rich, but limited setting. Electronic casebooks can integrate on-line information retrieval resources and computerized instructional tools that help students learn abstract concepts and give them practice applying the abstract concepts to solve problems.
> (c) Supporting pedagogical dialogues. Electronic casebooks can engage students in analytical dialogues about substantive legal domains and tasks. Conceivably, students can "talk" to electronic legal casebooks, and the casebooks can "talk" back.

To achieve these contributions, electronic casebooks will need to be integrated with intelligent tutoring environments like CATO. This article explains how that may be accomplished. It illustrates how CATO supports law students in realistic role-playing, makes abstract legal concepts explicit and manipulable, performs legal research and constructs arguments, and engages students in argumentation dialogues.

CATO is a prototype for integrating computerized instruction and electronic legal casebooks. It is presented to law students along with a traditional-appearing electronic casebook chapter. As in a regular casebook, argumentation and discussion questions follow the text's major cases. CATO offers a set of

---

3. *See* Vincent Aleven, Teaching Case-Based Argumentation Through a Model and Examples (1997) (Ph.D. dissertation, University of Pittsburgh) <http://www.cs.cmu.edu/~aleven/dissertation.html>.

4. *See* ETIENNE WENGER, ARTIFICIAL INTELLIGENCE AND TUTORING SYSTEMS 3–6 (1987).

5. *See generally* SEMANTIC INFORMATION PROCESSING (Marvin. L. Minsky ed., 1968).

6. *See generally* Edwina L. Rissland, *Artificial Intelligence and Law: Stepping Stones to a Model of Legal Reasoning*, 99 YALE L.J. 1957 (1990) (providing a very good introduction to and overview of research in AI and Law). *See also infra* note 116.

computerized instructional tools for helping students analyze and respond to these questions. Guided by a workbook, students use CATO's tools to pose and test hypotheses against a database of real legal cases, to construct multi-case arguments, and to compare their arguments with CATO's. In the classroom, law students encounter CATO as part of an electronic casebook published via the World Wide Web. Its chapters and workbooks are HTML documents with hypertext links to lead students from the end-of-case argumentation and discussion questions directly to CATO's tools.[7]

Ongoing improvements in CATO's ability to engage students in pedagogical dialogues point the way for electronic casebooks to talk back to students. Students will have a rich array of argument moves to make, and the electronic casebook will trump the student's point, concede, or sometimes introduce a new kind of argument in response. By reorganizing the electronic casebook's explicit information about cases and implicit knowledge of argumentation along the lines of CATO's knowledge sources, it is possible to orchestrate a real dialogue between a book and its reader.

This article develops and explains these claims. Part I describes the limitations and dangers of complex electronic casebooks and the need to guide students' use of the on-line resources. Part II presents the CATO program as a prototype for integrating intelligent tutoring and electronic casebooks. It describes CATO's curriculum of argumentation lessons and the environment in which students play the role of advocates making and responding to arguments. It explains how CATO's computational model of legal argument makes argumentation concepts explicit and manipulable. Part III focuses on how CATO talks back to students. Part IV presents empirical evidence that CATO teaches argumentation concepts as well as experienced human instructors. Part V summarizes CATO's research contributions and suggests further steps in designing interactive electronic legal casebooks.

I. PEDAGOGICAL OPPORTUNITIES OF ELECTRONIC CASEBOOKS

Although many electronic casebooks "mirror the organizational structure and content of the paper version" and fail to transform the "fundamental nature of the materials,"[8] such a casebook "can be constructed to permit any adopting law professor the ability to easily and professionally customize the casebook—to dissect, reshape, move, delete or add materials—prior to distribution to

7. In class, an instructor equipped with a computer projector may use CATO to demonstrate argumentation skills and examples, projecting the program's output onto a screen so that all students can follow along. Having seen how to work through a complex cycle of testing a legal hypothesis or constructing a multi-case argument, students can use CATO after class on their personal computers in a web browser environment to practice the in-class exercises and explore new examples.

8. Gary Neustadter, *Rethinking Electronic Casebooks, Lessons from the Web, JURIST: The Law Professor's Network* (last modified June 1, 1998) < http://jurist.law.pitt.edu/lesjun98.htm>.

students."[9] The "new" electronic casebook will be characterized by flexibility in presentation,[10] connectivity,[11] and interactivity.[12] However, the potential richness of the new electronic casebook may have a downside in complexity and demands on students.[13]

The challenge is to develop techniques for maximizing the pedagogical benefits of electronic casebooks' malleability and connectivity, while adequately focusing the student reader on the authors' intended lessons. Techniques that can help students take advantage of the richness of the new electronic casebook to learn and sharpen legal analytical skills include (a) constructing role-playing environments, (b) making abstract concepts explicit and manipulable, and (c) supporting pedagogical dialogues.

A. Role-Playing Environments

To be effective, instructional materials addressing analytical legal skills probably require a "strong real-world purpose or context."[14] Engaging students in playing the role of legal professionals in realistic exercises motivates them to learn and provides a concrete context in which to learn.[15]

Electronic legal publishing facilitates the creation of realistic professional role-playing environments. An electronic medium's flexibility in designing user interfaces lends itself to creating realistic virtual legal venues such as a law office or courtroom. More importantly, an electronic casebook's malleability and

---

9. *Id.*

10. As Professor Neustadter explains:

> The author of this type of electronic casebook will create the outline in a linear sequence which reflects her conceptual map, but that outline can easily be electronically reshaped by both professors (for their classes) and by students (for their outlines) without disturbing the links to the digital libraries. . . . Authors of these electronic casebooks can provide more than one possible outline, i.e. more than one template.

Id.

11. "[I]t is now possible to supplement these internal links with links to locations outside the book, on the World Wide Web. . . . Because electronic casebooks can easily accommodate vast quantities of data, they do not suffer the same length and volume constraints imposed upon conventional casebooks." *Id.*

12. "By virtue of 'News' [a location in which the author can post new materials for readers] and 'Contact Author' [a convenient email connection from the reader to the author] the book can become an active, ongoing collaborative learning experience in which users throughout the country can participate." *Id.* An electronic casebook can also be connected conveniently to on-line extra-class discussion groups such as those supported by WestGroup's TWEN network.

13. Professor Neustadter cautions:

> Enthusiasm about links to web sites must nonetheless be tempered with realism about limits on student time and endurance. Some students might ignore hypertext links to web sites, particularly if they feel overwhelmed by the endless stream of links to which the initial link can lead them. Others, inveterate web surfers, might pursue the links for hours, possibly at the expense of more careful attention to the assembled materials.

Id.

14. J. F. Stratman, *The Emergence of Legal Composition as a Field of Inquiry: Evaluating the Prospects*. 60 REV. EDUC. RES. 153, 213 (1990).

15. *Id.* at 215.

connectivity facilitate integrating the electronic writing environments and on-line information sources and retrieval tools that lawyers use in practice.[16]

An electronic casebook on Cyberspace and the Law, for example, might present a landmark case like *Reno v. American Civil Liberties Union*,[17] in a very different way from a traditional casebook. In that case, the Court struck down the indecency provisions of the Communications Decency Act (CDA) as unconstitutional on their face. While a student's first encounter with the decision in a traditional casebook would be the opinion itself, a new electronic casebook might first assign students a role to play. For instance, the electronic casebook might instruct the student: "Pretend you are a law clerk to Justice Stevens and that your assignment is to prepare an outline of a draft opinion that he can circulate among his colleagues." The electronic casebook might present hypertext links to the opposing parties' briefs,[18] the opinion of the three judge panel below,[19] and the transcript of the oral argument before the Court,[20] as well as URLs for some of the web site publishers who would be affected if the CDA provisions were upheld.[21] Students might compose an outline based on the linked materials. They might even orchestrate an on-line discussion group among themselves as if they were Justices discussing the decision.

It is questionable, however, whether many students would undertake such ambitious exercises on a regular basis in reading the new electronic casebooks. To perform exercises like these, students would need more guidance and feedback than is available in law school classes that have high student-teacher ratios and few opportunities for individualized student-teacher interaction. Intelligent tutoring environments such as CATO offer techniques for leveraging scarce law school teacher resources.

B. Making Abstract Concepts Explicit and Manipulable

Intelligent tutoring environments provide techniques for guiding students in learning how to apply concepts. For any legal concept, a traditional legal casebook typically presents a rather small set of legal opinions. Students read and compare a handful of cases. Ideally, Socratic discussion in the classroom leads them to formulate and test hypotheses about the legal concept. They identify underlying policies or principles, focus on a range of circumstances over which

---

16. Neustadter, *supra* note 8 ("[I]t is now possible to supplement these internal links with links to locations outside the book, on the World Wide Web. With these external links the author can bring the legal materials more fully to life . . . .").

17. 521 U.S. 844 (1997).

18. ACLU, *Brief for Appellee* (visited Apr. 11, 2000) <http://www.aclu.org/court/renovaclu.html>; Dep't of Justice, *Brief for Appellant* (visited Apr. 11, 2000) <http://www.ciec.org/SC_appeal/970121_DOJ_brief.html>.

19. American Civil Liberties Union v. Reno, 929 F. Supp. 824 (E.D. Pa. 1996).

20. *Transcript of Supreme Court Oral Argument, Reno v. American Civil Liberties Union, 521 U.S. 844 (1997)* (visited Apr. 11, 2000) <http://www.aclu.org/issues/cyber/trial/sctran.html>.

21. *See, e.g., Cyber-Liberties: American Civil Liberties Union Freedom Network* (visited May 8, 2000) <http://www.aclu.org/issues/cyber/trial/plantiff.html>.

the concepts have been applied, and learn to argue whether the concept applies in new fact situations.

An electronic casebook connected to on-line databases of cases could provide an opportunity to evaluate knowledge about a targeted legal concept over a wider range of fact scenarios. It is likely, however, that among the welter of real cases in an on-line database, there are so many conflicting examples and competing discussions that students may be overwhelmed, or, at least unable to test reliably their knowledge. An electronic casebook's links into the realistically complex world may be a spring board into the deep end of the pool for students who still need to learn to float.

An intelligent tutoring environment like CATO employs an intermediate-sized set of cases represented and indexed so that the tutoring system can help students retrieve and reason with them.[22] This set acts as a middle ground between the handful of cases a traditional legal casebook devotes to a concept and the on-line full-text databases of WESTLAW or LEXIS. The system's tools make the concepts explicit and manipulable. They assist students in applying the concepts to solve problems and make arguments. Students play professional roles focused on specific fact situations. Using the tools, students can retrieve cases illustrating targeted concepts, apply them to the fact situations, and test their models of the concept.

A new electronic casebook could incorporate an intelligent tutoring environment and an intermediate-sized case database to scaffold students' application of the casebook's legal concepts. Engaging students in testing hypotheses, solving problems, and making arguments over the case database can prepare them to tackle similar exercises using the full panoply of on-line resources. As students progress through the curriculum, scaffolding can be removed gradually.[23]

Another intelligent tutoring system technique involves dynamically modifying the order of presentation of a curriculum of materials in response to an assessment of a students' mastery of the subject matter.[24] An electronic casebook allows authors, instructors, and even students to organize their own outlines of the material. The intelligent tutor that engages students in applying analytical concepts in an intermediate-sized case database also could collect information about which exercises the student has tried and how the student has done. It could tailor the presentation of materials to the student's demonstrated need and ability, suggesting some paths and pruning others.[25]

---

22. CATO uses dozens or even hundreds of cases.

23. The example of CATO in the next section suggests how this might be accomplished.

24. *See* C. Conati & K. VanLehn, *A Student Model to Assess Self-Explanation While Learning from Examples*, 1999 PROC. SEVENTH INT'L CONF. ON USER MODELING 303.

25. Although CATO does not presently have a student modeling capability, it can collect data about a student's path through the material. As discussed in Part III, it could use these data to make decisions about when to introduce new material.

C. Supporting Analytical Dialogues

The interactivity of electronic casebooks offers hope of engaging students in analytical dialogues with fellow students, the instructor, or even the author. Two-way communication provides a concrete context for discussion, lends interest to the subject matter, and motivates mastery. In addition, engaging students in making and defending arguments about the subject matter can induce students to construct and revise their own models of the concepts.[26]

The pedagogical utility of such dialogues may depend, however, on the ability of instructors to monitor, moderate, and participate in the dialogues. Although a student responding to an instructor in class receives excellent feedback, most students are merely on-lookers. There is no record of the exchange except insofar as students take notes, which is itself a distraction. There are practical limits on the extent to which a human instructor can interact with each student in class and on the time available after class for individualized instruction.[27]

While it is not yet technologically feasible for a computer program to engage in free-form arguments or discussions of general topics, an intelligent tutoring environment like CATO can engage students in more constrained arguments like the one in Figure 1. In this dialogue, CATO teaches a student about the concept of distinguishing. The dialogue simulates an argument before a judge between two habitually contending attorneys, Perry Mason and Hamilton Burger, from a classic television series.[28] The student plays the role of Perry Mason, attorney for the defendant in a trade secret misappropriation case. In the sample dialogue, the problem case happens to involve a bar owner, Tony Mason (no relation to Perry), who sues the Jack Daniel Distillery for misappropriating his secret recipe for Lynchburg Lemonade. CATO controls the overall dialogue and, in particular, the words of the judge, of Hamilton Burger, opposing counsel for the plaintiff, and of Della Street, Perry's savvy assistant (who offers helpful hints). In particular, Burger makes arguments to which the judge directs Mr. Mason to respond. From a menu, the student selects argument moves for Mr. Mason to try, as well as specific information. Here, the student selects factual differences or factors.[29] CATO translates the student's menu selections into Perry Mason's argument and then responds appropriately on behalf of Burger and the judge. As illustrated,

---

26. *See* A.L. Stevens & A. Collins, *The Goal Structure of a Socratic Tutor,* 1977 PROC. NAT'L ACM CONF.; A. Collins & A.L. Stevens, *Goals and Strategies of Inquiry Teachers, in* ADVANCES INSTRUCTIONAL PSYCHOLOGY 65–119 (R. Glaser ed., 1982).

