### Compare and Contrast, A Test of Expertise 1

# Kevin D. Ashley and Edwina L. Rissland<sup>2</sup> Department of Computer and Information Science University of Massachusetts Amherst, MA 01003

#### Abstract

In this paper we present three key elements of case-based reasoning ("CBR") and describe how these are realized in our HYPO program which performs legal reasoning in the domain of trade secret law by comparing and contrasting cases. More specifically, the key elements involve how prior cases are used for: (1) Credit assignment of factual features; (2) Justification; and (3) Argument in domains that do not necessarily have strong causal theories or well-understood empirical regularities. We show how HYPO uses "dimensions", "case-analysis-record" and "claim lattice" mechanisms to perform indexing and relevancy assessment of past cases dynamically and how it compares and contrasts cases to come up with the best cases pro and con a decision.

#### I. Introduction

It is one thing for an expert to analyze a problem situation and another to compare it to similar situations and explain why they are the same or different. If a human expert could perform only the former task, we might well doubt his level of expertise. Critically comparing a situation to other cases - showing why they are the same or pointing out the crucial differences - is an important component of explaining, arguing and planning. One could not reason analogically without it. Only by focussing on important differences, as well as similarities, can one choose the best cases, avoid the worst cases or extrapolate from cases not so on point. Despite the importance of this crucial intellectual skill, most expert systems do not represent cases or have the control structure to facilitate comparing cases. Research in Case-Based Reasoning ("CBR") focusses on that deficit and how to correct it.

# II. CBR Involves Critically Comparing Cases

A case-based approach to reasoning has three basic elements:

- 1. Credit Assignment: A decision-maker decides a case because of some factual features and inspite of others. In other words, the decider assigns credit or blame to some of the case's factual features. In effect, the decision of a case: (a) Selects certain features that are important enough for purposes of credit assignment (Not all facts make a difference to the outcome.); (b) Clusters the selected features; and (c) "Weights" them. Features in the cluster that favor the decision are ranked higher than those against it. In this way a prior case represents "experience".
- 2. Precedential Justifications: That a prior case (i.e., a precedent) had a certain cluster of features, and that its decision was made because of some of those features and inspite of others, is treated as a basis for a justification for coming to the same conclusion in a future case with a similar combination of features. By assumption, a precedential justification is a reason for coming to a decision in a subsequent case (and in fact prior cases will be cited in support of an argument that the new case should be decided, or that conflicting vfeatures should be resolved, in the same way as in the prior case.) Since the experience represented by prior cases matters for future decision-making, those cases need to be accessible for analyzing future cases.
- 3. Arguments: CBR is inherently adversarial; there seldom is one right answer. Instead there are arguments based on prior cases. CBR generates arguments presenting the possibly inconsistent alternative justifications. Although there are criteria for preferring some justifications over others, for telling good arguments from bad, and for making decisions accordingly, CBR's recommendations always must be viewed as presenting alternatives.

Given its elements, it is essential that a CBR system facilitate comparing new cases against old. Searching for justifications for deciding a new case is like searching through a space of prior cases for relevant precedents where the criterion for assessing relevance must take into account how useful a prior case will be in an argument about the new case. To make the search feasible, a CBR system must represent and record cases and organize them for efficient selection and comparison. In a word, this means indexing. The cases in the CKB should be indexed by the same features that are involved in credit assignment.

<sup>&</sup>lt;sup>1</sup>This work was supported (in part) by: the Advanced Research Projects Agency of the Department of Defense, monitored by the Office of Naval Research under contract no. N00014-84-K-0017, and an IBM Graduate Student Fellowship.

 $<sup>^2\</sup>mathrm{Copyright}$  ©1986. Kevin D. Ashley & Edwina L. Rissland All rights reserved.

With its emphasis on comparing a new situation to prior cases and comparing prior cases among themselves to find those that make the best justifications, CBR yields some important advantages: First, it is useful in domains that do not have a strong model. In domains like law, strategic planning, philosophical inquiry and historical political analysis, experts make reasoned decisions in spite of the facts that the rules are incomplete, use predicates whose meanings are not well defined (the open textured problem) or lead to inconsistent results. In these domains the expertise is simply organized differently along casebased lines. To the outsider, legal decision-making may seem arbitrary and chaotic, but, with its doctrine of case precedent, the law is an organized chaos. See [Levi, 1949]. Second, even in domains with strong models, case-based approaches are better-suited for a number of reasoning tasks involving explanation, persuasion and planning. We expect experts to be able to: 1. Explain their analysis of a situation by giving examples and posing hypotheticals to demonstrate the critical features, which if different, would have lead to a different conclusion; 2. Persuade us to believe the conclusion by: comparing the current situation approvingly to previous cases; extrapolating from less-similar cases (e.g., by pointing out differentiating features of the current situation that warrant the desired conclusion even more strongly); and posing hypotheticals to illustrate the dire consequences if the proposed conclusion is not adopted. 3. Plan for contingencies by posing hypothetical scenarios (worst, best, most recent, most likely cases, etc.) that illustrate the consequences of and alternatives to a given course of action.

