A Model of Legal Reasoning with Cases Incorporating Theories and Values

Trevor Bench-Capon
Department of Computer Science
The University of Liverpool
Liverpool
UK
+151-794-3697
tbc@csc.liv.ac.uk

Giovanni Sartor
CIRSFID
Faculty of Law
University of Bologna
Italy
+39-051-229405
sartor@cirfid.unibo.it

Abstract

Reasoning with cases has been a primary focus of those working in AI and law who have attempted to model legal reasoning. In this paper we put forward a formal model of reasoning with cases which captures many of the insights from that previous work. We begin by stating our view of reasoning with cases as a process of constructing, evaluating and applying a theory. Central to our model is a view of the relationship between cases, rules based on cases, and the social values which justify those rules. Having given our view of these relationships, we present our formal model of them, and explain how theories can be constructed, compared and evaluated. We then show how previous work can be described in terms of our model, and discuss extensions to the basic model to accommodate particular features of previous work. We conclude by identifying some directions for future work.

1. Introduction

A primary focus of those interested in modelling legal reasoning in Artificial Intelligence and Law has been on reasoning with cases. Prominent examples of such work are McCarty's TAXMAN (McCarty and Sridaran, 1982, McCarty, 1995), HYPO (Rissland et al., 1984, Ashley, 1990), CABERET (Skalak and Rissland, 1991, Rissland and Skalak, 1991), BankXX (Rissland et al., 1996), CATO (Aleven, 1997) and GREBE (Branting, 2000). Attempts have also been made to capture reasoning with cases in rule based systems, (e.g. Schild and Herzog, 1993, Hage, 1997) and to model HYPO style reasoning in a rule based framework (Prakken and Sartor, 1998). In this paper we put forward a model of reasoning with cases which is intended to capture many of the insights to be found in this body of work.

A naive model of reasoning with cases, set up as a straw man in Frank, 1949, can be expressed as an equation, $R \times F = D$, intended to express that a decision, *D*, can be deduced by the application of a set of rules, *R*, to the facts of a particular case, *F*. Although the simplicity of this picture has its attractions, it is problematic in every respect. The facts of a case are not givens: cases need to be interpreted, and different lawyers will interpret them in different ways. The rules, intended to be derived from precedent cases, are also not in plain view; a case may interpreted in a variety of ways, and as Levi,1949, stresses, the interpretation of a precedent may change in the light of subsequent cases (see also Twining and Miers, 1991, p311 ff). Moreover, the rules that cases give rise to are inherently defeasible: when we come to apply them we will typically find conflicting rules pointing to differing decisions, so we need a means of resolving such conflicts. Thus none of describing the facts of the case, extracting rules from precedents and applying these rules is straightforward. To model reasoning with cases in a satisfactory way, we must account for all of the description of cases, the extraction of rules and the resolution of conflicts.

A better way of seeing reasoning with cases is to see it as a process of constructing and using a *theory*. As McCarty put it:

> “The task for a lawyer or a judge in a "hard case" is to construct a theory of the disputed rules that produces the desired legal result, and then to persuade the relevant audience that this theory is preferable to any theories offered by an opponent” (McCarty, 1995, p285).

We endorse this view, and the construction, evaluation and use of theories is the central point of our model. The arguments put forward when reasoning about cases can only be considered within a context: it is the theory constructed by the arguer that supplies this context.

Theory construction is intended to account for the interpretation required in determining the description of cases and the derivation of rules from precedents. But now we have the problem of how to deal with the conflicts amongst the rules that compose the theory. Since a decision *must* be made in every case, we need a way to prefer one rule to another. Where do these preferences come from? An answer can be found in the work of Berman and Hafner (1993, 2002). Their solution involves looking to the purposes of law. This idea was first mentioned in AI and Law in Gardner, 1987, drawing on jurisprudential work such as the Hart-Fuller debate (Hart, 1958 and Fuller, 1958). Gardener wrote (pp39-40)

"Every application of a predicate involves an ethical question as well as a question of meaning. To resolve the ethical question, it is insisted by Moore, Fuller and others that one must consult the purpose of the rule."

The basic idea is that the law is not arbitrary but exists to serve certain social ends. Rules derived from cases draw their justification from the fact that following them promotes some desirable end. Thus when rules conflict, we resolve this conflict through a consideration of the purposes served and their relative desirability. Precedent decisions record the ways in which conflicts have been resolved in the past and can be seen as revealing preferences amongst different purposes. Once revealed, these preferences can be used to resolve further disputes. This argument is also present in the jurisprudential work of Perelman, 1980. Perelman's stress is on the need to appeal to audience when presenting an argument, and that this appeal is grounded in the values which acceptance of the argument would promote or defend.

"If men oppose each other concerning a decision to be taken, it is not because they commit some error of logic or calculation. They discuss apropos the applicable rule, the ends to be considered, the meaning to be given to values, the interpretation and characterisation of facts" (Perelman, 1980 , p150).

These values, and the ordering of values, may vary from jurisdiction to jurisdiction, and may also change over time. One important role of judges is to articulate the values held by the society of which they are part, and their relative importance (for a fuller discussion of Perelman's ideas in the context of AI and Law, see Bench-Capon, 2001).

This is a second element that we wish to incorporate within our model, namely the grounding of rules on social values, which enables someone aware of these values to decide which argument should be preferred.

Throughout the paper we will illustrate our discussion with an example taken from Berman and Hafner, 1993, which consists of three cases involving the pursuit of wild animals. In all of those cases, the plaintiff (π) was chasing wild animals, and the defendant (δ) interrupted the chase, preventing π from capturing those animals. The issue to be decided is whether π has a legal remedy (a right to be compensated for the loss of the game) against δ or not. In the first case, *Pierson v Post*, π was hunting a fox on open land in the traditional manner using horse and hound when δ killed and carried off the fox. In this case π was held to have no right to the fox because he had gained no possession of it. In the second case, *Keeble v Hickeringill*, π owned a pond and made his living by luring wild ducks there with decoys, shooting them, and selling them for food. Out of malice δ used guns to scare the ducks away from the pond. Here π won. In the third case, *Young v Hitchens*, both parties were commercial fisherman. While π was closing his nets, δ sped into the gap, spread his own net and caught the fish. In this case δ won.

The organisation of the paper is as follows. In section 2 we will give a fuller informal explanation of our view of the relationships between cases, features of the cases, rules based on cases and values grounding those rules. In section 3 we will give a more formal account of our model of these relationships, and of the theory construction aspects of reasoning with cases. The model we present is intended to be fairly neutral with respect to previous work, incorporating common aspects of that work. Specifically the model is intended to capture the analysis of Berman and Hafner, 1993. In this section we will also consider how the competing theories that might be constructed against a given background can be used to explain decisions, and be compared and evaluated. We then show how our model can be used to understand previous work by considering how various proposed argument moves can be related to the model. In the next section we discuss how the model can be extended to capture particular aspects previous work, proposing extensions to accommodate the notion of dimensions found in early HYPO work, and a factor hierarchy expressing relations between factors as found in CATO. Finally we identify directions for future work and make some concluding remarks.

2. Levels of Justification

To give a better explanation of the role of theories in legal reasoning we can consider the ways in which people can disagree in a given case. Suppose we have a case: we may immediately say that it should be found for one of the parties, say the plaintiff (if we chose the defendant it would make no difference to the following). If our position is accepted, well and good. But if our intuition is not shared, we will have to give reasons for our view. Typically this will involve citing features of the case which we believe are reasons for deciding for the plaintiff. Such reasons are often called *factors* in AI and Law. Thus we describe the case using terms which tend to support a decision for our view. The person disagreeing with us may now describe the case using factors of his own, which will this time be reasons to decide for the defendant. Such descriptions do not come "written on" the cases: they involve a degree of interpretation. At this point it is possible to argue over the factors that should be used to describe the case, but let us suppose that we have resolved this. We now have a case with a number of reasons to decide it one way and a number of reasons to decide it in the other way. How do we justify our position in the face of this?

At this point we must ascend a level and introduce *precedent cases*. Precedents represent past situations where these competing factors were weighed against one another, and a view of their relative importance was taken. On the assumption that new cases should be decided in the same way as precedent cases, if we can find a precedent with the same factors as we have in the current case, then we can justify our choice using this precedent. If no precedents exactly match or subsume the current case, we argue about the importance of the differences. It is at this level that HYPO-like systems operate: but while they identify the differences, they do not justify acceptance or rejection of the significance of these differences[1].

To justify these preferences we must ascend a further level. At this level we ask why a factor is a reason for deciding for a given party. We argue that this is because deciding for that party where that factor is present tends to promote or defend some value that we wish to be promoted or defended. The conflict is thus finally stated in terms of competing values rather than competing cases or competing factors. At this point the solution may be apparent: our set of factors may relate to values which subsume our opponent's values, or be accepted by our opponent as having priority. Beyond this we can only argue about which values should be promoted or defended, and so move beyond positive law, into the realms of politics and general morality. Disagreement is still possible, but is no longer a purely legal matter. Laws apply to a community, and this community is held to have common priorities amongst values, and one role of the judge is to articulate these values. Communities can change their values, but to disagree with the values currently adopted by one's community is to commit to effecting such a change, which is beyond the scope of precedent-based legal argument.

The picture we see is roughly as follows: factors provide a way of describing cases. A factor can be seen as grounding a defeasible rule. Preferences between factors are expressed in past decisions, which thus indicate priorities between these rules. From these priorities we can abduce certain preferences between values. Thus the body of case law as a whole can be seen as revealing an ordering on values. Figure 1 gives a diagrammatic representation of the process.

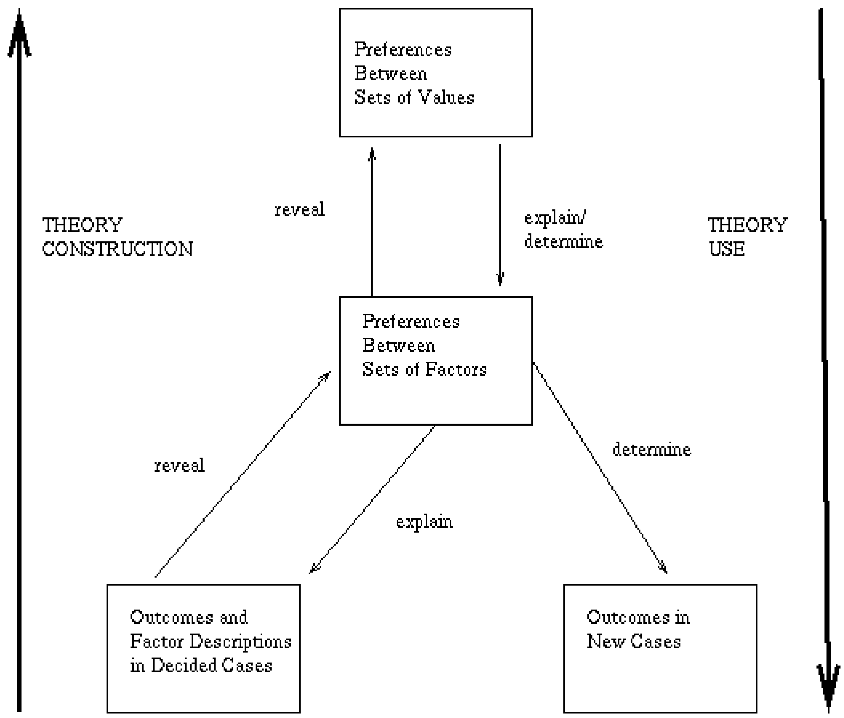

Figure 1: Construction and Use of Theories

Figure 1 depicts the three levels we need in our theory. Starting from decided cases (precedents), we construct the next levels by identifying the rule-preferences revealed in these cases, and the value preferences which these rule-preferences show. When the theory is constructed it can be used to explain the precedents and to yield a predicted outcome in new cases.

In the next section we will present our basic model of this process.

3. The Basic Model

[1] In CATO (Aleven 1997) an effort is made to supply some assessment of the significance of distinctions by introducing the notions of emphasising and downplaying distinctions. Even here, however, the arguments are indicated but the user is left to be persuaded or otherwise.

In this section we will describe the elements of a theory, provide a set of operators for constructing theories, and describe how theories can be used to explain past outcomes and predict new ones. Because it is possible to construct more that one theory, we need a way of comparing and evaluating theories. This topic will be discussed in section 3.4. We end this section by illustrating how our model can be used to illuminate previous work on reasoning with cases, by construing argument moves found in the HYPO and CATO systems in terms of our model.

3.1 Elements of a Theory

We assume that our theory construction process will start from the store of available knowledge, the *background*. This background will include six sets of elements: *cases*, *factors*, *outcomes*, *values*, *factor descriptions*, and *case factor-based descriptions*, which we denote respectively as C_{bg}, F_{bg}, O_{bg}, V_{bg}, Fds_{bg}, $Cfds_{bg}$.

The essential building blocks of the theories are decided cases. A case can be seen initially as a set of facts, together with a decision (an outcome) made on the basis of those facts. But this has not typically been found to be the most useful way of representing cases for case based reasoning purposes. Facts are in themselves neutral and not necessarily relevant to the outcome. Explanation of outcomes has usually therefore been in terms of *dimensions* (e.g. Ashley, 1990) or *factors* (e.g. Aleven, 1997). For discussions of the differences between factors and dimensions see Bench-Capon and Rissland, 2001 and Rissland and Ashley, 2002. We will specifically return to dimensions in section 4, but for the moment we will speak of factors, following Befman and Hafner, 1993, and take as our example the animals cases described in section 1. Factors are an abstraction from the facts, in that a given factor may be held to be present in the case on the basis of several different fact situations, and importantly factors are taken to strengthen the case for one or other of the parties to the dispute. In the above cases one such factor is whether plaintiff *p* can be deemed to have possession of his quarry. This abstracts from the hounds not yet having caught up with the fox, the ducks not yet having been shot and the fish still swimming in the sea rather than landed on the boat, to a single factor. That in none of the cases did *p* have contact sufficient to count as possession strengthens *d*''s position in each case. We make use of factors, and assume that a prior analysis of the cases has been carried out, which determines a set of applicable factors, and for each case whether the factor is present or absent. A variety of analyses of these example cases have been given in a number of papers, including Berman and Hafner, 1993, Hafner and Berman, 2002, Prakken, 2002, Sartor, 2002, Bench-Capon and Rissland, 2001, and Bench-Capon 2002.

In our example, we consider the cases described above (taken from Berman and Hafner, 1993): our cases background is

C_{bg} = {Pierson, Keeble, Young}.

As far as the set of outcomes O_{bg}, we consider only two possible outcomes: *P*, the outcome for *p*, indicating the recognition of a legal remedy to the plaintiff, and *D*, the outcome for *d*, indicating the denial of such a remedy. So our outcomes background is

$O_{bg} = \{P, D\}$.

For ease of later notation, let as denote as ~*o* the complement of outcome *o*, in particular when $o \in \{P, D\}$, ~*D*= *P*, and ~*P* = *D*.

As far as factors are concerned we identify four factors:

- *pLiv* = *p* was pursuing his livelihood (Keeble, Young), favouring *P*,
- *pLand* = *p* was on his own land (Keeble), favouring *P*,
- *pNposs* = *p* was not in possession of the animal (Pierson, Keeble and Young), favouring *D*
- *dLiv* = *d* was pursuing his livelihood (Young), favouring *D*.

So, our factors background is:

F_{bg} = {*pLiv, pLand, pNposs, dLiv*}.

We also need to link factors to values. We say that the reason a factor favours an outcome is because deciding for that outcome in a case where that factor is present promotes or defends some value, which it held that the legal system should promote or defend. In the example, following several of the analyses of the cases (e.g. Bench-Capon, 2002), the factor *pNposs* helps to promote clarity in the law and so discourage needless litigation; factor *pLand* helps promote the enjoyment of property; and factors *pLiv* and *dLiv* help to safeguard socially desirable economic activity. We thus have three values:

- *Llit* = Less Litigation

- *Prop* = Enjoyment of property rights
- *Mprod* = More productivity

So, our value background is:

$V_{bg} = \{Llit, Prop, Mprod\}$.

