A Predictive Role for Intermediate Legal Concepts

Kevin D. Ashley Stefanie Brüninghaus*
University of Pittsburgh - Intelligent Systems Program
Learning Research and Development Center and School of Law
Pittsburgh, PA 15260
{*ashley, steffi*}*@pitt.edu*

Abstract. Experiments described here demonstrate a role that intermediate legal concepts play in predicting the decisions of new cases. The experiments compare variations of each of an Issue-Based Prediction algorithm (IBP) and a CATO prediction algorithm in which the use of legal concepts are ablated in various ways. The results confirm those legal philosophical theories that assert that intermediate concepts in legal principles perform a guiding or extending role in deciding new cases.

1 Introduction

Legal philosophers have observed that intermediate legal concepts like "owner", "citizen", "territory", "trade secret", "breach of confidence" and others appear to serve as "vehicles of inference" between statements of legal grounds, on the one hand, and legal consequences, on the other. [11][p. 185]. According to the Scandinavian realist, Alf Ross, that was their only role [17]. In his famous example, he described a (fictional) South Seas island society with certain rules concerning *Tû-tû*, as in:

(1) If a person has eaten of the chief's food he is *tû-tû*.
(1′) If a person has killed a totem animal he is *tû-tû*. . . .
(2) If a person is *tû-tû* he shall be subjected to a ceremony of purification.

According to Ross, an assertion that someone is *tû-tû*, can be verified by proving the existence of one of the factual antecedents of the rules concluding *tû-tû* or whether the purification norm is applicable as a result. This is true, Ross observed, even though the word *tû-tû* is devoid of meaning, pure superstition and nonsense. In fact, one could omit using the term altogether! "It is plain that quite apart from what '*tû-tû*' stands for, or even whether it stands for anything at all, these two pronouncements, when combined in accordance with the usual rules of logic, will amount to the same thing as the following pronouncement:

(3) If a person has eaten of the chief's food he shall be subjected to a ceremony of purification." [17][p. 814].

Similarly, Ross argued that legal concepts like " 'ownership,' 'claim,' and other words, when used in legal language, have the same function as the word "*tû-tû*"; they are words without meaning, without any semantic reference, and serve a purpose only as a technique of presentation." [17][p. 814]. That is, they serve as a vehicle of inference while affording an economy of expression. For instance, there are many facts Fi from which one may infer

---

*We are grateful to Vincent Aleven for sharing his code for the CATO and HYPO-based prediction methods. This allowed us to reproduce his experiments and to generate examples for this paper.

Kevin D. Ashley and Stefanie Brüninghaus, 'A Predictive Role for Intermediate Legal Concepts' in D. Bourcier (ed.), *Legal Knowledge and Information Systems. Jurix 2003: The Sixteenth Annual Conference.* Amsterdam: IOS Press, 2003, pp. 153-162.

ownership O:

(I) F1 implies O, F2 implies O,. . . ,Fp implies O,

and many consequences Ci that follow from a conclusion of ownership O:

(II) O implies C1, O implies C2,. . . , O implies Cn.

Ross argued, " 'O' (ownership) merely stands for the systematic connection that F1 as well as F2, F3 . . . Fp entail the totality of legal consequences C1, C2, C3 . . . Cn. As a technique of presentation this is expressed then by stating in one series of rules the facts that 'create ownership' and in another series the legal consequences that 'ownership' entails." [17][p. 820]. Given this function, however, since the (logical) inferences are transitive, it does not matter what O stands for or whether it stands for anything at all. One could dispense with O altogether, but it is convenient not to do so given the n $\times$ p individual legal rules it abbreviates or "presents" with n + p rules [11][p. 193]; [18].

2 Intermediate Concepts as Guiding Decision of New Cases

Some legal philosophers have criticized Ross's theory for failing to account for the role that intermediate legal concepts play in guiding decisions of new or problematic cases. Recently, Lindahl has focused on the significant role that such concepts play in legal justification and argumentation. While conceding that some legal terms serve as vehicles of inference, others satisfying general structural schemes like (I) and (II) above are also accompanied by normative justifications. That is, they are "operative conditions" or "operative grounds" of jural relations (i.e., in the Hohfeldian sense such as having the duty to pay [11][p. 190]. See also [12]). So, for instance, Lindahl says, consider two empirical facts from which X's duty to pay Y may be inferred:

F1 "X says to Y 'I promise to pay you Z dollars', thereby making Y believe that X has a duty to pay him Z dollars."
F2 "X enters a bus run by Y without paying the stipulated fare of Z dollars."

