Reasoning Symbolically About Partially Matched Cases

Kevin D. Ashley and Vincent Aleven
Learning Research and Development Center, Graduate Program in Intelligent Systems Program, and School of Law
University of Pittsburgh
Pittsburgh, PA 15260 USA
+1(412)624-7496
ashley+@pitt.edu, aleven+@pitt.edu

Abstract

In teaching case-based argumentation skills, the CATO program, an intelligent learning environment, guides students' assessments of partial matches between problems and cases by generating alternative interpretations of the similarities and differences. CATO's Factor Hierarchy captures information about the significance of similarities and differences given the normative purposes of the domain classification. Its algorithms for emphasizing or downplaying significance tailor interpretations to the comparison context, block interpretations strongly contradicted by other factors and strategically determine how and how abstractly to characterize a difference. An empirical evaluation confirmed CATO's effectiveness in teaching basic argumentation skills.

1 Introduction

Traditionally, researchers in case-based and analogical reasoning have regarded as an obstacle the fact that cases can be described at various levels of abstraction. Carbonell abandoned a transformational approach to analogy in favor of a derivational approach because of the problems of recognizing analogous problems despite apparent differences in the language or levels of abstraction of their descriptions [Carbonell, 1983]. Describing cases at different levels of abstraction also frustrates isomorphism in structure-mapping approaches to analogical reasoning. Forbus acknowledges the obstacle as a limitation in MAC/FAC. Either the cases all have to be described in the same way and at exactly the same levels of abstraction or else the mappings will fail [Forbus, et al. 1994, p. 198].

In argumentational or interpretive CBR applications, however, multilevel case descriptions are not just an inconvenient obstacle for computerizing analogical reasoning. They are an integral part of why and how experts reason with cases. Experts assess cases and similarity in terms of a general domain model, to the extent there is one. They fit new cases into the domain model and into the body of cases, grouping a new case with similar examples and reconciling it with negative examples. The level of abstraction with which to describe a case is an important degree of freedom in fitting it into a general model and reconciling it with other cases. The importance of similarities and differences varies as cases are characterized abstractly in different ways and levels; the general model helps experts assess the importance of shared and unshared case features viewed more abstractly or less. By choosing the groupings and characterizations carefully, an expert may even extend the general domain model.

This is certainly true of legal case-based reasoning in which the general principles and policies of a legal domain inform a determination of what case features and which analogical similarities and differences are important [Burton, 1985]. Since quantitative weighting schemes are either inappropriate for or not sufficiently context-sensitive to be applied in legal argument [Ashley & Rissland, 1988], one needs to reason explicitly in terms of some model of why similarities and differences are significant. In law, however, the general model of a domain is often weak enough, and there is enough uncertainty about how cases fit the model and each other, that characterizing cases at different levels of abstraction is often done strategically; the level of abstraction an arguer chooses in asserting that cases are the same or different reflects the arguer's purpose.

In designing a program, CATO, to teach first year law students basic skills of making case-based legal arguments to classify new cases, we have partially addressed three of four problems associated with comparing cases at multiple levels of abstraction, the problems of: (1) representing cases at multiple levels, (2) reasoning symbolically about partially matched cases in light of the possible abstract descriptions, and (3) modeling the strategic uses of case description. Students learn how to argue that the complaining party should win (or lose) its claim in a new case by drawing analogies to past cases where the corresponding side won (or lost) and by distinguishing or otherwise counteracting similar cases with the opposite result. CATO teaches students, among other things, how to characterize a problem abstractly in arguing that a particular side should win and how to characterize the significance of particular differences between the problem and past cases so as to emphasize or downplay that significance. The lessons should help us address the remaining problem (4): identifying and retrieving cases described at different levels of abstraction.

In making competing arguments that a partially matched case is close enough to a problem or not, CATO alternatively interprets the differences in terms of more abstract characterizations contained in its domain model, the Factor Hierarchy. This graph captures information about the plausible significance of similarities and differences given the domain's normative purposes. It enables the program to determine which abstractions help an argument, hurt the argument, or are indifferent. Although CATO's case representations do not include multiple levels of abstraction, CATO does compare cases at multiple levels of abstraction. Its algorithms (for downplaying and emphasizing differences and for making issue-based arguments) select, from among all the abstractions at multiple levels and along multiple paths which may apply to a case, just the right ones (i.e., the focal abstractions) for characterizing the case in various argument moves. CATO's algorithm for selecting focal abstractions implements strategic criteria so that the resulting arguments avoid contradictions and exposure to counterattacks. When there is conflicting evidence whether an abstraction applies, CATO applies its general knowledge in the Factor Hierarchy to resolve the conflict by blocking certain interpretations, if possible. If not, its case-based arguments play out the conflict.