We need to associate with each factor the outcome favoured and the value promoted. We therefore represent information about factor f favouring outcome o in order to promote value v in the form of a *factor description* $<f, o, v>$. For simplicity, in this paper we assume that each factor promotes only one value, although the framework here introduced could, if desired, be straightforwardly extended to allow sets of values in factor-descriptions.

Definition 1. Factor description: *A factor description is a three tuple* $<f, o, v> \in F_{bg} \times O_{bg} \times V_{bg}$.

For the example, our background factor descriptions are:

$Fds_{bg} = \{<pLiv, P, Mprod>, <pLand, P, Prop>, <pNposs, D, Llit>, <dLiv, D, Mprod>\}$.

Note that Fds_{bg} contains, as it is typically the case, both factors favouring the plaintiff and factors favouring the defendant.

A way of representing cases can now be defined, which we call case factor-based descriptions, since cases are described through factors.

Definition 2. Case factor-based description: *A case factor-based description is a three tuple* $<c, F, o> \in C_{bg} \times Pow(F_{bg}) \times O_{bg}$.

Our background set of case factor-based descriptions is

$Cfds_{bg} = \{<Pierson, \{pNposs\}, D>, <Keeble, \{pLiv, pLand, pNposs\}, P>, <Young, \{pLiv, pNposs, dLiv\}, P>, <Young, \{pLiv, pNposs, dLiv\}, D>\}$.

Note that $Cfds_{bg}$ contains one description for each precedent (*Pierson* and *Keeble*) and two descriptions for the new case (*Young*) which has not been decided or is assumed to be so for the sake of the argument. This is, as we will see, to allow both parties to argue that *Young* should have the outcome they wish.

We can use these definitions to introduce some dependent notions. First of all, we view a rule as a connection between a set of factors and an outcome:

Definition 3. Rule: *A rule is a pair* $<F, o> \in Pow(F_{bg}) \times O_{bg}$.

In any rule $<F, o>$, we say that the set of factors F is the *antecedent* of the rule, and outcome o is its *consequent*. For example,

$<\{pLiv, pLand\}, P>$

is a rule having antecedent $\{pLiv, pLand\}$, and consequent P. A rule indicates that its antecedent (the presence of *all* factors it includes) is a reason for its consequent. We view rules as inherently defeasible. No suggestion that the presence of factors F conclusively determines outcome o is intended. By calling this connection between a reason and its output a "rule" we also do not intend to suggest that the rule prevents or excludes the consideration of other reasons (as in the notion of a rule used in Raz, 1975 and in Hage, 1997). Though those stronger, and more specific notions of a rule are frequently used in legal theory, and are relevant in many contexts, we do not need them to present our model.

In our model, rules are based on factors: the antecedent of a rule is formed from factors favouring the outcome which forms the consequent. From a given set of factor descriptions we can only construct those rules which link a set of factors having the same outcome, according to those descriptions, to that outcome. In particular, we consider that a rule is possible (or constructible) if and only if it is constructible from the given factor background.

Definition 4. Possible rule*: $<F,o>$ is a possible rule if and only if for each $f_i \in F$, $<f_i, o, v_i> \in Fds_{bg}$.*

Note that each factor in F must have the same outcome, but not necessarily the same value. We denote the set of possible rules as R_{poss} Among the possible rules, we call primitive rules those rules which correspond exactly to one factor (their antecedent is a set with only one element).

Definition 5. Primitive rule*: $<\{f\},o>$ is a primitive rule if and only if $<f, o, v> \in Fds_{bg}$.*

We now introduce a way of getting from rules to values. The idea is that following a rule promotes all values which are promoted by factors in the rule antecedent (when the outcome in the rule-consequent is followed).

Definition 6. rval*: The function rval: R_{poss} → $Pow(V_{bg})$, maps possible rules to sets of values: for all f ∈ F, v ∈ rval(<F, o>) if and only if there is a factor description <f, o, v> ∈ Fds.*

Thus following a rule *r* will promote all the values in the set returned by *rval(r)*. For example, *rval(<{pLiv, pLand},P>)* returns *{Mprod,Prop}, since* both *<pLiv, P, Mprod>* and *<pLand, P, Prop>* belong to Fds_{bg}.

We now define the notion of how a rule may attack another.

Definition 7. Attack*: A rule <F_1,o_1> attacks a rule <F_2,o_2> if and only if $o_1 = {\sim}o_2$.*

For example, <{*pLiv, pLand*}*, P*> attacks <{*dLiv*}*, D*>.

An attack may or may not succeed, depending on which rule is preferred. Preferences between rules are defined extensionally using the relation *rpref.*

Definition 8. Rule-preference: *A preference for rule r_1 over rule r_2, denoted as rpref(r_1,r_2,), is a pair <r_1, r_2> ∈ R_{poss} × R_{poss}.*

It is intended to be read as "r_1 is preferred to r_2". A rule-preference relation is a irreflexive transitive binary relation *Rpref* ⊆ R_{poss} × R_{poss}. One central feature of our theory construction model will be the analysis of the way in which parties build alternative preference relations. Note that preferences may exist between rules which do not attack one another. We can now define defeat:

Definition 9. Defeat: *A rule r_1 defeats a rule r_2 in regard to a set of rule preferences Rpref, if and only if r_1 attacks r_2 and not rpref(r_2,r_1).*

For example suppose that *rpref*(<{*pLiv, pLand*},*P*>, <{*dLiv*}, *D*>) ∈ *rpref.* Then <{*pLiv, pLand*},*P* > defeats <{*dLiv*}*, D*>.

Values are also preferred to one another. Moreover combinations of values can be preferred to other combinations of values.

Definition 10. Value preference: *A preference for value-set V_1 over value-set V_2, denoted as vpref(V_1,V_2,), is a pair <V_1, V_2> ∈ $Pow(V_{bg})$ × $Pow(V_{bg})$.*

A value preference relation is a irreflexive transitive binary relation *Vpref* ⊆ $Pow(V_{bg})$ × $Pow(V_{bg})$. Whether a rule is preferred to another rule or not depends on the values it promotes or defends.

Axiom 1. *rpref(r_1, r_2) if and only if vpref(rval(r_1),rval(r_2)).*

We are now in a position to define a theory:

Definition 12. Theory*: A theory is a five-tuple <Cfds, Fds, R, Rpref, Vpref>, where:*

- *Cfds* ⊆ $Cfds_{bg}$,
- *Fds* ⊆ Fds_{bg},
- *R* ⊆ R_{poss},
- *Rpref* ⊂ R_{poss} × R_{poss},
- *Vpref* ⊂ $Pow(V_{bg})$ × $Pow(V_{bg})$.

The theory thus contains descriptions of all the cases considered relevant by the proponent of the theory, descriptions of all factors chosen to represent those cases, all rules available to be used in explaining the cases, and all preferences between rules and values available to be used in resolving conflicts between rules. A theory is thus an explicit selection of the material available from the background, plus further components that are constructed from the selected background material.

3.2 Constructing Theories

We assume that at the outset all of *<Cfds,Fds,R,Rpref,Vpref>* are empty. The theory is then built up using a number of theory constructors. We will define these theory constructors in terms of their pre- and post-conditions. Essentially we need constructors to build up each element of the theory five-tuple. We begin by seeing how we can add cases.

Definition 13. Include-case:

Pre-condition:

- *current theory is <Cfds, Fds, R, Rpref, Vpref>, and*
- *<c, F, o> Î $Cfds_{bg}$*

Post-condition:

- *current theory is < $Cfds_{new}$, Fds, R, Rpref, Vpref>, with*
- *$Cfds_{new}$ = Cfds +[2] <c, F, o>.*

Essentially we can select any case in C_{bg}, and choose to include, from $Cfds_{bg}$, one of its possible descriptions. These are the cases that we aim to explain with our theory. Each party must include in his theory the current case, also called current situation, that is the case which is the object of the dispute. The current case has not yet been decided (or it is assumed so for the sake of the argument), and each party is claiming that it should be decided for their side. This is modelled here by assuming that two versions or the current case are contained in $Cfds_{bg}$, one with outcome *P* (to be included in *p*'s theories) and one with outcome *D* (to be included in *d*'s theories).

Cases bring with them factors, but we are not forced to consider in our theory all the factors associated with a case. We may believe some factors to be irrelevant. Levi, 1949 has shown that it is not always obvious which factors should be considered when describing a case. We must therefore explicitly include each of the factors we wish to consider.

Definition 14. Include-factor:

Pre-condition:

- *current theory is <Cfds, Fds, R, Rpref, Vpref>, and*
- *<f, o, v> Î Fds_{bg}*

Post condition:

- *current theory is <Cfds, Fds_{new} ,R_{new}, Rpref, Vpref>, with*
- *Fds_{new} = Fds + <f,o,v>*
- *R_{new} = R + <f, o>*

Note that a factor, if included in the theory, is always a reason for deciding for one party or the other. Therefore the factor brings with it its associated primitive rule.

Cases typically contain several factors favouring a given party. Therefore we need a way of extending primitive rules so that they can be tailored to particular cases. These rules will contain more antecedents, and thus in general represent more specific, and hence safer, reasons to decide for the favoured party than primitive rules. Factors can be merged only if they have the same outcome.

Definition 15. Factors-merging:

Pre-condition:

- *current theory is <Cfds, Fds, R, Rpref, Vpref>, and*
- *{<F_1,o> … <F_n,o>} Í R.*

Post-condition:

- *current theory is <Cfds,Fds, R_{new}, Rpref,Vpref>, with*
- *R_{new} = R + <{F_1È … È F_n},o>,*

Sometimes a case may lack some factors that were part of the antecedent of a rule used in a previous case. To make this rule applicable to the new case we must broaden it by dropping one or more of the antecedents. This is a common move in case based reasoning which we reflect in the following definition.

Definition 16. Rule-broadening:

Pre-condition:

- *current theory is <Cfds, Fds, R, Rpref, Vpref>, and*
- *<F_1, o> Î R,*
- *F_2 Ì F_1.*

Post-condition:

- *current theory is <Cfds,Fds, R_{new}, Rpref, Vpref>, with*
- *R_{new} = R + <F_2 ,o>.*

Note that the rule obtained by *rule-broadening* could also be built up from primitive rules using *factors-merging*. In a sense therefore, this theory constructor is superfluous. We have included it, however, because it represents a move very common in accounts of case based reasoning.

---

[2] We write $S + a$ to mean $S \cup \{a\}$, and $S - a$ to mean $S - \{a\}$.

A major role played by cases is to indicate preferences between rules. Assume that a theory T includes two conflicting rules, $<F_1, P>$ and $<F_2, D>$, with no preference between them, and a decided case $<c, F, P>$, to which both rules are applicable ($F_1 \subset F$, $F_2 \subset F$). As it stands, the theory cannot explain the decision, since the conflicting rules attack each other and, in the absence of preferences, the attack is successful. But we can now ask: what does the case tell us about the relative merits of the two rules? We believe that the case, interpreted in the light of theory T, tells us precisely that the first rule was preferred to the second in that case. This is what one must presuppose, if one believes that theory T was the basis of the decision in c, i.e. that it prompted the decision-maker of case c to decide for P. In other words, in the framework provided by T, one is authorised to assume or abduce that $rpref(<F_1, P>,<F_2, D>)$, since this is required if T is to explain the decision in c. This assumption is not arbitrary, but rather grounded on the *evidence* provided by precedent c (similar to the way in which scientific theories are grounded in the evidence provided by empirical observations). Accepting this preference between two rules also commits us to a preference for the values promoted by the preferred rule over those promoted by the defeated rule. We therefore introduce a theory constructor to include such abductions based on the evidence of previous decisions in our theories.

Definition 17*. *Preferences-from-case:

Pre-condition:

- *current theory is* $<Cfds,Fds,R,Rpref,Vpref>$, *and:*
- $<c, F, o> \in Cfds$,
- $<F_1, o> \in R$, *where* $F_1 \subseteq F$,
- $rval(<F_1,o>) = V_1$,
- $<F_2,\sim o> \in R$, *where* $F_2 \subseteq F$,
- $rval(<F_2, \sim o>) = V_2$.

Post-condition:

- *current theory is* $<Cfds, Fds, R, Rpref_{new}, Vpref_{new}>$ *with*
- $Rpref_{new} = Rpref + rpref(<F_1, p>,<F_2,\sim p>)$,
- $Vpref_{new} = Vpref + vpref(V_1,V_2)>$.

We can also use value preferences to derive rule preferences. If we know that a value is preferred to another value, we may deduce from Axiom 1 above, that rules promoting this value are preferred to rules promoting the other value.

Definition 18*. *Rule-preference-from-value-preference:

Pre-condition:current theory is

- $<Cfds,Fds,R,Rpref,Vpref>$ *and:*
- $\{r_1, r_2\} \subset R$
- $rval(r_1) = V_1$
- $rval(r_2) = V_2$
- $vpref(V_1,V_2) \in Vpref$.

Post-condition:

- *current theory is* $<Cfds,Fds,R,Rpref_{new},Vpref>$, *with*
- $Rpref_{new} = Rpref + rpref(r_1,r_2)$.

Sometimes we will simply wish to assert a preference between rules, even though this cannot be justified on the basis of previous cases, or existing preferences between values. In doing so we commit to expressing a preference amongst the corresponding values.

Definition 19. Arbitrary rule preference:

Pre-condition:

- *current theory is* $<Cfds,Fds,R,Rpref,Vpref>$, *and*
- $\{r_1, r_2\} \subset R$.

Post-condition:

- *current theory is* $<Cfds, Fds, R, Rpref_{new}, Vpref_{new}>$, *with*
- $Rpref_{new} = Rpref + rpref(r_1, r_2)$,
- $Vpref_{new} = Vpref + vpref(rval(r_1),rval(r_2))>$.

Similarly we may wish to assert a preference between values.

Definition 20. Arbitrary value preference:

Pre-condition:

- *current theory is* $<Cfds,Fds,R,Rpref,Vpref>$, *where*
- $\{<f_1, o_1, v_1>, <f_2, o_2, v_2>\} \subseteq Fds$.

Post-condition:

- *current theory is* $<C,Fds,R,Rpref, Vpref_{new}>$ *with*

- *Vpref* $_{new}$ *= Vpref+ vpref*(v_1,v_2).

These arbitrary preferences are often required to enable a theory to justify a position when no position is determined by previous cases. What they do is make quite explicit the preferences that are being used to justify that position. In so doing they can pinpoint points of disagreement between the disputants, which will be resolved when the case is decided.

The definitions 13 to 20 give us all we need to construct theories that can be advanced as explanations of particular case law domains.

3.3 Using Theories

The purpose of constructing a theory is to explain cases. We must therefore introduce the notion of explaining a case.

Definition 21. Explaining*: A theory <Cfds, Fds, R, Rpref, Vpref> explains a case c if and only if*

- $<c,F,o_1> \in Cfds$,
- $<F_1,o_1> \in R$,
- $F_1 \subseteq F$,
- *there is no rule* $<F_2,o_2> \in R$, *such that* $F_2 \subseteq F$ *and* $<F_2,o_2>$ *defeats* $<F_1,o_1>$.

Informally, the definition says that a case is explained if (*a*) we have a rule which allows us to conclude the outcome of the case on the basis of factors present in the case (as described in the theory) and (*b*) this rule is not defeated by any other rule in the theory whose antecedent is satisfied in the case. The overall aim of a disputant is to construct a theory that explains the current case, with the outcome desired by that disputant.

Let us illustrate this by constructing some theories to explain the three wild animal cases. We will suppose that *Young* has not yet been decided, that is, *Young* is our current case. If we wish to argue for the plaintiff, we will include the case with the outcome desired by the plaintiff, *<Young, {pLiv, pNposs,dliv}, P>*, in our theory, and then construct a theory which explains it. Conversely if we wish to argue for the defendant we will include *<Young, {pLiv, pNposs,dliv},D>* as the starting point of our theory.