The normative justification underlying the F1 inference is, "Responsibility for inducing legitimate expectations of a claim to get payment." For F2 it is, "Taking advantage of service, given only under the condition of payment." According to Lindahl, it is a mistake to subsume both justifications under an intermediate legal concept like "promising to pay" as in "F1 implies 'promises to pay'; F2 implies 'promises to pay'; 'Promises to pay' implies has duty to pay." Beside serving as a vehicle of inference, "promises to pay" is an operative condition; the kinds of empirical factual conditions it subsumes should be homogeneous from a normative or principled viewpoint. The principles underlying F1 and F2, however, are different. F1 involves inducing expectations. F2 does not. This leads to unfortunate results when judges start to speak of "implied" or "constructive" promises.

According to Lindahl, it matters therefore which intermediate legal concepts one chooses to serve as vehicles of inference. Collecting together various legal rules that embody the same principle not only facilitates generating transparent and normatively coherent explanations but the "justificatory rationale of legal norms, in its turn, provides guidelines for the handling of new and problematic cases, in arguments *ex analogia* and *e contrario*." [11][p. 199]. Similarly, Michael S. Moore has criticized Ross's view along somewhat similar lines. "If legal principles were abbreviatory of a mass of legal rules in the way Ross's analysis suggests, then [they] . . . cannot serve the creative or extending function." He also notes, "The same holds for reductive analyses in science: being able to derive a theory from an experimental law eliminates the power of the theory to generate new predictions and laws." [14][pp. 885-6].

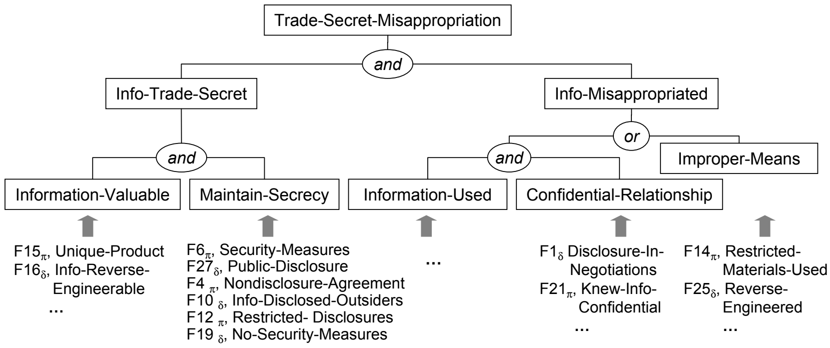

Figure 1: IBP's Domain Model

3 Using Prediction to Investigate Intermediate Legal Concepts

Some AI and Law researchers have had to grapple with the meaning of intermediate legal concepts in a very applied sense. They have developed techniques for representing such concepts and integrating them into their case-based models of legal reasoning. TaxmanII combined a template-like description of a legal concept, taxable income, and a set of possible mappings from one description into others [13]. HYPO represented case facts involving trade secret misappropriation claims in terms of Dimensions. It did not represent the claim's elements, in part because a subsequent court is free to reinterpret a prior court's holding on an issue as having been affected by other factors in the case not strictly related to that issue [3][pp. 238-241]. CABARET integrated logical reasoning with rules and dimension-guided comparisons of cases concerning the rules' terms into one legal reasoning process [16]. GREBE represented in semantic networks judges' explanations of why legal concepts, open-textured predicates extracted from statutory rules regarding workman's compensation law, applied to the facts of a case [8]. CATO's Factor Hierarchy provided legal reasons why trade secret factors mattered in terms of intermediate legal concepts, abstract factors based on the claim's issues [1]. Some complications in reasoning with factors within and across abstract factors, such as possible tradeoffs across issues, were discussed in [6]. A formal dialogue game framework has been presented that integrated cases, factors and logical reasoning with rules whose conclusions involved intermediate tax law concepts in [15].

We have investigated empirically the contribution intermediate legal concepts make to guiding, that is, predicting decisions of new cases. Originally developed for argumentation, the HYPO/CATO framework has been adapted to make predictions based on argumentation concepts [2]. The Issue-Based Prediction (IBP) algorithm [9] combines a logical representation of issues with a case-based reasoning (CBR) component for predicting and explaining case outcomes. Ablation experiments with these programs yield insights on the predictive contribution of legal concepts. An algorithm that takes legal concepts into account can be compared to a variant in which this knowledge has been turned off (i.e., ablated) with measurable results. In this way, one can begin to investigate empirically the role of intermediate legal concepts in guiding how to deal with new or problematic cases, the heart of Lindahl's and Moore's critiques.

