Precedent, Deontic Logic, and Inheritance

John F. Horty
Philosophy Department and
Institute for Advanced Computer Studies
University of Maryland
College Park, MD 20742
Email: horty@umiacs.umd.edu

Abstract

The purpose of this paper is to establish some connections between precedent-based reasoning as it is studied in the field of Artificial Intelligence and Law, particularly in the work of Ashley, and two other fields: deontic logic and nonmonotonic logic. First, a deontic logic is described that allows for sensible reasoning in the presence of conflicting norms. Second, a simplified version of Ashley's account of precedent-based reasoning is reformulated within the framework of this deontic logic. Finally, some ideas from the theory of nonmonotonic inheritance are employed to show how Ashley's account might be elaborated to allow for a richer representation of the process of argumentation.

1 Introduction

The purpose of this paper is to establish some connections between precedent-based argument as it is studied in the field of Artificial Intelligence and Law, particularly in the work of Ashley [2], and two other fields: deontic logic and nonmonotonic reasoning.

The deontic logic appealed to here is a formalism originally inspired by van Fraassen [14] and then developed in more detail in my [7, 9] for reasoning in the presence of conflicting norms. This logic is described in Section 2 of the present paper; although the logic is not new, the presentation has been streamlined considerably. Section 3 shows how to extend this framework to cover the kind of reasoning with conflicting norms found in case law. A simplified account of reasoning with precedents, based on Ashley's theory, is reformulated after the pattern of this deontic logic; the idea of the reformulation is to make literal sense of the intuition that past precedents show how current cases ought to be decided.

Ashley's theory has recently been criticized for capturing only a restricted account of the information provided by precedents, collapsing the entire process of reasoning contained in a precedent case into an immediate connection between the input facts of that case and its conclusion. In Section 4, I suggest one way in which techniques from the field of nonmonotonic reasoning can be used to enhance the theory so as to allow for a richer representation of the process of argumentation. Unlike similar work by Prakken and Sartor [12], this suggestion does not rely on a general defeasible logic, but instead on some ideas developed in the study of nonmonotonic inheritance reasoning [8]. Finally, Section 5 explores the question of whether the notion of preemption familiar from inheritance reasoning has applications in the theory of precedent-based argument.

ICAIL-99 Oslo, Norway Copyright (c) 1999 ACM 1-58113-165-8/99/6 $5.00

2 Deontic logic

Standard deontic logic is formulated under the assumption of normative consistency—the idea that any reasonable set of ought statements, representing what ought to be or what some agent ought to do, must be consistent, so that they can in principle all be satisfied at once. Of course, it often seems that we face conflicting oughts, and there are a number of vivid examples in philosophy and literature. One of the best known is Sartre's description [13] of a student during the Second World War who felt for reasons of patriotism that he ought to leave home in order to join the Free French, but who felt also, for reasons of personal devotion, that he ought to stay at home in order to care for his mother. The rationale behind standard deontic logic is that, in situations like this, the appearance of conflict is misleading—that such conflicts can always be reconciled, always resulting in a consistent set of oughts. It is also possible, however, to take these situations at face value, and to suppose that, at times, some of us really do confront genuine, and irreconcilable, normative conflicts. From the standpoint of deontic logic, the technical challenge presented by such a view is to design a formalism for reasoning coherently in these situations.

This section sketches such a formalism, a deontic logic inspired by van Fraassen [14], and developed in more detail in my [7, 9]. We focus here on a conditional version of this logic, where a conditional ought statement of the form 'It ought to be that B, given A' is symbolized as $\bigcirc(B/A)$.

Van Fraassen's approach is based on the idea that deontic logic can usefully be formulated against the background of a set of imperatives, representing the dictates of various sources of authority. Since we are concerned here with conditional oughts, we concentrate also on conditional imperatives, such as

When interest rates are high, buy bonds,

symbolized as $!(B/H)$, with the exclamation point as an imperative operator. Where i is such an imperative, we let $Ant(i)$ stand for its antecedent and $Con(i)$ its consequent; if i were the imperative displayed above, for example, we would have the statements H as $Ant(i)$ and B as $Con(i)$.

A conditional imperative can be fulfilled or violated only in situations in which its antecedent is satisfied, in which case it is said to be fulfilled if its consequent is also satisfied, and violated otherwise. Of course, since an agent might recognize conflicting sources of authority, and since even a single source of authority can issue conflicting imperatives, this picture must allow for the possibility that an agent might find himself constrained by a set of imperatives that cannot all be fulfilled at once. As an example, imagine that an agent is subject both to the imperative displayed above and also to the imperative

When inflation is projected, don't buy bonds,

represented as $!(\neg B/P)$. In that case, the agent would confront conflicting imperatives in any situation satisfying the formula $H \wedge P$—where interest rates are high but inflation is projected anyway—since one or the other of these two imperatives would then have to be violated.

Supposing, then, that $\mathcal{I}$ is the background set of imperatives governing some agent, how do we determine whether a particular conditional ought $\bigcirc(B/A)$ is supported? We answer this question in two stages, first identifying the subset of those imperatives belonging to $\mathcal{I}$ that might be thought to have some bearing or relevance under the antecedent condition A of the ought statement, and then specifying how the oughts dependent on this antecedent condition are calculated from these relevant imperatives.