A simple pro-defendant theory can be constructed using *include-case* to add *Pierson* and *include-factor* to add *pNposs* (for clarity we include the names of the theory components)*:*

***T*$_1$*:* *<cases: {<Young, {pLiv, pNposs,dliv}, D>, <Pierson, {pNposs}, D>},*
factors: {<pNposs, D, Llit},
rules: {<{pNposs},D>}
rule prefs: $\emptyset$*,*
value prefs: $\emptyset$*>*

This theory expresses the view that the plaintiff had no remedy (*D*) in *Pierson*, since he did not have possession of the animal (*pNposs*), which is indeed a reason for *D*, according to the rule *<{pNposs}, D>*, which is extracted from factor description *<pNposs, D, Llit>*. Exactly the same reasoning also explains why the plaintiff should have no remedy in *Young* also. No preferences are necessary: In T_1, *R* contains a single rule, and hence this rule is not attacked, and so cannot be defeated: it thus allows T_1 to explain both *Young* and *Pierson*.

The plaintiff can, however, produce a theory relying on *Keeble*, and subsuming T_1:

***T*$_2$*:* *<cases:* *{<Young, {pLiv, pNposs, dLiv}, P>,*
<Pierson, {pNposs}, D>,
<Keeble, {pLiv, pNposs, pLand}, P>},
factors: *{<pNposs, D, Llit>, <pLiv, P, Mprod>},*
rules: *{<{pNposs}, D>,<{pLiv}, P>},*
rule prefs: *{rpref(<{pLiv}, P>,<{pNposs}, D>)},*
value prefs: *{vpref(Mprod,Llit)}>.*

This theory is obtained, starting from T_1, by including *Keeble*, including factor *<pLiv, P, Mprod>* (*P* was pursuing his livelihood, favouring *P*, so as to promote value *Mprod*), and using *preferences-from-case* to get the required rule and value preferences from *Keeble*. Like T_1, T_2 implies that the plaintiff had no remedy in *Pierson* since he did not have possession of the animal. However, T_2 also implies that the plaintiff had a remedy (*P*) in *Keeble* since he was pursuing his livelihood (*pLiv*). Although the rule *<{pNposs}, D>* applies to *Keeble*, this rule is defeated, since *pLiv* supports *P* more strongly than not having possession of the animal (*pNposs*) supports

D (from the preference *rpref(<{pLiv}, P>,<{pNposs}, D>)*). According to the same reasoning, T_2 implies that *Young*, which shares with *Keeble* factors *pLiv* and *pNposs*, should also be decided for *P*.

Note that it is the rule-preference *rpref(<{pLiv}, P>,<{pNposs}, D>)* , derived from *Keeble*, which allows the rule *<{pLiv}, P>* to defeat the rule *<{pNposs},D>*. This means that the theory can explain why *Keeble* was decided for *P* and why *Young* should be decided in the same way. Note also that no description for the additional *D*-factor in *Young*, i.e. *dLiv*, has been included in T_2, and therefore this factor is not available to contest the explanation. Similarly, the theory does not consider the additional *P*-factor in *Keeble*, i.e *pLand* (*P* was on his own land). According to the proponent of T_2, neither of these factors is relevant.

The defendant can, however, make use of those factors and respond to T_2 in two different ways, depending on which of them he chooses to include. First he might add *Keeble* and factors *pLiv* and *pLand* to T_1 to get T_{3a}:

T$_{3a}$*: <cases: {<Young, {pLiv, pNposs, dLiv},D>,<Pierson, { pNposs},D>,*
<Keeble, {pLiv, pNposs, pLand}, P>},
factors: {<pNposs, D, Llit}, <pLiv, P, Mprod>, <pLand, P, Prop>},
rules: {<{pNposs}, D>,<{pLiv}, P>,<{pLand}, P>},
rule prefs: Æ,
value prefs: Æ>.

At this point, neither *Young*, nor *Keeble* is explained, since in the absence of preferences, rules attacking each other defeat each other (this is the case for *<{pNposs}, D*>, and either *<{pLiv}, P>*, or *<{pLand}, P>).* Clearly, the defendant does not want to explain *Keeble* as the plaintiff did, i.e. by using the rule *<{pLiv}, P>* with the preference *rpref(<{pLiv}, P>,<{pNposs}, D>).* This would lead, as we have just seen, to *Young* being decided for the plaintiff, on the basis of the same reasoning. He can, however, avoid that, by using *factors-merging* to add the rule *<{pLiv, pLand}, P>*, and *preferences-from-case* to add the preference derived from *Keeble*, taking into account these factors, *rpref(<{pLiv, pLand }, P>,<{pNposs}, D>*). In this way the theory distinguishes *Keeble* from *Young*: it explains why *Keeble* was decided for *P* without implying the same decision for *Young*. The plaintiff had a remedy (*P*) in *Keeble* since he was *both* pursuing his livelihood (*pLiv*) *and* on his own land (*pLand*), and the *combination* of these two factors supports *P* more strongly that not having possession of the animal (*pNposs*) supports *D* (according to the preference *rpref(<{pLiv, pLand},P>,<{pNposs},D>*)) . Note that the preference derived from *Keeble* is now different from that in the earlier theory: Keeble is explained by giving priority to the rule *<{pLiv, pLand},P>* rather than to the rule *<{pLiv},P>* . Therefore, the reasoning of *Keeble* cannot now be applied to *Young*, where there is only *pLiv* (and not *pLand*) to support decision *P*.

T$_{3b}$*: <cases: {<Young, {pLiv, pNposs,dLiv},D>, <Pierson, { pNposs},D>,*
<Keeble, { pLiv, pNposs, pLand},P>},
factors: {<pNposs, D, Llit}, <pLiv, P, Mprod>, <pLand, P, Prop>},
rules: {<{pNposs},D>, <{pLiv}, P>, <{pLand}, P>, <{pLiv, pLand},P>},
rule prefs: {rpref(<{ pLiv, pLand}, P>, <{pNposs}, D>)},
value prefs: {vpref({Mprod,Msec}, Llit)}>.

Unfortunately T_{3b} does not explain why *Young* should be decided for *D.* For this purpose, one would need the rule preference *rpref(<{pNposs},D>,<{pLiv},P>*), which would have to be either added arbitrarily or derived from the arbitrarily added value preference *vpref(Llit,Mprod).* (Remember that one's preference is arbitrary when it does not explain any precedent, but only supports the decision one wishes to have in current case.)

T_{3c}*: <cases: {<Young, {pLiv, pNposs,dLiv}, D>,<Pierson, { pNposs},D>,*
<Keeble, {pLiv, pNposs, pLand}, P>},
factors: {<pNposs, D, Llit}, <pLiv, P, Mprod>, <pLand, P, Prop>},
rules: {<{pNposs}, D>,<{pLiv}, P>,<{pLand}, P>, <{pLiv, pLand}, P>},
rule prefs: {rpref(<{ pLiv, pLand}, P>,<{pNposs}, D>), rpref(<{pNposs}, D>,<{pLiv}, P>)},
value prefs: {vpref({Mprod, Msec}, Llit), vpref(Llit, Mprod)} >.

T_{3c} suffices for the defendant, but the resort to arbitrary preferences is not desirable. A different tack for the defendant would be to ignore *pLand* and add *dLiv* instead to T_2.

T$_{4a}$*: <cases: {<Young, {pLiv, pNposs,dLiv}, D>, <Pierson, {pNposs}, D>,*
<Keeble, {pLiv, pNposs, pLand}, P>},
factors: {<pNposs, D, Llit}, <pLiv, P, Mprod>, <dLiv, D, Mprod>},
rules: {<{pNposs}, D>, <{pLiv}, P>, <{dLiv}, D>},
rule prefs: {rpref(<{pLiv}, P>,<{pNposs}, D>)},
value prefs: {vpref(Mprod, Llit)} >.

Now, by merging the primitive rules for *pNposs* and *dLiv*, introducing the value preference *vpref({Mprod,Llit}, Mprod)*, and using this to derive the rule preference *rpref(<{pNposs,dLiv}, D>,<pLiv, P>)*, an explanation of *Young* can be obtained.

T_{4b}: *<cases: {<Young, {pLiv, pNposs,dLiv}, D>, <Pierson, {pLiv}, D>,*
<Keeble, {pLiv, pNposs, pLand}, P>},
factors: {<pNposs, D, Llit}, <pLive, P, Mprod>, <pLand, P, Prop>},
rules: {<{pNposs}, D>, <{pLiv}, P>, <{dLiv}, D>, <{pNposs,dLiv}, D>},
rule prefs: {rpref(<{pLiv},P>, <pNposs}, D>), rpref(<{pNposs, dLiv}, D>,<{pLiv}, P>)},
value prefs: {vpref(Mprod, Llit), vpref({Mprod, Llit}, Mprod)}>.

Therefore, according to theory T_{4b}, *Young* should be decided for *D* since in *Young* the rule *<{pNposs,dLiv}, D>* is not defeated. This seems, according to Berman and Hafner, 1993, to be the theory used by the judges in *Young*. This explanation does rely on the introduction of a preference that is arbitrary, in the sense of not being supported by precedents. However it might be held that *vpref({Mprod, Llit}, Mprod)* is not entirely arbitrary on a different ground, namely since *{Mprod}*, is a subset of *{Mprod, Llit}*. The idea is that if all values are good, then a more inclusive set of values must be better that a less inclusive set (cf. Prakken, 2000 and Sartor, 2002)). This idea could be adopted into our framework by adding a theory constructor which allows one to introduce preferences for any set of values over its own proper subsets. We believe that this assumption is reasonable in many contexts, but possibly not in all, because of interferences between values: if two values are incompatible, then promoting only one of them can be better then promoting the two of them at the same time. So, we do not wish to enshrine this as a general and necessary feature of our approach, and since such preferences can always be introduced as arbitrary value preferences if desired, the relevant theory can still be constructed. None the less we would expect a preference of this sort to be acceptable in most cases, and for particular purposes we might want to use the additional constructor to allow such preferences to be distinguished from those which are merely arbitrary.

3.4 Evaluating Theories

In the above discussion we produced four theories, each of which would explain the decision in *Young*. How do we choose between them? Intuitively theories are assessed according to their coherence. We will not, however, even attempt to develop a precise notion of coherence in this paper. For coherence in law, there is a discussion in Alexy and Peczenik,1990 and for a general discussion of coherence and theory change, see Thagard (1992, 2001). For a recent attempt to develop some formal criteria with which to assess theories see Hage, 2000. In this paper we will do no more than indicate some considerations which might lead to one theory being preferred over another.

Firstly, we demand as much *explanatory power* as possible from our theories. In this context explanatory power can be approximately measured by the number of cases explained. More exactly, since different cases may have different weights (one case being more recent, or having been decided by a higher court, etc.) we should consider also the relative importance of the sets of cases that the competing theories can explain. We cannot consider here the details of the metrics for such a comparison, which is also dependant on the features of the legal system under consideration. At the very least, however, we can certainly say that theory T_1 has more explanatory power than theory T_2, if T_1 explains all precedents explained by T_2 and some others, so that the precedents explained by T_1 are a proper superset of the cases explained by T_2.

Secondly we can require theories to be *consistent*, in the sense that they should be free from internal contradiction. Note that we allow theories to include conflicting rules applicable to the same case, and we assume that these conflicts are solved through preferences. The contradictions we wish to avoid are those concerning rule and value preferences, i.e. the *rpref* and *vpref* relations. Thus we can require that theories do not contain both $rpref(r_1, r_2)$ and $rpref(r_2, r_1)$ in *Rpref*, and do not contain both *vpref(v,v')* and *vpref(v',v)* in *Vpref*. Such incoherence is explicit. There is also implicit incoherence when there is a value preference which would allow the introduction of a rule preference which would produce an incoherence in *Rpref*, or where the transitivity of the preference relations can be used to derive an explicit contradiction.

A third classically desirable feature of scientific theories is *simplicity*. This could be measured in terms of the number of factor descriptions in *F*. If we can explain a set of cases without introducing a given factor, this is a simpler theory than one which does include that factor. Suppose we extend T_{4b} above to include factor *pLand*.

T_5: *<cases: {<Young, {pLiv, pNposs,dLiv}, D>,<Pierson, {pNposs}, D>,*
<Keeble, {pLiv, pNposs, pLand}, P>},
factors: {<pNposs, D, Llit>, <pLiv, P, Mprod>, <dLive, D, Mprod>, <pLand, P, Prop>},
rules: {<{pNposs}, D>,
<{pLiv}, P>, <{dLiv}, D>,<{pLand}, P> , <{pNposs, dLiv}, D>,<{pLiv, pLand}, P>},
rule prefs: {rpref(<{pLiv, pLand}, P>,<{pNposs}, D>), rpref(<{pNposs,dLiv}, D>,<{pLiv}, P>)},
value prefs: {vpref({Mprod, Msec}, Llit), vpref({Mprod, LLit}, Mprod)}>.

Suppose we now have a new case in which the facts of *Keeble* are present, except that the plaintiff is hunting on common land. T_{4b} would explain a decision for the plaintiff, whereas T_5 would not explain either outcome. To explain an outcome for the plaintiff, T_5 would need the value preference *vpref(Mprod, Llit)* (T_{5a}), and to explain an outcome for the defendant, the value preference *vpref(Llit, Mprod)* (T_{5b}), so as to get the required preference between the rules *<{pLiv}, P>* and *<{pNposs}, D>*. In either case such an introduction would be arbitrary. We would therefore expect the plaintiff to rely on T_{4b}, whereas the defendant would advance the more complicated theory T_{5b}. If the case were to be found for the defendant, we could justify the complication of T_5 by its additional explanatory power, but if it were found for the plaintiff we should have no reason to complicate T_{4b}, since we get no gain in explanatory power. If decided for the plaintiff, there would be no reason to think that *pLand* was a relevant factor at all. Indeed Berman and Hafner, 1993 argues that *pLand* plays no significant role in the three cases under consideration.

An argument could, however, be mounted for preferring theories with *more* factors. Whenever a theory does not consider a factor that was present in one of its cases, that factor can be introduced, so jeopardising any rule (and value) preferences included in the theory based on that case, and so threatening its ability to explain its cases. The use of factor *pLand* in T_3 above to challenge T_2 is an example of this. Thus a theory is *safer* in accordance with the completeness of the factors it considers when using a case to derive a rule preference. Whether we should look for simplicity or safety depends on the status of the factors. If they have been used in the past decisions, completeness is desirable, but if, even though they do provide a reason, they have played no part in previous decisions, simplicity is to be preferred. Such a choice requires reference back to the full text of decisions, and cannot be settled in a general way.

Finally a theory is better in so far as less recourse to arbitrary preferences has been made. In moving from T_5 to T_{5a} and T_{5b} above it was necessary to add an arbitrary value preference. Such moves can only be justified externally to the theory, by an appeal to intuition or the like. In only one case does this seem to be entirely convincing, namely the arbitrary preference in T_{4b}, *vpref({Mprod,Llit},Mprod)*, does seem plausible because the preferred value is a superset of the other value. As we have said above, we might even wish to have an additional theory constructor legitimising the introduction of such value preferences.

3.5 Modelling Argument Moves in the Basic Theory

It is now interesting to relate the moves made in a HYPO-style argument to the above account of theories. A reconstruction of two of these moves, in terms of its own formalism, has been given in Prakken and Sartor,1998. Where appropriate, we will make comparisons with this work. A key element of our perspective on case based reasoning, is that reasoning with cases involves a number of related, but distinct, activities: namely first *constructing* a theory, then *using* the theory to explain cases, and finally *evaluating* competing theories, so as to adjudicate between competing explanations. The above discussion was structured around these three elements. Given this perspective, it is possible that argument moves in traditional case based systems, which do not make this distinction, conflate these elements.

3.5.1 Citing a Case

Citing a case just involves extending a theory with one additional precedent case. Typically, however, when this is done for a purpose, citing a case also involves expanding the theory with rules and preferences so that it can explain the cited case, and others included in the theory. An example above is T_1, which cites *Pierson* in support of the defendant in *Young* by introducing the case *<Pierson, {pNposs}, D>*, and a rule sufficient to explain it, that is *<{pNposs}, D>*. This citation is a particularly simple one, since the theory does not contain any rule which would require the case to have a different outcome. If the theory already includes such a rule, than the citation of a case also requires the introduction of a preference which explains why the case deserved the decision it had as a matter of fact, through the constructor *preferences-from-case*. As an example of this more complex type of citation, consider where the plaintiff *p* constructs theory T_2 by citing *Keeble*. At this stage *p* introduces, besides the case *<Keeble, {pLiv, pNposs, pLand}, P>* and the rule *<{pLiv}, P>}* also the preference *rpref(<{pLiv}, P>, <{pNposs}, D>)*, which enables the theory to explain *Keeble*. Pragmatically the best case to cite is the one which includes as many factors in common with the current case as possible. This allows the most specific, and thus safest, rule to be constructed, and thus pre-empts several possible challenges. Thus citing a case is essentially a move of theory construction, although considerations as to which is the *best* case to cite looks forward to the evaluation of the theory. Moreover, as implemented in HYPO, the criterion for choosing the best case favours safety over simplicity in theory evaluation.

