LEGAL THEORY AND CASE-BASED REASONERS: THE IMPORTANCE OF CONTEXT AND THE PROCESS OF FOCUSING

Kenneth A. Lambert
Department of Computer Science
Washington and Lee University
Lexington, Virginia 24450
"lambert.k.a@p9955.wlu.edu"

and

Mark H. Grunewald
School of Law
Washington and Lee University
Lexington, Virginia 24450
"grunewald.m@p9955.wlu.edu"

1 Introduction

It is widely urged that legal theory is important to the development of intelligent computer systems in law [Gardner 1987 and Susskind 1987]. Yet why this is so, and particularly where it has led researchers in AI and law, is not as obvious as one might assume. In this paper, we report briefly on our work on OPINE (Office Practice INquiry Expert) a generic case-based reasoner for use in legal domains. We concentrate on how our reaction to the treatment of legal theory in the AI and law literature has affected our approach.

2 Concepts from Legal Theory

Two concepts drawn from legal theory dominate most discussions of legal expert systems: first, the distinction between "easy" and "hard" cases, and second, the distinction between "shallow" and "deep" models of law [Gardner 1987 and McCarty 1984]. Researchers in AI and law who have used these concepts would probably concede that the terms underlying them are not easily defined, and that rather than describing binary categories, they represent extreme points on a continuum. If this view is accepted, it appears that the most important function these concepts serve is in setting a theoretical standard for judging the performance of expert systems. Under this view, the closer an expert system comes to dealing with "hard" cases or reflecting a "deep" model of a legal domain, the closer the system conforms to the behavior one would anticipate from a legal expert.

A closely related issue, but one not discussed as explicitly in the literature, is the question of the "determinacy" of law [Greenawalt 1990, Kress 1989, and Solum 1987]. At one level, determinacy is simply a matter of the extent to which legal outcome is predictable. At a deeper level, it is a question of whether law is coherent and constrained at all, or is more a form of politics. While posing this jurisprudentially contentious issue in the context of AI and law may seem odd, we believe it is critically important. If law is in some fundamental sense indeterminate, then in an equally fundamental sense, attempts to use formal computational methods to simulate legal reasoning not only face the well-documented problems flowing from complexity and open-texturedness but also are in a most basic sense misguided.

We believe, however, that the unease that this issue may pose for researchers can be accommodated without deciding this basic jurisprudential issue. Most importantly, we believe such accommodation requires being more explicit about the context in which intelligent systems would operate. We thus turn to the question of context generally and then report how, in our work to develop a generic case-based reasoner, context has influenced our choices, particularly the development of a technique we call "focusing."

3 Context

By context we refer primarily to three domain settings in which an understanding of law is required: (1) advice-giving, (2) advocacy, and (3) adjudication. In the advice-giving setting, the lawyer is performing a function that is predominantly predictive. The lawyer is using his understanding of the relevant body of law to predict a likely outcome on the basis of the facts that have been presented. While only the processes of advocacy and adjudication can finally fix the facts, the lawyer in the advisory context

operates from a set of facts we call "office facts"--the factual distillation (case description) upon which the advice will be based. The facts may present an "easy" case, in which event the advice will be straightforward and crisp; or the facts may present a hard case, in which event the advice will be hedged and possibly prolix. In either event, it is the predictive judgment that controls the advice. Similarly, the likely outcome may be one that could be explained by an extremely simple model of the domain or one that could be explained, if at all, only by an extremely complex model of the domain. While the complexity of the explanation may thus be influenced by the complexity of the assumed model, the essence of the advice is nevertheless controlled by the same predictive judgment. Thus, a degree of predictability in the law is at least a necessary condition for advice-giving. In neither other setting is the role of prediction as strong.