Of course, a CBR approach has costs: 1. Constructing and maintaining the index; 2. dealing with the combinatorics of large numbers of cases and the depth of inferencing necessary to invoke the index; and 3. coming up with evaluation criteria for assessing justifications and arguments. For examples of recent research on these issues, see [Kolodner, 1983, Kolodner, Simpson and Sycara-Cyranski, 1985, Hammond, 1986a, Hammond, 1986b, Carbonell, 1983a, Carbonell, 1983b].

## III. The HYPO Program and its Domain

HYPO is a case-based reasoning program which operates in the domain of trade secret law [Rissland, Valcarce and Ashley, 1984, Rissland and Ashley, 1986, Rissland and Ashley, 1987]. HYPO accepts a fact situation from its user, analyzes it, retrieves other relevant cases from its Case-Knowledge-Base ("CKB"), considers various assignments of importance to facts, "positions" the retrieved cases with respect to the curent case, selects important most-on-point and most-dangerous cases, suggests interesting or critical hypotheticals, proposes the skeleton of an argument, and justifies this argument with case citations in the form demanded in legal scholarship [Ashley, 1986, Ashley and Rissland, 1987].

In HYPO, the main sources of legal knowledge are contained in HYPO's CKB and its library of dimensions. Dimensions represent the legal relationship between various clusters of operative facts and the legal conclusions they support or undermine. Dimensions provide not only indices into lines of cases and their attendant analyses and arguments but also a mechanism by which to judge the strength, or weakness, of a fact situation with respect to that line of reasoning. For instance, one line of trade secret cases focusses on the degree to which the "cat (i.e., secret) has been let out of the bag", even by the complaining plaintiff, himself: that is, how many disclosures of the putative secret were there and of what kind? This way of looking at a trade secret case (captured by the Disclose-Secrets dimension) provides one approach to resolving a misappropriation dispute and was used in the Data General and Midland Ross cases discussed below. Another approach might emphasize the competitive advantage gained by the defendant at the plaintiff's expense or the switching of a key employee from the plaintiff to the defendant Rissland and Ashley, 1986. Each dimension has: prerequisites, expressed in terms of factual predicates, that tell whether a dimension applies to a case or not; focal slots that single out the particular facts making a case stronger or weaker along the dimension and range information that tells how a change in the focal slot affects that strength (e.g., for Disclose-Secrets, the focal slot is the number of disclosees. Increasing that number weakens the plaintiff's position.) See generally [Ashley, 1986].

### IV. HYPO's Reasoning Process

Here is how HYPO reasons about a new fact situation (call it the current fact situation or cfs, for short). First, in analyzing a new cfs, HYPO runs through the library of dimensions and produces a case-analysis-record that contains: (1) applicable factual predicates; (2) applicable dimensions; (3) near-miss dimensions; (4) potential claims and (5) relevant cases from the CKB. Near-miss dimensions are those for which some, but not all, of the prerequisites are satisfied. The combined list of applicable and near-miss dimensions is called the D-list. Figure 1 describes a cfs based, for purposes of illustration, on Crown Industries, Inc. v. Kawneer Co., 335 F.Supp. 749 (N.D.Ill., 1971). Figure 2 shows the case-analysis-record for the cfs.

Second, HYPO uses the case-analysis-record to construct the claim lattice, which is a lattice such that: (1) the root is the cfs together with its D-list; and (2) successor nodes contain pointers to cases that share a subset, usually proper, of the dimensions in the cfs's D-list. Figure 3 (a) shows the claim lattice actually generated by the HYPO program for analyzing the cfs of Figure 1 from the viewpoint of a trade secrets misappropriation claim. (There is a separate claim lattice for each possible claim.)