IBP [9] uses CATO's Factor model as its case representation. A case is represented by a set of Factors, abstract fact patterns that strengthen a plaintiff's claim or a defendant's response, called pro-p or pro-d Factors, respectively. (We use subscript π and δ, as well as (p) and (d) to indicate which side is favored by a Factor.) IBP's knowledge about the intermediate legal concepts is captured in a logical structure called the Domain Model, Fig. 1. The cases all involve claims for trade secret misappropriation, and the intermediate concepts are the elements of, or issues raised by, the claim. These issues have been identified in the Restatement of Torts First, Sec. 757 and in the closely related Uniform Trade Secret Act (UTSA). In order

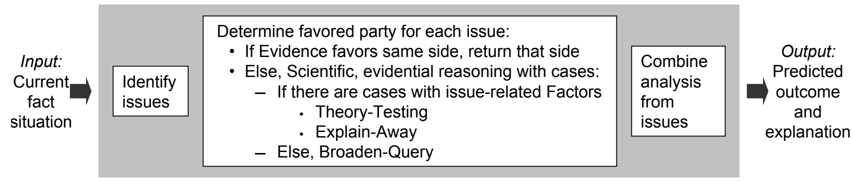

Figure 2: Overview of IBP

```
Prediction for MASON
  Factors favoring plaintiff: (F21 F15 F6)
  Factors favoring defendant: (F16 F1)
Issue raised in this case is SECURITY-MEASURES
  Relevant factors in case: F6(P) PLAINTIFF.
Issue raised in this case is CONFIDENTIAL-RELATIONSHIP
Relevant factors in case: F1(D) F21(P)
Theory testing has no clear outcome, try to explain away
exceptions.
Cases won by plaintiff:
  BOEING (F1 F4 F6 F10 F12 F14 F21)
  BRYCE (F1 F4 F6 F18 F21)
  DEN-TAL-EZ (F1 F4 F6 F21 F26)
  DIGITAL-DEVELOPMENT (F1 F6 F8 F15 F18 F21)
  FOREST-LABORATORIES (F1 F6 F15 F21)
  GOLDBERG (F1 F10 F21 F27)
  LASER (F1 F6 F10 F12 F21)
  LEWIS (F1 F8 F21)
  NATIONAL-INSTRUMENT (F1 F18 F21)
  VALCO-CINCINNATI (F1 F6 F15 F21 F10 F12)
Cases won by defendant:
  ECOLOGIX (F1 F19 F21 F23)
Trying to explain away exceptions won by DEFENDANT
  ECOLOGIX can be explained away because of
  the unshared ko-factor(s) (F23 F19).
Therefore, PLAINTIFF is favored.
[…]
Outcome of the issue-based analysis:
   For issue CONFIDENTIAL-RELATIONSHIP, PLAINTIFF is favored.
   For issue SECURITY-MEASURES, PLAINTIFF is favored.
   For issue INFO-VALUABLE, PLAINTIFF is favored.
=> Predicted outcome for MASON is PLAINTIFF, which is correct.
```

Figure 3: IBP's Output for Mason

to win a claim for trade secrets misappropriation, the plaintiff has to show that the information was a trade secret (Info-Trade-Secret in Fig. 1) and that the defendant misappropriated the information (Info-Misappropriated in Fig. 1). IBP's domain model follows the UTSA and the Restatement; to be a trade secret, the information has to have value for plaintiff's business (Info-Valuable) and the plaintiff has to take measures to keep it secret (Maintain-Secrecy). The Domain Model also represents that a defendant can misappropriate the trade secret by using information (Info-Used) conveyed in confidence (Confidential-Relationship), or by using improper means (Improper-Means) to gain access to the information. IBP's Domain Model does not logically connect the issues with case facts, represented by Factors. Rather, each issue has between 5 and 7 associated Factors favoring both sides. Case-based comparisons rather than logical inferences determine whether plaintiff or defendant is favored on the issue.

In IBP, prediction is guided by its knowledge of legal issues. When IBP predicts the outcome of a new case, it combines case-based and model-based reasoning (Fig. 2). It uses the Domain Model to identify the issues related to the case facts. It then determines for each of these issues, which side is favored. In this step, IBP applies CBR if necessary [9]. The algorithm makes its final prediction by combining the analysis of the issues following the logical structure in the Domain Model.

Each of the cases in our collection is represented as a list of the factors that, after a careful reading of the opinion, were deemed to apply in the facts of the case. With respect to both IBP and CATO-Arg, it should be noted that the case representation does not include a representation of the particular issues the judge identified in the opinion, only the applicable Factors. Thus, when either algorithm asserts that an issue is relevant in the case, it is a result of an inference drawn from IBP's Domain Model or CATO's Factor Hierarchy, respectively.

4 Prediction Algorithms

The experiments described below involve comparing two sets of prediction algorithms. The first set involves the issue-based prediction in the IBP algorithm and two variations, IBP-No-Issues and IBP-Model. The second set involves the case-based prediction in CATO-Arg and a non-issue variation, HYPO-Arg.