As an initial suggestion, it might seem natural to identify the imperatives relevant under some antecedent condition as those whose antecedents are themselves guaranteed to hold under that condition. Let $|A|$ represent the *model class* of the statement A, the set of models in which A is true. To say that B is guaranteed to hold under the condition A, is simply to say that B is true wherever A is, or that $|A| \subseteq |B|$. Our initial suggestion can thus be captured by taking the imperatives that are relevant under some condition to be the applicable imperatives, defined as follows.

Definition 1 (Applicable imperatives; $AP_{\mathcal{I}}(A)$) Where $\mathcal{I}$ is a set of conditional imperatives and A is a statement, the set of imperatives from $\mathcal{I}$ that are *applicable* under the condition A is $AP_{\mathcal{I}}(A) = \{i \in \mathcal{I} : |A| \subseteq |Ant(i)|\}$.

As it turns out, this initial suggestion is too liberal, forcing us to consider too many imperatives as relevant. To see this, imagine that the agent's background set consists of the two imperatives

If you're served asparagus, eat it with your fingers,
If you're served asparagus in a cream sauce, don't eat it with your fingers,

represented as $!(F/A)$ and $!(\neg F/A \wedge C)$. And suppose the agent is served asparagus in cream sauce in a restaurant, leading to the statement $A \wedge C \wedge R$ as an antecedent condition. In that case, both of the background imperatives are applicable—so both would be classified as relevant—since the antecedents of both imperatives are triggered by the antecedent condition.

It seems more natural, however, to say that only the second of these two imperatives should be classified as relevant under the circumstances: since the second imperative is based on more specific information, it is natural to suppose that, in any situation in which both are triggered, this second imperative should override the first. In order to capture this intuition—that one imperative can be overridden, rendered irrelevant, by another that is more specific but conflicting—we first order the imperatives according to the specificity of their antecedents. Where i and j are imperatives, let us say that j is more specific than i—written $i < j$—just in case $|Ant(j)| \subset |Ant(i)|$. We can then take the relevant imperatives as those that are applicable, but for which there are no more specific, conflicting imperatives that are also applicable; these can be collected together as the most applicable imperatives.

Definition 2 (Most applicable imperatives; $AP^{+}_{\mathcal{I}}(A)$) Where $\mathcal{I}$ is a set of conditional imperatives and A is a statement, the set of imperatives from $\mathcal{I}$ that are *most applicable* under the condition A is $AP^{+}_{\mathcal{I}}(A) = \{i \in \mathcal{I} : |A| \subseteq |Ant(i)|$ and there is no $j \in \mathcal{I}$ such that (1) $|A| \subseteq |Ant(j)|$,
(2) $i < j$, and (3) $\{A, Con(i), Con(j)\}$ is inconsistent$\}$.

Having identified the imperatives that are relevant under a particular condition as the most applicable imperatives, there are now two general approaches that might be followed in evaluating conditional oughts on the basis of these imperative, corresponding to the *credulous* and *skeptical* strategies in nonmonotonic reasoning. The first approach—much like that suggested in van Fraassen's paper, and corresponding to the credulous reasoning strategy—defines a conditional ought $\bigcirc(B/A)$ as following from a set of imperatives just in case, whenever the antecedent A holds, satisfying the consequent B is a necessary condition for fulfilling some maximal consistent set of relevant imperatives. In order to develop this idea formally, we first lift some of our basic concepts from individual statements to sets: where $\mathcal{F}$ is a set of statements, we let $|\mathcal{F}| = \bigcap\{|A| : A \in \mathcal{F}\}$; and where $\mathcal{I}$ is a set of imperatives, we let $Con[\mathcal{I}] = \{Con(i) : i \in \mathcal{I}\}$. The credulous approach can now be captured by defining the appropriate notion of consequence.

Definition 3 (Credulous consequence) A statement $\bigcirc(B/A)$ is a *credulous consequence* of an imperative set $\mathcal{I}$ if and only if $|\mathcal{F}| \cap |A| \subseteq |B|$ for some maximal consistent subset $\mathcal{F}$ of $Con[AP^{+}_{\mathcal{I}}(A)]$.

Broadly speaking, the idea underlying the credulous reasoning strategy is that a conclusion should be drawn from a body of information if one way of looking at things—in this case, one maximal consistent set of relevant imperatives—supports the conclusion. The idea underlying the skeptical strategy, by contrast, is that a conclusion should be drawn only when that conclusion is supported by every way of looking at things. Adapting this idea to the present setting, the skeptical approach can be set out as follows.

Definition 4 (Skeptical consequence) A statement $\bigcirc(B/A)$ is a *skeptical consequence* of an imperative set $\mathcal{I}$ if and only if $|\mathcal{F}| \cap |A| \subseteq |B|$ for each maximal consistent subset $\mathcal{F}$ of $Con[AP^+_{\mathcal{I}}(A)]$.

To illustrate these two approaches, let us suppose that

$$\mathcal{I} = \{!(B/H), !(\neg B/P), !(F/A), !(\neg F/A \wedge C)\}$$

is the agent's background set, containing all the imperatives set out so far. It is easy to see, first of all, that $AP^+_{\mathcal{I}}(H \wedge P) = \{!(B/H), !(\neg B/P)\}$—both financial imperatives are relevant under conditions of high interest and predicted inflation. Evidently, there are two maximal consistent subsets of $Con[AP^+_{\mathcal{I}}(H \wedge P)]$, both $\{B\}$ and $\{\neg B\}$. The credulous strategy thus yields both $\bigcirc(B/H \wedge P)$ and $\bigcirc(\neg B/H \wedge P)$ as consequences, since each of these statements is supported by one of these two maximal consistent subsets; the skeptical strategy yields neither, since neither statement is supported by both. Next, we can see that $AP^+_{\mathcal{I}}(A \wedge C \wedge R) = \{!(\neg F/A \wedge C)\}$—only the second asparagus imperative is relevant under the condition that the agent is served asparagus in cream sauce in a restaurant. Since $Con[AP^+_{\mathcal{I}}(A \wedge C \wedge R)]$ has only the single maximal consistent subset $\{\neg F\}$, both the credulous and skeptical approaches yield $\bigcirc(\neg F/A \wedge C \wedge R)$ as a consequence.