With respect to advocacy and adjudication, a sense of relative soundness of argument or result, based to an extent on prediction, is present, but the craft of persuasion or exercise of discretion, respectively, are the dominant characteristics of the expert's performance. Choosing the set of arguments to make or the evidence to present, giving priority to arguments and evidence, structuring the individual arguments or evidentiary presentations, and deciding what authority to use in support of the arguments or what type of witnesses to use for the evidentiary presentation are the essence of advocacy. These actions follow an expert judgment that a set of facts plausibly could be the basis for the sought-after outcome, but then, in a formal proceeding, are pursued without immediate concern for the likelihood that the outcome will be obtained. Stronger arguments may be emphasized; weaker arguments may be abandoned, but these and similar advocacy choices are guided by an expert judgment of persuasiveness rather than a judgment as to likely outcome.

Similarly, while adjudication should, in part, reflect the predictive judgments experts would have made at the outset about the case and, in part, the persuasiveness of the advocacy choices made, it will also in all but the most ministerial matters involve the exercise of some discretion. The discretion may lie in deciding what are the true facts, in choosing between two equally persuasive lines of argument, or in applying an inherently discretionary legal standard. Most importantly, the discretion must be exercised; the case must be decided; the adjudicator cannot decline to decide because the outcome could not have been predicted or because the advocacy failed to persuade clearly one way or the other. These and other aspects of adjudicatory discretion, even if substantially constrained, sharply distinguish the adjudicative setting from either the advice-giving or advocacy settings and ultimately are at the heart of the problem of predictability.

4 The Literature and Context

The AI and law literature reflects only limited sensitivity to how these issues of context bear upon the design of a system or an evaluation of its performance. Our experience in developing LESTER [Lambert and Grunewald 1989] and our current work on OPINE leads us to conclude that not only is the specification of context important as a development strategy, but also that failure to be explicit about context often tends either to overstate or understate the prospects for developing useful legal expert systems.

Bench-Capon considers a range of legal expert systems that have been developed or proposed [Bench-Capon 1989]. He distinguishes between systems that rest upon deep models and systems that use only shallow models of legal reasoning. Among the deep model systems are those whose knowledge is represented as a formalization of legislation. An example is the British Nationality Act project [Sergot 1986]. Such systems are considered deep because their rules are based upon direct connections to the statutory terms and because they reason from statutorily-described causal chains rather than empirical associations. A far more sophisticated variant of such a system would not only formalize the relevant legislation, but also relevant non-legislative terms and concepts. These terms and concepts might come from other related legal domains or from specialized non-legal domains such as psychology or from the commonsense world.

By contrast, there are systems whose knowledge is represented as the distillation of the rules of thumb that guide experts in the relevant field, including legislative fields. An example is the Latent Damage Advisor [Capper and Susskind 1988]. Such systems are considered shallow because their rules are based upon only empirical associations with legal concepts or practical considerations drawn from the expert's experience, rather than being isomorphic with the law source, such as legislative language. Bench-Capon suggests that systems resting on shallow models and systems resting on deep models may be equally effective in determining the outcomes of routine cases. He reasons that the correspondence of the model to the immediate law source--the legislation or the case law--is unimportant for such cases because in either instance decisionmaking involves no more than mechanical application of well-understood rules. Basing the rules upon associations drawn from the expert's experience as opposed to the terms of the legislation itself only affects such matters as maintainability. Yet when it comes to determining outcomes in non-routine cases, Bench-Capon argues that only deep model systems can be effective. In these instances, formalizing the relevant legislation is not enough. The reasoner must consider terms and concepts from extra-legal domains and therefore must have a deep model of one or more of these domains as well.

Bench-Capon provides some context for these propositions by drawing a distinction between high-level adjudicators who perform a law-making as well as a law-applying function, and low-level adjudicators who essentially are administrative functionaries who exercise no discretion. High-level adjudicators need deep models; low-level adjudicators can make do with shallow models. Adjudication is the context, but how the design of the system should be affected by that context is not clear. Ultimately, the distinction between high and low level adjudicators serves primarily as a basis for associating shallow models with routine matters and deep models with non-routine matters. Why adjudication should have these associations and why other lawyering contexts may or may not have them are open questions.