27. Email expands the medium of communications between student and teacher, but there are limits to the amount of time teachers can devote to interacting with students.

28. My apologies to Erle Stanley Gardner. Today's law students seem never to have heard of Perry Mason. (Even *L.A. Law* is old hat.) For a short refresher course, see *The Perry Mason Pages,* (visited Apr. 11, 2000) <http://www.ozemail.com.au/~jsimko/>. So far, I have resisted invitations to substitute personalities from *The Practice*, the O.J. Simpson trial, or *Law and Order*.

29. The "F" numbers in Figure 1 refer to factors. Defined in Part II.A, factors are CATO's representation of important factual similarities and differences among cases.

Burger is smart enough to take advantage of the weaknesses in Mason's (the student's) argument.

Court is in session . . .

1. **Della Street (whispering into Perry's ear):** Perry! Thank heaven you are here! The Judge wants to hear oral arguments on our motion to dismiss on behalf of defendant in *Mason v. Jack Daniel Distillery.*

2. **Judge:** Ah Mr. Mason, there you are. Your opponent, Mr. Burger, is just finishing up an argument for plaintiff.
   You may proceed, Mr. Burger.

3. **Mr. Burger for Plaintiff (CATO):** Your Honor, 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 *The Boeing Co. v. Sierracin Corp.*, 108 Wash.2d 38, 738 P.2d 665 (1987).

4. **Judge (to Defendant's counsel):** Your response, Mr. Mason?

5. **<STUDENT (selecting from menu):** Distinguish *Boeing* [F15] [F16] [F18] (Continue)>

   Mr. Mason for Defendant (Student): *The Boeing Co. v. Sierracin Corp.* is distinguishable, Your Honor, because in *Mason*, plaintiff's product information could be learned by reverse-engineering [F16]. This was not so in *Boeing*.
   Boeing is also distinguishable, because in *Mason*, unlike *Boeing*, defendant's product was identical to plaintiff's [F18] and plaintiff was the only manufacturer making the product [F15].

6. **Judge (to Plaintiff's counsel):** Mr. Burger?

7. **Mr. Burger for Plaintiff (CATO):** Your Honor, as Mr. Mason concedes, in the current problem, defendant's product was identical to plaintiff's [F18] and plaintiff was the only manufacturer making the product [F15]. *Boeing*, which was won by plaintiff, does not have these strengths on plaintiff's part but is otherwise similar to the problem at hand. Where these additional strengths are present, a decision in favor of plaintiff is warranted even more strongly than it was in *Boeing*.

8. **Judge (to Defendant's counsel):** I am sorry, Mr. Mason, but I am afraid I have to agree with Mr. Burger.

Figure 1. CATO-led Dialogue

As discussed in Part III, while the argument simulation continues, the student can still score some points against Burger. Thus, an electronic casebook can be coupled with an intelligent tutoring environment to engage students in pedagogical dialogues about the subject matter. CATO's dialogues are proof of that. Fully appreciating the dialogue in Figure 1 and something of the underlying technology will be easier after reviewing the next section, which discusses how the program works.

II. CATO AS PROTOTYPE

The CATO program demonstrates how the three pedagogical opportunities can be realized in an electronic casebook. As an adjunct to an electronic casebook chapter, the program engages students in role-playing in an environment that makes argumentation concepts explicit and usable over an intermediate-sized database of cases, and it supports students in the kind of argumentation dialogues shown in Figure 1.[30]

A. The CATO Environment

The CATO environment (Figure 2) consists of six computational tools, including a database of 150 trade secret cases, an electronic legal casebook chapter dealing with trade secret law, and a workbook. The electronic casebook chapter presents five important legal opinions. Each opinion is followed by argumentation problems.[31] The workbook instructs students in how to use CATO's tools to address the problems.[32]

CATO's legal cases are represented and indexed abstractly in terms of factors, so that the program can reason with them. Each factor represents a stereotypical collection of facts, which tends to strengthen or weaken a conclusion that a side should win a particular kind of legal claim.[33] Each case is represented with a textual summary, called a squib, and has been manually indexed by a list of applicable factors. In the instruction, students work with both the squib and factor versions of the cases. CATO's Factor Hierarchy represents why each factor makes a difference to the legal claim. It relates each factor to one or more legal issues representing the normative concerns of trade secret law through an intermediate layer of increasingly abstract factors.[34]

---

30. *See also* explanatory dialogues discussed in *infra* note 92.

31. Kevin Ashley & Vincent Aleven, *Casebook Chapter on the Law of Trade Secret Misappropriation, Parts I and II* (1996) (on file with the author).

In preparing the casebook chapter on trade secret law, I employed Paul Goldstein's text as a guide. PAUL GOLDSTEIN, COPYRIGHT, PATENT, TRADEMARK AND RELATED STATE DOCTRINES (3d ed., 1990). The casebook presents edited opinions of four cases discussed as major cases or in notes in this textbook.

32. Vincent Aleven, *CATO Workbooks 1-4* (1996) (on file with the author).

33. Factors are based on dimensions, a knowledge representation device invented for HYPO, a program that performed case-based legal reasoning in the domain of trade secret misappropriation law. *See generally* KEVIN ASHLEY, MODELING LEGAL ARGUMENT: REASONING WITH CASES AND HYPOTHETICALS (1990). As compared to factors, each dimension contains additional structure including tests for deciding if a dimension applies to a case or is a near-miss, and a focal slot for specifying the magnitude of a dimension in a case. A dimension's magnitude is not a measure of its weight. Instead, a magnitude indicates how extreme an example of the dimension the case was. HYPO has heuristics based on dimensions for posing meaningful hypothetical variations of target problems to strengthen or weaken the arguments for or against plaintiff's claim. CATO cannot pose hypotheticals.

34. *See generally* Aleven, *supra* note 3, at 44–49.

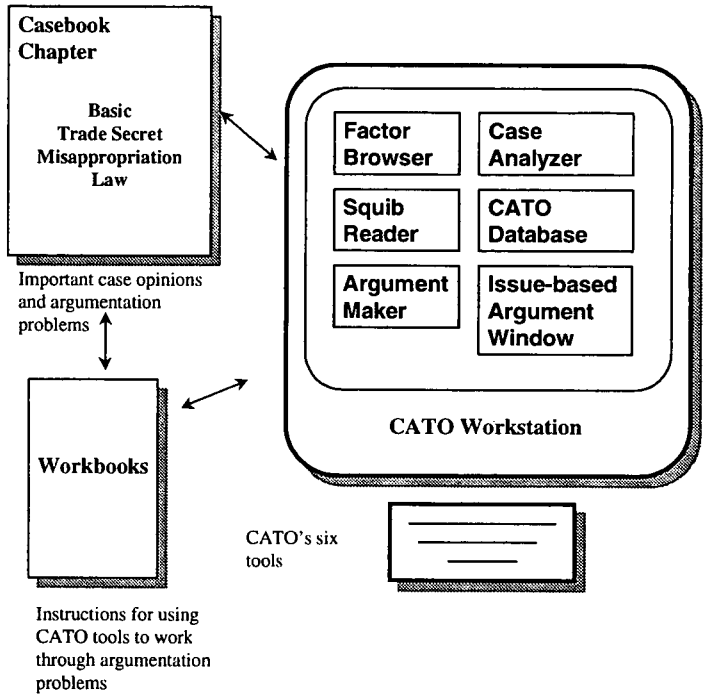

Figure 2. CATO Environment and Tools

CATO's instruction is based on a computational model of case-based legal argument. It draws legal inferences about target disputes by comparing them to source precedents and constructing legal arguments supporting the inferred conclusions based on the comparison. It compares a target problem and source cases in terms of the relative inclusiveness of the sets of factors the source cases share with the target problem.

The model enables CATO to perform seven basic argument moves, shown in Figure 3, with the cases in its database,[35] and to organize complex multi-case, multi-issue arguments by combining these moves. CATO generated all argumentation examples in its workbook and in this article from these basic moves. Associated with each argument move is a recipe or template that enables CATO to construct the move, where feasible, from the target problem and source cases specified by the user. The recipes and templates are defined in terms of the factors shared, or not shared, by the target problem and source cases and various relationships among those factors.

---

35. *See id.* at 19. The dissertation lists eight argument moves. It differentiates two counter-example moves, citing a "more on point" counterexample and an "as on point" counterexample. *See id.* For simplicity, I have consolidated these into one move.

Analogizing Move: analogizing a target problem to a relevant source case with a favorable outcome (i.e., a favorable source case).
Distinguishing Move: distinguishing a target problem from a relevant source case with an unfavorable outcome (i.e., an unfavorable source case).
Downplaying Move: downplaying the significance of a distinction between a target problem and a favorable source case.
Emphasizing Move: emphasizing the significance of a distinction between a target problem and an unfavorable source case.
Conflict-Resolution Move: citing a favorable source case to show that the target problem's factual strengths overcome its weaknesses.
Not Fatal Move: citing a favorable source case to argue that a target problem's factual weaknesses are not fatal.
Counterexample Move: citing a counterexample to an unfavorable source case cited by an opponent.

Figure 3. CATO's Basic Argument Moves

CATO can illustrate all of its argument moves and their related argumentation concepts in terms of examples it assembles from cases in its database. It employs small collections of cases, called Argument Contexts, to illustrate moves and concepts,[36] such as a source case's being on point, more on point than another case, or the most on point of a set of cases. Using these moves, it explains how to construct reasonable arguments supporting decisions about target problems by comparing them to source precedents. To some extent, it can help evaluate which case-based arguments are better than others in a given context. It also characterizes cases in terms of substantive legal issues, such as whether information is a trade secret, whether there is confidential relationship, and whether improper means were employed in obtaining information. Guided by the workbook and employing the six computational tools, students learn the argumentation concepts and how to employ them within CATO's case database.[37]

B. What CATO Teaches

CATO engages students in three types of argumentation and reasoning tasks: learning the factors, testing hypotheses, and making arguments. Students engage in some introductory tasks to help them learn about CATO's trade secret factors and gain experience in using its retrieval tools. They read case squibs and assign

36. Kevin Ashley & Vincent Aleven, *Generating Dialectical Examples Automatically*, 1992 PROC. TENTH NAT'L CONF. ON ARTIFICIAL INTELLIGENCE 654–60.

37. Students can access the casebook and workbooks on the web using a web browser, such as Netscape or Explorer, on a personal computer. They can access CATO's tools on a PC via a local area network connection to CATO running on a Unix Workstation. With advances in JAVA programming, it is feasible to make CATO's tools accessible directly from the web.

factors to them for practice. They also create hypothetical cases which are very strong or weak for a side by loading them with appropriate factors. By retrieving other cases that involve the same factors, students can confirm the factors' effects. In hypothesis-testing, students pose and assess hypotheses about trade secret law against the facts and results of actual cases. In making arguments, students engage in role-playing as an advocate for either plaintiff or defendant in trade secret disputes. They evaluate a target problem by comparing it to past cases and construct an argument based on the most relevant cases. This gives students practice in identifying a problem's strengths and weaknesses in terms of factors and issues, retrieving cases from CATO's database for purposes of making an argument, evaluating retrieved cases by comparing them to the problem, and producing written arguments that marshal the best cases.

1. *Learning Factors*

An example of CATO's factor model based on a real trade secret dispute is *Mason v. Jack Daniel Distillery*, the squib for which is shown in Figure 4. Students first encounter this and other squibs in the Squib Reader window. The format of each case squib is similar to that taught to first-year students in "briefing" cases: title of the case, citation information, date, parties, statement of the facts, issues, and holding.[38]

Deciding which factors apply to a target problem is a manual task for the student to perform. CATO's Case Analyzer tool helps students summarize the facts of the *Mason* case in terms of factors.[39] Students see a listing of factors and click on those they believe are present in the case described. For each factor a student selects, he may solicit CATO's feedback on whether that factor appears in CATO's representation of the case. The student also may find out which additional factors CATO's representation of the case includes.[40]

In this target problem, Mason developed and served a cocktail he dubbed "Lynchburg Lemonade." Because Mason took some measures to protect his recipe's secrecy, and because his was the only tavern producing this drink, factors F6 (Security-Measures) and F15 (Unique-Product) apply; both tend to favor the plaintiff. On the other hand, Mason disclosed his recipe in negotiations with a sales agent of the defendant, Jack Daniel Distillery, which started marketing the cocktail without compensating Mason. Thus, F1 (Disclosure-In-Negotiations) applies, a factor that tends to favor the defendant. The agent was aware, however, that the recipe was a "secret formula" and the distillery's cocktail was identical to Mason's, so F21 (Knew-Info-Confidential) and F18 (Identical-Products) also apply, tending to favor the plaintiff. Finally, because the

38. Unlike a law student's brief of an opinion, a squib does not contain the court's rationale.

39. *See* Aleven, *supra* note 3, at 20.

40. *See* Vincent Aleven, *CATO Workbook 1 and Reference Pages, Section 4* (1996) (on file with the author).

recipe could have been obtained by reverse engineering the cocktail, F16 (Info-Reverse-Engineerable) applies and favors the defendant.