These two theories, credulous and skeptical, exemplify the style of deontic logic that forms the background of this paper. In fact, both theories still leave a number of matters unsettled, as detailed in [9]. Nevertheless, they seem to provide a promising starting point for the development of a conditional deontic logic for reasoning with conflicting normative information, and I will not attempt here to refine them any further. Instead, I want to explore some ways in which the general framework underlying these theories might be modified to apply in a different domain, providing a formal account of reasoning with the patterns of conflicting norms found in case law.

The logics sketched here are based on a picture of imperatives as directives flowing from some source of authority; it is assumed that the agent might be subject to a number of different imperatives, possibly conflicting, which he must bring to bear on a particular situation in deciding what to do. As it turns out, it is also useful to think of past legal cases themselves as imperatives, whose normative force is provided by the rule of precedent, or *stare decisis*—that like cases should be decided alike. And from this perspective, a judge or decision making authority can likewise be viewed as applying the imperatives supplied by past cases to the present situation. Of course, there is nothing new in the mere suggestion that past cases can be viewed as a source of obligation in judicial decision. In his definitive treatment of the topic, for example, Cross relies on the notion of obligation in explicating the rule of precedent:

> When it is said that a court is bound to follow a case, or bound by a decision, what is meant is that the judge is under an obligation to apply a particular *ratio decidendi* to the facts before him in the absence of a reasonable legal distinction between those facts and the facts to which it was applied in the previous case [5, pp. 102–103].

The goal of the present paper is to try to make literal sense of this suggestion by setting out a formal account of the oughts generated by past cases.

3 Precedent

We begin with our representation of cases, a highly simplified version of the formalism introduced by Ashley in [1, 2] for describing the theory of legal argument underlying his HYPO system.

Any particular legal case is characterized by a number of incidental features—the particular individuals involved, their personal characteristics, and so on—most of which are not legally relevant. Those features of a case that are legally relevant are described as *factors*. Different kinds of cases exhibit different factors. A case in the domain of trade secrets law—the original application domain of HYPO—typically concerns the question whether the defendant has gained unfair competitive advantage over the plaintiff through the misappropriation of a trade secret; and here, for example, the factors involved might include whether the plaintiff took measures to protect the trade secret, whether the secret was disclosed to outsiders, and whether the defendant actually did gain a significant competitive advantage.

Factors have polarities, generally favoring one side of the other: in a trade secrets case, for instance, the presence of security measures favors the plaintiff, since it strengthens the claim that the information secured was in fact a valuable trade secret. In addition to polarities, many factors also have magnitudes: we might want to know, not simply whether security measures were present, but how extensive those measures were, with more extensive security measures providing greater support for the plaintiff's claim. Both polarities and magnitudes of factors are represented in Ashley's formalism, and play a central role in his theory of legal argument. Nevertheless, for simplicity, we adopt a more abstract treatment here, considering factors only as legally relevant properties, ignoring polarities and magnitudes.

In the present paper, a *precedent case* will be treated simply as a set of factors together with an outcome, a decision reached on the basis of those factors by some decision making authority. Formally, we can represent such a case as a pair $\langle\{f_1, \ldots, f_n\}, s\rangle$, where $f_1, \ldots, f_n$ are factors and s is the side in whose favor the case was decided. We take *Fac* and *Out* as functions mapping a case into its factors and its outcome respectively; if c were the above case, for example, we would have $Fac(c) = \{f_1, \ldots, f_n\}$ and $Out(c) = s$. Following Ashley, we will suppose to begin with that s is either π or δ, with π representing an outcome in favor of the plaintiff and δ an outcome in favor of the defendant. And where s is an outcome, we let $\overline{s}$ represent the opposite outcome: $\overline{\pi} = \delta$ and $\overline{\delta} = \pi$.

A *problem situation* is defined as a set of factors without an associated outcome. Let us suppose that Γ is a set of background cases from some particular domain, and that the deciding authority is faced with a problem situation $X = \{f_1, \ldots, f_n\}$ from that domain, for which he must render a decision in accord with the rule of precedent. Our task, then, is to define the conditions under which the case base Γ supports the conclusion that the situation X ought to be decided in favor of the side s—a conclusion we write as $\bigcirc(s/X)$.

As suggested earlier, the current analysis of precedent is modeled on our treatment of conditional oughts, but there is one crucial difference. In the previous account, we characterized an imperative i as applicable under a condition A only if any model satisfying the condition A was guaranteed also to satisfy the antecedent $Ant(i)$ of the imperative, so that the antecedent of the imperative could be thought of as a generalization of the condition. The most direct analogy to this previous account would take a background case c as relevant to a current problem X only if any situation exemplifying the factors belonging to X would also have to exemplify the factors $Fac(c)$ present in the case—that is, only if $Fac(c) \subseteq X$, so that, again, the factor description of the case could be thought of as generalizing that of the problem. This direct analogy, however, results in too severe a requirement for relevance. In case law, a precedent might be applicable to a problem situation, even if it does not provide a more general description of that situation, as long as it is similar to that situation in some legally relevant way. To capture this idea, we define the background cases that are relevant to a problem situation—the cases that are on point—as those that share some factor with that problem.