3.5.2 Counter Examples and Distinctions

HYPO permits two different responses to a cited case: providing a counter example and distinguishing the case. Providing a "trumping" counter example is the stronger move because it will include another case in an opponent's theory so as to licence rule preferences such that the resulting theory will explain both the counter example case and the cited case, besides giving the current case the result desired by the citing party. It thus wins on explanatory power. The use of *Keeble* in T_2 is an example of this move. Introducing counter examples is part of theory construction, but their strength derives from theory evaluation, in that an "as on point" counter example does no more that display a failure to explain certain cases on the part of the theory, whereas the trumping counter example gives rise to a new theory superior in explanatory power. In Prakken, 2000 the idea is that counter examples can be evaluated not in terms of on-pointness, but in terms of a comparison between the values promoted. A trumping counter example will always succeed because it promotes at least as many values as the case to which it is a counter example (in Prakken, 2000) a set of values is always preferred to its proper subsets). On the other hand, a non-trumping counter example both lacks a value present in the precedent and has a new value not present in the precedent, so whether it succeeds depends on how these values are compared. A counter example is dismissed if the required value preference cannot be added to the theory. Indeed the theory may already contain value preferences which show that the counter example is ineffective.

In addition to distinguishing cases according to differences along shared dimensions, which will be considered in section 4, there are two ways of distinguishing a case in HYPO. Either one points to a factor favourable to one's opponent present in the precedent and absent in the current case, or one points to a factor favourable to oneself present in the current case and absent in the precedent. Here we discuss only the first of these; similar considerations apply to the other.

One way of distinguishing a case involves introducing a new factor *f*, which is *in favour of the opponent*, and which is not already present in the opponent's theory. This factor is not contained in the current case, but is present in the precedent licensing the *preferences-from-case* move which produced the preference *rpref*($<F_1, o>, <F_2, \sim o>$), which allowed the opponent's theory to explain the current case. Once the new pro-opponent factor *f* is introduced, the old rule $<F_1, p>$, which explained why the precedent was decided for the opponent (and why the current case should be decided in the same way), is extended into $<F_1$ *È* $\{f\}, p>$, and a new preference *rpref*($<F_1$ *È* $\{f\}, o>, <F_2, \sim o>$) is provided to explain the precedent. The latter preference does not apply to the current case (which does not contain factor *f*). Moreover, once the new, more specific, preference is available, the old preference becomes unnecessary to explain the precedent, and so fails to provide a convincing ground for the decision of the current case.

The introduction of factor *pLand* in T_3 above exemplifies the *distinguishing* move: by introducing this additional factor, the defendant was able to transform the rule *<{pLiv}, P>* into the rule *<{pLiv, pLand}, P>*, which he then used to explain the case *<Keeble, {pLiv, pNposs, pLand}, P>*, according to the preference *rpref(<{pLiv, pLand}, P>,<{pNposs}, D>)*. The new rule (and the corresponding preference) are not applicable to the current case, *Young*, which has factors *pLiv, pNposs, dLiv*, and does not contain *pLand*, which is required if the new rule is to be applied. On the other hand, in this new theory (resulting from adding to T_3 the new rule and preference), the old rule *<{pLiv}, P>* and the corresponding preference *rpref(<{pLiv}, P>, <{pNposs}, D>)* can be dismissed as being redundant since they have no explanatory function. Therefore according to the new theory, a *P* decision in *Keeble* is consistent with a *D* decision in Young, which is what the defendant wanted to establish. The move is less powerful that a trumping counterexample because it does not form the basis for a different decision in the current situation, but merely blocks the rule which the opponent needs. In conclusion, this theory construction move involves a factor rather than a case. The effect of the move is to render the original theory weaker because it makes its rule preference arbitrary rather than grounded in a precedent.

An as-on-point counter example can also be seen as the combination of a distinguishing move together with a case which grounds a new alternative theory, based on different factors. This new theory can, of course, then be subject to a distinguishing move itself. We would then end up with two theories which both require arbitrary preferences in order to explain the current case. To be effective, the distinguishing factor must relate to a value which can be shown to be preferred, so that arbitrary preferences are not required. This is what happened above in T_{4b} when *dLiv* was used to distinguish *Young* from *Keeble*. This is an example of the second kind of distinguishing move (i.e. one introduces a new factor favourable to oneself), but its greater effect comes from the value associated with the distinguishing factor, not from it being an example of this other way of distinguishing a case.

3.5.3 Emphasising Strengths and Showing Weaknesses not Fatal

There are four other argument moves introduced in CATO (Aleven, 1997): *emphasise strengths*, *show weaknesses not fatal*, *emphasise a distinction* and *downplay a distinction*. The last two require an extension to the basic model and will be considered in 4.2.

The first of these simply corresponds to introducing more cases which are explained by the theory, with factors shared with current case, thus increasing the theory's explanatory power. Again these moves can be seen as constructing a theory which will be evaluated as better. Showing weaknesses not fatal is perhaps more interesting, in that it seems to suggest a different understanding of the rules derived from cases from that described above. For the absence of a factor to be fatal, it would have to be a necessary condition, and as we have described the situation above, case law can never give us such conditions, but only defeasible rules. The move would also involve including cases found for the desired side, but this time containing factors favourable to the other side which lead to defeated rules. In our terms therefore it can be seen as an attempt to increase the safety of the explanations in the theory, by anticipating and pre-empting the introduction of additional factors. It is also possible that such cases may licence the introduction of preferences which contradict preferences arbitrarily introduced by an opponent.

4. Extensions to the Basic Model

The theories constructed in the basic model given in the last section provide a very simple account of theory construction for reasoning with cases. In this section we consider two extensions to the basic model intended to capture insights of two important systems developed in this area, HYPO (Ashley and Rissland, 1988, Ashley, 1990), which takes a more sophisticated view of how cases should be described, and CATO (Aleven, 1997), which allows multi-step arguments through the use of a hierarchy of factors.

4.1 Dimensions

In section 3 we presented our model in terms of the approach used by Berman and Hafner, 1993. In fact there are considerable limitations in this approach. Consider the case of *Pierson*. Using the factors identified in Berman and Hafner, 1993 it would appear that the plaintiff had no case to present. But further consider the pro-defendant factor *pNposs* (the plaintiff has no possession of the animal), and assume that it can be applied whenever the plaintiff has not caught the animal. As set up, this is an all or nothing affair, in which either plaintiff has caught the animal (so that the factor does not hold), or has not done so (so that the factor holds). Under the first condition (the animal has not been caught) it does not matter whether the plaintiff has seen the animal, whether he was in hot pursuit of it, or even whether he has wounded it (perhaps mortally). All of these situations are treated by *pNposs* as being equivalent ways of realising the pro-*D* factor.

We do not, however, have to see the situation this way. We could see instead a range (discrete or continuous) of positions between seeing the animal and actually possessing it, and the points on this range as being progressively more favourable to *P* and less favourable to *D*. The factor-based perspective transforms this range into a binary alternative: according to *pNposs* having failed to catch the animal is a reason for finding for the defendant, whereas if the animal has been caught there is no such reason. However, factor *pNposs* is not the only way in which this transformation can take place. Instead of the pro-*D* factor *pNposs* we might have used a pro-*P* factor, *pChase,* which was intended to cover all cases in which the plaintiff had given chase: according to this choice, having pursued the animal is sufficient to establish a reason for finding for the plaintiff, and only failing to start a chase would not instantiate this reason. Note that the situation existing in *Pierson* (plaintiff was chasing the animal though it was not yet caught), would favour the defendant when seen from the perspective of factor *pNposs*, while it would favour the plaintiff when seen from the perspective of factor *pChase*. Consider also *pLiv* (the plaintiff was pursuing his livelihood). While the plaintiff in *Pierson* was not earning his living he might have been acting out of a number of progressively less favourable motives, such as altruism (foxes are vermin and a threat to farmers), pleasure, or even malice (if it was the defendant's pet fox). Perhaps the correct factor was one which would apply if the plaintiff was earning his living *or* acting out of concern for his neighbours. Had this factor been available, another pro-plaintiff factor would have been available in *Pierson*. Considerations such as these are present in the text of the judgement in *Pierson*. The judgement speaks of "caught *or mortally wounded*" and a dissenting opinion expressed the view that the social utility of the plaintiff's fox hunting was so great that the activity should be encouraged and protected by law.

The original conception of HYPO (e.g. Rissland et al., 1984, Ashley and Rissland, 1988, Ashley,1990) accommodated this kind of reasoning by using not factors but *dimensions*[3]. Dimensions were intended to be a spectrum of possible degrees for an aspect of the affair, and a given side was to be favoured according to the extent that the position on this spectrum approached the end favourable for that side. Thus for possession we could see a possible dimension *pControl*, representing the level of control which the plaintiff has over the animal, with possible degrees such as <*no-contact, seen, started, wounded, mortally-wounded, captured*>, which favours *P* according to the extent to which capture was approached, and favours *D* to the extent to which no-contact was approached.

Seen in this way, choosing a factor is not a matter of simply picking one property favouring one side of the dispute from a pre-existing background store of such properties, but rather involves selecting a significant point within a dimension from which a factor can be formed and linking that point to an outcome. This selection implies that the realisation of the dimension to that point is sufficient to favour the chosen outcome. Therefore for a plaintiff factor, all positions in the span from the chosen point towards the plaintiff extreme will also realise the factor, and for a defendant factor all positions in the span from the chosen point towards the defendant extreme will also realise the

[3] The differences between factors and dimensions were the subject of several conversations between the authors, Edwina Rissland and Kevin Ashley at the International Conference on AI and Law in St. Louis in 2001. Since this paper was written, other work has published on this topic. Bench-Capon and Rissland, 2001, argues for the need to use dimensions rather than factors and Rissland and Ashley, 2002, provides a useful discussion of these two notions.

factor. Dimensions have interesting connections with the notion of quantity spaces as found in qualitative reasoning (e.g. Forbus, 1984), and the points which determine factors to have similarities to the limit points of that theory. Exploring these connections further might enable exploitation of the mechanisms of qualitative reasoning such as qualitative proportionality to make more precise the notion of the influence of a fact on the outcome of the case. We must, however, leave such exploration for future work.

Dimensions also need to be related to values, as were factors. If a factor is a reason for deciding for a particular side because to do so would promote some value, then a dimension is an increasingly strong reason for deciding for a side as its position approaches its most favourable extreme because deciding for one side as the dimension goes in towards that side's extreme more probably or more strongly promotes some value. Thus we should see the positions of a dimension as progressively more certainly promoting some values as we move towards an extreme. Two types of value need to be distinguished: those which are more surely promoted by deciding for the plaintiff as we approach the plaintiff extreme and those which are more surely promoted by deciding for the defendant as we approach the defendant extreme.

This can be illustrated by considering *pControl,* with positions *<no-contact, seen, started, wounded, mortally-wounded, captured>*. This dimension can be seen as being supported by two values, reduction of litigation (*Llit*), towards the defendant's extreme (the beginning of the positions list), and property rights (*Prop*), towards the plaintiff's extreme (the end of the list). In fact, as we move towards the defendant extreme, i.e. when the plaintiff's control over the animal is more tenuous, we approach less clear cut situations. Deciding for the plaintiff in those situations would be more likely to encourage litigation in other similar cases, and would increasingly do so the less the plaintiff's control. If judges were to decide this way, hunters who missed the game they were pursuing, would, whenever they believed it was captured by other hunters, begin suing the latter, claiming to have been the first to wound, start, or even see the animal. Note that, from this perspective, if having wounded the animal is a form of control so tenuous that we have a reason to find for the defendant, mere pursuit will be a stronger reason to so find. Property rights, on the other hand, are more surely promoted by deciding for the plaintiff when he has a stronger control over the animal: in those cases deciding for the plaintiff would mean to give legal backing to the physical possession he has gained over the animal, and so recognise and encourage private appropriation. If merely starting a fox is a reason to find for the plaintiff, then mortal wounding will be a stronger reason.

Our discussion of the dimension *pControl* shows how one can extract from one dimension both pro-plaintiff and pro-defendant factors. In both cases we need to choose a starting point for the factor, but the behaviour of the factor will be different. Pro-plaintiff factors will promote a pro-plaintiff value, and will cover all positions in the range spanning from the chosen point to the pro-plaintiff extreme. Pro-defendant factors will promote a pro-defendant value, and will include all positions in the range spanning form the chosen point to the pro-defendant extreme.

The need to form factors from dimensions brings factor descriptions within the theory construction process. Let us see how we might formalise this. First we replace the background set of factors F_{bg}, with a background set of dimensions D_{bg}. Each dimension $d \in D_{bg}$ refers to a property that can be present in the cases to a range of different extents. The ways of realising one dimension are ordered in a spectrum, according to the extent in which they realise the dimension: we therefore refer to them as the possible *positions* in the dimension's spectrum. So, we have a background set of possible positions Pos_{bg} which indicate the possible ways in which dimensions can be realised. Dimension-descriptions can be defined as follows:

Definition 22. Dimension-description: *A dimension-description is a four tuple* $<d, <p_1 \ldots p_n>, <o^{\downarrow}, o^{\uparrow}>, <V^{\downarrow}, V^{\uparrow}>>$ *where*

- $d \in D_{bg}$,
- $<p_1 \ldots p_n> \in Pos_{bg} \times \ldots \times Pos_{bg}$, *is a spectrum of positions realising increasing degrees of d* (p_{i+1} *realises d more than* p_i),
- $<o^{\downarrow}, o^{\uparrow}> \in O_{bg} \times O_{bg}$, *is a pair of complementary outcomes, such that*
  - $o^{\downarrow}$, *the downward outcome, is increasingly favoured by decreasing degrees of d* ($o^{\downarrow}$ *is favoured by* p_i *more than it is favoured by* p_{i+1}),
  - $o^{\uparrow}$, *the upward outcome, is increasingly favoured by increasing degrees of d* ($o^{\uparrow}$ *is favoured by* p_{i+1} *more than it is favoured by* p_i),
- $<V^{\downarrow}, V^{\uparrow}> \in pow(V_{bg}) \times pow(V_{bg})$, *is a pair of sets of values, such that*
  - $V^{\downarrow}$, *the downward values, are more probably promoted by* $o^{\downarrow}$ *as d decreases* ($o^{\downarrow}$ *under condition* p_i *promotes each* $v \in V^{\downarrow}$ *more probably than* $o^{\downarrow}$ *under* p_{i+1} *does)*
  - $V^{\uparrow}$, *the upward values, are more probably promoted by* $o^{\uparrow}$ *as d increases* ($o^{\uparrow}$, *under condition* p_{i+1} *promotes each* $v \in V^{\uparrow}$, *more probably than* $o^{\uparrow}$ *under* p_i *does).*

Using the example of *pControl* given above, this would give the following dimension-based description:

Property: *pControl*
Spectrum: *<no-contact, seen, started, wounded, mortally-wounded, captured>*

Outcomes: <*D, P*>
Values: <*{Llit}, {Prop}*>

which we will write as

<pControl, <no-contact, seen, started, wounded, mortally-wounded, captured>, <D, P>, <Llit, Prop>>

(we drop parentheses on sets of values containing only one value).

As we said, the set of background factors description Fds_{bg} is now substituted with a set of background dimension descriptions Dds_{bg}. Cases can now be described in terms of dimensions rather than factors. Each case will be characterised by a set of dimensional qualifications *Dq*, where each dimensional qualification *<d, p>*, indicates that position *p* for dimension *d* was realised in the case.