Title: Mason v. Jack Daniel Distillery
Cite: 518 So.2d 130 (Ala. Civ. App. 1987).
Date: Aug. 5, 1987.
Parties: Plaintiff: Mason; Defendant: Distillery.
Claim: Misappropriation of trade secret.
Procedural setting: Trial court denied Distillery motion for directed verdict and jury awarded nominal damages to Mason. Mason appeals court decision barring punitive damages. Distillery cross-appeals denial of directed verdict.
Decision: For plaintiff; remanded.
Facts: In 1980, a restaurant owner named Mason developed a combination of Jack Daniel 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.
Issues: (1) Was there sufficient evidence to allow a jury to determine that the recipe for Lynchburg Lemonade was a trade secret? (2) Was there malice, willfulness, or wanton and reckless disregard of the rights of Mason allowing recovery of punitive damages?
Holding: (1) Yes. For plaintiff. (2) Possibly; remanded for determination by jury.

Figure 4. Case Squib of Mason vs. Jack Daniel Distillery

Assigning factors based on one's reading of a case is a subjective task. The goal has been to design CATO to minimize the consequences of this inevitable subjectivity. There is now a fairly stable list of factors to choose from. Also, the Factor Browser tool provides information specifying conditions under which each factor applies.[41] Although CATO's set of factors for trade secret law is fairly

41. Techniques have also been implemented to minimize the consequences of subjectivity in assigning factors. The representation of a case in the database has been prepared manually by a case enterer and represents that person's best effort to identify the plaintiff's factual strengths and weaknesses in terms of CATO's factors. Because interpreting the facts of a case is to some extent

complete, students may also identify a factual strength or weakness for which they cannot find a corresponding factor. Because the program does not support students' adding to CATO's list of factors, students are encouraged to use the closest corresponding factor they can find.[42]

2. *Testing Hypotheses*

In the hypothesis-testing phase of instruction, students use CATO's factor model to test hypotheses about a problem. Making an argument for the plaintiff in *Mason* is like proposing the hypothesis that "a defendant to whom confidential information was disclosed who knew that the information was confidential, is under an obligation not to use or disclose the information even though there was no written nondisclosure agreement." Beginning law students may not know how to deal with an abstract legal proposition like this. They may be able to articulate an opinion of its correctness in terms of the purpose of trade secret law, but often they are not able to articulate a process by which one could determine, as an empirical matter, whether the hypothesis is correct, or to carry out the legal research that such a process entails.

CATO makes legal hypotheses like this explicit and manipulable. It encourages students to (1) pose such a hypothesis and predict whether it is correct, (2) test the hypothesis against the cases in its database by running an appropriate query, (3) interpret the results of the query, and (4) if necessary, modify the hypothesis in light of the retrieved cases. For instance, a query can test the above hypothesis.[43] The query returns seven pro-plaintiff cases and one pro-defendant case.

Students can interpret their prediction in light of the query's results. In the Squib Reader window, they may read some or all of the cases returned.[44] For

---

subjective, students may select factors not included in CATO's list for a given case. In effect, students may disagree with the case enterer. If a student rejects the Case Analyzer's feedback, the student is encouraged to stand by his representation of a case. In subsequent processing, such as in generating argumentation examples about it, CATO will employ the student's representation of the case, not its own. *See id.*

42. *See id.*

43. The query "f1 f21 % f4" generates a list of all cases with factors F1 (Disclosure-in-Negotiations) (d)) and F21 (Knew-Info-Confidential (p)), but without factor F4 (Agreed-Not-To-Disclose (p)).

44. Students using CATO have ready access to the squibs of cases in the database. There are frequent warnings to students that they should read the squibs and not rely on the factor representation except as a guide.

In the current version of CATO, students do not have ready access to the opinions of cases in the database (unless the opinions appear in the casebook). The goal is to focus beginning students on comparing cases in terms of their facts rather than quoting rationales from case opinions and to apprise students of the processes of testing legal hypotheses and of making arguments. By simplifying the texts of the cases without compromising the essential features of the process, the need to deal with full-length opinions, which may obscure students' views of the overall process, is temporarily controlled. It is important, however, that convenient access to opinions is phased in fairly quickly.

example, a student who predicted that there was a lack of an obligation not to use or disclose the information might be surprised to see few cases won by defendant and many won by plaintiff because the results appear to be inconsistent with the hypothesis. Conversely, a student who had predicted that defendants *did* have an obligation not to use or disclose the information would be interested in examining and distinguishing the few cases won by the defendant. A student who can devise an alternative explanation for the decisions in those cases that contradict the hypothesis may rescue the hypothesis.[45] If not, these cases may force the student to modify or even abandon the hypothesis.[46]

Simplifying the task of legal research through the use of factor-based queries is one way that CATO makes abstract concepts explicit and manipulable. The complexity of legal research using manual indices and even on-line full-text tools like WESTLAW and LEXIS may sometimes mislead students into missing the point of their research. CATO's simplified research tools help students understand that they are undertaking the research to test a general hypothesis about a domain of law. The factor-based queries help them to relate the cases returned directly to the hypothesis they are testing. With a computer projector in class, a legal writing instructor could lead students through several cycles of posing, testing, interpreting, and modifying hypotheses much more easily than with traditional research tools.

3. *Making Arguments*

Posing and testing hypotheses is directly related to making arguments citing cases. As described below, CATO's factor-based queries help students compare retrieved source cases to the target problem. In the *Mason* problem, for instance, all of the seven pro-plaintiff cases are potentially good cases for the plaintiff to cite.

Opinions could readily be accessible via hypertext links to full-text retrieval services such as WESTLAW or LEXIS.

45. *See* Vincent Aleven, *CATO Workbook 1 and Reference Pages,* Section 7 (1996) (on file with the author).

46. *See* Aleven, *supra* note 3, at 114–18. Here is an example of a real law student's revising his hypothesis in light of cases he retrieved with CATO. The student hypothesized that a plaintiff who had disclosed his secrets to outsiders would lose a claim for trade secret misappropriation regardless of the fact that plaintiff took some security measures. He used CATO to test the hypothesis. His query sought all cases with Factors F6 (Security-Measures) and either F10 (Secrets-Disclosed-Outsiders) or F20 (Info-Known-To-Competitors). Upon seeing that the query returned four cases won by plaintiff and only one pro-defendant case, he exclaimed, "That boggles my mind!" When asked how he would resolve this, he said, "I would read the cases and see what's going on." In one of the cases, *Data General Corp. v. Digital Computer Controls Inc.*, 357 A.2d 105 (Del. Ch. 1975), the plaintiff won even though it had disclosed its "confidential" information to five thousand outsiders where the disclosures were subject to restrictions on use of the information (Factor F12, Outsider-Disclosures-Restricted (p)). Upon reading *Data General*, the student decided that he might save his hypothesis by limiting it to disclosures that were not subject to confidentiality restrictions, in other words, specifying that Factor F12 *not* apply.

CATO teaches argumentation through role-playing. It engages students in making arguments on behalf of one side and responding to those arguments on behalf of the other. Students focus particularly on factors relating to important legal issues and on making legal arguments about how to resolve the target problem's conflicting factual strengths and weaknesses. In the *Mason* problem, CATO's Argument Maker focuses students on three main issues including whether plaintiff's recipe is a trade secret. The Factor Hierarchy[47] records the information that factors F6 (Security-Measures (p)), F15 (Unique-Product (p)), F1 (Disclosure-In-Negotiations (d)), and F16 (Info-Reverse-Engineerable (d)), all relate to that issue. Only the pro-plaintiff F6 and F15 help plaintiff on that issue. F1 and F16 are potential weaknesses in the plaintiff's argument.

The Argument Maker teaches students strategies like "find cases to cover weaknesses on an issue." Students try queries like "(or f6 f15) (or f1 f16)" (in other words, "List all cases with either factors F6 (Security-Measures (p)), or F15 (Unique-Product (p)), and either F1(Disclosure-in-Negotiations (d)), or F16 (Info-Reverse-Engineerable (d))"). Such a query could retrieve from the database cases where courts have held that the strengths outweigh the weaknesses. *Boeing Co. v. Sierracin Corp.*[48] is one case that satisfies the above query. Figure 5 shows how it compares in terms of factors to the *Mason* fact situation. The comparison lists each case's factors and annotates them.

In the CATO model, analogizing a target problem and a source case involves making an argument that the target problem should be decided as the source case was, based on the relevant similarities between the two. Relevant similarities are defined as the set of factors that the target problem and source case have in common.[49] These similarities give rise to reasons for deciding the two cases in the same way.

Mason	*Boeing* (p)
= F1 Disclosure-In-Negotiations (d)	= F1 Disclosure-In-Negotiations (d)
= F6 Security-Measures (p)	* F4 Agreed-Not-To-Disclose (p)
F15 Unique-Product (p)	= F6 Security-Measures (p)
* F16 Info-Reverse-Engineerable (d)	F10 Secrets-Disclosed-Outsiders (d)
F18 Identical-Products (p)	* F12 Outsider-Disclosures-Restricted (p)
= F21 Knew-Info-Confidential (p)	* F14 Restricted-Materials-Used (p)
	= F21 Knew-Info-Confidential (p)
= shared factor * distinction	

Figure 5. Factor Comparison of Two Cases

Distinguishing a case involves making an argument that the target problem should *not* be decided as the source case was decided because of relevant

47. *See infra* Figure 10.

48. 738 P.2d 665 (Wash. 1987).

49. *See generally* Aleven, *supra* note 3, at 19–25, 58–61; *see also* ASHLEY, *supra* note 33, at 29–34.

differences. Relevant differences are certain factors that the cases do not share—those in the target problem that favor the distinguisher's side and those in the source case favoring its winner. The former strengthen the distinguisher's position in the target problem in ways not present in the source case. The latter strengthen the position of the source case winner in ways not present in the target.[50]

[1] Plaintiff's argument [by analogy] citing *Boeing*:
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 *The Boeing Co. v. Sierracin Corp.*, 108 Wash.2d 38, 738 P.2d 665 (1987).

[2] Defendant's response distinguishing *Boeing*:
Boeing is distinguishable, because in *Boeing*, defendant used materials that were subject to confidentiality restrictions [F14], plaintiff's disclosures to outsiders were subject to confidentiality restrictions [F12], and defendant entered into a nondisclosure agreement with plaintiff [F4]. This was not so in *Mason*. Also, in *Mason*, plaintiff's product information could be learned by reverse-engineering [F16]. This was not so in *Boeing*.

[3] Plaintiff's argument downplaying a distinction:
In *Boeing*, defendant used materials that were subject to confidentiality restrictions [F14]. This was not so in *Mason*. This however is not a major distinction. 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].

[4] Defendant's argument emphasizing the distinction: In *Boeing*, defendant used materials that were subject to confidentiality restrictions [F14]. This was not so in *Mason*. This is a marked distinction. It shows that in *Boeing*, defendant may have acquired plaintiff's information through improper means [F120]. This is supported by other facts in *Boeing* as well. For example, plaintiff imposed confidentiality restrictions in connection with its disclosures to outsiders [F12]. In *Mason*, on the other hand, defendant obtained or could have obtained its information by legitimate means [F120]. Plaintiff's information could be discovered by reverse engineering plaintiff's product [F16].

Figure 6. Examples of Four Argument Moves

50. *See generally* sources cited *supra* note 49.

C. Teaching Argumentation by Example

CATO's Argument Maker tool shows examples of how to use cases in argument moves and how to combine the moves into more complex arguments organized by legal issues.

1. *Argument Moves*

For instance, the Argument Maker can illustrate how to use *Boeing* in four of the Argument Moves of Figure 3 to argue whether Mason's recipe is a trade secret. The four argument moves are illustrated in Figure 6: (1) an argument that plaintiff should win analogizing *Mason* to *Boeing* (an argument that also resolves one of the weaknesses in favor of plaintiff), (2) defendant's argument distinguishing *Boeing*, (3) plaintiff's argument downplaying one of the distinctions, and (4) defendant's argument emphasizing the importance of that distinction. Part III illustrates how CATO incorporates these argument moves into realistic argumentation dialogues.

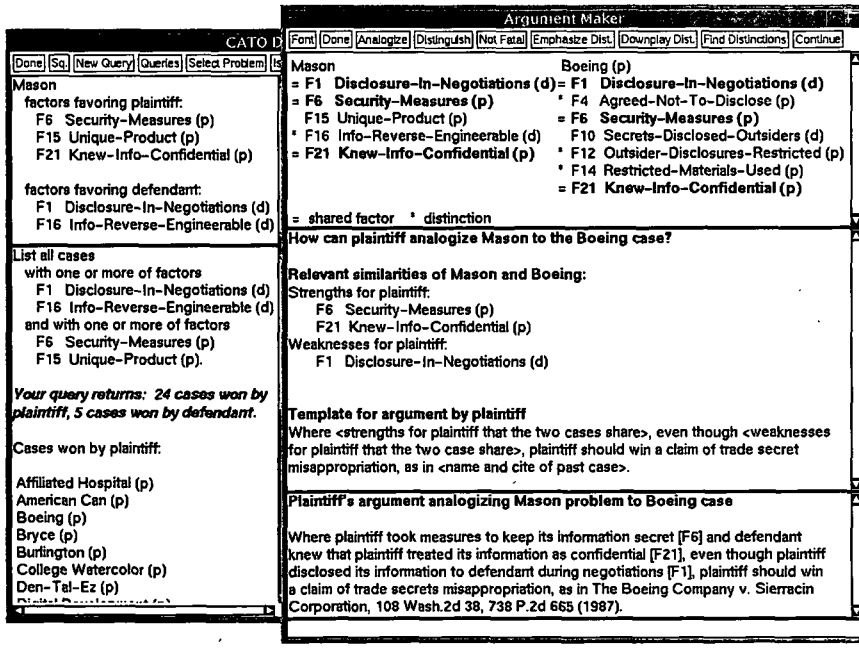

Figure 7. CATO's Argument Maker Window

As suggested in Figure 7,[51] each of the four sample argument moves appears in an Argument Maker window. The five buttons across the top correspond to

51. *See* Aleven, *supra* note 3, at 124.

five basic argument moves: Analogize, Distinguish, Not Fatal, Emphasize Distinction, and Downplay Distinction. Having selected a target problem and a source case from CATO's database, by pressing a button, a student may see how CATO would use that case in the corresponding argument move concerning the problem.