Definition 5 (On point cases; $OP_\Gamma(X)$) Where Γ is a set of cases and X is a problem situation, the set of cases from Γ that are *on point* relative to X is $OP_\Gamma(X) = \{c \in \Gamma : X \cap Fac(X) \neq \emptyset\}$.

Apart from this difference—that we require only overlap of factors, rather than subsumption, to establish the relevance of a case—the present account follows the previous treatment of conditional oughts rather closely. In our previous treatment, we found it necessary to focus, not simply on the imperatives that are applicable under a given condition, but on the most applicable imperatives; and it is likewise necessary to focus on the cases that are most on point with respect to the given problem situation, most relevant. Where c and d are cases and X is a problem situation, let us say that d is more on point than c to X—written $c <_X d$—whenever $(Fac(c) \cap X) \subset (Fac(d) \cap X)$; the idea, of course, is that d is more relevant than c if it is more similar to X, sharing more factors with X. It is then natural to define the background cases that are most on point relative to a problem as those cases that are on point, but for which there are no more on point cases supporting the opposite decision.

Definition 6 (Most on point cases; $OP^+_\Gamma(X)$) Where Γ is a set of cases and X is a problem situation, the set of cases from Γ that are *most on point* relative to X is $OP^+_\Gamma(X) = \{c \in \Gamma : X \cap Fac(c) \neq \emptyset$ and there is no $d \in \Gamma$ such that (1) $X \cap Fac(d) \neq \emptyset$, (2) $c <_X d$, and (3) $Out(d) = \overline{Out(c)}\}$.

Because our representation of cases is developed within a formalism that is so much simpler than the full propositional language underlying conditional oughts, our treatment of the oughts generated by cases can be simpler as well; and in particular, since there are no logical interactions among the outcomes of different cases, we need only consider single cases as reasons for reaching a conclusion, rather than something like maximal consistent sets. The credulous approach thus tells us that some problem X ought to be decided in favor of the side s whenever there is some most on point case that favors s; the skeptical approach tells us that X ought to be decided in favor of s whenever each most on point case favors s.

Definition 7 (Consequence: cases) A statement $\bigcirc(s/X)$ is a *credulous consequence* of a case base Γ if and only if $Out(c) = s$ for some case $c \in OP^+_\Gamma(X)$. A statement $\bigcirc(s/X)$ is a *credulous consequence* of a case base Γ if and only if $Out(c) = s$ for each case $c \in OP^+_\Gamma(X)$.

Setting aside differences of notation and scope, it can be seen that the notion of skeptical consequence defined here corresponds to Ashley's argument evaluation criterion from Section 9.3 of [2]: the current approach defines $\bigcirc(s/X)$ as a skeptical consequence just in case the side s has, in Ashley's sense, a stronger argument.

To illustrate these various definitions, let Γ be a simple case base containing only two cases: $c_1 = \langle\{f_1, f_2, f_3\}, \pi\rangle$ and $c_2 = \langle\{f_1, f_4, f_5\}, \delta\rangle$. Consider first the problem situation $X = \{f_1, f_6\}$. Here, it is easy to see that $OP^+_\Gamma(X) = \{c_1, c_2\}$. Since both cases are classified as most on point, and they support different outcomes, the credulous approach yields both $\bigcirc(\pi/X)$ and $\bigcirc(\delta/X)$ as conclusions, while the skeptical approach yields neither. Next, consider the situation $Y = \{f_1, f_4, f_6\}$. Since $c_1 <_Y c_2$, we now have $OP^+_\Gamma(Y) = \{c_2\}$, so that both the credulous and skeptical approach yield $\bigcirc(\delta/Y)$ as the unique conclusion. Readers familiar with the literature on nonmonotonic inheritance will note that the X presents a situation like the classic Nixon Diamond, where the cases c_1 and c_2 provide incomparable reasons for conflicting conclusions, while Y is like a Tweety Triangle, where c_2 is thought to provide better information than c_1. Notice that, in contrast to ordinary inheritance reasoning, which relies on specificity to adjudicate between conflicting rules, c_2 is here taken to provide better information about Y than c_1 even though c_2 is not itself more specific than c_1, but only more similar to Y.

4 Extended arguments

The simple account of precedent-based reasoning developed here is subject to some severe limitations. Apart from relying on a rudimentary treatment of similarity as factor overlap, it involves, as noted earlier, a view of factors that abstracts away from the polarities and magnitude that give Ashley's account much of its texture. Although I believe it would not be difficult to introduce polarities and magnitudes into the present account, I now want to focus instead on a different limitation, which this account shares with Ashley's.

Both accounts reflect a view of precedent much like that of Goodhart [6], according to which the meaning, or *ratio decidendi*, of a precedent case is exhausted by the material facts of that case together with its outcome, and is not affected by any of the intermediate reasoning steps that may have led to that outcome. This hard-headed perspective, approaching that of legal realism, still has some adherents. However, it is often argued—for instance, by Branting [3, 4] and by Prakken and Sartor [12]—that an accurate model of legal reasoning must be based on a more liberal perspective, which accords at least some meaning to the intermediate steps through which the outcome in a precedent case is determined.