Definition 23. Case dimension-based description*: A case dimension-based description is a three tuple <c, Dq, o> where:*

- *c* Î C_{bg}, *and*
- *for every <d, <*$p_1 \ldots p_n$*>, <o¯, o->, <V¯, V->>* Î Dds_{bg} *there is at most one pair <d, p>* Î *Dq, with p* Î *<*$p_1 \ldots p_n$*>.*

Let us denote the set of all case dimension-based descriptions available in the background as $Cdds_{bg}$. Now we will consider how to go from dimensions to factors, extracting factors from dimensions and transforming dimension-based descriptions of cases into factor-based descriptions.

Factor descriptions can be constructed out of dimensions by choosing one of the positions on the spectrum, one outcome, and one of the values promoted by that outcome. If the outcome is *o-*, then that positions and all position with an index higher than that position will mean that the factor is present. Similarly if the outcome is *o¯*, that position and all positions with an index less than that position will mean that the factor is present. Let us assume that our background also contains a set of factor names Fn_{bg}. A factor description thus becomes:

Definition 1b. Factor description*: A factor description is a five tuple <f, d, p, o, v>, where:*

- *f* Î Fn_{bg},
- *<d, <*$p_1 \ldots p_n$*>, <o¯, o->, <V¯, V->>* Î Dds_{bg}, *p* Î *<*$p_1 \ldots p_n$*>, and*
- *either o = o¯ and v* Î *V¯, or o = o-, and v* Î *V-.*

For example, given the dimension *pControl,* described above, one could construct the pro-plaintiff factor *pSureCatch* (*p* is sure of the catch), with description *<pSureCatch, pControl, mortally-wounded, P, Prop>*, or the pro-defendant factor *pNposs* (*p* has no possession) with description *<pNposs, pControl, mortally-wounded, D, Llit>*. We call the set of all factor descriptions of this sort which are constructible from the background dimensions Fds_{poss} (the possible factors).

The construction of a factor description *<f, d,* p_i*, o, v>* from a dimension *<d, <*$p_1 \ldots p_n$*>, <o¯, o->, <V↓, V↑>* amounts to saying that the realisation of the chosen position p_i, supports the chosen outcome *o*, so as to promote the indicated value *v*. This has the following implications:

a) p_i cannot support the outcome complementary to *o* unless appeal is made to a different value (since one single feature cannot be the ground for two complementary outcomes considered with respect to a single value),
b) if *o = o¯*, than any p_j such that $j < i$ also more strongly supports *o*,
c) if *o = o-*, then any p_j such that $j > i$ more strongly supports *o*,

It is important to stress that a factor *<f, d,* p_i*, o, v>* applies not only to the cases that exhibit the dimensional position p_i, but also to the cases exhibiting a position that more strongly favours *o* along the dimensional spectrum. In other words, p_i is the lowest bound for the realisation of the factor, which is also realised by more *o*-favourable positions. For example, according to the dimension *pControl*, with <*no-contact, seen, started, wounded, mortally-wounded, captured*> and outcomes <*D, P*>, the pro-plaintiff factor *pSureCatch* is realised not only when the animal was mortally wounded, but *a fortiori* when the animal was captured, whereas the pro-defendant factor *pNposs* is realised not only when the animal was mortally wounded, but *a fortiori* in all positions preceding mortally-wounded in the dimensions list. We can next define the notion of a factor subsuming a dimensional qualification (i.e. the qualification being a way of realising the factor).

Definition 24**. *Subsuming:*** *A factor f with description <f, d,* p_i*, o, v>* Î Fds_{poss}*, subsumes the dimensional qualification <d,* p_j*> if and only if* $p_i = p_j$ *or* p_j *is more favourable to o then* p_i *(i.e if o = o- then j ³ i, and if o = o¯ then j £ i).*

For example, factor *SureCatch* above subsumes both *pControl(mortally wounded)* and *pControl(captured)* while it does not subsume *pControl(wounded)* nor *pControl(started).* Note that building factors out of dimensions gives a degree of discretion: it requires setting the bound at which one outcome is supported along one dimension. Different choices in this regard would lead to different interpretations of the cases. So while the factor *pSureCatch* favours outcome *P* only from the point where the animal is mortally wounded, a factor *pContact (P* had contact with the animal), with description *<pContact, pControl, started, P, security of possession>* would imply that just starting the animal (and *a fortiori* wounding it, even though not mortally) supports the outcome *P*.

Now given the set *Fds* of all factor descriptions so far constructed, we can transform a case described via dimensions into that case described via factors. We now define a function, *factorise(Dq, Fds)*, which takes a set of dimensional qualifications *Dq*, and a set of factor-descriptions *Fds*, and returns the set of factors described in *Fds* that subsume dimensions in *Dq*.

Definition 25. Factorise: *Factor f Î factorise(Dq, Fds) if and only if there are <d, p_i> Î Dq and <f, d, p_j, o, v> Î Fds, such that f subsumes <d, p_i>.*

For example assume the following dimensions:

- *<pControl, <no-contact, seen, started , wounded, mortally wounded, captured>, <D, P>, <Llit, Prop>>*
- *<pLand, <dProperty, dLease, other people's property, communal property, pLease, pProperty>, <D, P>, <Freedom, Prop>>.*

where *pLand* expresses the connection between the plaintiff and the land where he was chasing (which is most tenuous when he is on the defendant's property, and strongest when he is on his own property). Assume also to have constructed the following factors from the dimensions above:

- *<pNposs, pControl, mortally-wounded, D, Llit>* and
- *<pOwns, pLand, pLease, P, Prop>.*

Now, given what we said above, how should we translate two dimensional qualifications, such as *pControl(mortally wounded)* (in regard to his control over the animal, *p* had wounded it) and *pLandcommunal property)* (in regard to his connection to the land, *p* was on a communal property) into factors? The result is given by

factorise({pControl(mortally wounded), pLand(communal property)}, {<pNposs, pControl, mortally wounded, D, Llit> , <pLand, pOwns, pLease, P, Prop>} = {pNposs}.

Transforming the dimension-based description *<c, Dq, o>* of a case into its factor-based description *<c, F, o>*, requires factorising the dimensional qualifications *Dq* into factors *F.* The factorisation of cases will, of course, be relative to *Fds,* the factor descriptions so far constructed. To achieve this we define a function *FactoriseCase(<c, Dq, o>, Fds*), which takes the dimension-based description *<c, Dq, o>* of a case and a set of factor descriptions *Fds,* and returns the factor-based description *<c, F, p>* of that same case, which results from factorising *Dq* into *F:*

Definition 26. FactoriseCase: *FactoriseCase(<c, Dq, o>,Fds) = <c, factorise(Dd, Fds), o>*

For example, *factorise(<c1, {pControl(mortally wounded), pLand(pLease)}, P>, {<pNposs, pControl, mortally-wounded, D, Llit> , <pOwns, pLand, pLease, P, Prop>})* returns *<c1, {pNposs, pOwns}, P>,*

We also provide a function which factorises a set of cases only in regard to one factor. This is the function *ApplyFactor(Cfds, <f, d, p_i, o, v>),* which takes as input a set of case factor-based descriptions *Cfds*, and a factor description *<f, d, p_i, o, v>).* It returns the new set of case factor-based descriptions which results from adding factor *f* to each case factor-based description *<c, F, o> Î Cfds,* whenever *f* subsumes a dimensional qualification *<d, p_i>* of case *c.*

Definition 27. ApplyFactor: *ApplyFactor(Cfds, <f, d, p_i, o, v>) is the set of all <c, F, o> such that <c, F', o> Î Cfds, and*

- *<c, F, o> = <c, F' È f, o> when f subsumes <d, p> where <d, p> Î <c, Dq, o>,or*
- *<c, F, o> = <c, F', o> otherwise.*

We can now revisit the definitions of section 3 supposing that we start from a set of dimensions and a set of cases described through dimensional qualifications, rather then with a set of factors and of cases described through factors.

The first point to note is that factor descriptions are now local to a theory, rather than being available globally. Also, much of what was originally in the background to a theory is now dependent on these factor descriptions. Suppose that *Fds* is the set of factor descriptions in theory *T*, and F_{Fds} is the set of the names of all those factors. Now the set of possible case factor-based descriptions which are obtainable with those factors are

$$Cfds_{Fds} = C_{bg} \times Pow(F_{Fds}) \times O_{bg}$$

(each case description contains the name of the case, a set of the constructed factors, and one outcome). The set of possible (constructible) rules is now relative to the constructed factors.

Definition 4b. Possible rule*: <F, o> is a possible rule, given factor descriptions Fds, if and only if for each f Î F, <f, o, v> Î Fds.*

Let us denote the set of rules which are possible relative to a set of factor descriptions *Fds*, as $R_{Fds.}$ Let us similarly denote preferences constructible from a rule set R_{Fds}, as $Pref_{Fds} = R_{Fds}$ ´ R_{Fds}.

Definition 12b. Theory*: A theory is a five-tuple <Cfds, Fds, R, Rpref,Vpref,>, where:*

- *Fds is a set of factor descriptions of the form <f, d, e, p, v>;*
- *Cfds Í* $Cfds_{Fds}$,
- *R Í* R_{Fds},
- *Rpref Í* $Rpref_{Fds}$,
- *Vpref Í* $Vpref_{bg}$.

Let us now see how we can add a factor to a theory. Adding a factor now requires building it from some dimension description. However, we wish to block the use of dimensions to produce two factors based on the same dimensional qualification *with the same value*. Were this allowed, we would have the possibility of explaining cases using multiple factors based on the same dimension and value, which we regard as undesirable as representing counting a feature of the case twice, and intolerable where a case satisfies both a pro-plaintiff and a pro-defendant factor based on the same dimension with the same value. If we have two factors based on the same dimension and value available in the theory, we need to choose which we wish to use. We thus modify definition 14:

Definition 14b: Include-factor:

Pre-condition:

- *current theory is <Cfds, Fds, R, Rpref, Vpref>, and*
- *<d, <$p_1 \ldots p_n$>, <o -, o¯>, <V-, V¯>> Í Dds,*
- *<f, d, p, o, v> Î Fds,*
- *p Î <$p_1 \ldots p_n$> and either o = o - and v Î V-, or o = o¯ and v Î V¯,*
- *there is no <f', d, p', o, v> Î Fds,*

Post-condition:

- *current theory is is <$Cfds_{new}$, Fds_{new},R,Rpref,Vpref>, with*
- *$Cfds_{new}$ = <applyFactor (Cfds, <f, d, p, o, v>),*
- *Fds_{new} = Fds È {<f, d, p, o, v>} .*

Notice that, when a set of factors *Fds1* is expanded into a larger set *Fds2*, the sets of the constructible rules and of the constructible preferences will also be expanded: if $Fds1 \subset Fds2$, then R_{Fds1} Ì R_{Fds2} and $Rpref_{Fds1}$ Ì $Rpref_{Fds2}$. Besides *include-factor* we need to rephrase *include-case*, since the background information only contains dimension-based description of cases, which need to be transformed into factor-based descriptions.

Definition 13b. Include-case:

Pre-condition:

- *current theory is <Cfds, Fds, R,Rpref,Vpref>, and*
- *<c, F, p> Î* $Cdds_{bg}$,

Post-condition:

- *current theory is is <$Cfds_{new,}$ Fds, R, Rpref, Vpref>, with*
- *$Cfds_{new}$ = Cfds + factoriseCase(<c,F,p>,Fds).*

All the other definitions, subject to the relativity of the notions to *Fds*, can remain essentially unchanged.

A suitable set of example dimensions for our example cases might be the following. (Note that when one dimension goes only one way, favouring no outcome at one extreme, we put 0 for missing outcome):

- *<pControl, <no-contact, seen, started, wounded, mortally-wounded, captured>, <D, P>, <Llit, Prop>>,*
- *<pLand <dProperty, dLease, otherPeopleProperty, communalProperty, pLease, pProperty>, <D, P>, <MFreedom, Prop>>,*
- *<pMotive, <pMalice pSport, pLivelihood>,<0, P>, <0, Mprod>>,*
- *<dMotive, [dMalice, dSport, dLivelihood], <0, D>, <0, Mprod>>.*

The examples in section 3.3 above all still apply, supposing that we have first used four applications of make-factor to produce *Fds,* so that it includes the following factors based on these dimensions:

- *<pNposs, pControl, mortally wounded, D, Llit>,*
- *<pLand, pOwn, pProperty, Prop>,*
- *<pLiv, pMotive ,pLivelihood, P, Mprod>,*
- *<dLiv, dMotive, dLivelihood, D, Mprod>}.*

This section represents quite a significant extension to the simple model. It does give an important gain in that it allows us to explore, if desired, the creation of factors rather than taking them simply as given, which is useful since the available factors can significantly bias our view of a case. It also allows the possibility of an additional argument move, recognised in HYPO but not in CATO, or any other approach which ignores dimensions.

In HYPO when citing a case no comparison of strength along dimensions is made. Thus in a new case similar to *Pierson* except that the fox had been wounded, *Pierson* would be cited for the defendant, who would provide a theory explaining Pierson according to factor *<pNposs, pControl, mortally-wounded, D, Llit>*, which also provides a ground why the new case should also be decided for the defendant, according to the rule *<{pNposs}, D>*. In the response, however, the plaintiff can take a different position within the dimension into account, and so would be able to point out that in this dimension the current situation is more favourable to him. For example, he could build a factor *pContact,* meaning that *P* had contact with the animal, with description *<pContact, pControl, wounded, P, Llit>*, and a factor *pNoContact,* with description *<pNoContact, pControl, started, D, Prop>*. The first factor is satisfied in the new case (the factorisation of which includes *pContact*) and produces a rule *<{pContact}, P>* , while the second factor is satisfied in Pierson, and produces the rule *<pNoContact, D>*, which contributes to explaining Pierson's outcome.

What is happening here is that in the citation the factor is chosen so as to explain *both* the current case and the precedent case, whereas in the response a *different* factor is chosen from the dimension which will still explain the precedent, but which will not apply to the current case. Essentially the disputants are making and including different factors in their different theories. Similar moves are possible with respect to counter examples and their rebuttals. This is a very important type of move, but one which obviously requires dimensions.

Factors have been found sufficient for some of the analyses we wish to subsume, and for those the simpler model will suffice. We have, however, shown how we can treat the richer analyses of the original HYPO system in a similar fashion. If we wished, we could take a further step back, and bring the choice of dimensions into theory construction also, by removing D_{bg} from the background and replacing it with a set of pairs of attributes and unordered positions for these attributes, which would need to be turned into dimensions by choosing a sub-set of the possible positions, and ordering them according to some social value or values. We will not, however, pursue this further here.

4.2 CATO and a Hierarchy of Factors

After HYPO, Ashley began work, with Aleven, on the CATO system, most fully reported in Aleven, 1997. CATO did not use dimensions, but used factors like those in the basic model. It did, however, make a different refinement to factors by organising them into a hierarchy, with the presence of factors contributing to or detracting from more abstract factors. This extra organisation permitted the introduction of two new argument moves, *emphasising distinctions* and *downplaying distinctions*. To represent the factor hierarchy we need to modify the notion of factor as given in Definition 1, but differently from the modification given to accommodate dimensions in Definition 1b. A fully satisfactory account of these moves also requires a more elaborated notion of a case being explained than that provided in Definition 21, so as to allow for arguments to be chained to an arbitrary length, with the possibility of conflicts at different points in the chain. This is impossible in the framework we have so far presented, since we do not allow chaining of rules. A logic which would provide the necessary support is given in section 4.2.1. In 4.2.2 we show how this logic can be used to model the operation of the factor hierarchy as used in CATO, and to allow for the new moves of emphasising and downplaying distinctions.

4.2.1 An Extended Logic for Using Theories

The notions of explanation we proposed in definition 21 above only allows for one step inferences, where the final outcome of a case is directly supported by the factors in the case description. To deal with abstract factors we adopt a very simplified variant of the argumentation-based system proposed by Prakken and Sartor,1996, but other logics would be equally appropriate, if they can deal appropriately with prioritised conflicting rules. Let us introduce a few simple notions (those notions can be expanded to take into account issues such as those of undercutting, (Pollock, 1995) or pre-emption (Horty,2001), but they are sufficient for our purposes).

For simplicity let us view all elements in our knowledge representation as instances of the same syntactic structure, which we call a conditional.