While we recognize that Bench-Capon did not intend his categorizations as design criteria, we believe system developers must pursue the questions he and others leave open. Do the characteristics of adjudication at any level guide the choice between shallow and deep models? Do the characteristics of adjudication make the distinction between routine and non-routine (easy and hard) cases an important design consideration? Do the characteristics of legal activity in other contexts--most notably advising and advocating--entail different design considerations?

Before turning to these questions, we note that the systems Bench-Capon discusses are all rule-based systems--the dominant mode in legal expert systems. The natural fit in many legal domains of analogic reasoning from cases, as well as concern about depth-of-model limitations in rule-based systems, has lead some researchers to propose and develop case-based systems, such as HYPO [Ashley 1990 and Rissland and Ashley 1987]. More recently, members of the case-based reasoning community have proposed multiparadigm systems which incorporate rule-based reasoning as well as case-based reasoning. [Oskamp 1989 and Rissland and Skalak 1989]. The multiparadigm system seeks to take advantage of the efficiency of rule-based systems for addressing cases for which there is a well-defined set of necessary and sufficient conditions (easy cases) while employing a form of case-oriented analogic reasoning for addressing those for which there is no such set of conditions (hard cases). Some of the work of Ashley and Rissland is quite context specific. For example, HYPO is clearly intended for use in the advocacy setting as an argument generator. On the other hand, Skalak and Rissland's multiparadigm system, CABARET, does not appear to have as specific a context orientation and is described as a system for "statutory interpretation" [Skalak 1989]. We point out this divergence in some of the better known case-based work only to indicate our view that the need for context specification extends to case-based systems as well as rule-based systems.

5 Advice, Advocacy, Adjudication: Determinacy in Law

Earlier we suggested that an expert system intended to operate in the advisory setting is more dependent upon predictability in law than a system intended to operate in either the advocacy or adjudicatory setting. We now elaborate that position. First, truly easy cases--cases about which there would be no disagreement among experts as to outcome--pose no issue of relative predictability at all. By saying we can advise with absolute assurance how a judge will decide, we must also be saying there is no known technique of persuasion or non-corrupt exercise of discretion that could conceivably produce a different outcome. Passing the question of whether such cases even involve the skills that ordinarily would be classified as lawyering, we agree with Bench-Capon that any number of system designs could handle such cases. Second, assuming that one must move somewhat beyond such cases in order to have an "interesting" system, we believe the move is not to hard cases--cases about which experts would sharply disagree--but to cases somewhere between easy and hard. This middle area of cases is as much or more a part of the legal expert's professional diet than either easy or hard cases. In our view, this area of cases is both interesting and potentially fruitful for legal expert systems. And it is the area in which, for the advice-giving function, relative predictability is more important than in the advocacy or adjudication functions. Finally, for truly hard cases, we also agree with Bench-Capon that only the deepest models (deep models that incorporate other deep models) would hold promise, but we believe these systems, at best, lie in the distant future and that the enormous challenge that these systems present should not discourage the development of more modest, but nevertheless useful, systems [Berman and Hafner 1989].

To be considered interesting, a case-based reasoner in the advisory context, thus, must be capable of dealing meaningfully with a middle class of cases, but need not accomplish more. Rather than simply suggesting that this class of cases lies somewhere between the end points "easy" and "hard", we would define it as those cases as to which an expert in the domain would be comfortable in predicting outcome, subject to error resulting from unanticipated persuasive advocacy (argument or evidence) or arbitrary exercises of discretion. This definition seeks to take account of the contingencies associated with nature of advocacy and adjudication and goes beyond the class of cases in which all or virtually all experts would agree.