In the remainder of this paper, I will sketch one way in which the simple account set out so far of the oughts generated by cases can be extended to allow for intermediate reasoning steps. Because the present account treats cases as defeasible rules, the theory developed here is closest in conception to that of Prakken and Sartor. There is, however, a significant difference between the two theories. Prakken and Sartor attempt to analyze precedent-based reasoning using a fully expressive logical language, the formalism for general defeasible reasoning developed in their [11]. The present account relies on a more limited formalism especially tailored to the task at hand, but closely patterned on the theories of defeasible inheritance reasoning surveyed in [8]. These two accounts thus exhibit the contrast that is often seen between top-down and bottom-up approaches in knowledge representation.

We begin by relaxing our notation so that the case base Γ can contain items of the form $\langle\{f_1,\ldots,f_n\}, f_k\rangle$, carrying the intuitive meaning that the presence of the factors $f_1,\ldots,f_n$ in a situation supports the presence of the factor f_k. Adapting Branting's phrase, we refer to such an item as a *precedent constituent*, noting that, if the outcomes π and δ are considered as special factors, the class of precedent constituents can be thought of as including the precedents. We extend our *Fac* and *Out* notation in the natural way, so that, where c is the above precedent constituent, we have $Fac(c) = \{f_1,\ldots,f_n\}$ and $Out(c) = f_k$. And as before, we take f and $\overline{f}$ as opposite factors, with $\overline{\overline{f}}$ identical to f.

When an authority confronts a problem situation X and decides in favor of a side s, we suppose that the force of his decision is to supplement the case base Γ with a number of precedent constituents, but we make no assumptions concerning the exact nature of this supplementation: it might include the simple precedent $\langle X, s\rangle$, but it might include, instead or in addition, a number of precedent constituents representing the intermediate reasoning steps that led the authority to this conclusion. Our goal is show how the entire body of precedents and precedent constituents present in Γ can be used to justify a decision in a new problem situation.

A fundamental notion in our analysis is that of an argument based on a particular situation, analogous to the notion of a reasoning path in an inheritance network, and defined as follows.

Definition 8 (Argument) An argument based on a situation X_0 is a sequence of the form

$$X_0 \underset{c_1}{\rightarrow} X_1 \underset{c_2}{\rightarrow} X_2 \rightarrow \cdots \underset{c_n}{\rightarrow} X_n,$$

where $0 \leq n$, and where (1) $X_i \cap Fac(c_{i+1}) \neq \emptyset$ and (2) $X_{i+1} = X_i \cup \{Out(c_{i+1})\}$.

The idea here is that X_0 represents the original problem situation, which the reasoning agent then successively augments with factors taken from relevant precedent constituents, arriving at more and more extensive characterizations of that situation. Each link $X_i \underset{c_{i+1}}{\rightarrow} X_{i+1}$ in the argument indicates that an appeal to the precedent constituent c_{i+1} is involved in moving from the characterization X_i to the augmented characterization X_{i+1}. Clause (1) tells us that c_{i+1} must be on point with respect to the factors established in X_i; clause (2) tells us that X_{i+1} results from augmenting that characterization with the outcome factor of c_{i+1}. Where f is a factor belonging to the final node X_n, the argument as a whole is said to *support* the conclusion that the factor f can be included in the characterization of the original situation X_0

We take a schema of the form $\alpha(X_0, X_1, \ldots, X_{n-1}, X_n)$ to refer to an arbitrary argument based on X_0 and leading to the augmented characterization X_n, which passes through at least the intermediate characterizations $X_1,\ldots,X_{n-1}$. Not all intermediate nodes in the argument need be displayed; for example, the schema $\alpha(X, U, Y)$ is taken to represent any argument based on X and terminating in Y, which at some point passes through the intermediate node U, though it may pass through others as well. For notational convenience, we allow the argument schema $\alpha(X, X)$ to represent also the *degenerate* argument X, which does not involve any further elaboration beyond the mere statement of the facts from the initial problem situation. An argument of the form $\alpha(X, U) \underset{c}{\rightarrow} Y$ which does elaborate the original problem situation we refer to as *non-degenerate*.

Where Γ is a case base and Φ is a set of arguments based on the situation X, we define a triple of the form $\langle\Gamma, X; \Phi\rangle$ as a *reasoning context*. The intuitive picture is that the agent starts with a case base Γ and an initial problem situation X as inputs; after a certain amount of reasoning, he is then led to accept the set Φ as an *interpretation* of the case base applied to that problem situation—a set of arguments supporting the conclusion that various further factors can be included in the characterization of that situation. With respect to any given context, certain arguments can be classified as *forcible*, or convincing. This notion of a forcible argument is a central concept in the theory presented here, and is defined through the three preliminary ideas of constructibility, conflict, and trumping.

Constructibility is a form of chaining, a means of extending an already accepted argument with the information provided by some relevant precedent constituent from the case base.

Definition 9 (Constructibility) An argument of the form $\alpha(X, U) \underset{c}{\rightarrow} Y$ is constructible in the context $\langle \Gamma, X; \Phi \rangle$ if and only if $\alpha(X, U) \in \Phi$ and $c \in \Gamma$.

Suppose, for example, that Φ contains the argument

$$\{f_1, f_2\} \underset{c_1}{\rightarrow} \{f_1, f_2, f_3\},$$

and that Γ contains the constituent $c_2 = \langle\{f_3\}, f_4\rangle$. It then follows that the argument

$$\{f_1, f_2\} \underset{c_1}{\rightarrow} \{f_1, f_2, f_3\} \underset{c_2}{\rightarrow} \{f_1, f_2, f_3, f_4\}$$

is constructible in the context $\langle \Gamma, X; \Phi \rangle$.