As compared with either CASEY or GREBE, CATO employs an alternative approach and a different kind of model to guide interpretation of partial matches. CASEY reasons symbolically about differences between a problem and an explained, diagnosed case in terms of the strong causal diagnostic model used to generate the past case's explanation and evidence principles for characterizing whether differences are important. [Koton, 1988]. GREBE represents cases at multiple levels of abstraction in a structural sense. Each case's explanation is like a small structural model indicating which features are relevant to which conclusions. GREBE attempts to bridge gaps in the mapping of explanations by importing inferences from other case explanations and legal rules [Branting, 1991].

Neither CASEY nor GREBE have anything like a Factor Hierarchy that represents *how* features strengthen or weaken a conclusion and *why,* information which informs CATO's arguments emphasizing or downplaying differences. Although CATO's model cannot solve problems by itself as CASEY's can, that is to be expected given the differences between CASEY's medical domain and CATO's legal one. While GREBE may compare cases at multiple levels of structural abstraction, it does not make sense to ask whether it strategically selects a level of abstraction at which to characterize similarity or differences as CATO does in selecting focal abstractions. GREBE uses a different kind of model to make a different kind of argument.

2 CATO's Factor Hierarchy

Students use CATO to analyze argumentation problems in a traditional casebook chapter on trade secret misappropriation, a legal regime protecting confidential commercial information from unfair access and use by competitors. Students may retrieve cases from CATO's database of 147 trade secret cases. A special query language helps students formulate constraints that a case must satisfy to be useful in an argument and translate them into queries to retrieve relevant cases.

2.1 Factor Representation

In CATO, legal cases are represented by a short textual description of the facts and decision (i.e., the case squib) and a set of indexing factors. Factors represent stereotypical collections of facts which tend to strengthen or weaken a plaintiffs legal claim [Ashley, 1990]. CATO's representation is illustrated in Figure 1, showing a squib's description of the facts of the *MBL* case,and Figure 2, box 1 indicating the set of six factors indexing the case.

Facts: Chemi-Flex (a division of plaintiff) had developed, over a number of years, a molding process for manufacturing urethane belts. Defendant Diekman was a former employee. When he first worked for plaintiff, Diekman had signed an employer-employee agreement which contained a confidentiality and non-competition clause. However, when he was rehired in 1979, he refused to sign such an agreement. Following his second tenure with Chemi-Flex, defendant formed two companies and began to design centrifugal molding machines and designs for other belt-producing equipment. Plaintiffs employees were not aware of what information plaintiff considered to be confidential. Some employees, but not all, signed nondisclosure agreements. Plaintiffs plans, designs and customer data were not locked up; customer names and orders were not marked confidential; process formulas and machines were shown, without restriction, to employees, outside consultants and others; no licensing or confidentiality agreements were signed by outside parties with access to the process and formulas; and the entire process, formula and machinery were open to a team of engineers from a stockholder corporation of plaintiff. Customer names and specifications were on orders and requisitions located in various places in the plant. The use of plaintiffs process was known in the industry, although the process is not specifically discussed in any industry literature.

Figure 1: Squib's Facts for *MBL (USA) Corp.* v. *Diekman*

Typically, legal disputes involve some factual strengths for the plaintiffs claim, the pro-plaintiff factors, and some weaknesses, the pro-defendant factors. In *MBL,* some factors favor the plaintiff (p): the defendant had entered into nondisclosure and noncompetition agreements (at least, at one time), F4 and F13, evidence that plaintiff took at least one kind of measure to protect the security of its confidential information, F6. On the other hand, the agreement did not specifically indicate what information was regarded as secret, F5, some of the information had been disclosed to outsiders, F10, and in any event, some of the information was already known in the industry, F20.

In law, there is no algorithmic or statistical technique for combining strengths and weaknesses to come to a decision. Instead, lawyers make arguments by analogy to past cases presenting the same combinations of strengths and weaknesses, arguing for a similar result in the problem.

2.2 Constructing Arguments

CATO's recipe for constructing such an argument is illustrated in Figure 3, which shows on the right side a set of steps for justifying a favorable decision for a side on a legal issue associated with his claim. On the left is an outline of an argument in favor of plaintiffs trade secret misappropriation claim in the *MBL* case focusing on one of two issues: that the information is a trade secret. Pursuant to the recipe, each issue argument attempts to capitalize on the related factual strengths and overcome the weaknesses. Accomplishing each goal requires finding and citing appropriate cases satisfying certain constraints. As the capitalized text indicates, in order to emphasize the strengths and downplay the weaknesses, the arguer needs to find cases which satisfy the specified constraints.