Definition 28. Conditional*: A conditional is a couple <L, l> where L, the antecedent, is a (possibly empty) set of literals (an atomic formula, or the negation of such a formula) and l, the consequent, is a literal.*

In particular, any rule <*F*, *b*> can be viewed as a conditional, and in the same way we can represent conditioned preferences. The unconditioned assertion that a factor occurs, or that a certain rule or value preference is the case, can be viewed as a conditional with empty antecedents: *<Æ, j >, <Æ, rpref(r_1, r_2)>, <Æ, vpref(r_1, r_2)>*. We use the consequent of such degenerate conditionals as their abbreviations: rather then *áÆ, j ñ, áÆ, rpref(r_1, r_2) ñ, áÆ, vpref(r_1, r_2)ñ*, we write respectively *j , rpref(r_1, r_2), vpref(r_1, r_2)*.

Conditionals can be chained together to form arguments, where an argument is sequence of conditionals such that each literal in the antecedent of any conditional occurs previously in the argument, as the consequent of some conditional.

Definition 29. Argument*: A sequence of conditionals A = <<L_1, l_1> … <L_n , l_n> >, is an argument, if and only if any <L_i , l_i> Î A is such that for each l_j Î L_i, there is a <L_j, l_j> Î A with j < i.*

This means that the consequent of any conditional in the argument can be derived, via a sequence of modus ponens inference-steps, from previous elements in the inference. We therefore say that all such consequents are *the conclusions* of the argument, and denote the conclusions of an argument *A* with *Conclusions(A)*. For example argument <*a*, <{*a*}, *b*>> has conclusions *a* and *b*, that is *Conclusions(A)* = {*a*, *b*}.

Definition 30. Defeat*: An argument A_1 defeats A_2 if there are conditionals y_1 Î A_1 and y_2 Î A_2 such that:*

- *y_1 attacks y_2, and*
- *A_2 does not have consequence rpref(y_2, y_1).*

For example, arguments *B_1 = <a_1, <{a_1}, b>>* and *B_2 = <a_2, <{a_2}, Øb>>* defeat one another, since they contain conditionals (*<{a_1}, b>,<{a_2}, Øb>*) attacking each other, and there are no preferences adjudicating the conflict.

Definition 31. Strict Defeat: *A_1 strictly defeats an argument A_2 if A_1 defeats A_2, but A_2 does not defeat A_1.*

For example argument *B_3 = <a_3, <{a_3}, b>, rpref(<{a_3}, b>,<{a_2}, Øb>)>* defeats *B_2*, but *B_3* is not defeated by *B_2*, since *B_3* contains the preference *rpref(<a_3, b>, <a_2, Øb>)* which prevents defeat.

Our purpose is to establish which arguments and therefore which conclusions are justified within a certain premises set. For an argument *A* being justified within a certain premises set *S* (with *A Í S*), we mean that *A* has no valid defeater within *S*: so long as we accept as our premises set *S*, we need to endorse *A* and all of its conclusions. For computing whether an argument is justified, the consideration of defeat between single arguments is not sufficient, since a defeated argument can be reinstated when its defeaters are strictly defeated by further arguments. For example, consider the three arguments:

B_1 = <a_1, <{a_1}, b_1>, <{b_1}, g>>,
B_2 = <a_2, <{a_2}, b_2>, <{b_2}, Øg>>,
B_3 = <a_3, <{a_3}, Øb_2>, rpref(<{a_3}, Øb_2>, <{a_2}, b_2>)>.

B_1 is defeated by argument *B_2*, but *B_2* is strictly defeated by *B_3*. *B_1* though being defeated, is justified within the premises set *S = B_1 È B_2 È B_3* since it has no valid defeaters in *S*: its only defeater, *B_2*, is strictly defeated by *B_3*, which has no defeaters within *S*.

Here we will only provide a procedural notion of "being justified": an argument is justified if and only if there is a proof for it. And, in the context of a premises set *S*, the proof that an argument is justified takes the form of a tree of arguments belonging to *S*. In this tree, nodes located at an even level (the level of a node being the distance from the root) attack, directly or indirectly (i.e., by attacking its supporters), the root argument, while nodes located at an odd level support, directly or indirectly (i.e., by attacking its attackers) the root argument. The notion has considerable similarities to well-founded support in truth maintenance systems and the well-founded semantics for logic programs (e.g. Dean et al., 1995, van Gelder et al., 1991), and with other formalisms used in artificial intelligence and logic programming (see in particular Dung, 1993, and for other references to the relevant literature, Prakken and Sartor, 1997). Here is a more exact definition:

Definition 32. Proof tree: *A proof tree for argument A, within premises set S, is a tree of arguments from S, such that*

- *A is at 0-level (root),*
- *each A_i at an even-level node is followed by all of A_i's defeaters,*
- *each A_i at an odd-level node is followed by an A_j such that, A_j strictly defeats A_i,*
- *odd level arguments are not repeated in the same branch of a tree.*

A proof is a proof tree where no attack against the root argument *A* was successful (since every branch terminates with a node supporting *A*), and that no further attacks are possible:

Definition 33. Proof: *A proof of argument A within premises set S, is a proof tree for A within S, such that:*

- *each branch of the tree terminates with an even-level node,*
- *it is not possible to add further nodes.*

Definition 34. Justified argument: *An argument A is justified within premises set S if and only if there is a proof for A within S.*

Let us denote as *JustArg(S)*, the set of arguments which are justified within premises set S. For example, let us assume we have the following set of premises:

- factors $\{a_1, a_2, a_3\}$,
- rules *{<{a_1}, b_1>, <{b_1}, g>, <{a_2}, b_2>, <{b_2}, Øg>, <{a_3}, Øb_2>}*,
- preferences {rpref(<{α_3}, $\neg\beta_2$>, <{α_2}, β_2>)}.

Then we can build a proof that argument <α_1, <{α_1}, β_1>, <{β_1}, γ>> is justified. This would be the corresponding proof tree:

0. <α_1, <{α_1}, β_1>, <{β_1}, γ>>

1. <α_2, <{α_2}, β_2>, <{β_2}, $\neg\gamma$>>

2. <α_3, <{α_3}, $\neg\beta_2$>, *rpref*(<{α_3}, $\neg\beta_2$>, <{α_2}, β_2>)>

It is now clear when a theory explains a case.

Definition 21b. Explaining*: A theory <Cfds,Fds, R, Rpref, Vpref> explains a case c if and only if:*

- *<c, F,o> Î Cfds,*
- *there is an argument A such o Î Conclusions(A) and A Î JustArgs(F È R È Rpref).*

4.2.2 Modelling CATO

In this section we use the logic developed above to give a model of CATO and its argument moves.

First, to convey the required hierarchy information, we must allow our background knowledge to include factors favouring intermediate results, besides the final outcomes *{P, D}* of the dispute. To model this, we will need to allow for more then one factor description for a given factor: if factor *f* promotes a certain final outcome o_1, via the intermediate outcomes $o_2, \dots, o_n$, it will have description descriptions <f, o_1, v_1> … <f, o_n, v_1>. Note that we assume that the values promoted by the factor remain the same. The intermediate outcomes will, in their turn, be factors favouring further outcomes, which may still be intermediate, or may represent one outcome for the dispute. This means that our background set of factors F_{bg} will include two sets: the set of concrete factors Fc_{bg}, which are to be used in describing the cases, and the set of the abstract factors Fa_{bg}. Abstract factors will be both factors and outcomes: Fa_{bg} Í F_{bg} and Fa_{bg} Í O_{bg}. By linking each factor to the outcome it produces, we obtain a tree where each intermediate node (each node, except the root and the leaves of the tree), is both an intermediate outcome and an abstract factor.

Abstract factors do not appear in the representation of the cases, but they will be used in multi-step arguments, where the final outcome of a case is explained through a sequence of chained rules. For example, suppose we have the hierarchy shown in Figure 2. If a case contains factor f_6 and had decision *o*, its explanation may be based on the argument <f_6, <{f_6}, f_3>,<{f_3}, f_1>, <{f_1}, o>>. Note that any factor supports not only its parent node: via its parent, it gives support to all of its ancestors (to avoid clutter, we represent this only for factor f_6 in figure 2). So, the factors background besides factor-descriptions connecting each factor to its parent, e.g. <f_6, f_3, v>, will also contain factor-descriptions connecting each factor to all of its ancestors, shortcutting the intermediate links, e.g. $<f_6, f_1, v>$, $<f_6, o, v>$ (we may alternatively assume that these short cuts do not need to be in the background knowledge, but can be introduced at the theory construction stage).

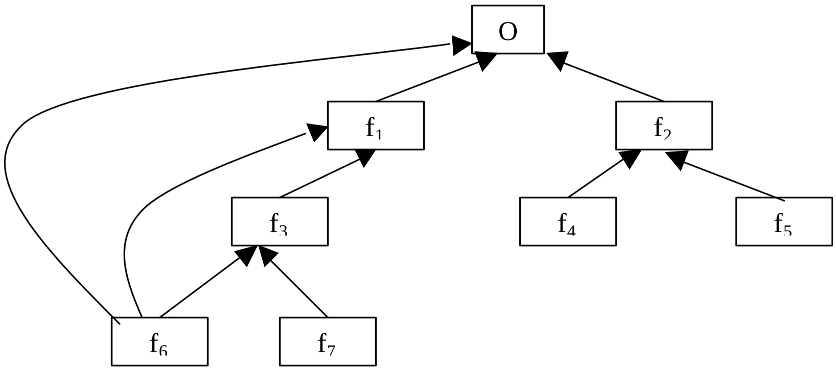

Figure 2: A factor hierarchy

A factor with multiple ancestors, has multiple factor descriptions, and so provides the opportunity for the reasoning move that Ashley and Aleven call "downplaying a distinction". As we have seen above, distinguishing involves (a) adding a new factor which applies to the precedent c_{prec} and does not apply to the current case c_{curr}, and (b) explaining the precedent by a rule including this factor. Downplaying the distinction of a precedent consists in providing an explanation for both c_{prec} and c_{curr} through a rule including an *abstract factor*, which is an immediate consequence *both* of the distinguishing factor, and of a different factor which can be established in the current case. Let us now go through the process of distinguishing and downplaying the distinction. We assume the following factor hierarchy:

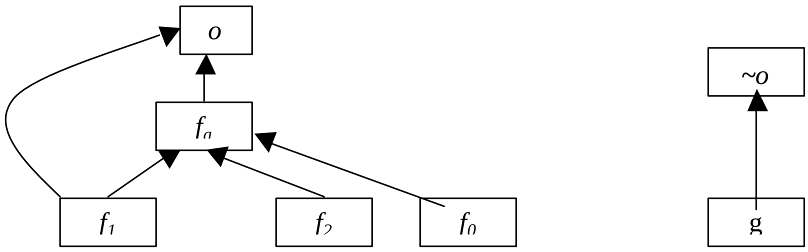

Figure 3: Example Factor Hierarchy

Suppose the plaintiff's theory $T_1 = <Cfds_1, Fds_1, R_1, Rpref_1, Vpref_1>$ is such that:

1. $Cfds_1 = \{<c_{prec}, F_{prec}, P>, <c_{curr}, F_{curr}, P>\}$, with $F_{prec} = \{f_0, f_1, g\}$, $F_{curr} = \{f_0, f_2, g\}$,
2. $Fds_1 = \{<f_0, P, v>, <g, D, v'>\}$,
3. $R_1 = \{<\{f_0\}, P>, <\{g\}, D>\}$,
4. $Rpref_1 = \{rpref(<\{f_0\}, P>, <\{g\}, D>\})$.

According to this theory, c_{prec} has the following explanation: c_{prec} had outcome P since rule $<f_0, P>$ applies, which is stronger then rule $<\{g\}, D>$. In other words, in premises set $F_{prec} \cup R_1 \cup Rpref_1$, argument $A_1 = < f_0, <f_0, P>, rpref(<\{f_0\}, P>, <\{g\}, D>) >$ is justified, having no defeaters. The same explanation also holds for c_{curr}: A_1 has no defeaters also in $F_{curr} \cup R_1 \cup Rpref_1$.

The defendant can reply by distinguishing c_{prec} from c_{curr}. For this purpose she needs to transform theory T_1 into the following theory $T_2 = <Cfds_2, Fds_2, R_2, Rpref_2, Vpref_2>$ which satisfies the following

1. $Cfds_2 = Cfds_1$,
2. $Fds_2 = Fds_1$ È $<f_1, P, v_1>$,
3. $R_2 = R_1$ È $\{<\{f_1\}, P>, <\{f_0, f_1\}, P>\}$,
4. $Rpref_2 = (Rpref_1 - rpref(<\{f_0\}, P>, <g, D>) + rpref(<\{f_0, f_1\}, P>, <g, D>)$

Note that after removing the preference $rpref(<\{f_0\}, P>, <g, D>)$, the new preference $rpref(<\{f_0, f_1\}, P>, <g, D>)$ can be added by "preferences from case". Theory T_2 allows the defendant to explain why c_{prec} had decision P, but does not support the conclusion that c_{curr} should also have decision P. In fact, c_{prec} is explained in T_2 by appealing to a rule ($<\{f_0, f_1\}, P>$), which is not applicable to c_{curr} (since c_{curr} does not include f_1). More exactly, the justified argument $<f_0, f_1, <\{f_0, f_1\}, P>, rpref(<\{f_0, f_1\}, P>, <g, D>>$, having conclusion P, is available in $F_{prec} \cup R_2 \cup Rpref_2$, but is not available in $F_{curr} \cup R_2 \cup Rpref_2$, since $f_1 \notin F_{curr}$.

Let us now consider the move "downplay distinction". This move also exploits the factors hierarchy, and in particular the fact that both factors f_1 in F_{prec} and factor f_2 in F_{curr} favour the abstract factor f_a, which favours P. To downplay the distinction the plaintiff can build the theory:

$T_3 = <Cfds_3, Fds_3, R_3, Rpref_3, Vpref_3>$ where

1. $Cfds_3 = Cfds_2$,
2. $Fds_3 = Fds_2 \cup \{<f_1, f_a, v>, <f_2, f_a, v>, <f_a, P, v>\}$,
3. $R_3 = R_2 \cup \{<\{f_1\}, f_a>, <\{f_2\}, f_a>, <\{f_a\}, P>, <\{f_0, f_a\}, P>\}$
4. $Rpref_3 = (Rpref_2 - rpref(<\{f_0, f_1\}, P>, <g, D>) + rpref(<\{f_0, f_a\}, P>, <g, D>)$

The plaintiff has removed from theory T_2 above the unwanted preference $rpref(<\{f_0, f_1\}, P>, <g, D>)$, and has added, instead, the new preference $rpref(<\{f_0, f_a\}, P>, <g, D>)$ (by preferences from case). This preference again allows both c_{prec} and c_{curr} to be explained according the same reasoning: in both cases rule $<\{f_0, f_a\}, P>$ is applicable, and prevails over rule $<g, D>$. This rule and the corresponding preference, in fact occurs in two arguments, which explains why P should be the decision of c_{prec} and c_{curr} respectively:

- $A_{prec} = <f_0, f_1, <\{f_1\}, f_a>, <\{f_0, f_a\}, P>, rpref(<\{f_0, f_a\}, P>, <g, D>)>$, *with* $A_{prec} \in JustArg(F_{prec} \cup R_3 \cup Rpref_3)$
- $A_{curr} = <f_0, f_2, <\{f_2\}, f_a>, <\{f_0, f_a\}, P>, rpref(<\{f_0, f_a\}, P>, <g, D>)>$, *with* $A_{curr} \in JustArg(F_{curr} \cup R_3 \cup Rpref_3)$

So, the plaintiff has achieved the result of disarming the defendant's attempt at distinguishing c_{prec} from c_{curr}: in his new theory a preference revealed by its explanatory role in the precedent can also be used to support the outcome he wants in the current case.