In presenting a sample argument move, the Argument Maker employs three vertically stacked window panes, shown on the right in Figure 7. The top pane presents shared and unshared factors. The middle pane provides a general template or recipe for generating the argument move. The bottom pane demonstrates the actual argument created by filling in the template or following the recipe using the top pane's factor comparison information concerning the target problem and source case.

The templates, or recipes, for generating the four argument moves are shown in Figure 8. The Argument Maker uses these templates to explain how the corresponding arguments[52] were generated from the comparison of the target problem and source case.[53]

Downplaying and emphasizing moves recharacterize distinctions more abstractly to downplay or emphasize their legal significance.[54] Downplaying suggests that the extra factual strengths, which the distinguisher has focused on (in [c] of the distinguishing template), or the extra factual weaknesses (in [d]), are not very important in the target problem. Emphasizing argues that these distinctions are significant.[55]

2. *Multi-case, Issue-based Arguments*

CATO also helps students construct multi-case, issue-based arguments on behalf of either side in a trade secret misappropriation claim. Given a problem like *Mason*, students first select the important legal and factual issues for the plaintiff and defendant. For comparison, the Issues tool shows them those issues CATO deems important and an outline of how CATO organizes an argument for the plaintiff. The outline relates the plaintiff's factual strengths and weaknesses to the legal issues and provides notes specifying the types of cases needed for support. Students are then asked to "outline the strongest overall argument on behalf of the plaintiff in the *Mason* problem" using the best cases they have encountered so far. After the students have finished their arguments, they can compare them with the argument CATO makes using the same set of cases the student used.

---

52. *See supra* Figure 6.

53. *See supra* Figure 5.

54. As discussed *infra* Part II.D.2, the distinctions are recharacterized in terms of more abstract factors in the Factor Hierarchy.

55. Part II.D.2 discusses these moves in more detail.

Template for Analogizing/Conflict-Resolution Move:
[a] Where
<Insert strengths for plaintiff/defendant shared by target problem and source case>,
[b] even though
<Insert weaknesses for plaintiff/defendant shared by target problem and source case>,
<Insert plaintiff/defendant> should win a claim of trade secret misappropriation, as in
<Insert name and cite of source case>.

Template for Distinguishing Move:
<Insert name of the source case> is distinguishable. It is stronger for <Insert plaintiff/defendant> than is the current problem.
[c] In the <Insert name of the source case>,
<Insert extra strengths for plaintiff/defendant in the source case>.
This was not so in the current problem.
[d] Also, in the current problem,
<Insert extra weaknesses for plaintiff/defendant in the target problem>.
This was not so in <Insert name of the source case>.

Recipe for Downplaying Move:
1. Recite reasons why distinction *d* might matter.
2. Recite factors in target problem which contrast with *d* (i.e., undercut *d*; they matter for exactly the opposite reason).
3. Recite factors in the source case which are similar to *d* (i.e., matter for the same reason).

Recipe for Emphasizing Move:
1. Recite reasons why distinction *d* matters which are not contradicted in the target problem and not present in the source case.
2. Recite factors in target problem which are similar to *d* (i.e., matter for the same reason).
3. Recite factors in the source case which contrast with *d* and which are not present in the target problem.

Figure 8. Templates and Recipes for Four Argument Moves

CATO's argument on behalf of plaintiff in *Mason* is shown in Figure 9. It employs three cases students already will have encountered in the workbooks: *USM* (p),[56] *Forest Laboratories* (p),[57] *Television* (p),[58] and *Boeing* (p). CATO composes its argument with multiple applications of the basic argument moves. For instance, it employs *Boeing* in three argument moves: Analogizing, Conflict-Resolution, and Not Fatal moves. The first two of these are very much like the analogizing move shown in Figure 6. CATO generates it by plugging *Mason* and

56. USM Corp. v. Marson Fastener Corp., 393 N.E.2d 895 (Mass. 1979).
57. Forest Lab., Inc. v. Formulations, Inc., 299 F. Supp. 202 (E.D. Wis. 1969), *rev'd in part*, 452 F.2d 621 (7th Cir. 1971).
58. Television Telecomm. Sys., Inc. v. Saindon, 522 N.E.2d 1359 (Ill. App.2d 1988).

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 (First) of Torts § 757, and comment b, factors 1-6 (1939). In the current fact situation, plaintiff adopted security measures [F6] and plaintiff's product was different from products made by competitors [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 *Televation Telecommunication Systems, Inc. v. Saindon*, 522 N.E.2d 1359 (Ill. App. 2d Dist. 1988), *USM Corp. v. Marson Fastener Corp.*, 379 Mass. 90, 393 N.E.2d 895 (1979), *Forest Laboratories, Inc. v. Formulations, Inc.*, 299 F. Supp. 202 (E.D. Wis. 1969), and *The Boeing Co. v. Sierracin Corp.*, 108 Wash.2d 38, 738 P.2d 665 (1987), there was similar evidence that plaintiff's information is a trade secret [F101], and plaintiff won. In *Televation*, *USM*, and *Forest Laboratories*, as in the current problem, plaintiff took security measures to protect the information [F6] and plaintiff's product was unique on the market [F15], and plaintiff won. *Boeing* held for plaintiff, where, as in the current case, plaintiff took measures to keep its information secret [F6]. [Could flesh out these analogies some more.]

The fact that plaintiff disclosed its information to defendant during negotiations [F1] does not preclude a conclusion that plaintiff's information is a trade secret [F101], especially where plaintiff took measures to keep its information secret [F6]. *Forest Laboratories*, *Boeing*.

The fact that plaintiff's information could be discovered by reverse engineering plaintiff's product [F16] does not preclude a conclusion that plaintiff's information is a trade secret [F101]. *Televation*.

A confidential relationship existed between plaintiff and defendant [F114]

A confidential relationship existed between plaintiff and defendant [F114]. Restatement (First) of Torts § 757 (1939). In the current fact situation, defendant knew that plaintiff intended its information to be kept confidential [F21]. This shows that defendant was on notice that using or disclosing the information would be a breach of confidentiality [F115]. In *Televation*, *Forest Laboratories*, and *Boeing*, which held for plaintiff, there was similar evidence that a confidential relationship existed between plaintiff and defendant [F114]. *Televation*, *Forest Laboratories*, and *Boeing* held for plaintiff, where, as in the current problem, defendant knew that plaintiff's information was confidential [F21]. [Could flesh out these analogies some more.]

The fact that plaintiff conveyed its information to defendant in the course of negotiations [F1] does not preclude a conclusion that a confidential relationship existed between plaintiff and defendant [F114]. *Boeing*, *Forest Laboratories*.

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

Defendant acquired plaintiff's information through improper means [F110]. Restatement (First) of Torts § 757 (1939).

The fact that plaintiff conveyed its information to defendant in the course of negotiations [F1] does not preclude a conclusion that defendant acquired plaintiff's information through improper means [F110]. *Boeing*, *Forest Laboratories*. [Could flesh out this analogy some more.]

Figure 9. A Multi-case, Issue-Based Argument By CATO

Boeing into its analogizing/conflict-resolution move template,[59] but here it is done in the context of the larger argument. CATO inserts some notes to the student, such as "[Could flesh out these analogies some more]." The notes are intended to prompt students to read the squibs for the cited source cases and incorporate more textual support from them to elaborate on the analogy with the target problem.

Organized around three legal issues, the argument claims that plaintiff's information is a trade secret, a confidential relationship existed between plaintiff and defendant, and the defendant acquired plaintiff's information through improper means. It relates the plaintiff's factual strengths and weaknesses to each of these issues using the cases as support. Part II.D.2 explains how the Factor Hierarchy helps construct such arguments. The "F" numbers refer to the numbers of the applicable factors and abstract factors in the Factor Hierarchy. Briefly, CATO's algorithm employs information in the Factor Hierarchy to characterize a relevant similarity more abstractly for purposes of relating it to the appropriate issue. For instance, as already noted, *Boeing* shares the following similarity with *Mason*: factor F21 (Knew-Info-Confidential). CATO characterizes the significance of *Boeing* and this relevant similarity more abstractly in terms of: (1) the higher level legal issues to which it relates (F114, Confidential-Relationship); and (2) intermediate abstract factors (i.e., F115, Notice-Of-Confidentiality). As shown in Figure 11, the Factor Hierarchy relates F21 to F114 through F115.

In the final exercises of the CATO curriculum, students outline a response on behalf of the defendant distillery to the issue-based argument for plaintiff generated by CATO or by another student. The responses make an initial argument relating defendant's issues, factual strengths, weaknesses, and supportive cases. They also respond to the cases cited by the plaintiff, either by distinguishing them or citing counterexamples. After producing their responses, students use the Argument Maker window to compare CATO's argument for the defendant.

D. How CATO Works

The templates and recipes of the previous section provide a window into how CATO works. The simplest explanation is this: CATO "knows" when and how to fill out the argumentation templates and recipes. Its argumentation algorithms interact with its database of cases indexed by factors and the Factor Hierarchy to fill out the templates and recipes appropriately. Its algorithms enable it to determine whether an argumentation template or recipe is appropriate and what information to plug into the template's slots or to select in response to the recipe's directions. It can select this information appropriately by virtue of its coordinated knowledge representation.[60]

---

59. *See supra* Figure 8.

60. For a full explanation, see Aleven, *supra* note 3, at 41–148, 248–52; Vincent Aleven & Kevin Ashley, *Teaching Case-Based Argumentation Through a Model and Examples: Empirical*

CATO employs interrelated techniques for representing knowledge. CATO's argumentation templates, recipes, and algorithms interact with its set of factors, database of indexed cases, and the Factor Hierarchy. The interaction captures selected aspects of a complex reality faithfully enough to enable a computer program to manipulate its representations in a meaningful way.

1. *The Factor Hierarchy*

The Factor Hierarchy is central. As noted above, CATO's Factor Hierarchy represents the underlying meaning of factors, that is, why a factor's stereotypical collection of facts makes a difference to the legal claim. More specifically, the Factor Hierarchy represents the connections between factors and those legal issues to which they are relevant and which give them significance.[61] As shown in figures 10 through 12, the Factor Hierarchy is a collection of graphs. Each graph is an assemblage of nodes connected by links. A given child node may have more than one parent node. The root nodes, located at the top of each graph, represent the main legal issues in a claim of trade secret misappropriation, set forth in Table 1.[62] The leaf nodes, found at the bottom of the graphs, are the factors.

Layers of intermediate nodes link the leaf nodes, or issues, to the root nodes, or factors. As one moves upward from a leaf node, each intermediate node represents an increasingly abstract characterization of the legal significance of the factors below that node. This leads ultimately to the legal issue nodes to which those factors are relevant.[63] The issues and the abstract factor characterizations will all be referred to as "abstract factors." The Factor Hierarchy contains 26 leaf node factors and 16 abstract factors, five of which are top-level legal issues.[64]

Evaluation of an Intelligent Learning Environment, in ARTIFICIAL INTELLIGENCE IN EDUCATION, 1997 PROC. WORLD CONF. ON ARTIFICIAL INTELLIGENCE IN EDUC. 87 (August. B. du Boulay & R. Mizoguchi eds.); Kevin Ashley & Vincent Aleven, *Reasoning Symbolically About Partially Matched Cases*, 1997 PROC. FIFTEENTH INT'L JOINT CONF. ON ARTIFICIAL INTELLIGENCE 335 (Martha Pollack ed.); Vincent Aleven & Kevin Ashley, *How Different Is Different? Arguing About the Significance of Similarities and Differences*, 1996 PROC. THIRD EUROPEAN WORKSHOP ON CASE-BASED REASONING 1 (I. Smith & B. Faltings eds.).

61. *See generally* Aleven, *supra* note 3, at 44–49.

62. Aleven, *supra* note 3, at 239.

63. *See id.* at 44–49.