1. *MBL* = F4 Agreed-Not-To-Disclose (p) * F5 Agreement-Not-Specific (d) = F6 Security-Measures (p) * F10 Secrets-Disclosed-Outsiders (d) = F13 Noncompetition-Agreement (p) * F20 Info-Known-To-Competitors (d)	2. *Elcor* (p) = F4 Agreed-Not-To-Disclose (p) = F6 Security-Measures (p) = F13 Noncompetition-Agreement (p) * F15 Unique-Product (p) * F18 Identical-Products (p)
3. *MBL* = F4 Agreed-Not-To-Disclose (p) * F5 Agreement-Not-Specific (d) = F6 Security-Measures (p) * F10 Secrets-Disclosed-Outsiders (d) F13 Noncompetition-Agreement (p) * F20 Info-Known-To-Competitors (d) = shared factor * distinction	4. *Sperry* (p) * F2 Bribe-Employee (p) F3 Employee-Sole-Developer (d) = F4 Agreed-Not-To-Disclose (p) = F6 Security-Measures (p) * F7 Brought-Tools (p) * F15 Unique-Product (p) * F18 Identical-Products (p) * F26 Deception (p)

Figure 2: Factor Comparison of *MBL* Case with the *Elcor* Case (Boxes 1 , 2) and with the *Sperry Case* (Boxes 3, 4)

CATO's Factor Hierarchy provides information relating factual strengths and weaknesses (i.e., factors) to those legal issues for which they are relevant. CATO uses this information in constructing its descriptions of the constraints on cases which would be useful in an argument (i.e., the capitalized text in Figure 3.) Excerpts of the Factor Hierarchy are shown in Figure 4. Base-level factors at the bottom of the Hierarchy are linked through intermediate legal concerns to high-level legal issues at the top. The nodes representing intermediate concerns and top-level issues are called "abstract factors". Each represents two opposite conclusions, one favoring plaintiffs and one favoring defendants. The links represent a (defeasible) support relation between the nodes. Links may be strong (thick) or weak (thin), indicating the level of support they provide, and are marked as to whether they support the same side as the primary conclusion of the parent node (+) or the opposing conclusion (-). CATO uses the link strength to determine whether to block certain inferences (discussed below.) Currently, the Factor Hierarchy contains 26 base-level factors for trade secret law, 16 abstract factors (5 of which are legal issues) and 50 links.

CATO's query language helps students translate argument constraints directly into queries for relevant cases from its database. For example, students could find cases relevant for plaintiffs argument on the trade secret issue using queries such as "(and F4 F6)" and "(or F4 F6) (or F10 F20)". Among the cases returned by the queries are the *Elcor* and *Sperry* cases. Figure 2 shows a factor comparison between the *MBL* case and each of the *Elcor* and *Sperry* cases which confirms that they satisfy the first query's constraints and thus are eligible to be used in the argument of Figure 3.

Deciding whether to include these cases in the argument necessitates reasoning about partial matches. As Figure 2 indicates, each case only partially matches *MBL.* Each case has numerous distinctions with respect to *MBL.* Distinctions are those unshared factors that tend to make the cited cases *{Elcor* or *Sperry)* stronger for plaintiff than *MBL.* In particular, if the case is employed in the argument, the opponent could respond to it by distinguishing it, that is, pointing out these relevant differences. Are the problem and cited case really the same or different? Are the distinctions between them really important? Do they warrant not including the cases in the argument?

3. Reasoning Techniques for Partial Matches

CATO's Factor Hierarchy and techniques for emphasizing or downplaying distinctions help students frame answers to these questions, and thus reason about partial matches, by generating examples of arguments to consider. Figure 5 shows several arguments CATO makes to emphasize or downplay a particular distinction (F15 Unique-Product (p)) between *MBL* (or versions thereof) and the *Elcor* and *Sperry* cases. In both *Elcor* and *Sperry,* the plaintiffs products were unique, no other competitors marketed products like that. Since pro-plaintiff F15 was not in *MBL,* the defendant could distinguish these cases, pointing out that *MBL* was correspondingly weaker for plaintiff. This distinction is, indeed, important. As CATO's arguments in boxes 1 and 3 of Figure 5 indicate, CATO can make strong arguments on behalf of defendant emphasizing the distinction but cannot make any argument for the plaintiff downplaying it.

In emphasizing a distinction, one has to find abstract interpretations to use as "focal points" for characterizing the two cases as very different. One interprets the distinction's significance in a case, shows factors in that case which support the interpretation (i.e., "corroborating factors"), and shows factors in the other case supporting a contrary characterization (i.e., "contrasting factors"). In downplaying a distinction, by contrast, the goal is to dismiss the distinction, finding an abstract interpretation to use as a focal point for characterizing the two cases as the same.