Like constructibility, the treatment of conflict is also straightforward, specifying two arguments as conflicting whenever they support the presence of opposing factors in the characterization of the same initial situation, and then defining an argument as conflicted in a context if it conflicts with another argument already present in that context.

Definition 10 (Conflict, conflicted) Arguments of the form $\alpha(X, Y)$ and $\alpha(X, Z)$ are said to *conflict* if and only if there is some property f such that $f \in Y$ and $\overline{f} \in Z$. The argument $\alpha(X, Y)$ is *conflicted* in the context $\langle \Gamma, X; \Phi \rangle$ if and only if Φ contains some argument that conflicts with it.

As an example, the argument

$$\{f_1, f_2\} \underset{c_1}{\rightarrow} \{f_1, f_2, f_3\} \underset{c_2}{\rightarrow} \{f_1, f_2, f_3, \overline{f_4}\}$$

conflicts with the argument displayed above, and would be conflicted in any context $\langle \Gamma, X; \Phi \rangle$ in which Φ already contained that argument.

We now turn to the notion of trumping—the idea that, in case of conflict, arguments based on some constituents should be preferred to arguments based on others because those constituents are more on point, more relevant to current situation. Of course, when our attention is restricted to single-level arguments, as in Ashley's theory, the treatment of relevance between a constituents and the situation at hand is unproblematic: there, as we have seen, a constituents d can be classified as more relevant than a constituents c to the original problem situation X whenever the factors that d shares with X extend those shared by c—formally, whenever $c <_X d$. The matter is more complicated, however, when we turn to extended arguments. Here it seems clear that derived factors should be included along with those belonging to the original problem situation in assessing the relevance of constituents, but this decision still does not determine a unique treatment of trumping.

Consider an argument of the form $\alpha(X, U) \underset{c}{\rightarrow} Y$, which begins with an initial situation X, extends the characterization of this situation to U, and finally appeals to the constituents c to extend the characterization still further to Y. And suppose the case base contains a constituents d whose outcome is opposite to that of c: $Out(d) = \overline{Out(c)}$. Perhaps the simplest idea, then, is to say that the argument $\alpha(X, U) \underset{c}{\rightarrow} Y$ should be trumped whenever d is more relevant than c on the basis of the derived characterization U—that is, just in case $c <_U d$. Unfortunately, this idea is too simple, as we can we with an example.

Imagine an academic department in which graduate students receive financial support, but where their presumption of continued support is frequently re-evaluated, often on a case by case basis, in a way that is sensitive to precedent. Suppose the department's case base in these matters includes precedent constituents reflecting the following decisions: a student who is late fulfilling requirements is denied further support; support is continued for a student who is late but has an excuse; illness is classified as an excuse. Let us abbreviate the factors involved in these decisions as follows: f_1 = the student is late; f_2 = the student is ill; f_3 = the student has an excuse; f_4 = the student receives support. The constituents themselves can then be represented as $c_1 = \langle\{f_1\}, \overline{f_4}\rangle$, $c_2 = \langle\{f_1, f_3\}, f_4\rangle$, and $c_3 = \langle\{f_2\}, f_3\rangle$.

Now suppose we are presented with a problem situation X in which a student is late but ill: $X = \{f_1, f_2\}$. It is possible, of course, to advance the argument

$$\{f_1, f_2\} \underset{c_1}{\rightarrow} \{f_1, f_2, \overline{f_4}\},$$

suggesting on the basis of c_1 that the student should be denied support because he is late. But from an intuitive point of view, it seems much more natural instead to endorse the argument

$$\{f_1, f_2\} \underset{c_3}{\rightarrow} \{f_1, f_2, f_3\} \underset{c_2}{\rightarrow} \{f_1, f_2, f_3, f_4\},$$

suggesting on the basis of c_3 that the student has an excuse because he is ill, and then on the basis of c_2 that support should be continued because, though late, there was an excuse. From an intuitive point of view, it seems that the application of c_1 in the first argument should be trumped by c_2. Notice, however, that the simple treatment of trumping sketched above does not allow this, since c_2 in the second argument is not applied to the same characterization of the situation as c_1 in the first, but instead to a more extensive characterization, containing additional factors.

In order to allow for trumping in examples such as this, we must amend the simple treatment so that the application of a constituents c to a situation under a characterization U can be trumped by an opposing constituents d that is more relevant to the situation, not necessarily under that very same characterization, but perhaps only under another characterization V that is more extensive. This idea is reflected in the following definition.

Definition 11 (Trumping) An argument of the form $\alpha(X, U) \underset{c}{\rightarrow} Y$ is *trumped* in the context $\langle \Gamma, X; \Phi \rangle$ if and only if there is an argument $\alpha(X, V) \in \Phi$ and a precedent constituent $d \in \Gamma$ such that (1) $U \subseteq V$, (2) $Out(d) = \overline{Out(c)}$, and (3) $c <_V d$.