64. For ease of reference, the leaf nodes have numbers below 30; abstract factors have numbers above 100. *See id.*

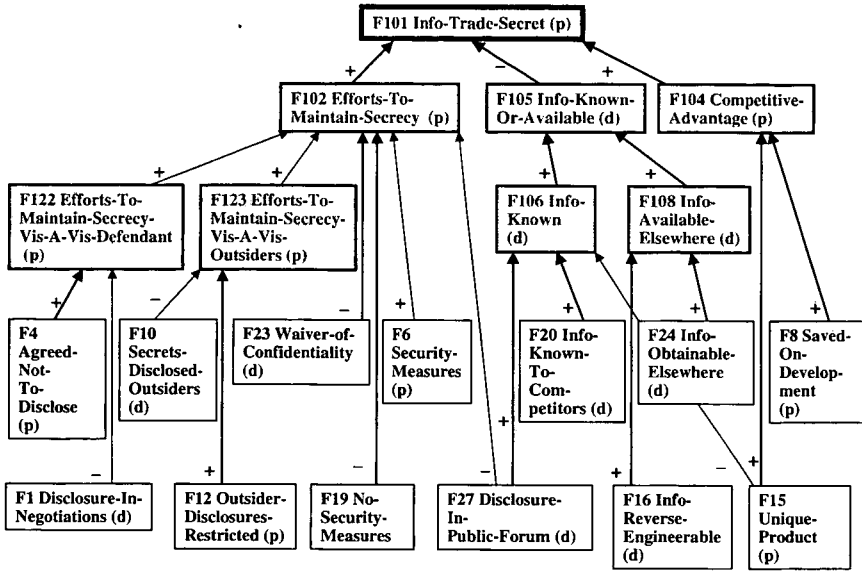

Figure 10. CATO's Factor Hierarchy for Trade Secret Misappropriation Issue: Whether Plaintiff's Information Was a Trade Secret

F Number	Short Title (side favored)	Meaning
F101	Info-Trade-Secret (p)	Plaintiff's information is a trade secret.
F114	Confidential-Relationship (p)	A confidential relationship existed between plaintiff and defendant.
F110	Improper-Means-Conclusion (p)	Defendant acquired plaintiff's information through improper means.
F112	Info-Used (p)	Defendant may have used plaintiff's information and usurped a competitive advantage.
F124	Defendant-Ownership-Rights (d)	Defendant may have ownership rights in the information.

Table 1. Legal Issues in Top-Level of Factor Hierarchy

Each abstract factor represents two possible, opposing conclusions about its legal effect in a problem, one favoring the plaintiff and the other favoring the defendant.[65] The name of the abstract factor refers to one of the two conclusions and the side it favors, plaintiff (p) or defendant (d). For example, factor F115 (Notice-Of-Confidentiality) in Figure 11, represents two possible conclusions:

65. *See id.*

"Defendant was on notice that using or disclosing the information would be a breach of confidentiality" or "Defendant was not on notice that using or disclosing the information would be a breach of confidentiality." Each factor or abstract factor is linked to at least one parent abstract factor and supports one or other of the conclusions associated with the parent. A child node's pro-plaintiff factor or abstract factor supports the pro-plaintiff conclusion in the parent node and in all of its ancestor nodes. A pro-defendant factor supports the pro-defendant conclusion in the parent node and all of its ancestor nodes. In Figure 11, for example, all of the factors and abstract factors linked to F115 provide evidence for or against the conclusions associated with F115. This provides evidence concerning the more abstract issue of whether a confidential relationship existed between plaintiff and defendant, F114.[66]

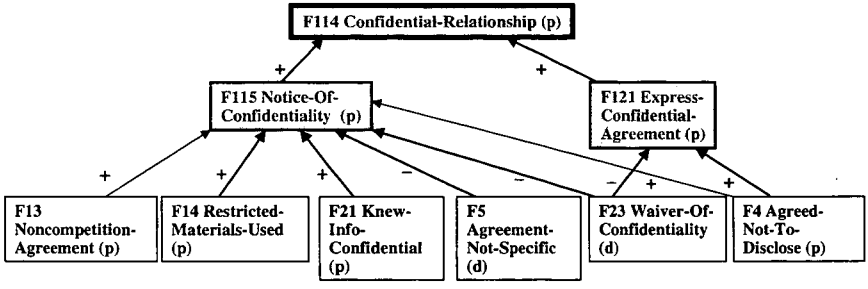

Figure 11. CATO's Factor Hierarchy for Trade Secret Misappropriation Issue: Whether a Confidential Relationship Existed Between Plaintiff and Defendant

A particular factor may relate to more than one abstract factor (including more than one issue). This is represented as a factor's having more than one parent, including a parent in a different issue's Factor Hierarchy. For instance, factor F14 (Restricted-Materials-Used (p)) is related not only to F115 (Notice-Of-Confidentiality (p)) in Figure 11, but also to F111 (Questionable-Means) in Figure 12. This is the Factor Hierarchy for the issue whether defendant acquired plaintiff's information through improper means (F110) or, alternatively, whether the information was legitimately obtained or obtainable (F120). CATO makes use of this property in generating alternative arguments emphasizing or downplaying the significance of a distinction.

66. *See id.*

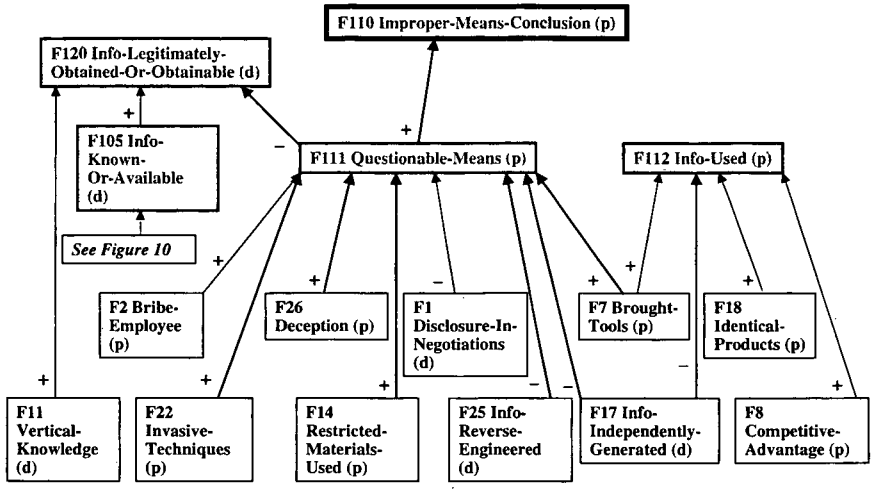

Figure 12. CATO's Factor Hierarchy for Trade Secret Misappropriation Issue: Whether Defendant Obtained Plaintiff's Information Using Improper Means

The factors, issues, and other abstract factors have been gleaned from a variety of sources including § 757 of the *Restatement (First) of Torts*,[67] which many jurisdictions have adopted as an authoritative statement of the law of trade secrets. The factors underneath F101 (Info-Trade-Secret (p)) in the Factor Hierarchy, Figure 10, elaborate the six important factors listed in the *Restatement*'s definition of a trade secret.[68] Other factors come from the opinions, treatises, and law review articles.[69] Secondary sources tend to group cases together that illustrate the effect of particular factual strengths and weaknesses or illustrate counterexamples, where a court reaches a result in spite of the strengths or weaknesses.[70]

---

67. RESTATEMENT (FIRST) OF TORTS § 757 (1939) was not included in the RESTATEMENT (SECOND) OF TORTS. Many courts continue to quote it approvingly, however, including the language of comment b. More recently, the RESTATEMENT (THIRD) ON UNFAIR COMPETITION §§ 39–45 (1993) provide an authoritative restatement of trade secret misappropriation law. Those sections appear to be consistent with § 757.

68. RESTATEMENT (FIRST) OF TORTS § 757 cmt. b (1939).

69. *See, e.g.*, 12 ROGER M. MILGRIM, MILGRIM ON TRADE SECRETS (1995).

70. Guided by § 757's statement of General Principle, Vincent Aleven organized the factors into the Factor Hierarchy in a process of trial and error. This Factor Hierarchy is not the only reasonable interpretation of the important issues, concepts, and fact-patterns in trade secret law.

2. *Factor Hierarchy's Four Roles*

The Factor Hierarchy has four roles. It helps to identify issues in a case or problem, explain the significance of factors, organize an argument by issues, and reason in a context-sensitive way about similarities and differences.

CATO's multi-case, issue-based argument in Figure 9 illustrates the first three roles. Following a fairly complex argument-generation recipe, for each top-level legal issue, CATO has determined from the Factor Hierarchy which factors in the *Mason* problem are related to the issue, and it has found abstract factors in the Factor Hierarchy to characterize the significance of the factors that strengthen the plaintiff's position on the legal issue.[71] It has selected from among the four specified input cases those relevant to the issue (i.e., those whose factors, shared with the target, are related to that issue). It has determined from the Factor Hierarchy if some of the strengthening factors compensate for some of the weakening factors with respect to the issue (i.e., if they share an intermediate legal concern). It has checked which relevant cases can be used to justify a conclusion that plaintiff should prevail on that issue using the argument moves Analogizing, Conflict-Resolution, or Not Fatal. It has checked, when arguing for defendant, for counterexamples among the cases and cites any it finds.[72] Finally, it has composed the selected materials into a textual argument that plaintiff should prevail.

As shown in Figure 6, items [3] and [4], the Factor Hierarchy's fourth role, which alternatively characterizes the importance of a distinction, is illustrated in CATO's arguments by downplaying and emphasizing a distinction between *Mason* and *Boeing*. The arguments show the numbers of the factors and abstract factors employed to recharacterize the target and source case. CATO uses the Factor Hierarchy to categorize the cases more abstractly in terms of the higher level abstract factors or legal issues.

The idea underlying the recipe for downplaying a distinction *d* is to argue that, at a more abstract level of description, the target problem and source case are alike. One way to do this is by pointing out undercutting factors in the target problem. These are factors in the target that tend to contradict the reason why *d* matters. They permit one to argue that the distinguishing factor is not a concern in the target problem. Another downplaying strategy is to show similar factors in the source case (i.e., factors which matter for the same reason as *d*) from which it may be inferred that the presence of *d* does not make the cases significantly different.[73] Figure 6 shows how plaintiff downplays the distinction that in *Boeing*, defendant used materials that were subject to confidentiality restrictions F14.

---

71. These abstract factors will sometimes be referred to as "intermediate legal concerns." In Figure 9, the references to abstract factors such as F114 and F115 in the section dealing with "[a] confidential relationship . . ." mark where CATO has inserted into the argument information about intermediate legal concerns from the Factor Hierarchy.

72. When arguing for plaintiff, CATO does not check for counterexamples.

73. *See* Aleven, *supra* note 3, at 67–69.

From the plaintiff's viewpoint, CATO characterizes *Mason* and *Boeing* broadly, as both involving defendants on notice that using or disclosing the information would be a breach of confidentiality. Here, it applies abstract factor F115, which can be seen in the Factor Hierarchy in Figure 11.

Conversely, the emphasizing recipe argues that the two cases are very different. The student must give a reason why the distinction matters in the target problem, preferably a reason that does not apply in the source case.[74] If possible, the student also should point to other facts in the target case that matter for the same reason.[75] In Figure 6, the resulting argument suggests that *Mason* is a very different kind of case from *Boeing*. Responding for the defendant and applying F120, CATO recharacterizes *Boeing* as a case where the defendant used improper means. CATO maintains that in *Mason*, on the other hand, the defendant could have obtained its information by legitimate means.

In reinterpreting the cases and the significance of factor F14 (Restricted-Materials-Used), which is present in *Boeing* but not *Mason*, CATO follows the alternate paths upward from F14 in Figure 11 and Figure 12 to different abstract factors. CATO's algorithms for downplaying and emphasizing guide it as to which paths to follow in the Factor Hierarchy, how high to climb, and how to group other applicable case factors in an argument supporting the interpretation of the case.

3. *Argument Contexts*

While the argumentation examples given here have focused on a handful of cases, CATO can automatically generate any of the argument moves and issue-based arguments with any subset of cases from its database. This makes it easy for students to compare their arguments to CATO's for the same subset. Thus, CATO is extraordinarily flexible in retrieving small collections of cases that are related to each other in pedagogically valuable ways.

CATO's computational model of legal argument provides working definitions of important concepts. The program implements these concepts computationally. It can retrieve collections of cases that satisfy queries expressed in terms of the concepts. There are working definitions of argumentation concepts including "on point," "as on point," "more on point," "most on point," "best case to cite," and "trumping counterexample." These argumentation concepts are defined in terms of more basic concepts like "relevant similarity," "relevant difference," "citable," and "conflict resolution case." These definitions, in turn, are based on certain primitive concepts including "case," "factor," "side," "outcome," "applicable factor," and "favors."[76]

---

74. The distinction may not apply because of the presence of contrasting factors in the source case or, at least, the absence of similar factors.

75. *See id.* at 70–72.

76. Many of these concepts were defined originally in ASHLEY, *supra* note 33. Pursuant to an idea of Vincent Aleven, the concepts were redefined in terms of a knowledge representation language

Defining concepts and computationally implementing their definitions are feasible because CATO has been implemented with a specific kind of knowledge representation system, Loom,[77] a "structured inheritance" knowledge representation system (or "KL-ONE-style" system), that supports deductive queries. In this kind of system, one represents knowledge by providing definitions for concepts and relations and by specifying facts about individual instances of the concepts. One may then submit queries to the knowledge base in a form similar to a first-order logical formula, but one that includes variables. In a process of exhaustive search, Loom attempts to satisfy the query with the contents of the knowledge base. It tries to find particular instances, or sets of instances, that when substituted for the variables, result in a valid formula. In doing so, it retrieves the logical consequences of all of the knowledge base's facts and definitions.[78]

```
(defrelation more-on-point
    (satisfies (?c1 ?c2 ?cfs)  ; case c1 is more on point than case c2
                               relative to the problem situation cfs
    (:and
        (Case ?c1) (Case ?c2) (Case ?cfs)
        (:for-all ?f           ; if c1 has all the factors that c2 shares with the cfs
            (:implies
                (:and (Factor ?f)
                    (shared-factor ?c2 ?cfs ?f))
                    (applicable-factor ?c1 ?f)))
        (:for-some ?f          ; and shares at least one additional factor with the
                               cfs
                (:and (Factor ?f)
                    (shared-factor ?c1 ?cfs ?f)
                    (:not (applicable-factor ?c2 ?f)))))))
```

Figure 13. Implemented Definition of "More On Point"

An example of a working definition of the concept "more on point"[79] is shown in Figure 13. Focusing on the comments gives a sense of how the concept definition works.[80] In the example, Loom compares two cases in terms of the relative inclusiveness of the set of cases that each shares with the problem situation.