In CATO's algorithm for emphasizing a distinction, the key is selecting the focal point for emphasizing a distinction D in case CI in favor of side S (either plaintiff or defendant.) A focal point is a 3-tuple comprising an abstract factor P to use as a focal abstraction for characterizing the distinction's significance, a set X of contrasting factors, and a set Y of corroborating factors. Set X comprises con-S factors in the other case C2. Set Y comprises pro-S factors in CI. P is the most specific common ancestor in the Factor Hierarchy of D and the factors in X and Y.

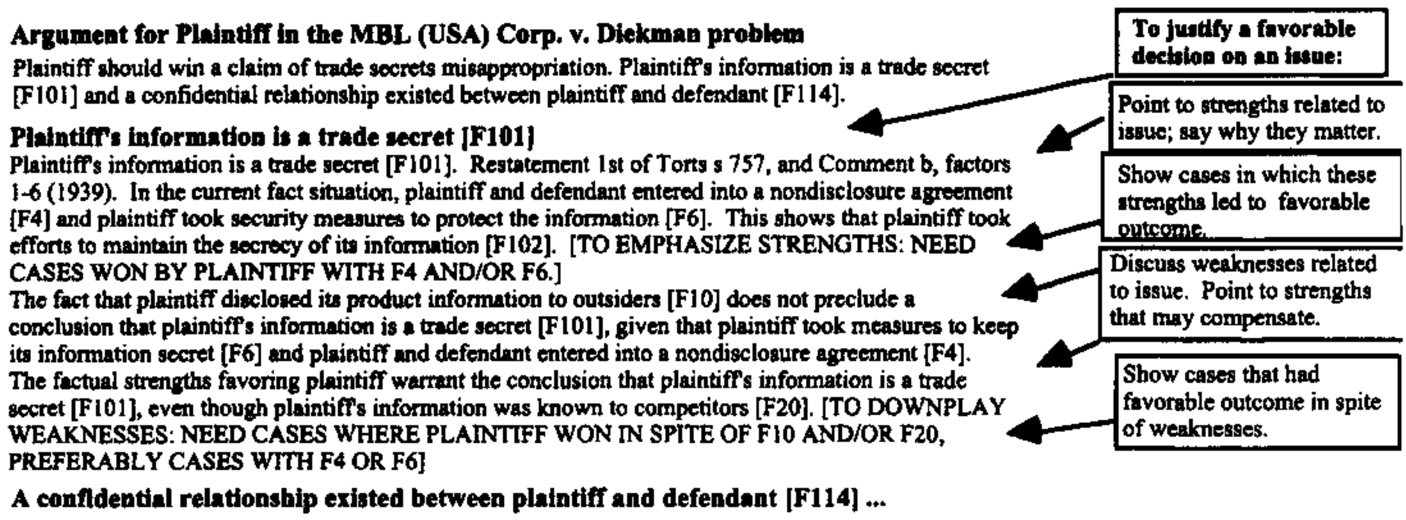

Figure 3: Sample CATO Issue-based Argument (left) and Recipe for Making It (right). (Factors and Abstract Factors are indicated in square brackets.)

To emphasize distinction D of case CI as compared to C2 (where D favors side S]:

1. Select candidate focal points to emphasize D:
   A. Find all combinations of a focal abstraction P and a set of contrasting factors X in C2 such that Pis a most specific common ancestor of D and the factors in X.
   B. Find all combinations of a focal abstraction P and a set of corroborating factors Y in CI such that P is a most specific common ancestor of D and the factors in Y.
2. Organize candidate focal points
   A. Join candidate focal points with same focal abstraction P.
   B. Filter unsuitable factors and focal points. For each candidate:
   • Check if contrast exists between C1 and C2 for P's ancestors.
   • If candidate has no contrasting factors, check if closed-world assumption is appropriate that pro-S conclusion associated with P is absent in C2.
   • Remove contrasting or corroborating factors shared by CI and C2. (Shared factors not useful to draw contrast.)
   C. Consolidate candidate focal points whose focal abstractions are ancestor and descendant.
   D. Order the focal points by estimated strength: (1) X is not empty. (2) Y contains other factors beside D. (3) The rest.
3. Generate text for surviving focal points with templates.

The algorithm (and the one for downplaying a distinction in [Aleven and Ashley, 1996]) supports three "smart" techniques for reasoning symbolically about the significance of similarities and differences: (1) Tailor interpretations to fit the context of comparison. (2) Block interpretations strongly contradicted by other factors. (3) Strategically interpret similarities and differences.

Tailoring Interpretations: CATO achieves a measure of context sensitivity in comparisons by virtue of the Factor Hierarchy's structure in which individual factors may relate to more than one abstract factor and by the Emphasize Distinction Algorithm's identifying abstract factors (P) as focal points which are the most specific common ancestors (msca) of the distinction to be emphasized (D), the contrasting factors, and the corroborating factors. (See [Kolodner, 1993] for other uses of msca.) By "context sensitivity", we mean knowing which similarities and differences are most salient in different circumstances and why: which should a reasoner focus upon, how should it characterize them, and which should it ignore. Although a normative concern may imply generally that a particular factual circumstance is significant, in the context of a particular problem and case, that assessment may be affected by the co-occurrence of other factual circumstances and other concerns, the arguer's rhetorical viewpoint and the dialectical role in which the arguer is engaged.