In order to illustrate IBP, we apply it to *Mason v. Jack Daniel Distillery*. In this case, the plaintiff, a restaurant owner, had developed a mixed drink. The drink was a popular item at his restaurant. In order to keep the recipe exclusive, the plaintiff shared the recipe only with his bartenders. He also instructed them to mix the drink outside of customers' view. As a result, one could not order the drink in any other restaurant, even though skilled bartenders probably would have been able to duplicate it. When the defendant, a sales representative for Jack Daniel, visited plaintiff's restaurant, he tasted the drink. Plaintiff told defendant that the recipe was a secret and then disclosed it to defendant. Plaintiff sued for trade secret misappropriation after he learned that the Distillery marketed his drink nationwide. The court decided that the defendant was liable for trade secret misappropriation, but only awarded plaintiff symbolic damages of one dollar. In our case representation, Mason has pro-p Factors $F6_{\pi}$, Security-Measures; $F15_{\pi}$, Unique-Product and $F21_{\pi}$, Knew-Info-Confidential, and pro-d Factors $F1_{\delta}$, Disclosure-In-Negotiations; and $F16_{\delta}$, Info-Reverse-Engineerable.

When Mason's Factors are given as input to IBP, it first uses its Domain Model to identify the relevant issues, Security-Measures, Confidential-Relationship, and Info-Valuable. These issues have a strong impact on the rest of IBP's procedure; the program analyzes each issue in turn and focuses on the issue-related Factors, rather than the entire set of Factors.

IBP has several techniques for analyzing issues. If the issue-related Factors favor the same side, this side is returned. In *Mason*, the only Factor related to Maintain-Secrecy is $F6_{\pi}$, Security-Measures (Fig. 3, Nr. 1). Thus, IBP concludes that plaintiff is favored for the issue.

If there are issue-related Factors for both sides, like $F1_{\delta}$ and $F21_{\pi}$ for Confidential-Relationship in *Mason* (Fig. 3, Nr. 2), IBP uses a kind of evidential scientific reasoning with cases to resolve the conflict. In its function Theory-Testing, the program tries to find cases where these conflicting issue-related Factors apply.

If all precedents were won by the same side, there is good reason to believe that in the current problem, this side is favored on the issue. The implied hypothesis is that the conflicting issue-related factors should be resolved in favor of the same side in the problem as, apparently, in the precedents. When there are examples but no counterexamples among the cases, the hypothesis is taken as confirmed.

Theory-Testing may retrieve cases favoring both sides. In *Mason*, Theory-Testing for Confidential-Relationship retrieves 10 cases won by plaintiff and 1 case, *Ecologix*, won by defendant (Fig. 3, Nr. 3). Here, IBP does not take a vote and conclude that the same side is favored on the issue that prevailed in the majority of cases. Rather, it formulates a hypothesis that the majority side should be favored (here the plaintiff) and tests the hypothesis against the retrieved cases. When it finds *Ecologix*, it tries to salvage the hypothesis by explaining away the counterexample. The applicable Factors in *Ecologix* indicate that the plaintiff did not keep the information secret, Factor $F19_{\delta}$, No-Security-Measures. $F19_{\delta}$ applies neither in *Mason* nor in the cases that support the hypothesis. It is a very dominant Factor; (almost) all cases where it applies were won by defendant. In IBP, Factors like $F19_{\delta}$ that strongly dominate case outcomes are called Knockout, or KO-Factors, and are used to explain away exceptions like *Ecologix* in Theory-Testing. If all exceptions can be explained away, IBP concludes that its hypothesis was confirmed, otherwise it abstains. In *Mason*, Theory-Testing is successful in confirming hypotheses regarding issues Confidential-Relationship and Info-Valuable (not shown). When IBP has analyzed all issues, it combines the analysis for its final prediction (Fig. 3, Nr. 4).

If the fact situation is too specific, there may not be precedents that share the whole set of conflicting issue-related Factors to test a hypothesis as above. Here, IBP applies a technique called broadening a query. That function determines if cases can be found to test a more general hypothesis from which the more specific but untestable hypothesis would follow *a fortiori*. An example of broadening can be found in the output of an IBP variation called IBP-No-Issues for *Scientology*, discussed next.

For these experiments, we set up an ablated version of IBP, called IBP-No-Issues, that did not include knowledge about issues. Instead of IBP's hierarchical Domain Model (Fig. 1), IBP-No-Issues knows only one "issue," whether plaintiff should win the claim of trade secret misappropriation, an "issue" that relates to all Factors. Apart from a minor technical modification for broadening a query, all functions in IBP-No-Issues are identical to IBP, and it uses the same cases. The *Scientology* case in Fig. 5 illustrates how IBP-No-Issues works. Its initial query for Theory-Testing retrieves no cases. It is broadened to retrieve precedents for a weaker hypothesis by dropping Factors. The hypothesis is that plaintiff is favored despite the pro-d Factors even when pro-p Factors are dropped. This weaker hypothesis cannot be confirmed, however; the program finds only cases won by defendant. A similar attempt to broaden the query for defendant fails, and the program abstains.