---

that would support deductive retrieval. *See* Kevin Ashley & Vincent Aleven, *A Logical Representation for Relevance Criteria*, *in* FIRST EUROPEAN WORKSHOP: TOPICS IN CASE-BASED REASONING 342 (Stefan Wess et al. eds., 1993).

77. *See* R. MacGregor, *The Evolving Technology of Classification-Based Knowledge Representation Systems*, *in* PRINCIPLES OF SEMANTIC NETWORKS: EXPLORATIONS IN THE REPRESENTATION OF KNOWLEDGE 385 (John F. Sowa ed., 1991).

78. *See* Ashley & Aleven, *supra* note 36, at 658.

79. *See* Ashley & Aleven, *supra* note 76, at 343.

80. The notation "?c1" specifies a variable to be filled in with, that is, "bound to," a source case. "?cfs" is bound to a case serving as the target problem. "?f" will be bound to a factor.

Once defined, such concepts can be used in defining other concepts or in queries to retrieve cases that satisfy the concepts. The concept of a "most-on-point" case, for instance, can be defined very succinctly once "more-on-point" is defined. A source case is most on point relative to a target problem if, no other case is more on point. Loom "knows" how to apply such working definitions when searching for cases that satisfy it.

The ability to specify variables to stand in for cases or factors in queries permits one to design very abstract queries for retrieving a set of cases related to each other in specified ways without ever needing to specify further any of the cases or factors. In satisfying a query, Loom will try *all* combinations of a set of cases to see if they satisfy the constraints. For instance, once "more-on-point" is defined, one can query the database for all of the cases *x* that are more on point with respect to the *Mason* problem than *Boeing*,[81] or all of the cases *x* that are more on point with respect to some case *y* than *Boeing*,[82]or all of the cases *x* that are more on point with respect to some case *y* than some other case *z*.[83]

With this retrieval power, there are a number of useful query concepts for retrieving small collections of cases that are related to one another in interesting ways. For instance, one can retrieve cases for illustrating the strategy of "covering the bases." The aim of "covering the bases" is to find a small set of cases that together cover all of a side's weaknesses in a target problem. Given a target problem, cover-the-bases returns all sets of two cases that cover, or resolve, all of the opponent's strengths in the target problem.[84] For instance, as shown in Figure 9, two cases that cover the bases for the plaintiff in the *Mason* problem are *Boeing* and *Televation*.[85] Such pairs of conflict-resolution cases can be useful in teaching students to seek a parsimonious way of covering a side's weaknesses, one in which each case takes minimal advantage of an advocate's strengths while the cases as a pair effectively cover the opponent's strengths.

Casebook authors, students, and practitioners would find many uses for more complex queries like these. Even a professor already familiar with the opinions of one or two dozen cases in an area would find it a mental feat to retrieve such collections of cases. Unlike a human, the computer has no difficulty "remembering" which cases had which features and methodically trying out many combinations to see which sets of cases satisfy the constraints.[86]

---

81. Here, the search command is (more-on-point ?x boeing mason).

82. Again, the command is (more-on-point ?x boeing ?y).

83. The command is (more-on-point ?x ?z ?y).

84. *See* Ashley & Aleven, *supra* note 36, at 657, 659.

85. This is one of the pairs of cases returned by the following query: (cover-the-bases mason ?c2 ?c3).

86. In the CATO instruction and in the Database tool, we deliberately have provided only a simple query language for students to use. It allows Boolean combinations of factors with "and," "or," and "not." Although more complex logical expressions, like those in the concept definitions above, are probably beyond the capabilities of most law students, more advanced users could be given the power to define more complex queries.

Given CATO's ability to retrieve complex examples and its Factor Hierarchy, an important next step in improving its retrieval capabilities will be to combine the two. CATO should support retrieving cases, not just in terms of their applicable factors, but also in terms of their applicable issues. It should be possible to search for cases that have combinations of abstract factors, for instance, cases where plaintiffs were weak on protecting the security of a trade secret but defendants employed deceptive means.

4. *Generating Texts*

CATO can generate natural language text in the form of arguments and dialogues. Generating English text under constraints is much easier than understanding such texts. Sophisticated natural language generation requires a program to plan what it is going to say, but CATO finesses the need for text planning by employing textual templates. For each argument move, multi-case, issue-based argument, and dialogue type, a textual template provides a general framework for constructing the text. The textual template for the Emphasizing Move, for instance, indicates where to insert the textual descriptions of the factor to be emphasized and where to insert abstract factors to recharacterize the significance of that distinction. The program fills out the template by retrieving snippets of "canned" text associated with each of the factors and abstract factors, but CATO's use of canned text sometimes produces a stilted or repetitive effect.[87]

III. AN ELECTRONIC CASEBOOK THAT TALKS BACK

An interactive casebook could engage readers in analytical dialogues. In teaching the argumentation concept "distinguishing," CATO anchors the lesson in specific role-playing exercises involving concrete case examples and engages students in argumentation dialogues[88] to produce a deeper understanding of the concept of distinguishing cases. A controlled study has shown that first-year students found distinguishing relevant from irrelevant facts the most difficult of the skills tested. Additionally, the study found that instructing novice legal readers in basic rehearsal routines for reading and comprehending cases improved some legal reasoning skills.[89]

First-year students often do not realize that not all differences between two cases amount to distinctions. Only those unshared factors that favor the

---

87. To make CATO's textual arguments more natural, more sophisticated text planning techniques will be required. *See* Johanna D. Moore, *Discourse Generation for Instructional Applications: Making Computer-Based Tutors More Like Humans*, 7 J. ARTIFICIAL INTELLIGENCE EDUC. 181 (1996).

88. *See supra* Figure 1. While we take some poetic license for purposes of pedagogy, a courtroom dialogue like this might occur if a trial judge entertained oral arguments on defendant's motion to dismiss plaintiff's claim and motion for summary judgement. Such an argument might also occur on appeal of a judgment on the merits.

89. *See* Stratman, *supra* note 14, at 214, *citing* M. A. Lundeberg, *Metacognitive Aspects of Reading Comprehension: Understanding in Legal Case Analysis*, 22 READING RES. Q. 407 (1987).

distinguishing side's position or that weaken the opponent's position are true distinctions. Yet, students sometimes do not pay sufficient attention to which side a difference favors. They may even point to differences that hurt, rather than help, their side's argument.[90]

Although CATO does not have the capability to process free-form arguments, it has a rich, but limited, argumentation model that enables the reader to participate in fairly sophisticated dialogues by selecting argument moves and values.[91] Arguments in CATO are composed of citations and factors and a characterization of a factor's significance in terms of the underlying legal issues (as represented in CATO's Factor Hierarchy). As a result, the reader can select from menus the components of an argument that CATO then restates in text.[92]

Argumentation dialogues take place in the context of the Argument Maker tool, shown in Figure 7. Students select an argument move at the top of the screeen. In the window pane's comparison of the target and source cases, students click on factors. The distinguishing move focuses on the factors students select. In Step 5 of Figure 1, for instance, the student chooses to distinguish *Boeing* based on factors F15, F16, and F18. CATO translates that into Perry Mason's argument for Defendant in Step 5. CATO responds to that argument on behalf of Mr. Burger in Step 7, who exploits the opening the student provides. Factors F15 and F18 are differences, but they help the plaintiff, not the defendant. In Step 8, the judge (CATO) registers agreement with Burger's position.

The argument dialogue continues in Figure 14. In light of the judge's reaction, the student asks for help in Step 9. Della's hint, Step 10, explains the problem that two of the factors, though unshared, do not give rise to reasons for deciding the two cases differently. The student tries again, selecting other differences with better results.[93]

---

90. A related pedagogical goal is teaching students to explain why the problem scenario and the cited case should have a different outcome in light of the relevant factual differences. A student should be able to justify an assertion that a factual difference has legal significance.

91. *See supra* Figure 1.

92. *See* Figure 1, Step 5.

93. *See* Aleven, *supra* note 3, at 141. At the time CATO was evaluated, it engaged students in an explanatory "mini-dialogue" where the program critiques a student's attempts at distinguishing. The program invited a student to make an argument distinguishing a case and, where appropriate, explained why the student's response is either a strong one or not as strong as it could be. Below is the mini-dialogue version of the dialogue of Figure 1. CATO sets up the context of the lesson and then invites the student to select the factors which he would emphasize to distinguish *Boeing* on behalf of the defendant. The student selected two differences, factors F15 and F18, which, though not shared between the two cases, actually strengthen the argument that plaintiff should win in the problem situation. CATO makes this point in step 5.

1. CATO: Assume you are arguing on behalf of the defendant in *Mason*. Suppose plaintiff cites the *Boeing* 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 with defendant [F1], plaintiff should win a claim of trade secrets misappropriation, as in *The Boeing Co. v. Sierracin Corp.*, 738

9. <**STUDENT**: Help! (Continue)>

10. Della Street (whispering into Perry's ear): Here is a hint. Even though factors F18 and F15 are differences, defendant cannot use them to distinguish *Boeing*. Defendant needs to focus on other differences, namely, on those differences that make *Mason* look stronger for defendant than *Boeing*.

11. <**STUDENT (selecting from menu)**: Distinguish *Boeing* [F4][F12] [F14] [F16] (Continue)>

Mr. Mason for Defendant (Student): *Boeing* is distinguishable, because in *Boeing*, defendant used materials that were subject to confidentiality restrictions [F14], plaintiff's disclosures to outsiders were subject to confidentiality restrictions [F12], and defendant entered into a nondisclosure agreement with plaintiff [F4]. This was not so in *Mason*. Also, in *Mason*, plaintiff's product information could be learned by reverse-engineering [F16]. This was not so in *Boeing*.

12. Judge (to Plaintiff's counsel): Mr. Burger?

13. Mr. Burger for Plaintiff (CATO): Uhh. Well, as you know, Your Honor, *Boeing*, is not the only case we cite in our brief.

14. Della Street (whispering into Perry's ear): Brilliant job, Perry!

Figure 14: Continuation of Argument Dialogue in Figure 1

P.2d 665 (1987).
How would you respond on behalf of the defendant?

. . .

2. STUDENTS: [F15] [F16] [F18] (Continue)
3. CATO: Suppose defendant made the following argument:
The Boeing Co. v. Sierracin Corp. is distinguishable, because in *Mason*, defendant's product was identical to plaintiff's [F18] and plaintiff was the only manufacturer making the product [F15]. This was not so in *Boeing*.
Would this really help defendant?
When you have answered this question for yourself, please click on "Continue."
4. STUDENTS: (Continue)
5. CATO: Plaintiff could respond to defendant's argument, as follows:
Your Honor, as defendant's counsel correctly pointed out, in the current problem, defendant's product was identical to plaintiff's [F18] and plaintiff was the only manufacturer making the product [F15]. *Boeing*, which was won by plaintiff, does not have these strengths on plaintiff's part but is otherwise similar to the problem at hand. Where these additional strengths are present, a decision in favor of plaintiff is warranted even more strongly than it was in *Boeing*.
In short, even though factors F18 and F15 are differences, defendant cannot use them to distinguish *Boeing*. Defendant needs to focus on other differences, namely, on those differences that make *Mason* look stronger for defendant than *Boeing*.
Please click on "Continue."

See id. at 141 figure 5-25(a), 142 figure 5-25(b). Currently, new dialogues for other argument moves are being designed. One new dialogue, shown in Figure 15, deals with the argument move of emphasizing distinctions. The dialogue picks up near the end of the dialogue of Figure 14.

Transitions from one lesson in the curriculum to the next can be marked by Mr. Burger's suddenly "getting smarter." Previously, in Step 13 of Figure 14, when the judge turned again to plaintiff's counsel, Mr. Burger, simply conceded. This gives the program a chance to reinforce the student's good performance in differentiating between mere differences and real distinctions. Once a student has demonstrated proficiency in that lesson, however, it is time to move on to the next lesson. One way to implement such transitions in the curriculum is to pose a variation on a problem students already have solved, for instance, by having Mr. Burger try a more advanced argument move.

In the dialogue of Figure 15, instead of conceding, Mr. Burger downplays one of Perry Mason's distinctions. The student who has not seen this move before asks for help in Step 15. Della suggests emphasizing the distinction, thus moving on to a more advanced topic. The transition to a new lesson has been made. At Step 17, CATO translates the student's selection into an argument using the recipe for emphasizing shown in Figure 8.

In generating the argument dialogues of figures 1, 14, and 15, CATO follows an algorithm.[94] First, it makes the argument analogizing the target problem and source case. It then:

- Invites the student to identify the distinctions
- Diagnoses the student's answer in terms of three categories:
  1. Correctly identifies a distinction.
  2. Incorrectly selects a similarity rather than a difference.
  3. Incorrectly selects a difference that helps the opponent's argument.
- Responds appropriately.
  - If number 1, it solicits more distinctions until all have been correctly identified and then the opponent concedes.
  - If numbers 2 or 3, it shows the opponent's argument in response, offers help, and invites the student to try again.

In making the arguments, it uses the same argument moves shown in Figure 6 and Figure 8, plugging *Mason* and *Boeing* into its templates as appropriate in the context of the dialogue. Ultimately, CATO will generate dialogues for all of its argument moves.