The effect of tailoring an interpretation of a difference may be seen in Figure 5. Each box shows an argument emphasizing the same distinction involving F15 between the *MBL* case (and two variations of *MBL)* and the *Elcor* and *Sperry* cases, respectively. The arguments focus on different interpretations of the distinction in terms of abstract factors F106, F120, or F104. Depending on the context, one or other of these abstract factors is the focal abstraction. For instance, in Box 1, CATO interprets F15's significance in *Elcor (C\)* in terms of abstract factor F106 as showing that in *Elcor,* plaintiffs information was not known outside its business, whereas in *MBL* (C2) it was. In Box 3, F15 is significant in *Sperry* (CI) because it suggests the defendant in *Sperry* got the information through improper means, unlike *MBL* (C2). In comparing *MBL* and *Elcor,* the algorithm finds that F106 is the most specific common ancestor of F1S and the contrasting factors in *MBL,* F10 and F20. In comparing *MBL* and *Sperry,* however, the algorithm finds that by selecting the more abstract F120, additional corroborating factors can be brought to bear in a broader contrast distinguishing *Sperry* as a case turning upon improper means, namely F2, F7, and F26. The decision, to adopt F120 as the interpretation, is made in the algorithm's consolidation step, 2C.

Other evidence of CATO's context sensitive interpretations is in Figure 5. Comparing Boxes 1 and 2 shows the effect on the *MBLIElcor* comparison of making *MBL* stronger for plaintiff by adding F8, so that the defendant saved product development time or expense by accessing plaintiffs information. CATO now makes an argument downplaying the F15 distinction interpreting it in terms of F104, a new focal abstraction indicating that the information was valuable. Boxes 3 and 4 show the effect on the *MBL I Sperry* comparison of making *MBL* stronger for plaintiff by adding F7, so that the defendant brought plaintiffs product-development tools with him to the defendant. CATO again says the F15 distinction shows possible improper means in *Sperry,* but this time CATO downplays the distinction, arguing on the basis of F7 that the *MBL* hypo and *Sperry* are the same.