In a second variation of IBP, called IBP-Model, access to the cases was disabled, only the Domain Model is used for predicting outcomes. IBP-Model does not reason with precedents, thus it does not carry out Theory-Testing, Explain-Away and Broaden-Query. For each issue in a new problem, IBP-Model tests whether the issue-related Factors favor the same side. If so, it infers that this side is favored for the issue; otherwise it abstains for the issue.

The HYPO/CATO-based algorithms take a different approach to prediction. Given a problem, HYPO-Arg determines whether all of the most persuasive cases to cite in an argument (i.e., the best untrumped cases or BUCs) had the same outcome. If so, it predicts that outcome; otherwise it abstains. BUCs are the most relevant cases a side can cite without fear of the opponent's responding with a more relevant (i.e., trumping) counterexample. Here, relevance is measured by the inclusiveness of the set of factors a case shares with the problem [3].

The CATO-Arg algorithm improves upon HYPO-Arg by incorporating issues. Using a more refined relevance criterion that takes Factors and issues into account, it focuses on only cases not significantly distinguishable from the problem. That is, it focuses only on relevant cases whose distinctions from the problem can be downplayed and not emphasized. Given this reduced set of cases, it applies the BUC criterion to find the best cases on which to base its prediction. CATO can downplay distinctions, unshared Factors that underlie reasons for deciding cases differently, by trying to find alternative rationales to explain the conclusion favored by the distinction. It can also emphasize distinctions by pointing out alternative rationales that would support the opposite conclusion. These alternative rationales correspond to alternate paths from a distinguishing factor to an issue in its Factor Hierarchy [2], whose top-level issues are intermediate legal concepts based on the Restatement provision.

5 Experiments and Discussion

As noted above, we ran experiments comparing the predictive accuracy of the different algorithms. We used a database of 186 trade secret misappropriation cases of which 106 were won by plaintiff, the rest by defendant. All algorithms were tested in a leave-one-out cross-validation. For details on the experiment setup see [9]. Fig. 4 summarizes the results. The best performance in our experiment was achieved by IBP's issue-based prediction. IBP correctly predicted the outcome of 169 cases, made 14 mistakes and abstained once, which corresponds to an accuracy of 91.8% (Fig. 4, Nr. 1). The observed differences with the other algorithms were statistically significant [9].

A number of two-way comparisons between methods illustrate how adding knowledge about intermediate legal issues improves prediction. Since all algorithms in our experiment use the same Factor representation, results can be compared across algorithms.

a. IBP v. IBP-No-Issues: As discussed above, IBP-No-Issues does not have knowledge of the issues, whereas IBP makes inferences about cases using its Domain Model. In our

Nr.	Algorithm	Correct	Abstain	Errors	Accuracy
1	IBP	169	1	14	91.8%
2	CATO	143	22	19	77.7%
3	IBP-No-Issues	144	30	10	78.3%
4	HYPO	125	50	9	67.9%
5	IBP-model	132	38	14	71.7%
6	Baseline	106	0	78	57.6%