From 1962 to 1964, Crown Industries, Inc., the plaintiff  $(\pi)$ , developed a hydraulic power pack, PX-121, for automatic door openers. Crown complained that defendant  $(\delta)$  Kawneer Co. developed a competing product, PX-125, by misappropriating  $\pi$ 's trade secrets. Crown's power packs had been sold to and installed in five public retail establishments. Crown made disclosures about the power pack to a third party, and in 1963 and 1965 a Crown employee made disclosures concerning the pack to Kawneer. PX-121 did not have any unique features not generally known to the prior art. It took Kawneer six years to develop PX-125, from 1962 to 1968.

Figure 1: Current Fact Situation (cfs) based on Crown Industries, Inc. v. Kawneer Co.

The ordering scheme enables claim lattices to capture a sense of closeness to the cfs of cases in the CKB. Those sharing more dimensions are nearer to the cfs. Those nodes closest to the root whose subsets of the cfs's D-list do not contain near-miss dimensions can be considered most-on-point-cases "mopc's" to the cfs; leaf nodes are the least-on-point. All of the cases displayed are relevant to the cfs because they all share some legally important strengths or weaknesses with the fact situation as represented by the dimensions shared with the cfs.

Third, HYPO uses the claim lattice to identify the competing parties' mopc's. There are two pro-defendant (" $\delta$ ") mopc's in Figure 3 (a): Midland-Ross and Yokana. Since mopc's share the most legally important strengths and weaknesses with the cfs (i.e., mopc's are the closest analogies to the cfs), Midland Ross and Yokana are the most persuasive cases HYPO could cite for the defendant. (Crown Industries is also a mopc, but that is the very case on which the cfs is based. Eventhough it would be silly to cite a case in an argument about itself, it makes sense that HYPO regards a case as most on point to itself.)

#### **Applicable Factual Predicates:**

exists-corporate-claimant, exists-confidential-info, exists-disclosures ...

Applicable Dimensions: Disclose-Secrets

#### Near-Miss Dimensions:

Restricted-Disclose, Competitive-Advantage, Vertical-Knowledge

Potential Claims: Trade Secrets Misappropriation

Relevant CKB cites: See claim lattice, Figure 3 (a)

Figure 2: Case-Analysis-Record for CFS

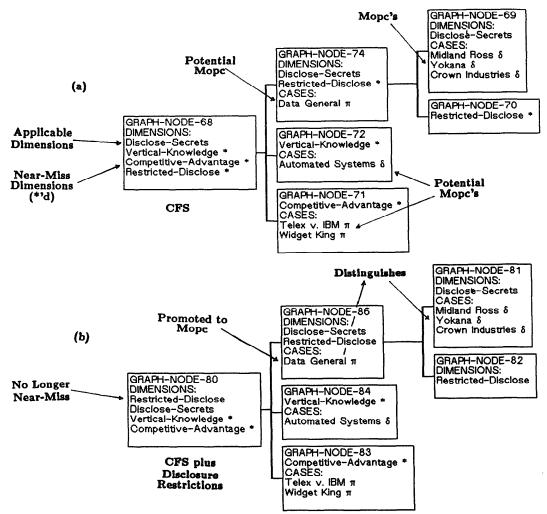
There are no pro-plaintiff (" $\pi$ ") mope's in Figure 3 (a). Data General, for example, is not a mope because, although it is very close to the root, the Restricted-Disclose dimension, which applies to Data General, and which would help  $\pi$  if it applied to the cfs, is only a near-miss for the cfs. (Restricted-Disclose is a near-miss because the cfs does not have the prerequisite factual predicate that some disclosees agreed to keep  $\pi$ 's secrets confidential. Note that Restricted-Disclose is \*'d in Figure 3 (a).) Although not a mope, the Data General case is potentially a mope for  $\pi$ . A potential mope is very similar to the cfs, except that some dimensions that apply to it are near-misses with respect to the cfs; Potential mopc's reside in nodes closest to the root. As shown below, if it were true that the disclosees had agreed to keep  $\pi$ 's confidential information secret, Data General would become a very important case to the plaintiff.

Fourth, HYPO uses the cases in the claim lattice to make and respond to precedent-citing arguments about the cfs. Different major branches of the lattice indicate different ways to argue the case, effectively one way for each group of mope's. HYPO a three-ply argument starting with a point for side 1, a response for side 2 and, possibly a rebuttal for side 1 again.