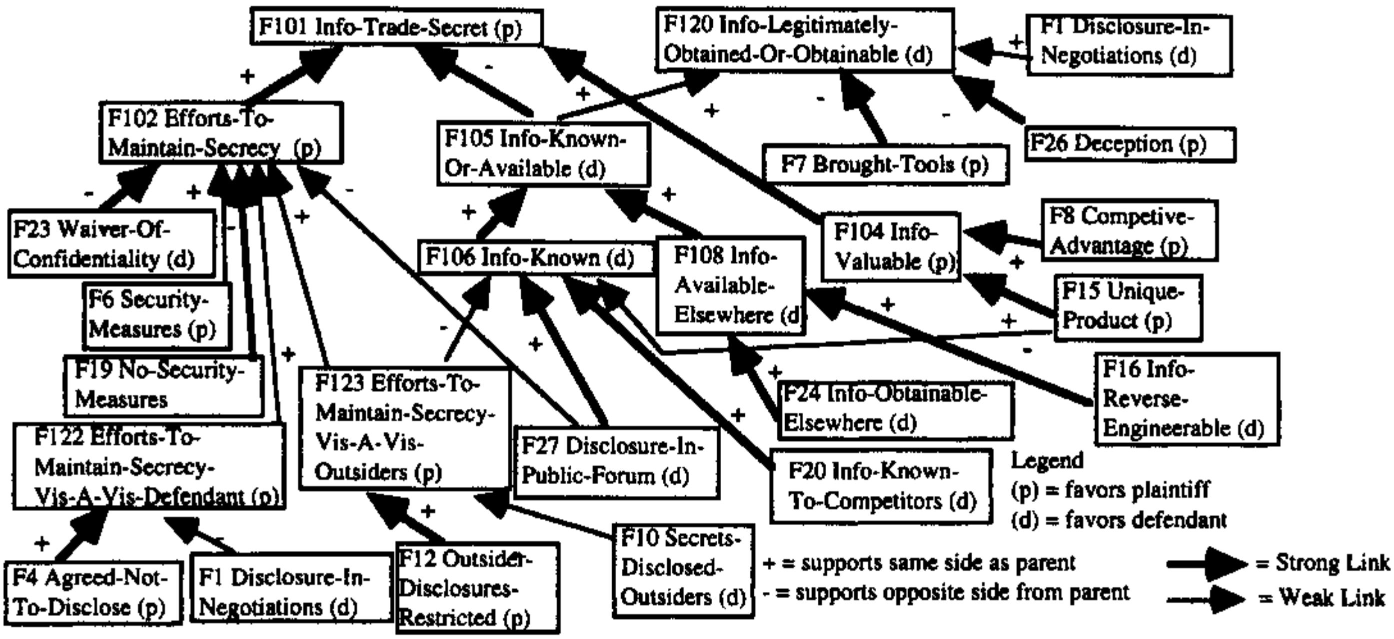

Figure 4; Excerpts from CATO's Factor Hierarchy

Blocking Interpretations: Link strengths in the Factor Hierarchy are designed to block certain interpretations in appropriate circumstances. The Hierarchy will characterize a similarity or difference in terms of only those abstractions not strongly contradicted by other factors. Evidence of blocking can be seen in Figure 5, Box 4, involving the *MBL* problem as modified by adding F7. As a result of blocking, CATO, arguing for defendant, does not characterize the *MBL* factors as evidence that defendant's means were not improper, as it does in Box 3. Where defendant brought plaintiffs product development tools, one does not want to argue that no improper means were used (F120) solely on the basis of plaintiffs disclosures to outsiders (F10) and competitors' having knowledge of the information (F20). F7 is itself evidence of improper means. F7 should block the interpretation that F10 and F20 are evidence of no improper means. In the Factor Hierarchy, Figure 4, a strong negative link from F7 to F120 blocks the weak positive link from F105 through which pro-defendant factors F10 and F20 in *MBL* are related to F120.

Strategic Interpretation: CATO's algorithm for emphasizing distinctions (and for downplaying them) takes strategic considerations into account in deciding which abstraction path to follow in the Factor Hierarchy and how abstractly to characterize a distinguishing factor. One wants to characterize a distinction broadly enough to take advantage of all relevant contrasting and corroborating factors. One also wants to focus on distinctions that hold up at more abstract levels and avoid characteriziations so broad as to destroy the contrast or allow obvious objections.

An example of a strategic choice using the Factor Hierarchy is choosing a narrower interpretation of a distinction where a broader interpretation does not lead to a contrast, as shown by comparing Boxes 3 and 4 in Figure 5. After hypothetically modifying the *MBL* problem so that the defendant took product development information (F7), CATO can still emphasize the F15 distinction but this time does not focus on the improper means aspect (F120), as it did before in the *MBL/Sperry* comparison (Box 3). This would not be an effective way to draw a contrast, since in the modified problem there is also evidence of improper means (as CATO now points out in the new argument downplaying F15). Instead, CATO again interprets F15 in terms of F106, an abstract factor "lower down" in the Factor Hierarchy (see Figure 4). CATO eschews a broader interpretation in favor of a narrower one that more clearly contrasts with the other case, even though that means certain corroborating factors in the broader interpretation cannot be used.

How does CATO use strategic considerations to decide which path to use and how far up or down the Factor Hierarchy to go in interpreting a distinction? The answer lies in heuristic policies embodied in CATO's algorithm for emphasizing a distinction. The policies ensure that a sufficient contrast between the cases exists, with respect to both the focal abstraction itself (steps 1.A and 1.B) and its ancestors in the Hierarchy (step 2.B).

As a result, CATO decides not to use F120 in Box 4 for emphasizing the distinction, because there is no suitable contrast with respect to this abstract factor. F120 must either not be supported for plaintiff in *MBL,* or supported for defendant as well. However, it is supported for plaintiff in *MBL* (by F7), but not supported for defendant. As noted above, pro-plaintiff factor F7 in *MBL* blocks the pro-defendant inference from F10 and F20 to F120. Abstract factor F106, however, does satisfy the requirements of both heuristic policies. The two cases have opposing factors related to F106 itself (pro-plaintiff F15 in *Sperry,* pro-defendant F10 and F20 in *MBL).* Also, they contrast suitably with respect to ancestors of F106: The path from F106 to F105 to F101 does not contain abstract factors supported for plaintiff in *MBL,* satisfying the second heuristic policy.