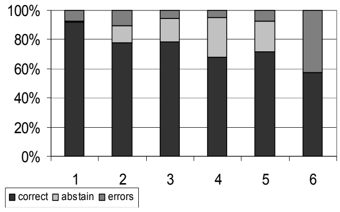

Figure 4: Experiment Results

experiments, IBP's performance is significantly better than IBP-No-Issues' (Fig. 4, Nrs. 1 and 3): accuracy is 91.8% compared to 78.3%. Because IBP abstains less frequently, it also has a much better coverage of the data set. This coverage comes at the cost of few extra errors; IBP has 4 more errors than IBP-No-Issues, but it more than compensates for this by the much higher rate of correct predictions (169 vs. 144). Both methods rely on the same set of cases, and use the same functions for reasoning with cases; the only difference is IBP's Domain Model. Thus, we conclude that adding knowledge about intermediate legal concepts leads to more accurate predictions.

```
Prediction for SCIENTOLOGY
  Factors favoring plaintiff: (F12 F6 F4)
  Factors favoring defendant: (F20 F10)
Issue raised in this case is INFO-TRADE-SECRET          1
  Relevant factors in case:
  F20(D) F4(P) F6(P) F12(P) F10(D)
Theory testing retrieved no cases, broadening query.
Query can be broadened for the plaintiff:
Dropping factor: F12
  Theory testing with Factors
  (F6 F4 F10 F20) gets the following cases:
  (MBL DEFENDANT F4 F5 F6 F10 F13 F20)
  (CMI DEFENDANT F4 F6 F10 F16 F17 F20 F27)
  In broadened query, DEFENDANT is favored.
Dropping factor: F6
  Theory testing with Factors
  (F12 F4 F10 F20) retrieves no cases.
[…]
Trying to broaden by dropping two Pro-P-factors
Dropping F6 F4
  Theory testing with Factors (F12 F10 F20) retrieves no cases.
Broadening query for PLAINTIFF retrieves cases
won by DEFENDANT, so the algorithm abstains.            2
Query can be broadened for defendant
Dropping factor: F10
  Theory testing with Factors (F20 F12 F6 F4) retrieves no cases.
Dropping factor: F20
  Theory testing with Factors (F10 F12 F6 F4)
  gets the following cases:
  (TRANDES PLAINTIFF F1 F4 F6 F10 F12)
  (FMC PLAINTIFF F4 F6 F7 F10 F11 F12)
  (BOEING PLAINTIFF F1 F4 F6 F10 F12 F14 F21)
  In broadened query, PLAINTIFF is favored.
Broadening query for DEFENDANT only retrieves cases
won by PLAINTIFF, so the algorithm abstains.
Trying to broaden for defendant: NIL
Trying to broaden for plaintiff: NIL                    3
Outcome of the issue-based analysis:
   For issue INFO-TRADE-SECRET, NIL is favored.
=> Predicted outcome for SCIENTOLOGY is ABSTAIN
```

Figure 5: IBP-No-Issues' Output for *Scientology*

The *Scientology* case shows how IBP's knowledge about intermediate issues focuses Theory-Testing on issue-related Factors, rather than on unproductive attempts to find cases that share all Factors, as in IBP-No-Issues. IBP's processing for this case (not shown) is straight forward. It identifies four issues. For three of them, the related Factors favor one side only. For the fourth issue, Security-Measures, Theory-Testing is used. The issue analyses are combined into a correct prediction for defendant. Fig. 5 shows IBP-No-Issues' output. This time, Theory-Testing does not find any cases with all of the Factors (Fig. 5, Nr. 1). IBP-No-Issues tries to broaden the query, which is not successful: dropping pro-p Factors retrieves cases won by defendant (Fig. 5, Nr. 2), and dropping pro-d Factors retrieves cases won by plaintiff (Fig. 5, Nr. 3). The more general hypotheses can not be confirmed. In sum, knowledge of intermediate legal concepts in IBP helps formulate testable hypotheses.

b. CATO-Arg v. Hypo-Arg: As discussed in Section 4, CATO-Arg employs additional knowledge about intermediate and higher-level issues in trade secret law represented in the Factor Hierarchy. This knowledge allows CATO-Arg to perform better and make more correct predictions; its accuracy is 77.7%, compared to 67.9% for Hypo-Arg. CATO-Arg also has much better coverage. However, this performance improvement comes at the cost of more errors, about twice as many as Hypo-Arg. Although this appears to put CATO-Arg at a disadvantage, that is not necessarily so. By abstaining frequently, Hypo-Arg can limit its predictions to the easier-to-predict cases, thereby avoiding mistakes for cases with harder factual conflicts. This may be a reasonable strategy in an experiment, but it is not very acceptable in practice. Thus, CATO-Arg's better coverage and accuracy outweigh the higher error rate.

The *Chicago-Lock* case illustrates how CATO-Arg's additional knowledge about Factors and issues allows it to base its predictions on more relevant cases, those with no significant distinctions and leads to better accuracy and coverage [2]. Hypo-Arg finds four cases that satisfy its relevance criteria, two won by defendant and two by plaintiff, *K&G* and *Technicon*, and therefore abstains. By contrast, CATO-Arg uses its knowledge of intermediate legal issues to conclude that neither *Technicon* nor *K&G* are relevant precedents. Both have significant distinctions that make them much stronger for plaintiff than *Chicago Lock*. In both cases, there was evidence that defendant knew the information was confidential ($F21_{\pi}$) and that defendant used materials that were explicitly marked as confidential ($F14_{\pi}$). CATO-Arg bases its correct prediction on the two remaining BUC cases without significant distinctions, *Secure-Services* and *Flotec*, both won by defendant.

c. IBP v. CATO-Arg: Perhaps the most interesting comparison is between IBP and CATO-Arg. The algorithms have a lot in common. Both use CBR techniques related to HYPO's claim lattices (Ashley 1990). Both have a representation of intermediate legal issues in trade secret law, IBP's Domain Model and CATO's Factor Hierarchy. The algorithms, however, use this knowledge and the CBR techniques in different ways.