HYPO, for instance, can argue the case for  $side\ 1$ , the defendant (" $\delta$ ") in the cfs, by citing a pro-defendant mope, as in Figure 4 [a]. Recall that in Figure 3 (a) there are two such mope's,  $Midland\ Ross$  and Yokana. HYPO justifies the point expressly by drawing the analogy between the cfs and the cited cases by reciting the facts associated with dimension they have in common, Disclose-Secrets, namely that in both cases, plaintiff disclosed its secrets to some outsiders.

HYPO responds to points like that of Figure 4 [a] by distinguishing the cited case using three basic methods: (1) Comparing the strengths of cfs and cited case along the dimensions they share in common; (2) Finding strengths or weaknesses, represented by dimensions, that the cfs and cited case do not share. (3) Finding other cases that are more on point than the cited case. Figure 4 [b] is an example of the first method. HYPO distinguishes Midland-Ross on behalf of side 2, the plaintiff, by comparing values in the cfs and cited case of the focal slots of the shared dimension. HYPO knows from the claim lattice and the range information about the Disclose-Secrets dimension, that  $\mathit{Midland\text{-}Ross}$  presents a stronger case for  $\delta$  because  $\pi$  disclosed the confidential information to 100 outsiders; in the cfs, Crown disclosed to only five. HYPO supports the response by citing the Data General case where the plaintiff won despite having made many more disclosures than in Midland-Ross.

For a rebuttal, HYPO distinguishes any case cited in the response, as in Figure 4 [c]. Using the second method of distinguishing HYPO points out the pro- $\pi$  strength whose absence from the cfs makes Data General only a potential mope, namely that the disclosures were subject to restrictions to maintain confidentiality (a feature captured by the Restricted-Disclose dimension that applies to Data



The root node of claim lattice (a) represents cfs in Figure 1 and its D-list. (Dimensions that are near-misses for cfs have \*'s.) Successor nodes contain pro-plaintiff  $(\pi)$  or pro-defendant  $(\delta)$  trade secrets cases that are on point to cfs. Nodes closest to root that do not have near-miss dimensions contain mope's; otherwise they may contain potential mope's. Leaf nodes are least-on-point. Each major branch of lattice that contains mope's represents one way of arguing the cfs. Mope's distinguish cases in successor nodes. Potential mope's suggest fruitful hypothetical variants of cfs like that in (b). (b) is lattice for same cfs as (a) plus fact that disclosees agreed to treat  $\pi$ 's secrets as confidential. Argument for  $\pi$  is stronger in (b) than (a) because Data General: (1) has been promoted to being pro- $\pi$  mope (Restricted-Disclose dimension is no longer near miss in (b)); (2) is more on point than  $\delta$ 's mope's.

Figure 3: Two Claim Lattices.

General but is only a near-miss for the cfs.)

The fifth step in HYPO's reasoning process is to generate hypotheticals that are useful for testing the strengths and weaknesses of a party's position. HYPO uses its knowledge of how a case may be distinguished to suggest hypothetical modifications of the cfs that would strengthen or weaken the plaintiff's position [Rissland and Ashley, 1986]. For example, HYPO uses the relative positions of  $\pi's$  potential mope Data General and  $\delta$ 's mope Midland-

Ross in the claim lattice of Figure 3 (a) to suggest a hypothetical variant of the cfs in which  $\pi$ 's disclosures were made on a restricted basis. Then Data General can be used to distinguish Midland-Ross using the third method of distinguishing, significantly improving  $\pi$ 's position. Figure 3 (b) shows the claim lattice that would result for the modified cfs. The basic differences between the two claim lattices are that the Restricted-Disclose dimension, a nearmiss in Figure 3 (a) is an applicable dimension in Fig-

[b]  $\leftarrow$  For Side 2: ( $\Pi$ 's response to [a])

Distinguish: Midland-Ross (In Midland-Ross, II disclosed to 100 outsiders. II in cfs disclosed to only 7 outsiders.)

Cite: Data General  $(\Pi \text{ in } Data \text{ General } won \text{ eventhough } \Pi \text{ disclosed to 6000 outsiders, more than in } Midland-Ross.)$ 

[c]  $\hookrightarrow$  For Side 1: ( $\triangle$ 's rebuttal to [b])

Distinguish: Data General (In Data General disclosees agreed to keep secrets but not so in cfs.)

Figure 4: Citing & Distinguishing Precedents: 3-Ply Arguments

ure 3 (b) and that *Data General* has become  $\pi$ 's real mope and one that is more on point (i.e., closer to the root) than  $\delta$ 's mope's.