1. *MBL* / *Elcor* / F15 Unique-Product (p) **Plaintiff's argument downplaying distinction F15 in *Elcor*:** None. **Defendant's argument emphasizing distinction F15 in *Elcor*:** In *Elcor*, plaintiff was the only manufacturer making the product [F15]. This was not so in *MBL*. This is a marked distinction. It shows that in *Elcor*, the information apparently was not known outside plaintiff's business [F106], whereas in *MBL*, plaintiff's information was known outside plaintiff's business [F106]: Plaintiff disclosed its product information to outsiders [F10] and plaintiff's information was generally known in the industry [F20].	**2. *MBL* + F8 Competitive-Advantage (p) / *Elcor* / F15 Unique-Product (p)** **Plaintiff's argument downplaying distinction F15 in *Elcor*.** In *Elcor*, plaintiff was the only manufacturer making the product [F15]. This was not so in *MBL*. However, this does not amount to an important distinction. In *MBL*, defendant's access to plaintiff's information enabled it to develop its product in less time or at lower cost [F8]. It follows that in both cases, plaintiff's information was valuable for plaintiff's business [F104]. **Defendant's argument emphasizing distinction F15 in *Elcor*:** (Same as in Box 1)
3. *MBL* / *SPERRY* / F15 Unique-Product (p) **Plaintiff's argument downplaying distinction F15 in *Sperry*:** None. **Defendant's argument emphasizing distinction F15 in *Sperry*.** In *Sperry*, plaintiff was the only manufacturer making the product [F15]. This was not so in *MBL*. This is a clear and striking distinction. It shows that in *Sperry*, defendant may have acquired plaintiff's information through improper means [F120]. Other facts in *Sperry* also support this: Defendant offered a salary increase or bonus to induce plaintiff's employee to work for them and bring plaintiff's trade secrets [F2], plaintiff's former employee brought product development information to defendant [F7], and defendant obtained plaintiff's information through deception [F26]. By contrast, in *MBL*, defendant obtained or could have obtained its information by legitimate means [F120]: Plaintiff made disclosures to others outside its business [F10] and plaintiff's information was generally known in the industry [F20].	**4. *MBL* + F7 Brought-Tools (p) / *Sperry* / F15 Unique-Product (p)** **Plaintiff's argument downplaying distinction F15 in *Sperry*.** In *Sperry*, plaintiff was the only manufacturer making the product [F15]. This was not so in *MBL*. However, this difference is not significant. In *MBL*, plaintiff's former employee took documents, blueprints, or tools to defendant [F7]. In both cases, therefore, defendant may have acquired plaintiff's information through improper means [F120]. **Defendant's argument emphasizing distinction F15 in *Sperry*.** In *Sperry*, plaintiff was the only manufacturer making the product [F15]. This was not so in *MBL*. This distinction is highly significant. It shows that in *Sperry*, the information apparently was not known outside plaintiff's business [F106]. In *MBL*, by contrast, plaintiff's information was known outside plaintiff's business [F106]: Plaintiff made disclosures to others outside its business [F10] and plaintiff's information was known to competitors [F20].

Figure 5: CATO's Arguments Emphasizing / Downplaying Distinctions

4 Evaluation

An experiment with 30 first-year law students compared CATO's instructional effectiveness to that of an accomplished legal writing teacher teaching the same material in a traditional way [Aleven & Ashley, 1997]. The subjects were randomly assigned to experimental and control groups. All subjects read a traditional trade secret law casebook chapter.

The control group attended six classroom session of 50 minutes in small groups of about eight students. The human instructor discussed the casebook cases and presented a framework for analyzing trade secret problems. During two sessions, students made oral arguments about two problems in a moot court setting in which the instructor played the role of "judge". Students prepared for the moot court sessions in two 75 minute practice sessions outside of class.

In the experimental group's nine 50-minute sessions, students learned to use CATO's tools to address the argumentation problems at the end of each section of the casebook following instructions in four workbooks. Students worked with CATO in pairs. After sessions introducing students to CATO's tools and to factors and their interpretations, students used CATO to analyze problems in terms of factors, retrieved cases to test certain hypotheses about the importance of certain factors, and then worked with a legal dispute much like the *MBL* case. They determined which factors apply, retrieved and compared cases as in Figure 2, practiced argument moves with the problem like those in the argument recipe of Figure 3 and in emphasizing and downplaying distinctions as in Figure 5, organized and wrote a multi-case argument and compared it with CATO's argument (as in Figure 3 but incorporating actual cases.)

All subjects took a pre-test and post-test of Basic Argument Skills. We also administered a more advanced take-home post-test involving a memo-writing assignment, considerably beyond the sophistication of the CATO instruction, using a previous semester writing assignment grade as a control. The legal writing instructor graded the exams in a blind test.

	Basic Argument Skills				Memo Writing		
	Pre-Test		Post-Test		Prev.	Post-Test	
Exper. Gr. Avg.	60	C-	70	C+	63	70	B-
Control Gr. Avg.	55	D	68	C	63	79	B+

Table 1: Pre-test and post-test scores (maximum is 100)

As shown in Table 1, on the Basic Argument Skills tests, in both groups the improvement from pre- to post-test was statistically significant (t-test, $p < .05$). There was no significant difference between the two groups, and their improvement scores (i.e., post-test score minus pre-test score) did not differ significantly between the two groups (t-test, $p > .05$). On the memo-writing assignment however, the control group did better, and the difference was statistically significant. The previous semester's assignment showed no statistically significant difference.