Returning to our example, let us take $U = X = \{f_1, f_2\}$ and $V = \{f_1, f_2, f_3\}$. Then where Γ is a case base containing our three precedent constituents, it is clear that the first argument $\{f_1, f_2\} \underset{c_1}{\rightarrow} \{f_1, f_2, \overline{f_4}\}$ is trumped in the context $\langle\Gamma, X; \Phi\rangle$ as long as Φ contains the argument

$$\{f_1, f_2\} \underset{c_3}{\rightarrow} \{f_1, f_2, f_3\},$$

since (1) $U \subseteq V$, (2) $Out(c_2) = \overline{Out(c_1)}$, and (3) $c_1 <_V c_2$.

Having considered the three preliminary notions involved—constructibility, conflict, and trumping—we can now define the central concept of a forcible argument.

Definition 12 (Forcibility; $\mid\!\sim$) An argument α is defined as *forcible* in a context $\langle\Gamma, X; \Phi\rangle$—written, $\langle\Gamma, X; \Phi\rangle \mid\!\sim \alpha$—by cases, as follows: (1) if α is a degenerate argument, then $\langle\Gamma, X; \Phi\rangle \mid\!\sim \alpha$ if and only if $\alpha = X$; (2) if α is a non-degenerate argument, then $\langle\Gamma, X; \Phi\rangle \mid\!\sim \alpha$ if and only if α is constructible but neither conflicted not preempted in the context $\langle\Gamma, X; \Phi\rangle$.

The interested reader is invited to compare this definition with the treatment of inheritability from [8].

An interpretation of a case base applied to a problem situation, we recall, is simply a set of arguments based on that situation. But of course, from an intuitive point of view, not every interpretation is coherent; there is nothing to prevent an interpretation from containing conflicting arguments, for example, or a trumped argument. We define a coherent interpretation as a fixed point of the forcibility relation—an interpretation containing exactly those arguments that are forcible in the context determined by that interpretation.

Definition 13 (Coherent interpretation) The set Φ is a *coherent interpretation* of the case base Γ applied to the situation X if and only if $\Phi = \{\alpha : \langle\Gamma; X, \Phi\rangle \mid\!\sim \alpha\}$.

Such an interpretation can be thought of as a sensible way of bringing the case base to bear on a problem situation—an internally coherent set of arguments constructed from the precedents contained in the case base.

Just as a theories in certain nonmonotonic logics allows for different extensions, a case base might allow for different coherent interpretations in its application to a given situation—a possibility that is particularly attractive from a legal point of view, where the different coherent interpretations might represent different argumentative standpoints. Because the theory developed here allows for multiple coherent interpretations, it is again natural to specify both credulous and skeptical notions of consequence, with the credulous notion sanctioning any conclusion that is supported by some argument belonging to some interpretation, and the skeptical notion sanctioning a conclusion only if it is supported by some argument that belongs to every interpretation.

Definition 14 (Consequence: extended arguments) A statement $\bigcirc(f/X)$ is a *credulous consequence* of a case base Γ if and only if there is some coherent interpretation Φ of Γ applied to the situation X and some argument $\alpha(X, Y)$ with $f \in Y$ such that $\alpha(X, Y) \in \Phi$. A statement $\bigcirc(f/X)$ is a *skeptical consequence* of a case base Γ if and only if there is some argument $\alpha(X, Y)$ with $f \in Y$ such that $\alpha(X, Y) \in \Phi$ for each coherent interpretation Φ of Γ applied to the situation X.

Three technical remarks. First, the notion of skeptical consequence defined here is a strong notion, sanctioning a conclusion only if that conclusion is supported by the same argument in each coherent interpretation. It is also possible to introduce a weaker notion—allowing for "floating conclusions" [10]—that sanctions a particular conclusion only if that conclusion is supported by some argument in each coherent interpretation, but not necessarily the same argument. Second, this treatment of consequence is a conservative extension of the previous treatment from Definition 7, generating the same consequences when the case base Γ is limited to cases proper, with only π and δ as outcomes.

Finally, just as certain cyclic inheritance networks fail to have extensions, the current theory allows for "cyclic" case bases that simply do not have coherent interpretations when applied to particular problem situations. To illustrate, consider the case base Γ containing only the single case $c = \langle\{f_1\}, \overline{f_1}\rangle$ applied to the problem situation $X = \{f_1\}$. It is easy to see that Γ has no coherent interpretation applied to this problem situation, for suppose there were such an interpretation, a set Φ containing all and only those arguments forcible in $\langle\Gamma, X; \Phi\rangle$. Of course, the degenerate argument $\{f_1\}$ would have to belong to Φ, by the definition of forcibility. But what about the argument $\{f_1\} \underset{c}{\rightarrow} \{f_1, \overline{f_1}\}$? Suppose it does not belong to Φ. Then it is constructible, and neither preempted nor trumped, and so forcible; so there is a forcible argument not contained in Φ. On the other hand, suppose it does belong to Φ. Then it is conflicted, and so not forcible; so Φ contains an argument that is not forcible.

A longer version of this paper will describe acyclicity conditions that guarantee the existence of a coherent interpretation of a case base; the conditions are natural, ruling out only peculiar case bases such as that described here.

5 Preemption

Once we move from the simple setting of Section 3 to a more general setting that allows also for extended arguments, as in Section 4, certain notions from the simple setting, such as that of trumping, become more complicated. But the more general setting also suggests the introduction of entirely new relations among arguments that are not present at all in the simple setting. One of these relations, familiar from the theory of defeasible inheritance, is that of preemption—a preference, other things being equal, for arguments based on more specific information.