The essential system function, then, is to determine how similar a new case is to cases in the casebase with known outcomes, and to reason from there in an informally probabilistic fashion to "likely" outcome. Only if law, beyond easy cases, is determinant to an extent that such

predicted outcomes would be widely recognized by experts as sound will such a system succeed. Thus, the ultimate success of our work depends not only on the design of the system, but also on the existence of this level of determinacy in law in the domains in which the system would operate.

The key design objective is the capacity to accurately position a new case within a relatively small casebase that represents the knowledge that an expert in the particular legal domain would use to make at least an office evaluation of likely outcome, using "office facts" as the case description. While positioning is essentially a test of similarity, to make the most effective use of the "feel" for the domain the casebase represents and to simulate most realistically a central aspect of the reasoning process a legal expert uses, we alternately broaden and narrow the test for similarity to arrive at the "best" position for evaluating likely outcome. We refer to this process as "focusing." When the process is concluded, the system advises as to the likely outcome or reports that reliable likely outcome cannot be determined on the given facts.

6 OPINE: A Generic Case-based System for Evaluating Likely Outcome in an Advisory Context

OPINE has a single function--to provide evaluation of likely case outcome--and it is set in a particular legal context--that of advice-giving. The system employs paradigm-based reasoning to perform this task. As we have argued elsewhere [Lambert and Grunewald 1989], paradigm-based reasoning subsumes reasoning in legal domains that have either strong or weak notions of precedent. Thus, the casebase can be structured to reflect the added significance of a high level precedent in hierarchical schemes for decisionmaking or can be structured from purely hypothetical cases that collectively reflect the special significance of certain fact occurrences in non-hierarchical schemes.

The first task of such reasoning is to pick, from the infinite number of respects in which cases can be similar and dissimilar, a manageable set of respects or features that could support a conclusion that one case is so similar to another that it will likely have the same outcome. Without such constraint, one would be faced with the commonsense impracticability, if not the jurisprudential impossibility, of defining the entire set of features that any case in the domain can have, together with the full ranges of possible values that those features could take.

Taking this arbitrarily restricted, but practically necessary, case structure, one could in principle generate the complete set of cases belonging to the domain and produce thereby a case base containing one case that would be exactly the same as any possible test case in the domain. But the size of such casebases would be intractable for case structures possessing more than an extremely small number of features. Therefore, the next task of the reasoner is to install in the case base a set of cases considered typical of those one is likely to encounter in the domain. These can be either real or hypothetical cases that an expert concludes collectively capture the essence of the domain.

In essence, the process of carving out from the theoretical wholeness of a domain a constrained set of features and choosing a representative case base to operate on those features is the process of moving from the extent of legal indeterminacy that exists in that domain to an artificial determinacy that, if the system succeeds, will be the source of reliable and useful advice. The reliability and usefulness of that advice compared to the advice experts would give is, then, in a sense both a measure of the extent of determinacy in that domain and of the degree to which a computational process can accommodate it in a particular setting.

A tractable case base, however, is purchased at the price of some uncertainty. Only a small percentage of the possible cases in a domain will be exactly the same as a particular case in the case base. The reasoner is again faced with the task of finding cases similar enough to support a conclusion that their outcomes are similar. To accomplish this task we use a method of determining similarity, we call "focusing."

Assuming an exact case match cannot be found, the process of focusing first loosens the criteria of similarity in several ways to produce sets of cases that are similar in a broader sense. The outcomes of these cases are then evaluated using certain hueristics that reflect expert analysis, at this level of generality, to determine whether there is a pattern of outcomes sufficiently reliable for a prediction of likely outcome in the test case. Recognizing, however, that the determinacy that has been imposed on the domain by the limitations of the case structure itself distorts even the loose similarity a case set shares with the test case, the criteria of similarity are next tightened through various forms of weighting to produce case sets that are similar in a much narrower sense. Again heuristics are used to determine whether a sufficiently reliable pattern of outcomes is present. The presumably tighter sense in which these case sets are similar to the test case is, of course, diminished as a jurisprudential matter by the necessary arbitrariness of any weighting function. [Ashley and Rissland 1988]. Despite these weaknesses, the outcome patterns from the two levels of analysis may confirm one another with a sufficient degree of reliability to permit determination of a likely outcome for the test case. If the patterns are not confirming, the evaluation focuses on cases common to case sets at both levels of analysis. This