5 Discussion and Conclusion

CATO instruction lead to a significant improvement in students' basic argumentation skills, comparable to that achieved by a legal writing instructor teaching small groups. The subjects were all from a special program for students judged most in need of individualized attention, and the human instructor, the director of that program, was experienced and successful in teaching such students.

The memo-writing post-test shows a limitation on CATO's efficacy. While both groups learned basic argumentation skills, the control group integrated these skills better in a complex assignment. The human taught skills indirectly and holistically by engaging students in oral arguments without focusing on elements of argumentation. By contrast, CATO taught students basic elements with examples gradually combining more elements. CATO's exercises to help students integrate the components came at the end; few students had enough time to practice them. As a result, we need additional time and techniques for integration. For example, with an LCD projector, instructors could demonstrate complex arguments and techniques with CATO in class which students could then practice with CATO at home.

CATO tackles some general problems of abstraction in case description. It illustrates, and the evaluation lends some support to the utility of, three techniques in comparing partially matched cases at multiple levels of abstraction, the use of: (1) factors to represent relevant similarities and differences, (2) a Factor Hierarchy to represent the significance of factors abstractly in terms of the purposes of the classification, and (3) a set of criteria for selecting focal abstractions. Implemented in CATO's algorithms for emphasizing and downplaying differences and for making issue-based arguments, these criteria enable CATO to select the right path and level of abstraction for characterizing a case.

CATO's approach helps frame the remaining problems of representing, recognizing and retrieving cases described at different levels of abstraction. Given the range of choice in characterizing a case abstractly, and the underlying strategic considerations, any multilevel case representation must preserve flexibility of interpretation. Multilevel descriptions also complicate the problem of evaluating which candidate cases are best. The value of a case lies in its uses in arguments. Multilevel characterizations expand the ways a case may be used in arguments and responses. Those possible uses need to be folded into the retrieval process. BankXX's heuristically-guided argument search employing complex argument evaluation functions [Rissland, et al. 1996], or factoring adaptation cost into retrieval [Smyth and Keane, 1994] may provide models. In an educational context like CATO's, however, the uses of retrieved cases in arguments need to be played out explicitly for students to understand and learn from them.

References

[Aleven & Ashley, 1997] V. Aleven and K. Ashley "Teaching Case-Based Argumentation through a Model and Examples: Empirical Evaluation of an Intelligent Learning Environment" To appear in *Proc. AI in Education.* Kobe, Japan. August.

[Aleven and Ashley, 1996] V. Aleven and K. Ashley. How Different is Different? Arguing About the Significance of Similarities and Differences. In Smith and Faltings, ed. *Advances in CBR.* Lecture Notes in AI 1168. pp. 1-15. Springer: Berlin.

[Ashley, 1990] K. Ashley, *Modeling Legal Argument, Reasoning with Cases and Hypotheticals.* MIT Press: Cambridge.

[Ashley and Rissland, 1988]. K. Ashley and E. Rissland. Waiting on Weighting: A Symbolic Least Commitment Approach. In *Proc, AAA1-88.* pp. 239-244. Morgan Kaufmann: San Mateo, CA.

[Branting, 1991]. Branting, L. K. *Integrating Rules and Precedents for Classification and Explanation: Automating Legal Analysis.* Ph.D. Dissertation. AI Lab., Tech. Rep. AI90--146. The University of Texas at Austin.

[Burton, 1985] Steven J. Burton. *An Introduction to Law and Legal Reasoning.* Little, Brown: Boston.

[Carbonell, 1983] Carbonell, J. G. "Derivational Analogy and Its Role in Problem Solving". In *Proc. AAAI83.* pp. 64-69. William Kaufmann: Los Altos, CA.

[Forbus, et al. 1994] K. Forbus, D. Gentner, and K. Law "MAC/FAC: A Model of Similarity-based Retrieval", *Cognitive Science* 19, 141-205.

[Kolodner, 1993] J. Kolodner. *Case-Based Reasoning.* Morgan Kaufmann: San Mateo, CA.

[Koton, 1988] P. Koton. *Using Experience in Learning and Problem Solving.* Ph.D. Thesis, MIT.

[Rissland, et al. 1996] E. Rissland, D. Skalak, and M. Friedman. "BankXX: Supporting Legal Arguments through Heuristic Retrieval". *AI and Law.* V. 4 No. 1. pp. 1-71. Kluwer: Dordrecht.

[Smyth & Keane, 1994] Smyth, B. and Keane, M. "Retrieving Adaptable Cases: The Role of Adaptation Knowledge in Case Retrieval". In *Topics in Case-Based Reasoning.* S. Wess, K-D. Althoff, M. Richter, eds. pp. 209-220 Lecture Notes in AI series. Springer: Berlin.