Obviously, CATO controls these dialogues. Despite following an algorithm, however, the dialogues are not "hard wired." First, the program flexibly responds to the range of factors on which students choose to focus. Second, such a dialogue can be instantiated for any pair of cases in the CATO database. CATO adjusts the dialogue to the facts of the paired cases. Of course, it makes a better dialogue if the target and source case have both similarities and distinctions.

Third, CATO can model the student's apparent grasp of the lesson and decide when to make transitions to new lessons. As a student chooses CATO's argument moves and values, CATO records the student's activities. For instance, during the dialogue of Figure 1, it records that the student is using the Argument

94. *See id.* at 252.

Maker and has selected factors F15, F16, and F18 from *Mason* as distinctions. CATO will be able to use data like this, collected over the course of a student's interaction with the program, to note when a student is no longer making mistakes and decide to move to the next lesson, as illustrated in Figure 15.

Court is in session . . .

12. Judge (to Plaintiff's counsel): Mr. Burger?

13. Mr. Burger for Plaintiff (CATO): Your Honor, in *Boeing*, defendant used materials that were subject to confidentiality restrictions [F14]. This was not so in *Mason*. This however is not a major distinction. 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].

14. Judge (to Defendant's counsel): Mr. Mason? What do you say to that?

15. <STUDENT: Help! (Continue)>

16. Della Street (whispering into Perry's ear): Mr. Burger just downplayed one of the distinctions. Here is something that may help. Try emphasizing F14. Try emphasizing F16.

17. <STUDENT (selecting from menu): Emphasize [F14] [F16] (Continue)>

Mr. Mason for Defendant (Student): In *Boeing*, defendant used materials that were subject to confidentiality restrictions [F14]. This was not so in *Mason*. This is a marked distinction. It shows that in *Boeing*, defendant may have acquired plaintiff's information through improper means [F120]. This is supported by other facts in *Boeing* as well: Plaintiff imposed confidentiality restrictions in connection with its disclosures to outsiders [F12]. In *Mason*, on the other hand, defendant obtained or could have obtained its information by legitimate means [F120]: Plaintiff's information could be discovered by reverse engineering plaintiff's product [F16].

18. Judge (to both counsel): I think I have heard enough to help me make my decision.

Figure 15: Example of New Dialogue: Transition to a New Lesson as Mr. Burger "Gets Smarter"

With dialogues like these, CATO can engage students in fairly realistic legal arguments concerning cases and problems in an electronic casebook chapter. Its integrated resources make a set of legal argumentation concepts explicit and manipulable. Students can apply a rich, but limited, set of argument moves over its set of cases. The cases are represented and indexed so that the program can use them to generate argumentation examples in the dialogues.

IV. WILL IT WORK?

Electronic casebooks' flexibility, connectivity, and interactivity may have a positive pedagogical effect or may simply overwhelm students and confuse the author's intended message. CATO's three techniques may provide valuable guidance and constraints by constructing role-playing environments, making abstract concepts explicit and manipulable in a rich, but limited, setting, and supporting pedagogical dialogues. This section presents evidence supporting the techniques' pedagogical effectiveness and discusses some reasons to believe that they may be applied to other substantive legal domains.

A. Empirical Evidence that CATO Teaches

An experiment was undertaken to test whether law students can learn basic argumentation skills with CATO's computer-based environment. We compared the effectiveness of CATO to small group instruction by an experienced legal writing instructor. The February 1996 experiment involved a first-year, second semester legal writing class. Thirty law student volunteers were assigned randomly to an experimental group (16 students) or control group (14 students).[95] The course of instruction extended over a three-week period. The trade secret law casebook chapter was the basis for instruction of both groups. Students in each group were asked to read the casebook chapter as homework.[96]

The control group spent about 7.5 hours on this task. Instead of attending regularly scheduled legal writing classes, three subgroups of four to ten students from the control group attended six special classroom sessions. The instructor focused on cases and argumentation exercises from the casebook. In two of the sessions, students made in-class oral arguments based on those exercises. They prepared for their arguments in small teams.[97]

During the same three weeks, instead of attending their regular legal writing classes, the experimental group students received 7.5 hours of instruction with CATO in nine fifty-minute sessions conducted in a specially prepared CATO lab at the law school. Students worked in pairs using the workbooks to help analyze many of the argumentation problems at the end of each section in the casebook chapter. All students covered the seven argument moves and had some exposure to the lessons on constructing multi-issue, multi-case legal arguments.[98]

Students were given written examinations before and after the instruction.[99] The legal writing instructor graded the tests without knowing the identities of the

---

95. *See id.* at 151–52.

96. *See id.* at 152–53.

97. *See id.*

98. Students who covered the previous material more quickly had greater exposure to the later lessons. Only half the students had a chance to employ the mini-dialogue on distinguishing. *See supra* note 92.

99. Each test consisted of two question sets involving a problem situation and two short cases. The students made one-page arguments about the problem using the cases. *See id.*

students or the nature of the instruction they had received. (However, he could tell which were the pre-tests and the post-tests.) Unknown to the instructor, we also included for grading, answers generated by the CATO program that were transcribed into handwriting and presented in the same format as a student's answer.[100]

In addition, students were given a take-home writing assignment. They had about a week to write a maximum of six pages about a trade secret problem using six pre-selected cases.[101] This task was more demanding than the CATO exercises and the two in-class tests. It presented a more complex configuration of cases to integrate and required more composition and rhetorical skills.[102] The instructor graded the memoranda blindly.[103]

The grading criteria for the in-class tests included how well the student formulated arguments comparing the target problem and source cases, drew important factual similarities, explained the similarities' legal significance for the student's proposed decision of the target problem, and responded to such arguments by distinguishing the cited source cases, pointing out factual differences and their legal significance for deciding the target problem. Additional negative criteria included the use of conclusory statements, making assertions not in the facts of the problem, attempting to distinguish in terms of differences that were not really distinctions, and unnecessary repetition.[104] These grading criteria also were used for the memorandum, as were additional criteria focusing on organization.[105]

The instructor was not familiar with the CATO model. In teaching the control group, he addressed the casebook exercises by describing a general model of analogical legal reasoning without presenting any specific templates.

The scores on the in-class tests are shown in Table 2. Each of the pre- and post-tests had two questions. The overall numerical grade equals 0.75 Q1 + 0.25 Q2. From pre-test to post-test, there was a significant improvement within each group.[106] Each group, in other words, appeared to have improved as a result of their instruction (or other influences during the period of instruction). On the pre-test and post-test, there was no significant difference between the experimental

---

100. *See* Aleven, *supra* note 3, at 158.

101. Unlike the in-class, fifty-minute exams, students had ample time to complete the take-home writing assignment.

102. *See id.* at 154–55.

103. Again, we included a memorandum generated by CATO and presented it as a student's answer. This memorandum was similar to the arguments organized by issues shown in Figure 9. *See id.* at 158.

104. *See id.* at 156 figure 6-3.

105. The criteria were devised after consultation with the instructor, who agreed that the final list consisted of skills that were important from a legal pedagogical viewpoint and appropriate for this group of students from the second semester of the first year. *See id.* at 157 figure 6-4.

106. *See id.* at 159–60. Significance is based on a correlated samples t-test assuming equal variance in each group and using two-tailed tests; experimental group: $t(15) = -3.4$, $p = .004$; control group: $t(13) = -3.7$, $p = .002$.

and control groups' performance.[107] In addition, the mean change in scores (post-test grade minus pre-test grade) does not differ significantly between the experimental group and control group.[108]

The instructor also ranked CATO's answers among the best student answers. He said that he had not recognized that the answers were not generated by a student.

	Pre-Test Overall (Num.)	Pre-Test Overall (Letter)	Pre-Test Q1	Pre-Test Q2	Post-Test Overall (Num.)	Post-Test Overall (Letter)	Post-Test Q1	Post-Test Q2
Experimental Gr. Avg.	60	C-	75	15	70	C+	80	41
Control Gr. Avg.	55	D	70	10	68	C	81	30
CATO Answers	81	B+	82	79	87	A-	88	85
CATO Answers Ranking	1		4	1	3		6	2

Table 2: Results of In-Class Basic Argument Skills Tests

The results of the take-home assignment are shown in Table 3. The control group performed significantly better than the experimental group.[109]

	Memo Writing (Num.)	Memo Writing (Letter)	Previous Semester (Num.)	Previous Semester (Letter)
Experimental Gr. Avg.	70	B-	64	B/B+
Control Gr. Avg.	79	B+	63	B
CATO Answer	62	C		

Table 3: Results of Take-Home Memorandum Assignment

---

107. *See id.* Significance is based on an independent samples t-test using two-tailed tests; pre-test: $t(28) = 1.56$, $p = .13$; post-test: $t(28) = 0.57$, $p = .58$.

108. *See id.* at 158–62. No significant difference based on independent samples t-test using two-tailed tests; $t(28) = -0.66$, $p = .51$.

109. *See id.* at 165. Significance was tested using a two-tailed independent samples t-test; $t(25) = -2.38$, $p = .03$. On a memorandum writing assignment of the previous semester for the same students, there was no significant difference between the groups. *See id.*; $t(25) = 0.05$, $p = .96$. The instructor ranked CATO's answer well below average.

The results suggest both strengths and limitations in the CATO approach. The basic argument skills test administered in class indicates that the CATO instruction was associated with almost as much improvement in student performance as that achieved by an experienced and dedicated legal methods instructor with small groups.[110] In addition, the subjects were drawn from a special section, comprised of those judged most in need of support in legal writing and analysis.

The take-home memorandum results, on the other hand, indicate that CATO's instruction was less effective in improving performance on longer assignments. Several explanations are possible. For one thing, the 7.5 hours were simply too short a time to cover all of the material in the curriculum. CATO's instruction focused on the seven basic argument moves. By the time students reached the point of learning how to integrate the basic argument moves into more complex, multi-issue, multi-case arguments, the available time for instruction had all but run out.[111] Yet, the memorandum called for making an even more complex multi-case, issue-based argument than the examples in the CATO curriculum.

A second explanation might lie in CATO's model of the format of a multi-issue, multi-case argument. CATO's examples of multi-case arguments did not reflect the format and organization the grader expected.[112] Apparently, CATO is not as adept at integrating the discussion of a large number of cases as it needs to be. CATO's routines for generating an argument are geared toward supporting an argument for one side at a time (like an appellate brief). At the top level, CATO's arguments are organized in terms of plaintiff's argument, defendant's argument, and plaintiff's response. The legal writing instructor, on the other hand, expected to see a more balanced discussion (like an intra-office memorandum) organized around enumerated issues. Significantly, the students had considerable knowledge of this expectation. They had learned the preferred format in a first-semester writing exercise with the same instructor. CATO did not have this knowledge. Perhaps the experimental group students who scored better grades than CATO drew on their prior knowledge.[113]

---

110. The legal writing instructor is the director of that program as well as all of the law school's legal writing instructors. An experienced, dedicated teacher, he enjoys an excellent rapport with his students. He took seriously the task of instructing the control group. For instance, in the two sessions of oral arguments concerning casebook problems, he played the role of an active interlocutor. *See id.* at 164.

111. It took longer than expected for the experimental group students to work through the CATO curriculum. It took longer than expected to introduce a number of the students to CATO's window environment. Some of the students had much less experience with computers, windows or word processors than expected. It is expected that students will have more experience in the standard conventions of accessing information on the web.

112. He said the CATO "memo was 'on track,' but that there were 'serious problems with organization and presentation in general' and that the analysis was 'too fragmented.'" *See id.* at 174–75.

113. *See id.* at 175. The instructor did not criticize CATO's argument for linking the facts to the wrong issues. Since the Factor Hierarchy guided these linkages, it appears to have cleared an

The evaluation of CATO did not focus on the pedagogical effectiveness of dialogues. At the time of the evaluation, CATO only generated the more explanation-oriented mini-dialogues,[114] not the argumentation dialogues of figures 1, 14, and 15. In addition, only half of the experimental group tried the mini-dialogues.[115]

B. Applicability to Other Legal Domains

A second limitation of CATO is that its techniques may not apply to electronic casebooks in all legal domains. Some techniques depend on the feasibility of creating intermediate-sized databases of cases represented and indexed so that a program can reason with them. While CATO demonstrated the techniques for some first-year argumentation concepts and trade secret law factors, electronic casebooks deal with a far greater range of substantive legal concepts and legal domains, including statutory domains very different from trade secret law.

Applying the techniques to all domains covered by electronic casebooks would require new research in computationally representing and reasoning with legal knowledge. On the other hand, AI and Law practitioners have been developing representations in both common law and more statutory legal domains.[116] Factors have been used in AI and Law research to model the legal

---

important hurdle. *See id.* Concerning CATO's organization of the discussion by issues, however, the instructor thought that it addressed too many issues. He did not see how the discussion of each issue added up to an overall conclusion that the plaintiff should or should not win. And, he felt that CATO's argument was too much like an outline. He expected to see more specific descriptions of the factual similarities and differences among the cases. On the other hand, when he learned that CATO had drafted the particular memo, he expressed surprise that it had been written by a computer program. He exclaimed, "Wow, I'm impressed that the computer can do that actually. Wow, I knew they could write music and be chess champions but I didn't know that they could write legal memos." *Id.*

114. *See supra* note 92.