set of cases can also be enlarged or contracted by adjustment of the acceptable feature values or weights to permit evaluation of alternative case sets for reliable patterns of outcome. Utimately, focusing seeks to arrive at "optimal" similarity by comparing cases at both a level generality and a level of specificity that we believe simulates in meaningful ways the reasoning process used by legal experts to make predictive judgments in the advisory context.

7 Conclusion

We have conducted limited tests of OPINE using a modified version of the case structure and casebase from LESTER and are now developing a case structure and case base in a different legal domain to further test OPINE. Our experience in transferring the LESTER knowledge base leads us to conclude that with further refinement OPINE can be used to construct legal expert systems that will provide reliable and useful case evaluation for the advice-giving function. We are continuing this work.

References

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

[Ashley and Rissland 1988] Ashley and Rissland, Waiting on Weighting: A Symbolic Least Commitment Approach, in Proceedings AAAI-88, American Association for Artificial Intelligence, 1988.

[Bench-Capon 1989] Bench-Capon, Deep Models, Normative Reasoning, and Legal Expert Systems, Proceedings of the Second International Conference on Artificial Intelligence and Law, 1989, pp. 37-45.

[Berman and Hafner 1989] Berman and Hafner, The Potential of Artificial Intelligence to Help Solve the Crisis in Our Legal System, Comm. of the ACM, Vol. 32, No. 8, p. 928-38 (1989).

[Capper and Susskind 1988] Capper and Susskind, Latent Damage Law-- The Expert System, Butterworths, 1988.

[Gardner 1987] Gardner, An Artificial Intelligence Approach to Legal Reasoning, MIT Press, 1987.

[Greenawalt 1990] Greenawalt, How Law Can Be Determinate, 38 U.C.L.A. L. Rev. 1 (1990).

[Kress 1989] Kress, Legal Indeterminacy, 77 Calif. L. Rev. 283 (1989).

[Lambert and Grunewald 1989] Lambert and Grunewald, LESTER: Using Paradigm Cases in a Quasi-Precedential Legal Domain, Second International Conference on Artificial Intelligence and Law, 1989, pp. 87-92.

[McCarty 1984] McCarty, Intelligent Legal Information Systems: Problems and Prospects, 9 Rutgers Comp. and Tech. L. J. 265-94 (1984).

[Oskamp 1989] Oskamp, Walker, Schrickx, and van den Berg, Prolexs, Divide and Rule: A Legal Application, Second International Conference on Artificial Intelligence and Law, 1989, pp. 54-62.

[Rissland and Ashley 1987] Rissland and Ashley, A Case-Based System for Trade Secrets Law, Proceedings of the First International Conference on Artificial Intelligence and Law, 1987, pp. 60-66.

[Rissland and Skalak 1989] Rissland and Skalak, Interpreting Statutory Predicates, Second International Conference on Artificial Intelligence and Law, 1989, pp. 46-53.

[Skalak 1989], Taking Advantage of Models for Legal Classification, Second International Conference on Artificial Intelligence and Law, 1989, pp. 234-41.

[Sergot 1986] Sergot, Sadri, Kowalski, Kriwaczek, and Cory, "The British Nationality Act as a Logic Program," Comm. of the ACM, Vol. 29, No. 5, pp. 370-86 (1986).

[Solum 1987] Solum, On the Indeterminacy Crisis: Critiquing Critical Dogma, 54 U. Chi. L. Rev. 462 (1987).

[Susskind 1987] Susskind, Expert Systems in Law: A Jurisprudential Inquiry, Oxford Press, 1987.