Let us now consider a concrete example. Assume a new case where a patient was cured by a doctor in a hospital, without there being a contract, and the doctor omitted to give a therapy to the patient, so damaging his health: the factors are: *Hospital*, *Omission*, *HealthDamage*). The issue to be decided is whether the doctor is liable for her omission: we still use P and D to mean the outcomes respectively favouring the plaintiff and the defendant, but now P means "the doctor is liable", and D means "the doctor is not liable". Assume that there is a precedent c_{prec} where a doctor was considered to be liable for the damage suffered by a patient, a contract being in place, even if the damage was due to an omission (the factors in c_{prec} were: *Contract*, *Omission*, *HealthDamage*). Now, *HealthDamage* favours outcome P in accordance with the value of *Health*, while *Omission* favours outcome D in accordance to the value of *Liberty* (which seems to require that nobody is punished for not taking an initiative). Assume also that the patient explains both this case and the current situation (with output P) according to the theory that causing a health damage produces the liability of the doctor, even when the damage is caused by omitting a therapy (rather then by providing a wrong therapy). In other words, he builds a theory:

$T_1 = <Cfds_1, Fds_1, R_1, Rpref_1, Vpref_1>$ such that:

1. $Cfds_1 = \{<c_{prec}, F_{prec}\ P>, <c_{curr}, F_{curr}, P>\}$, with $F_{prec} = \{Contract, Omission, HealthDamage\}$ and $F_{curr} = \{Hospital, Omission, HealthDamage\}$,
2. $Fds_1 = \{<HealthDamage, P, Health>, <Omission, D, Liberty>\}$
3. $R_1 = \{<\{HealthDamage\}, P>, <\{Omission\}, D>\}$,
4. $Rpref_1 = rpref(<\{HealthDamage\}, P>, <\{Omission\}, D>\}$,
5. $Vpref_1 = vpref(Health, Liberty)$

This theory allows the patient to provide the same explanation for the precedent and the current case: in both the rule that a doctor is liable for causing health damages applies, and it prevails over the rule that there is no liability for omissions. In fact, argument $A_1 = <HealthDamage, <\{HealthDamage\}, P>, rpref(<\{HealthDamage\}, P>, <\{Omission\}, D>>$, belongs to both $Justarg(F_{prec} \cup R_1 \cup Rpref_1)$ and $JustArg(F_{curr} \cup, R_1 \cup Rpref_1)$, since it strictly defeats its only attacker, that is $<Omission, <\{Omission\}, D>>$.

The doctor may now distinguish the precedent from the current situation by claiming that in the precedent there was a contract, i.e. she proposes a theory where a doctor's liability is explained by two factors, causing a health damage, and being bound by a contract to provide adequate care (which requires a liability upon the defaulting party, to advance the value of trust): only the *combination* of these two factors may prevail over the principle that there should be no liability for omissions. This is done by producing the following theory:

$T_2 = <Cfds_2, Fds_2, R_2, Rpref_2, Vpref_2>$, where:

1. $Cfds_2 = \{<c_{prec}, F_{prec}\ \ P>, <c_{curr}, F_{curr}, P>\},$
2. $Fds_2 = \{<HealthDamage, P, Health>, <Omission, D, Liberty>, <Contract, P, Trust>\},$
3. $R_2 = \{<\{HealthDamage\}, P>, <\{Omission\}, D>, <\{Contract\}, P>, <\{HealthDamage, Contract\}, P>\},$
4. $Rpref_2 = \{rpref(\{HealthDamage, Contract\}, P>, <\{Omission\}, D>)\},$
5. $Vpref_2 = vpref(\{Health, Trust\}, Liberty).$

In theory T_2 there is an explanation why c_{prec}, had decision P, which is not applicable to c_{curr}: this is given by the argument:

$A_2 = <HealthDamage, Contract, <\{HealthDamage, Contract\}, P>, rpref(\{HealthDamage, Contract\}, P>, <\{Omission\}, D>)>.$

Note that $A_2 \in Justarg(F_{prec} \cup R_1 \cup Rpref_1)$ but A_2 is not available in the current case $F_{curr,}$ since F_{curr} does not contain factor *Contract.*

The patient may downplay this distinction, by claiming that the existence of the contract implied that the doctor was warranting a careful performance, and that this was the real reason why a doctor should be held liable under a contract, according to the value of trust. He may also claim that the same warranty is also implicitly given by the practice of the medical profession in a hospital, regardless of the existence of a contract, so that the doctor in the current situation would still be liable for the same reasons (causing a health damage after warranting an adequate performance) that he was liable in the precedent. Here is the new theory of the patient:

$T_3 = <Cfds_3, Fds_3, R_3, Rpref_3, Vpref_3>$, where:

1. $Cfds_3 = \{<c_{prec}, F_{prec}\ \ P>, <c_{curr}, F_{curr}, P>\},$
2. $Fds_3 = \{<HealthDamage, P, Health>, <Omission, D, Liberty>, <Contract, P, Trust>, <Contract, Warrant, Trust>, <Hospital, Warrant, Trust>, <Warrant, P, Trust>\},$
3. $R_3 = \{<\{HealthDamage\}, P>, <\{Omission\}, D>, <\{Contract\}, P>, <\{Contract\}, Warrant>, <\{Hospital\}, Warrant>, <\{Warrant\}, P>, <\{HealthDamage, Warrant\}, P>\},$
4. $Rpref_3 = \{rpref(<HealthDamage, Warrant\}, P>, <\{Omission\}, D>)\},$
5. $Vpref_3 = \{vpref(\{Health, Trust\}, Liberty)\}.$

This theory allows the patient to explain both c_{prec} and c_{curr} by using the same rule $<\{HealthDamage, Warrant\}, P>$, and preference $rpref(<HealthDamage, Warrant\}, P>, <\{Omission\}, D>)$. This is done through the following justified arguments, available respectively in $F_{prec} \cup R_3 \cup Rpref_3$ and in $F_{curr} \cup R_3 \cup Rpref_3$:

- $A_{3,1} = <HealthDamage, Contract, <\{Contract\}, Warrant>, <\{HealthDamage, Warrant\}, P>, rpref(<HealthDamage, Warrant\}, P>, <\{Omission\}, D>)>$
- $A_{3,2} = <HealthDamage, Hospital, <\{Hospital\}, Warrant>, <\{HealthDamage, Warrant\}, P>, rpref(<HealthDamage, Warrant\}, P>, <\{Omission\}, D>)>$

After downplaying, it is still possible to re-introduce a distinction, by claiming that the distinguishing factor also causes another intermediate consequence, favourable to oneself, which does not hold in the current situation. Remember that downplaying takes place after one party distinguished precedent c_{prec} from the current situation c_{curr}. This is obtained by arguing that a factor f, which is only present in the precedent, was necessary for producing the outcome o of the precedent, according to a rule $<A \cup \{f\}, o>$. As we have just seen, the opponent can downplay the distinction by showing that p was produced via an abstract factor f_a, and that f_a follows both from the factor f in the precedent and the factor f' in the current situation. At this stage, however, the distinguishing party can try to reinstate the distinction, by showing that f also produces another intermediate factor f_{a2}, which is not produced by f'.

Assume that the doctor, to reinstate the distinction, claims that the contract in the precedent also implied that there was a consideration (which requires liability according to the value of reciprocity), and that both consideration and warranty are required to ground liability for health damages in cases of omissions. This is done by the following theory $T_4 = <Cfds_4, Fds_4, R_4, Rpref_4, Vpref_4>$, where:

1. $Cfds_4 = \{<c_{prec}, F_{prec}\ \ P>, <c_{curr}, F_{curr}, P>\},$
2. $Fds_4 = \{<HealthDamage, P, Health>, <Omission, D, Freedom>, <Contract, P, Trust>, <Contract, Warrant, Trust>, <Hospital, Warrant, Trust>, <Warrant, P, Trust>, <Contract, Consideration, Reciprocity>, <Consideration, P, Reciprocity>\},$
3. $R_4 = \{<\{HealthDamage\}, P>, <\{Omission\}, D>, <\{Contract\}, P>, <\{Contract\}, Warrant>, <\{Hospital\}, Warrant>, <\{Warrant\}, P>, <\{Contract\}, Consideration>, <\{Consideration\}, P>, <\{HealthDamage, Warrant, Consideration\}, P>\},$
4. $Rpref_2 = \{rpref(<HealthDamage, Warrant, Consideration\}, P>, <\{Omission\}, D>)\}$

This theory again allows the doctor to explain c_{prec} via a rule *<{HealthDamage, Warrant, Consideration}, P>}* which is not applicable to c_{curr}.

The dispute may then go on, with the patient still trying to downplay this further distinction (e.g. for example by claiming that the doctor was going to be paid for her work at the hospital in any case, so that in a sense there was a consideration), and the doctor trying to introduce further distinctions, still based on the absence of a contract in the current situation.

An alternative, although more restrictive, way of downplaying distinctions, given in Prakken, 2000, simply requires that some factor in the current case promote the same value as the distinguishing factor. If this is so, although the factors differ, the competing rules return the same values when given as arguments to *rval*, and thus can enjoy the same preference relations.

In contrast, emphasising a distinction does not give rise to new theories: it only draws attention to the non-availability of the downplaying move, and the consequent need for the opponent to resort to arbitrary preferences to repair his theory. In Prakken, 2000 this is expressed in terms of a difference between the values associated with the two sets of factors. The very difference in values alerts us to the significance of the distinction, which requires a consideration of the value preferences to resolve. The move is of course most effective, if the distinction relates to a more highly prized value.

A recent paper by Roth, 2000 suggests some further moves that can be made to augment the notions of downplaying and up-playing distinctions. These mostly turn on a richer account of intermediate moves in arguments: for example a factor may only promote an intermediate factor in the presence of some other factor. His example is that the potential to find another job may be an intermediate factor, but whether a particular job promotes this factor depends on the current state of the labour market. Such moves would require a more sophisticated logic to be used to apply the theories. It would also be an interesting exercise to see what extensions to our basic formalism would be necessary to represent the information needed to make such moves.

5. Future Work

Given the above framework, we can identify a number of issues that we think that it is important to explore. In this section we will briefly describe some of these.

5.1 Metrics for Theory Coherence

In section 3.4 we discussed the notion of theory coherence in qualitative terms, identifying a number of considerations that might lead us to think that one theory is better than another. Such comparison is unsatisfactorily vague, however, and it would be interesting to see whether the comparison might be made more precise. We have reported some steps towards this in Bench-Capon and Sartor, 2001. In that paper we draw on the work of Thagard, 1992, which developed a model for assessing competing scientific theories (for a more recent statement of Thagard's view in coherence, and for references to the coherence literature, see Thagard, 2001). The essential idea is to represent the evidence to be accounted for by a theory and the tenets of a theory as nodes connected by links representing support and conflict. A set of initial values (between 1 and -1) is assigned to these nodes, and these values are then propagated, support links increasing the values of nodes, and conflict links decreasing them. Moreover, links are subject to a rate of decay so that isolated nodes decrease in value. This propagation is continued through a number of cycles, until the values of the nodes stabilise. In Thagard's interpretation of this process, nodes which end with a high activation can be considered part of a coherent, and hence acceptable theory, while those with a low activation do not form part of that coherent theory, and so should be rejected.

We adapted this approach to our theories of bodies of case law taking cases as providing our evidence, and taking cases, rules, and preferences as nodes. These nodes can support one another in several ways: rules may be applicable to cases, rules may give rise to intermediate conclusions, rules can explain cases. Also rules may conflict with cases if they appear to be applicable, and yet would suggest the opposite outcome for the case. To deal with preferences, however, we needed to extend Thagard's approach. The effect of a rule preference is to render the less favoured rule inapplicable in a case. Thus the preference does not conflict with the rule it disfavours, but rather with the ability of that rule to conflict with the case to which it is deemed inapplicable. The preference thus should be seen not as decreasing the value of the rule node, but a decreasing the value of the link which adversely propagates its value to the case in question. In our model therefore the weights of links are not fixed, but can be affected by the propagation process. In particular preferences are in conflict with the links from the less favoured rules to cases where the preference applies. We modelled our theories, such as those given in section 3 above, as a set of connected nodes, assign initial values to the nodes, and then propagated these values until they stabilise, and which point we can see the average level of activation as an indicator of the coherence of the theory.

These preliminary experiments yielded some encouraging results, and showed that the numbers which emerged from the process accorded largely with our intuitions. A number of technical issues were raised by the experiments, for which the interested reader is referred to the original paper. Additionally the important question arises as to whether the approach has any cognitive validity with respect to the ways in which lawyers evaluate theories. Once the technical questions above have been answered, some kind of empirical study will be required to see how far the judgements on theories given by this approach can be seen as reflecting a legal consensus.

5.2 Comparison of Values

A central tenet of our account of theories for case based reasoning is that preferences between defeasible rules are justified in terms of preferences between sets of values. This means that we need a principled way of comparing sets of values. This has not so far been much studied, but is something that needs to be investigated. In this section we will identify some of the issues concerning this topic.

At least three questions need to be asked about values:

- Are values scalar? In the foregoing we have tended to see values as either promoted or not. It could, however, be argued that values can be promoted to different extents, according to which factor is involved, or, especially if we consider dimensions rather than factors, the degree to which the factor is realised.
- Can values be ordered at all? We have assumed that they can, but it is possible to see values as simply incommensurable.
- How can sets of values be compared? Is it always better that more values are promoted? Can values conflict so the promotion of their combination is worse than promoting either separately? Can several less important values *together* overcome a more important value?

We have, in our model of section 3, avoided these questions to some extent by requiring that *Vpref* be defined extensionally. Certainly we do assume an ordering on values, and it is important to us that different factors can promote the same value, and that the extent of such promotion can be ignored. In order to explain *Young* we rely on the promotion of *Mprod* by *pLiv* and *dLiv* to be "cancelled" when comparing sets containing them. Indeed the utility of introducing value preferences as an explanation of rule preferences relies on the ability of *different factors* to promote the *same values*, so that the set *Vpref* is smaller than *Rpref*, allowing a preference revealed in a case concerning one set of factors to applied in a case with different factors.

Two recent papers address these issues. Prakken, 2000 puts forward a formal account of teleological reasoning for case based systems which has considerable similarity with the basic thrust of our account. In particular Prakken explains rule preferences in terms of value preferences. Prakken assumes an ordering on values, does not recognise different degrees of promotion, and takes a position on comparison of sets.

Prakken first identifies an ordering on the values recognised within the theory, which he terms f_{valord}. Because rules promote sets of values, he then requires a way of comparing sets of values. He does this with a rule Val_{comp}, which expresses that a set of values, V_1 *is* better that a set of values V_2 if for every value present in V_2 but missing in V_1 there is a better value present in V_1 but missing in V_2. Essentially the most important value in a set is the primary determinant of its status: lesser values are used as tiebreakers with sets which contain exactly the same more highly rated values. Val_{comp} thus incorporates some choices: it means that a set of values is always better than any of it strict sub-sets, but it also means that if both sets being compared contain values not in the other set, the set containing the most preferred of these values is always preferred, no matter how many values from the other set it may be missing. For example given the ordering used in section 3,

Prop> Mprod> Llit

a set containing *Prop* will always be considered better that one without it. For example, it will give *{Prop}> {Mprod, Llit}*. In fact, although none of the example cases can test this preference, it seems reasonable enough: deciding a case in favour a person poaching on another's land who had caught a saleable animal would promote *{Mprod, Llit}*, whereas deciding against him would promote *{Prop}*. Since poaching is generally held to be undesirable, the preference would seem vindicated. It is unclear, however, whether this principle would hold in general; it does not seem impossible that a combination of less preferred values might be sufficient to defeat a single stronger value in some cases. We would not, however, wish to advance as a general notion that a larger set always defeats a smaller set: the poaching case provides a counter example to this.

Hage, 2001 offers a different approach to comparing sets of values (actually Hage speaks of *reasons* rather than values, but we take his reasons as including our values). Hage's approach does not insist on an initial ordering on values, but does introduce the notion of degrees of strength with which values are promoted. Hage provides a set of rules for comparing sets, but these are not complete; certain sets of values will be incommensurable. Thus Hage, gives us a relatively "safe" account, identifying those cases with which anyone can agree, but remaining silent on cases that might be disputed. Prakken, in contrast, offers a "bold" account, able to determine all questions, but not necessarily fitting with our intuitions in every situation.

We believe that the way forward, given that the role of values has only recently been identified, is to conduct some investigation into the way in which they are used in actual legal practice. This will give us a handle on how we should resolve the questions posed at the start of the section, and this in turn will inform the design of any machinery we wish to develop to express the behaviour we identify.