115. In conjunction with the Center for Interdisciplinary Research on Constructive Learning Environments, a research center located at the University of Pittsburgh and Carnegie Mellon University, an experiment is planned to compare the relative pedagogical effectiveness of the two types of dialogues and to test whether students trained on the argument dialogues will not only be more highly motivated to learn, but better at distinguishing cases. The Center's aim is to build and test a new generation of computer tutoring systems that encourage students to construct the target knowledge instead of having it told to them. CATO's argumentation dialogues may be one means of encouraging students' construction of knowledge without requiring very sophisticated natural language and dialogue planning capabilities. *See CIRCLE's Home Page* (last modified Apr. 20, 2000) <http://www.pitt.edu/~circle/>.

116. In addition to factors, AI and Law programs have employed a wide range of knowledge representation techniques. Some representative examples are:

> *Logical rules*: Logical rules have been employed to represent analyzing problems raising the issue of British citizenship under the British Nationality Act. *See* Marek J. Sergot et al., *The British Nationality Act as a Logic Program*, 29 COMM. ACM 370–86 (1986). Translating statutory rules into logical formulations raises a number of issues of resolving logical ambiguities. *See* Layman E. Allen & Charles S. Saxon, *Some Problems in Designing Expert Systems to Aid Legal Reasoning*, 1987 PROC. FIRST INT'L CONF. ON ARTIFICIAL INTELLIGENCE & L. 94–103. Other problems are discussed in Donald Berman & Carole Hafner, *Obstacles to the Development of Logic-Based Models of Legal*

domains of the home office tax deduction[117] and the Bankruptcy "Good Faith" domain.[118] Both applications use factors explicitly in connection with statutes. The factors and indexed cases help to represent the meaning of open-textured statutory terms such as "principal place of business" or "proposed in good faith." A number of other statutes also explicitly invoke factors, such as the Fair Use provision of the Copyright Act[119] and may be especially suitable for this approach.

We are developing databases for additional topics law students might encounter in the first-year curriculum. We have constructed a small database using factors and cases to represent the difference between the agency law

---

Reasoning, in COMPUTER POWER AND LEGAL LANGUAGE 185, 185–214 (C. Walter ed., 1986).

Heuristic rules: Waterman employed heuristic production rules derived from legal experts to evaluate asbestos-related tort claims. The rules covered substantive areas like product liability and negligence. *See* D. A. Waterman & M. Peterson, Models of Legal Decisionmaking 155 (1981) (unpublished technical report, Rand Corp.)

Prototypes and deformations: The TAXMAN II program employed a representation of "prototypes and deformations" to represent legal concepts in the field of corporate tax. The representation includes template-like descriptions of a legal concept (e.g., taxable income) and a set of possible mappings from one description into other possible ones. The mappings can be applied adaptively in arguments. *See* L. Thorne McCarty & N. S. Sridharan, The Representation of an Evolving System of Legal Concepts: II. Prototypes and Deformations (1981) (unpublished technical report, Laboratory for Computer Science Research, Rutgers University). *See also* L. Thorne McCarty, *Reflections on TAXMAN: An Experiment in Artificial Intelligence and Legal Reasoning*, 90 HARV. L. REV. 837 (1977) [hereinafter *Reflections*]. A more general deontic language for legal discourse (LLD) and planning is outlined in L. Thorne McCarty, *A Language for Legal Discourse: I. Basic Features*, 1989 PROC. SECOND INT'L CONF. ON ARTIFICIAL INTELLIGENCE & L. 180.

Augmented Transition Network: Gardner employed an Augmented Transition Network (ATN) for representing a kind of legal grammar of rules for "parsing" events having to do with offer and acceptance. With each new event, such as telephone enquiry or receipt of a letter, the ATN determines the legal "state of affairs" as to whether there is a binding contract. *See* ANNE GARDNER, AN ARTIFICIAL INTELLIGENCE APPROACH TO LEGAL REASONING 1–16, 33–66, 84–188 (1987). More recently, researchers have applied logical rules interpreting the United Nations Convention on the International Sale of Goods to deduce the legal state of affairs as other kinds of events in a contract dispute occur. *See* Hajime Yoshino, *Logical Structure of Contract Law System*, 2 J. ADVANCED COMPUTATIONAL INTELLIGENCE 2 (1998). Another program employs a kind of augmented transition network to guide inferences about property distributions in connection with divorce. John Zeleznikow et al., *Project Report: Split-Up—A Legal Expert System Which Determines Property Division Upon Divorce*, 3 ARTIFICIAL INTELLIGENCE & L. 267 (1995).

Semantic Networks: A semantic network comprises a set of nodes connected by arcs. The nodes represent objects, concepts, or events. The arcs represent relations such as has-part, isa, and subset. TAXMAN I employed a semantic network representation of legal concepts concerning tax treatment of corporate reorganizations. *See* McCarty, *Reflections*, *supra*, at 837–93. GREBE employed semantic networks to represent Workman's Compensation cases. *See* L. Karl Branting, *Building Explanations from Rules and Structured Cases*, 34 INT'L J. MAN-MACHINE STUD. 797 (1991).

Connectionist Networks: A connectionist or neural network is a system of many nodes connected to other nodes by weighted links. Using a set of training examples, the network is trained (i.e., the weights associated with links are adjusted pursuant to a training rule) so that the network can classify new instances correctly. A number of programs are hybrids of connectionist networks and other representations. *See, e.g.*, Daniel E. Rose & Richard K. Belew, *A Connectionist and Symbolic Hybrid for Improving Legal Research*, 35 INT'L J. MAN-MACHINE STUD. 1 (1991); Zeleznikow et al., *supra*.

117. IRC §280 A (c)(1). *See* Edwina L. Rissland & David B. Skalak, *CABARET: Statutory Interpretation in a Hybrid Architecture*, 34 INT'L J. MAN-MACHINE STUD. 839 (1991).

118. 11 U.S.C. §1325(a). *See* Edwina L. Rissland et al., *BankXX: Supporting Legal Arguments Through Heuristic Retrieval*, 4 ARTIFICIAL INTELLIGENCE & L. 1 (1996).

119. 17 U.S.C. §107 (1976).

concepts of an "employee" and an "independent contractor." We are also researching a role for factors in dealing with the enforceability of restrictive covenants in connection with real property and employment.

CATO and the other AI and Law models[120] capture diverse aspects of legal reasoning in a range of substantive legal domains. They probably do so well enough to create tutoring environments as adjuncts to electronic casebooks. A number of these models make important legal concepts explicit and manipulable in intermediate-sized case databases and conceivably can be used to engage students in pedagogically useful exercises and dialogues. Consider, for instance, how an electronic casebook in Cyberspace and the Law might treat *ACLU v. Reno*.[121] Much of the Supreme Court and district court opinions make comparisons to other cases involving different media, in which similar regulations have either been struck down or sustained. The cases focus on concepts such as overbreadth and vagueness. The comparisons among the media cases relevant to these concepts can readily be characterized in terms of such factors as: (1) height of the barriers to entry into the medium, (2) whether the barriers to entry are identical for speakers and listeners, (3) diversity of content available on the medium, (4) extent to which the medium is a scarce public good, (5) intrusiveness of the medium, and (6) whether the restriction is a prohibition or a zoning regulation. These factors could form the nucleus of an intermediate-sized database of cases that could engage students in useful exercises and dialogues.

Extending CATO to other legal domains would be easier if there were tools for automatically representing and indexing legal opinions according to factors. In related work, we are developing a machine learning program, SMart Index LEarner (SMILE), for this purpose. SMILE has been trained on CATO's manually classified case squibs and learns to classify new squibs presented as raw texts according to some of the factors that apply.[122] SMILE does not read an opinion text. It is trained, in effect, to recognize patterns of words and concepts from which it can infer that a factor applies. To make learning more effective, we are experimenting with techniques for pre-processing the training examples and adding linguistic information about negation and phrases. We hope to extend the approach to classifying full text opinions of trade secret misappropriation and other claims.[123]

---

120. *See supra* note 116.

121. American Civil Liberties Union v. Reno, 929 F. Supp. 824 (E.D. Pa. 1996).

122. *See* Stefanie Brüninghaus & Kevin D. Ashley, *Toward Adding Knowledge to Learning Algorithms for Indexing Legal Cases*, 1999 PROC. SEVENTH INT'L CONF. ON ARTIFICIAL INTELLIGENCE & L. 9.

123. The SPIRE program locates passages likely to contain information about legally relevant features in full-text judicial opinions. SPIRE deals with ten features of bankruptcy cases relevant to the statutory requirement that a debtor's plan to satisfy creditors must have been submitted in "good faith." Some features are quite specific, such as whether the debtor's future income will increase; others are very general such as any special circumstances that excused the debtor's situation. *See* Jody J. Daniels, Retrieval of Passages for Information Reduction (1997) (unpublished Ph.D. dissertation, University of Massachusetts, Amherst.); Jody J. Daniels & Edwina L. Rissland, *Finding Legally*

Examples of arguments and dialogues illustrate how an intelligent tutoring environment like CATO may help an electronic casebook teach abstract concepts. CATO provides a context in which students may practice the method of comparison and difference that underlies the case system. The context is an intermediate-sized set of cases, represented and indexed so that the tutoring system can retrieve and reason with them. The system's computerized instructional tools make the concepts explicit and manipulable. Students retrieve cases illustrating targeted concepts and test their model of a concept on the cases in the database. In engaging students in professional role-playing activities focused on specific problems, the tutoring environment scaffolds students in using the tools to practice applying the concepts to solve problems and make arguments.[124]

Once in place, the tools, concepts and case database serve as a buffer between the main opinions in an electronic casebook and the enormous full-text case databases such as WESTLAW or LEXIS. The transition to the full-text databases can be smoothed in a variety of ways. These techniques also could improve the usefulness of legal information retrieval tools like WESTLAW or LEXIS. A link from CATO to Shepard's could invite students to update important cases. Factor-based queries could be translated automatically into natural language queries by substituting canonical descriptions of the factors in the query. Although the results of the full-text database are not guaranteed to satisfy the constraints of the CATO query, the cases returned are likely to have one or more of the relevant factors. Automated text processing techniques such as SMILE and SPIRE[125] may help students analyze the opinions to identify factors and other important features. Arguments could be hyperlinked directly to cases, statutes, and Restatements.[126]

Relevant Passages in Case Opinions, 1997 PROC. SIXTH INT'L CONF. ON ARTIFICIAL INTELLIGENCE & L. 39.

124. From a technical viewpoint, there are five key requirements to creating an explicit and manipulable set of targeted concepts and generating arguments and dialogues like CATO's. One needs:

(1) To represent cases so that a program can reason with them. CATO represents cases in terms of factors. Its Factor Hierarchy represents the factors' legal significance.

(2) Methods for deductively retrieving small collections of cases related to each other in interesting ways. CATO employs a structured inheritance knowledge representation system that can deductively retrieve such collections (i.e., Argument Contexts), using its working definitions of important argumentation concepts.

(3) Templates and algorithms to implement a set of basic argument moves and to compose those moves into more complex arguments. CATO's templates and algorithms enable it to characterize the significance of factors within the context of the argument move at an appropriate level of abstraction using the Factor Hierarchy.

(4) A user interface that presents the role-playing dialogue and enables students to select argument moves and values. CATO's Argument Maker is organized around "hot buttons" for argument moves and menus presenting lists of cases and factors to employ.

(5) Natural language generation techniques. CATO employs a canned text approach, but planning techniques could help generate more natural texts.

125. *See supra* Part IV.B.

126. When students find and cite materials not represented and indexed in a manner CATO can

Beyond demonstrating the feasibility of an electronic casebook that talks back, the CATO research has contributed to cognitive science and jurisprudence. From the viewpoint of cognitive science, CATO's dialogues provide a medium through which students construct and test the targeted knowledge rather than simply being told it. CATO's approach encourages activity that is somewhat similar to self-explanation. Students who explain examples to themselves learn better and use analogies more economically.[127] Using CATO to retrieve cases for testing hypotheses and making arguments, students search for cases that satisfy explicit constraints imposed by the targeted concepts. They are led to focus actively and explicitly on the targeted concepts, detailed constraints, techniques for satisfying them, and uses of the relevant cases thus retrieved. In a controlled experiment, we found that students taught by CATO learned the targeted concepts as well as those taught by a human instructor. In an additional experiment involving a more complex legal writing assignment, the students taught by the human instructor did better, but that experiment has provided us with guidance for improving CATO's instruction.

CATO models a phenomenon on which jurisprudential analysis has not focused. CATO generates multiple interpretations of the legal significance of differences among similar cases. CATO's detailed model of how to generate multiple interpretations of a cited case, by downplaying or emphasizing the legal significance of distinctions, explicates an important feature of arguing by analogy. The requirements of an argument lead the program to decide which characterizations of the importance of relevant similarities and differences to employ and how abstractly to characterize them. One can point to exactly which knowledge sources are employed and detailed algorithms by which the program chooses a plausible interpretation to support an argument. The program's algorithms for downplaying and emphasizing distinctions guide a kind of "conceptual ascent" as CATO decides which paths upward through the Factor Hierarchy to pursue and how high to go. It can perform these tasks for any small collection of a problem scenario and source cases drawn from its database of cases.

Integrating intelligent tutoring environments and electronic casebooks in diverse legal domains will not be easy. But there is reason to believe that CATO's methods can be adapted to teach substantive legal concepts in other legal domains such as the elements of a legal claim or the meaning of open-textured legal predicates. Success in adapting CATO and other AI and Law representations to more diverse legal domains would increase the likelihood that standardized techniques may someday be developed for designing electronic casebooks that talk back.

---

process, of course, the program cannot comment on the results. That is in the nature of the transition to the real world beyond the electronic casebook.

127. K. VanLehn et al., *A Model of the Self-explanation Effect*, 2 J. LEARNING SCI. 1 (1992).