Our experiments show that IBP's techniques are more successful for prediction. IBP has higher accuracy, 91.3% vs. 77.7%. At the same time, IBP has better coverage and fewer errors. We think that IBP is more accurate because of how it uses its knowledge about issues. In order to illustrate this hypothesis, consider the *Mason* case. By breaking up the case into issues, IBP can focus on subsets of conflicting Factors, which facilitates finding relevant precedents. As Fig. 3 shows, the cases retrieved by Theory-Testing are related to the respective issues. CATO-Arg, on the other hand, has a complementary strategy that considers the case overall in a gestalt-like way. CATO's relevance criteria can consolidate evidence from across issues, which may sometimes lead to reasonable arguments but cause problems for prediction. For instance, in the *Mason* example, CATO-Arg retrieves seven pro-plaintiff BUCs, but let us suppose, it also retrieves a hypothetical case *c*, won by defendant with the following factors: $F1_{\delta}$, $F4_{\pi}$, $F16_{\delta}$, $F18_{\pi}$, $F19_{\delta}$, and $F27_{\delta}$. CATO-Arg does not find any significant distinctions for case *c*, even though *c* has two strong pro-d Factors, $F19_{\delta}$ and $F27_{\delta}$. These distinctions, CATO-Arg finds, can be downplayed with the following cross-issue argument: In *Mason* and *c*, since the shared Factors $F1_{\delta}$ and $F16_{\delta}$ show that the information could have been discovered by proper means, plaintiff is not significantly worse off in *c* than *Mason* even though it publicly disclosed the information ($F27_{\delta}$). From a normative viewpoint this is a reasonable argument and, given the overwhelming support for plaintiff's position, quite an ingenious one for defendant to make. From the viewpoint of empirical prediction, however, it causes CATO-Arg to abstain from making a prediction for *Mason*.

d. IBP v. IBP-Model: In the previous comparisons, we argued that the Domain Model makes IBP more accurate, because it guides the process of making a prediction by focusing on the relevant issues in a case. This leads to the question whether the Domain Model by itself leads to high prediction accuracy. IBP-Model took only the intermediate legal concepts into account; it did not use precedents for making predictions. Its performance, however, was significantly worse than IBP's, with an accuracy of 71.7% (Fig. 4, Nrs. 1 and 5). IBP-Model abstains for 38 cases, including *Mason* and *Scientology*. In all these cases, the Factors related to an issue favor both sides, which can not be resolved using IBP's Domain Model alone. Interestingly, IBP's prediction is correct for all cases where IBP-Model abstains. This indicates that using knowledge about issues helps focus the case-based methods, in particular framing hypotheses in Theory-Testing on the relevant Factors in a case.

6 Conclusions

To summarize, our experiments support the following conclusions about the role of knowledge of intermediate legal concepts for prediction:

1. Using knowledge about issues to focus hypothesis-testing with precedents on the issue-related conflicts leads to better prediction. (IBP vs. IBP-No-Issues)
2. Using knowledge about issues to improve case comparison leads to better prediction. (CATO-Arg vs. Hypo-Arg)
3. Using knowledge about issues to focus hypothesis-testing with precedents on the issue-related conflicts rather than to improve case comparison leads to better prediction. (IBP vs. CATO-Arg)
4. Knowledge about issues does not by itself lead to a strong predictive model; its role in guiding hypothesis-testing with precedents is important. (IBP vs. IBP-Model)

It is, perhaps, no mystery why the integration of intermediate legal concepts improves predictive accuracy. In deciding trade secret misappropriation cases, judges are aware of the legal concepts in the Restatement and UTSA provisions. While the representations of the cases used to test the prediction algorithms do not include information about which concepts were most at issue according to the judge's opinion or how conflicts were resolved, IBP's Domain Model (and CATO's Factor Hierarchy) make it possible to reconstruct reasonable rationales relating factors to intermediate concepts to outcomes. In AI terms, the background knowledge about the intermediate legal concepts provides the "right" conceptual guidance for formulating hypotheses about which party should be favored. Without this knowledge-based guidance, the hypotheses are not as successful for prediction. Often they are not specific enough nor do they correspond to the likely significance and meaning of the Factors. Given the difficulties of representing judge's rationales concerning why intermediate legal concepts apply in a case as attempted in GREBE, it is interesting that such accurate predictions can be obtained even without representing the rationales.

The empirical evidence thus appears to be consistent with Lindahl's and Moore's arguments in reaction to Ross's *Tû-tû* example: intermediate legal concepts play an important role in guiding and predicting the outcomes of new cases.