5.3 Changes in the social context.

In the model above, an important aspect is missing, namely, an account of the dynamics of case law, as it depends on the evolution of the socio-political context. These dynamics seems to undermine the very possibility of constructing a coherent theory of a case-law domain: how is it possible to fit in a single theory cases which were decided differently, even in the presence of the same constellations of factors and dimensions, since different decisions were required by different contexts?

One way of approaching this issue (cf. Sartor, 2002) is to focus on the link between factors and values, which is represented in our factor descriptions. Let us recall that a factor description had the form $<f, v, o>$, and meant that by responding to factor f with outcome o, we would promote value v. This makes both an evaluative judgement, the judgement that v is a value, a socially beneficial goal, and also a factual or empirical judgement, to the effect that by practising the factor-outcome link we promote the value.

The evaluative judgement involves deep and controversial philosophical issues. Are values objective, conventional or merely subjective? Are they eternal and universal or relative to particular times and places? The question of whether and how much certain values are going to be advanced through certain practices concerns an empirical connection, which undoubtedly is dependent upon changing socio-economical conditions. Even if ultimate legal values remain unchanged, the ways in which the practice of a specific rule impacts on them may change over time (a similar change would also concern instrumental values, but I will not consider them here).

For example, it may be argued that under the circumstances prevailing in modern industrialised countries, hunting has lost its ancient economic function: rather then contributing to productivity, it may detract from it. This may be true especially when hunting hinders some forms of recreation (watching wild animals, hiking, etc.) and so jeopardises the livelihood of those involved in the corresponding economical activities (hotel personnel, tour operators, tourist guides, etc.). In such a context, a factor description such as $<pChase, P, Mprod>$ (the fact that plaintiff is chasing a wild animal favours an outcome for his side, since this decisional practice, by facilitating hunting, promotes more productivity) is inappropriate: hunting does not promote social productivity, but rather impairs it.

To model this phenomenon, we need to temporalise factor descriptions (and dimension descriptions), so as to be capable of building theories which can explain conflicting decisions, adopted on the basis of the same set of factors, but taken at different times, when the impact of (the practice of) the rules using those factors on the relevant values has changed. This requires some changes and refinements in the definitions we have provided, but seems fully compatible with the bases of our approach. For an attempt in this direction, see Sartor, 2002.

5.4 Links to Texts of Decisions

In the debate on precedent, formalistic (strict) and anti-formalistic (sceptic) approaches are frequently opposed (cf. e.g. MacCormick, 1987, p. 157; Twinings and Mier, 1991, p. 311). The first approach construes the binding meaning of the precedent on the basis of the text of the opinion and the plausible intention of the judge. The record of the case includes therefore the detailed argumentation with was developed at the time when it was decided. The meaning of the case is reduced to one rule (the *ratio decidendi*) which can be extracted from that argumentation.

The latter approach looks beyond the text and its author, by considering interpretations given by subsequent judges, and more generally, by providing a holistic interpretation of the development of case law. The record of the case is therefore basically limited to the facts of the decision plus its outcome (according to the jurisprudential model proposed by Goodhart, 1959). It is up to the interpreter, using all materials involved, to provide an explanation in the framework of the body of the case law.

In our approach we have mainly taken the anti-formalistic side: the explanation of a case is not given by the expressed opinion of the judge, but by the theory which more coherently explains the body of the case law. This does not imply, however, that expressed opinions are necessarily irrelevant in a coherence-based approach, so that we can take on board the concerns that underlie the formalistic approach.

In fact, we may expand the background knowledge available to the parties, for example, with information concerning the statements of the judges and the context of their utterance. This would lead to a further theory-construction profile : the need to make sense of the "history" of the case, and in particular of the judges' opinions, in the circumstances where they were stated. So, the record of a case, besides the dimensional qualifications (or the factors) and the outcome, may also include the rules and arguments asserted by the judges. An additional profile of the coherence of a theory would be the way in which the theory can successfully incorporate those rules and arguments into the explanations it provides. How to implement this profile, and how to relate it to the aspects of coherence we indicate above, and to balance it with them, will be the object of further research.

5.5. Implementation

We have implemented the framework defined in Definitions 1-20 above in PROLOG. The process of implementation was relatively straightforward as the definitions have a fairly direct mapping into PROLOG. For example, the implementation of Definition 19 is:

```
arbitraryRulePref(Theory,R1,R2):-
     retract(theory(Theory,TCases,
       TFactors,TRules,TPrefs,TValPrefs)),
     member([R1,[F1,O]],TRules),
     member([R2,[F2,O2]],TRules),
     ruleval(F1,V1), ruleval(F2,V2),
     asserta(theory(Theory,TCases,TFactors,
          TRules,[pref(R1,R2)|TPrefs],
               [vp(V1,V2)|TValPrefs])).
```

Having implemented the definitions we embedded them in a menu system. The menu simply provides access to calls to the theory constructors, together with options to display the theory, list the cases explained by the theory, and to list the background cases and factor definitions. The program is perhaps not very exciting, but it does allow a user to construct theories ensured to be correct in accordance with the definitions, and to check that the expected consequences can be delivered[4].

We are currently expanding the implementation along the lines described above: extracting factors from dimensions, testing the coherence of the developed theories, and allowing for abstract factors. The ultimate aim would be to construct a program which provides an effective help in producing the best theory for a given side in a given new case against a given background of precedents, factors or dimensions. Such a program would require both heuristics to pick the theory constructors to use, and to evaluate the theories constructed.

6. Summary and Conclusions

The main object of this paper has been to present a formal account of reasoning with case law as building, evaluating and using theories. We have articulated the elements of such theories, provided constructors to build them, and suggested ways in which they can be evaluated and used. We have illustrated the formalism using a well known example. We have shown how such theories could be used to argue for and against positions with respect to as yet undecided cases, and to account for a body of decided cases. We have shown how our model can be expanded to accommodate some special features of particular systems. Finally we have shown how the basic moves in the currently most developed implementations of reasoning with case law can be expressed in our formalism. We believe that the work provides insight into reasoning with cases in the following ways:

- We see reasoning with cases as involving all of the construction of a theory, the application of that theory, and the evaluation of competing theories. The above gives a precise account of all these elements, while maintaining a clear conceptual separation between them.
- Our formal account is sufficient to reconstruct the reasoning and argument moves common to several existing case-based reasoning systems, making clear at which points theory construction, application and evaluation are involved.
- Our formal account can be extended to accommodate additional features built into some systems, such as dimensions in HYPO and the factor hierarchy of CATO, which go beyond reasoning with a flat set of factors.

---

[4] The initial Prolog program has recently been re-engineered in JAVA to provide a robust, fully functioning implementation of the basic model with a graphical user interface. Additionally this programme has the capacity to generate executable Prolog code corresponding to the theory. This work is described in Chorley and Bench-Capon, 2003.

- A problem with many models of legal reasoning is that they indicate what arguments can be made, but fail to account for why an argument might be persuasive. Our account distinguishes two levels of persuasion. Firstly we can show some arguments to be persuasive in terms of values: a value preference derived from a rule preference exhibited in a decided case can be used to derive a new rule preference applicable to other cases. Secondly arguments can be found persuasive by a comparative evaluation of the competing theories from which they derive. We have suggested some criteria which could be used to compare theories.
- We have identified a number of topics for future work, especially regarding comparison of values, and a quantifiable measure of theory coherence.

Of course, much work is needed to capture the full richness of current case based systems. None the less we think we have provided some firm foundations on which such work can be built.

ACKNOWLEDGEMENTS

This paper is a revised and much extended version of Bench-Capon and Sartor, 2001. We would like to thank all those who have discussed with us and commented on that paper, both during its preparation, and since its publication. We especially wish to thank Henry Prakken, Jaap Hage, Kevin Ashley and Edwina Rissland, and the anonymous referees of this paper.

REFERENCES

[1] Aleven, V. 1997. *Teaching Case Based Argumentation Through an Example and Models.* PhD Thesis. The University of Pittsburgh.
[2] Alexy, R., and A. Peczenik 1990. The Concept of Coherence and Its Significance for Discursive Rationality. *Ratio Juris* 3: 130-147.
[3] Ashley, K.D., and E.L. Rissland 1988. A Case-Based Approach to Modelling Legal Expertise. *IEEE Expert*, Fall 1988: 70-77.
[4] Ashley. K.D. 1990. *Modelling Legal Argument*. Bradford Books, MIT Press: Cambridge (Mass.).
[5] Bench-Capon, T.J.M. 2002. The Missing Link Revisited: The Role of Teleology in Representing Legal Arguments. *Artificial Intelligence and Law, Vol 10, 1-3, pp79-94.*
[6] Bench-Capon, T.J.M., and G. Sartor 2000. Theory Based Dialectics for Resolving Disagreement in Law. *ECAI workshop on Computation Dialectics*, Berlin, August 2000.
[7] Bench-Capon, T.J.M., and G. Sartor 2000. Using Values and Theories to Resolve Disagreement in Law. In J. Breuker, R. Leenes, and R. Winkels (eds*), Legal Knowledge and Information Systems: Jurix 2000,* 73-84. IOS Press, Amsterdam.
[8] Bench-Capon, T.J.M., and G. Sartor 2001. Theory Based Explanation of Case Law Domains. In *Proceedings of the Eighth International Conference on AI and Law,* 12-21. ACM Press: New York.
[9] Bench-Capon, T.J.M. 2001. Review of George C. Christie, The Notion of the Ideal Audience in Legal Argument, *Artificial Intelligence and Law* 9: 59-71.
[10] Bench-Capon, T.J.M., and E.L., Rissland, 2001. Back to the Future: Dimensions Revisited. In B. Verheij, A.R. Lodder, R.P. Loui, and A. J. Muntjewerff (eds) *Legal Knowledge and Information Systems*, IOS Press, Amsterdam, pp41-52.
[11] Bench-Capon, T.J.M., 2002. The Mising Link Revisited: The Role of Teleology in Representing Legal Argument, *Artificial Intelligence and Law*, volume 10, 1-3, pp 79-94.
[12] Berman, D.H., and C.L. Hafner. 1993. Representing Teleological Structure in Case Based Reasoning: The Missing Link. In *Proceedings of the Fourth International Conference on AI and Law,* 50-59. ACM Press, New York.
[13] Branting, L.K. 2000. *Reasoning with Rules and Precedents: A Computational Model of Legal Analysis.* Kluwer: Dordrecht.
[14] Chorley, A.H., and T.J.M. Bench-Capon, 2003. Developing Legal Knowledge Based Systems Through Theory Construction. Sunmitted to Ninth International Conference on Artificial Intelligence and Law.
[15] Dean, T., J. Allen, and Y. Aloimodis (1995). *Artificial Intelligence Theory and Practice*, Benjamin Cummings: Redwood City, California.
[16] Dung, P.M. 1995. On the Acceptability of Arguments and Its Fundamental Role in Nonmonotonic Reasoning, Logic Programming, and N-Person Games. *Artificial Intelligence* 77: 321-357.
[17] Forbus, J.D., 1984. Qualitative Process Theory. *Artificial Intelligence* 24, pp85-168.
[18] Frank, J. 1949. *Courts on Trial*, Princeton University Press: Princeton.
[19] Fuller, L., 1958. Positivism and Fidelity to Law: A Reply to Professor Hart *Harvard Law Review,* volume 71, pp630-672.
[20] Gardner, A. v.d.L., 1987. *An Artificial Intelligence Approach to Legal Reasoning*. Bradford Books, MIT Press, Cambridge, Mass.
[21] Van Gelder, A., K. A. Ross, and J.S. Schlipf (1991). The Well-Founded Semantics for General Logic Programs, Journal of the ACM, Vol 38, No 3.
[22] Goodhart, A.L. 1959. The Ratio Decidendi of a Case. *Modern Law Review* 21: 117.

[23] Hafner, C.L., and D.H. Berman, 2002. The Role of Context in Case-Based Legal Reasoning: Teleological, Temporal and Procedural. *Artificial Intelligence and Law*, Volume 10, 1-3 pp 19-64.

[24] Hage, J.C. 1997. *Reasoning With Rules*. Kluwer Academic Publishers, Dordrecht.

[25] Hage, J.C. 2000. Goal Based Theory Evaluation. In J. Breuker, R. Leenes and R. Winkels (eds*), Legal Knowledge and Information Systems: Jurix 2000,* 59-72. IOS Press, Amsterdam.

[26] Hage, J.C. 2001. Formalising Legal Coherence. In *Proceedings of the Eighth International Conference on AI and Law,* 22-31. ACM Press: New York.

[27] Hart, H.L., 1958. Positivism and the Separation of Law and Morals. *Harvard Law Review*, volume 71, pp593-629.

[28] Horty, J. F. 2001. Argument construction and reinstatement in logics for defeasible reasoning. *Artificial intelligence and Law* 9: 1-28.

[29] Levi, E.H. 1949. *An Introduction to Legal Reasoning*. University of Chicago Press: Chicago.

[30] MacCormick, N. 1987. Why cases have rationes and what these are. In Goldstein, L. (ed.) 1987. *Precedent in Law*, 155-182. Oxford: Oxford University Press.

[31] McCarty, L.T., and M.S. Sridharan 1982. *A Computational Theory of Legal Argument*. Technical Report LRP-TR-13, Computer Science Department, Rutgers University.

[32] McCarty, L.T. 1995. *An Implementation of Eisner v Macomber*. In Proceedings of the Fifth International Conference on AI and Law, 276-286. ACM Press: New York.

[33] Perelman, C. 1980. *Justice, Law and Argument*. Reidel: Dordrecht.

[34] Pollock, J.L. 1995. *Cognitive Carpentry: A Blueprint for How to Build a Person*. New York: MIT.

[35] Prakken, H., and G. Sartor 1996. A Dialectical Model of Assessing Conflicting Arguments in Legal Reasoning. *Artificial Intelligence and Law* 4: 331-368.

[36] Prakken, H., and G. Sartor 1997. Argument-based Extended Logic Programming with Defeasible Priorities. *Journal of Applied Non-Classical Logics* 1-2: 22-75.

[37] Prakken, H., and G. Sartor 1998. Modelling Reasoning with Precedents in a Formal Dialogue Game. *Artificial Intelligence and Law* 6: 231-287.

[38] Prakken, H. 2000. An exercise in formalising teleological case based reasoning. In J. Breuker, R. Leenes and R. Winkels (eds*), Legal Knowledge and Information Systems: Jurix 2000,* 49-57. IOS Press, Amsterdam.

[39] Raz, J. 1975. *Practical Reason and Norms.* London: Hutchinson.

[40] Rissland, E.L., E.M. Valcarce, and K.D. Ashley. 1984. Explaining and Arguing with Examples. In *Proceedings of AAAI 1984*, 284-294.

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

[42] Rissland, E.L., D.B.Skalak, and M.T. Friedman, 1996. BankXX: Supporting Legal Arguments Through heuristioc Retrieval. *Artificial Intelligence and Law* 4: 1-71.

[43] Rissland, E.L., and Ashley, K.D., 2002. A Note on Factors and Dimensions. *Artificial Intelligence and Law*, volume 10, 1-3, pp65-77.

[44] Roth, B. 2000. New Reasoning Patterns in Analogical Case Based reasoning: An Informal Investigation. In J. Breuker, R. Leenes and R. Winkels (eds*), Legal Knowledge and Information Systems: Jurix 2000,* 113-122. IOS Press: Amsterdam.

[45] Sartor, G. 2002. Telelogical Arguments and Theory Based Dialectics. *Artificial Intelligence and Law, Vol 10 1-3, pp95-112*.

[46] Schild, U.J., and S. Herzog 1993 The Use of Meta-Rules in Rules Based Legal Computer Systems. In *Proceedings of the Fourth International Conference of AI and Law,* 100-109. ACM Press: New York, pp.

[47] Skalak, B.B., and E.L. Rissland 1991. Arguments and Cases: An Inevitable Intertwining. *Artificial Intelligence and Law* 1: 3-44.

[48] Thagard, P. 1992. *Conceptual Revolutions.* Princeton(NJ): Princeton University Press.

[49] Thagard, P. 2001. *Coherence in Thought and Action*. Cambridge (MA): MIT Press.

[50] Twining, W., and D. Miers 1991. *How to do Things with Rules.* London: Butterworth.