HYPO illustrates the new strength in the plaintiff's position by replaying the three-ply argument. Given the facts of the modified hypothetical in Figure 3 (b), HYPO can now generate a stronger response to the point in Figure 4 [a]:

[d] ← For Side 2: (II's response to [a])

Cite: Data General

( $\Pi$  should win because in *Data General*,  $\Pi$  won where  $\Pi$  disclosed secrets and disclosees agreed to keep disclosures secret.)

Distinguish: Midland-Ross, Yokana (Data General is more on point than these cases where disclosees did not agree to keep disclosures secret.)

### V. Comparing and Contrasting Cases in HYPO

Using information contained in the case-analysis-records and claim lattice, HYPO expressly compares and contrasts cases at three levels: (1) Facts; (2) Justifications; and (3) Arguments.

At the level of facts, HYPO compares the cfs to relevant cases from the claim lattice by focusing on the important facts they share as indicated by the dimensions they have in common. As we have seen, in making points, HYPO draws the analogy between the cfs and various cases by reciting these facts. HYPO contrasts cases when it responds to points by distinguishing the cited cases. Using

the first two methods of distinguishing (i.e., focussing on differing strengths along shared and unshared dimensions), HYPO is able to point out factual differences that justify *not* treating the cfs like a cited case.

At the level of justifications, HYPO compares relevant cases to each other using the claim lattice to see which make better precedents for deciding the cfs. Cases are compared in terms of: how on point they are relative to the cfs (mopc's vs. less on point cases); how useful they are in a legal argument about the cfs (e.g., using the third method of distinguishing to contrast a cited case with a more on point opposing case.); and how potentially useful they would be in a legal argument about the cfs (e.g., finding pro-opponent cases that can be used to distinguish mopc's).

HYPO makes comparisons at the arguments level by comparing the claim lattices. In moving from the cfs, Figure 3 (a) to the variant in (b), there has been a big shift in the balance of the argument in favor of the plaintiff, a comparative legal conclusion that HYPO can infer from a simple comparison of the claim lattices. One of HYPO's evaluation functions for comparing claim lattices involves simply comparing mope's. In Figure 3 (a) there are pro- $\delta$ mopc's but no pro- $\pi$  mopc, indicating a strong argument for the defendant. In Figure 3 (b), beside the same pro- $\delta$ mope's, there is a new pro- $\pi$  mope, Data General, which is more on point (i.e., closer to the root) than Midland-Ross or Yokana, indicating a strong argument for plaintiff. In other words, claim lattices can be used to evaluate the arguments in favor of a proposition, essentially by comparing the relationships of the pro and con mopc's.

# VI. Assessing HYPO's Performance

In its selection of *Midland-Ross* as defendant's best case, HYPO agreed with what the court actually did in its opinion in the case on which the cfs is based, *Crown Industries, Inc. v. Kawneer Co.* The court said,

Even though the Plaintiff's power packs, exemplified by PX-121, might have had to be rendered inoperative and examined by an engineer in order to discover the alleged trade secrets contained therein, the sale of the power packs nevertheless constituted a public disclosure which defeats a claim founded upon alleged misappropriation of the trade secrets allegedly contained in the power packs. *Midland-Ross Corp. v. Sunbeam Equipment Co.*, 316 F. Supp. 171, 177 (W.D.Pa. 1970), affirmed, 435 F.2d 159 (3d Cir. 1970).

HYPO's analysis of a cfs by comparing and contrasting it with mope's is similar to that actually performed by courts. Consider the opinion of the court in another case with similar issues to our cfs, National Rejectors, Inc. v. Trieman 409 S.W.2d 1, 40-42 (Sup. Ct. Mo., 1966):

[W]e do find some significant parallels between the facts of this case and those of Midland-Ross Corporation v. Yokana (D.C. N.J.), 185 F.Supp. 594 | The Yokana case involved the same plaintiff as Midland-Ross Corp. v. Sunbeam Equipment Co. and the same defense that plaintiff had disclosed its secrets to outsiders].... Thus the claim of trade secrets by National and by plaintiff in Midland-Ross have essentially the same basis. ... What was lacking in Yokana as in this case, was any evidence that, prior to defendant's competition, plaintiff considered the information which Yokana sought to use trade secrets. The court pointed out that plaintiff's blueprints in Midland-Ross were furnished plaintiff's suppliers and customers and potential customers. The court found an absence of precautions on the part of plaintiff to keep secret information regarding its machines.