Of course, it may be objected that Lindahl and Moore were speaking about normative guidance in deciding new cases, not merely facilitating better empirical predictions. Running experiments to isolate normative guidance provided by intermediate concepts in legal principles must await new developments. Currently no AI & Law representation relates intermediate legal concepts directly to abstract legal principles. IBP and CATO represent the meanings of intermediate legal concepts only through (1) the explicit relations between legal issues and Factors represented in IBP's Domain Model or CATO's Factor Hierarchy, and (2) the database of cases whose facts are represented by Factors and whose decisions are interpretable in terms of issues using these hierarchical models. SIROCCO demonstrates how cases operationalize the meanings of abstract normative principles [5]. Sartor and Bench-Capon's recent formalizations of legal argument as theory construction involve using cases and values to derive preferences among rules expressed in terms of Factors [7]. They do not appear to deal explicitly with intermediate legal concepts, but see [10]. Once such normative representations are available, researchers will still need to do more work in reading, representing and building databases of larger numbers of real-world legal cases incorporating the normative representations.

In the meantime, the experimental results in this paper imply that the guiding or extending function that Lindahl and Moore see for intermediate legal concepts reflects not only a normative component but a cognitive one, as well. For one thing, Lindahl was concerned about heterogeneous normative justifications; we observed cross-over instances where a case's strength in one issue may compensate for or interfere with the predictive influence of weakness in another issue. For another, the observation that intermediate legal concepts provide the right conceptual guidance for framing predictive hypotheses to test against actual cases should be important for certain theories of analogical legal reasoning. As explicated in [4], Sunstein's and Brewer's theories of relevance in legal analogies focus on the defining role of lower-level legal principles abduced in the process of comparing cases, and which may be informed by or related to more abstract normative principles. Parenthetically, these accounts seem to tie in well with Lindahl's focus on operative conditions in legal principles as guiding reasoning *ex analogia*. Conceptually guiding the framing of predictive hypotheses plays an important role in that process of abduction and testing.

References

[1] Aleven, V. (1997) *Teaching Case-Based Argumentation through a Model and Examples.* PhD Dissertation, University of Pittsburgh.

[2] Aleven, V. (2003) Using Background Knowledge in Case-Based Legal Reasoning: A Computational Model and an Intelligent Learning Environment. *Artificial Intelligence* In Press.

[3] Ashley, K.D. (1990) *Modeling Legal Argument, Reasoning with Cases and Hypotheticals.* MIT-Press.

[4] Ashley, K.D. (2002) An AI Model of Case-Based Legal Argument from a Jurisprudential Viewpoint. *Artificial Intelligence and Law* 10(1-3):163-218.

[5] Ashley, K.D. and McLaren, B. (2001) An AI Investigation of Citation's Epistemological Role. In *Proc. ICAIL-01*, 32-41.

[6] Bench-Capon, T. (1999) Some Observations on Modelling Case Based Reasoning With Formal Argument Models. In *Proc. ICAIL-99*. 36-42.

[7] Bench-Capon, T. and Sartor, G. (2001) Theory based explanation of case law domains. In *Proc. ICAIL-01*. 12-21.

[8] Branting, L.K. (1991) Building Explanations from Rules and Structured Cases. *International Journal of Man-Machine Studies* 34(6):797-837.

[9] Brüninghaus, S. and Ashley, K.D. (2003) Predicting the Outcome of Case-Based Legal Arguments. In *Proc. ICAIL-03*. 234-242.

[10] Greenwood, K., Bench-Capon, T. and McBurney, P. (2003) Towards a Computational Account of Persuasion in Law. In *Proc. ICAIL-03*. 22-31.

[11] Lindahl, L. (2003) Operative and Justificatory Grounds in Legal Argumentation. *Associations* 7(1):185-200.

[12] McCarty, L.T. (2002) Ownership: A Case Study in the Representation of Legal Concepts. *Artificial Intelligence and Law* 10(1-3):135-161.

[13] McCarty, L.T. and Sridharan, N.S. (1981) The Representation of an Evolving System of Legal Concepts: II. Prototypes and Deformations. LRP-TR-11. Laboratory for Computer Science Research. Rutgers University.

[14] Moore, M.S. (1997) Legal Principles Revisited. *Iowa Law Review* 82:867.

[15] Prakken, H. and Sartor, G. (1997) Reasoning with Precedents in a Dialogue Game. In *Proc. ICAIL-97*, 1-9.

[16] Rissland, E.L. and Skalak, D.B. (1991) CABARET: Statutory Interpretation in a Hybrid Architecture. *International Journal of Man-Machine Studies* 34(6):839-887.

[17] Ross, A. (1957) *TÛ-TÛ*. *Harvard Law Review* 70(March):812-825.

[18] Wedberg, A. (1951) Some Problems in the Logical Analysis of Legal Science. *Theoria* 17:246-275.