Although the following cases do not parallel the present case as closely as *Yokana* our conclusion here is consistent with that reached in: [citing and describing other cases.]

Not only are the facts of *Midland-Ross Corporation v. Yokana* comparable to those in this situation, but we find the relief afforded in that case also appropriate in this....

#### VII. Conclusion

In this paper, we have presented three key elements of case-based reasoning (CBR): 1. That prior cases select and assign credit to factual features and weight conflicting features; 2. That prior cases are justifications for deciding a new fact situation (cfs) with similar combinations of features; and 3. That CBR yields arguments how to decide the cfs based on these potentially conflicting justifications. We have reviewed our indexing scheme based on "dimensions" that organizes cases in the Case-Knowledge-Base (CKB). HYPO performs indexing and relevancy assessment of past cases dynamically by (1) analyzing how prior cases can be viewed from the point of view of the cfs and (2) determing what aspects of these prior cases apply, and how strongly, to the cfs. This sort of analysis - accomplished through HYPO's dimensions, "case-analysis-record" and "claim lattice" mechanisms - allows HYPO to promote some prior cases over others as precedents for interpreting and arguing the cfs. HYPO compares and contrasts the cfs and prior cases at the levels of facts, justifications and arguments to come up with the best cases pro and con a decision and to pose instructive hypothetical variants of the cfs.

#### References

[Ashley, 1986] Kevin D. Ashley. Modelling Legal Argument: Reasoning with Cases and Hypotheticals - A Thesis Proposal. Project Memo 10, The COUNSELOR

- Project, Department of Computer and Information Science. University of Massachusetts, 1986.
- [Ashley and Rissland, 1987] Kevin D. Ashley and Edwina L. Rissland. But, See, Accord: Generating "Blue Book" Citations in HYPO. In Proceedings: First International Conference on Artificial Intelligence and Law, Northeastern University, 1987.
- [Carbonell, 1983a] J. G. Carbonell. Derivational Analogy and its Role in Problem Solving. In Proceedings of the Third National Conference on Artificial Intelligence, American Association for Artificial Intelligence, Washington, D.C., 1983a.
- [Carbonell, 1983b] J. G. Carbonell. Learning by Analogy: Formulating and Generalizing Plans from Past Experience. In Michalski, J.G. Carbonell, and T. Mitchell, editors, Machine Learning: An Artificial Intelligence Approach, Tioga Publishing, CA, 1983b.
- [Hammond, 1986a] Kristian J. Hammond. CHEF: A Model of Case-based Planning. In Proceedings of the Fifth National Conference on Artificial Intelligence, American Association for Artificial Intelligence, Philadelphia, PA, 1986a.
- [Hammond, 1986b] Kristian J. Hammond. Learning to Anticipate and Avoid Planning Problems through the Explanation of Failures. In Proceedings of the Fifth National Conference on Artificial Intelligence, American Association for Artificial Intelligence. Philadelphia, PA, 1986b.
- [Kolodner, 1983] Janet L. Kolodner. Maintaining Organization in a Dynamic Long-Term Memory. Cognitive Science, 7(4):243-280, 1983.
- [Kolodner, Simpson and Sycara-Cyranski, 1985] Janet L. Kolodner, Robert L. Simpson, and Katia Sycara-Cyranski. A Process Model of Case-Based Reasoning in Problem Solving. In Proceedings of the Ninth International Joint Conference on Artificial Intelligence, International Joint Conferences on Artificial Intelligence, Inc., Los Angeles, CA, 1985.
- [Levi, 1949] Edward H. Levi. An Introduction to Legal Reasoning. University of Chicago Press, 1949.
- [Rissland and Ashley, 1987] Edwina L. Rissland and Kevin D. Ashley. A Case-Based System for Trade Secrets Law. In Proceedings: First International Conference on Artificial Intelligence and Law, Northeastern University, 1987.
- [Rissland and Ashley, 1986] Edwina L. Rissland and Kevin D. Ashley. Hypotheticals as Heuristic Device. In Proceedings of the Fifth National Conference on Artificial Intelligence, American Association for Artificial Intelligence. Philadelphia, PA, 1986.
- [Rissland, Valcarce and Ashley, 1984] Edwina L. Rissland, E.M. Valcarce, and Kevin D. Ashley. Explaining and Arguing with Examples. In Proceedings of the Fourth National Conference on Artificial Intelligence, American Association for Artificial Intelligence, 1984.