To illustrate, imagine that the governing board of a condominium community is empowered to adjudicate certain matters not explicitly treated in the condominium bylaws. And suppose the board's decisions regarding animals in the building include the following precedent constituents: a seeing eye dog is classified as a medically aid animal; a medical aid animal is classified as a pet (and is therefore subject to a variety of rules governing pets, such as vaccination and registration requirements); a pet is not allowed in the common areas of the building, such as the lobby; a medical aid animals is allowed in the common areas in the building. Let us abbreviate the factors involved in these precedent constituents as follows: f_1 = the animal is a seeing eye dog; f_2 = the animal is a medical aid animal; f_3 = the animal is pet; f_4 = the animal is allowed in the common areas. The various precedent constituents themselves can then be represented as $c_1 = \langle\{f_1\}, f_2\rangle$, $c_2 = \langle\{f_2\}, f_3\rangle$, $c_3 = \langle\{f_3\}, \overline{f_4}\rangle$, and $c_4 = \langle\{f_2\}, f_4\rangle$.

Suppose the question facing the condominium board concerns whether a particular seeing eye dog should be allowed in the common areas—that is, the current fact situation is $X = \{f_1\}$, and the issue is whether this fact situation should be elaborated to include f_4 or $\overline{f_4}$. Faced with such a problem, one might advance the argument

$$\{f_1\} \underset{c_1}{\rightarrow} \{f_1, f_2\} \underset{c_2}{\rightarrow} \{f_1, f_2, f_3\} \underset{c_3}{\rightarrow} \{f_1, f_2, f_3, \overline{f_4}\},$$

supporting the conclusion that the dog should not be allowed in the common areas, since it is a medical aid animal, and therefore a pet, and it was decided that a pet should not be allowed; but from an intuitive point of view, it really seems much better to endorse the argument

$$\{f_1\} \underset{c_1}{\rightarrow} \{f_1, f_2\} \underset{c_4}{\rightarrow} \{f_1, f_2, f_4\},$$

which supports the conclusion that the dog should be allowed, since it is a medical aid animal, and it was decided that a medical aid animals should be allowed. Notice, however, that the first argument is not trumped, under any analysis of trumping: the factors that c_4 shares with the problem situation can never include those shared by c_3, no matter how the problem situation is eventually characterized. Instead, it seems to be a matter of preemption: it is preferable to draw conclusions based on more specific information—that the dog is a medical aid animal, rather than an ordinary pet.

Of course, it would be possible to deal with this difficulty by suggesting that what the condominium board really meant by its decision that medical aid animals are allowed in the common areas is that pets which are medical aid animals are allowed. In that case, the constituent c_4 would have to be reformulated as $c_4' = \langle\{f_2, f_3\}, f_4\rangle$, and the first argument above would indeed be trumped. But what if the language of the board's decision says quite explicitly that medical aid animals should be allowed, with no mention of pets at all, so that c_4 really is the most accurate representation? As a general point of methodology, it seems best to develop a theory that gives correct results based on the precedent constituents as they are actually stated, rather than being forced to reformulated these constituents so as to yield the desired results before applying the theory.

How should the notion of preemption be defined in the present setting? In the setting of defeasible inheritance, where priority in an argument path corresponds to specificity, a tentative extension of some argument path is preempted whenever conflicting information can be derived from some node that lies on some accepted argument from the initial to the final element of that path. The most straightforward adaptation of that idea to the present setting would define an argument $\alpha(X, U) \underset{c}{\rightarrow} Y$ as preempted in a context $\langle\Gamma, X; \Phi\rangle$ whenever there is an argument of the form $\alpha(X, V, U)$ in Φ and a constituent d in Γ such that (1) $V \cap Fac(d) \neq \emptyset$ and (2) $Out(d) = \overline{Out(c)}$. The intention here is that the argument $\alpha(X, V, U)$ positions V as a more specific characterization than U, while the clauses (1) and (2) tell us that the constituent d is applicable on the basis of V and also that the outcome of this constituent conflicts with that of the c.

It is easy to see that this treatment results in preemption of the objectionable argument displayed above: just take $U = \{f_1, f_2, f_3\}$, $V = \{f_1, f_2\}$, $c = c_3$, and $d = c_4$. Unfortunately, however, this straightforward adaptation of the notion of preemption from inheritance theory to the present setting is not accurate in general. The reason for this is that, while priority in the kind of arguments allowed by simple inheritance theory really does seem to correspond to specificity, the present setting allows for arguments in which prior nodes need not represent more specific characterizations of the problem situation.

To illustrate, consider a new example in which the case base contains the four precedent constituents $c_1 = \langle\{f_1\}, f_2\rangle$, $c_2 = \langle\{f_1\}, f_3\rangle$, $c_3 = \langle\{f_2\}, \overline{f_4}\rangle$, and $c_4 = \langle\{f_3\}, f_4\rangle$, and suppose the original problem situation is $X = \{f_1\}$. Now consider the argument

$$\{f_1\} \underset{c_1}{\rightarrow} \{f_1, f_2\} \underset{c_2}{\rightarrow} \{f_1, f_2, f_3\} \underset{c_4}{\rightarrow} \{f_1, f_2, f_3, f_4\}.$$

According to the simple treatment sketched above, this argument will be preempted: take $U = \{f_1, f_2, f_3\}$, $V = \{f_1, f_2\}$, $c = c_4$, and $d = c_3$. But of course, this is incorrect. The constituent c_3 is applicable on the basis of the factor f_2, while c_4 is applicable on the basis of f_3; but the factor f_2, although introduced into the argument earlier than f_3, is not more specific than f_3, and does not provide a reason for the introduction of f_3—as far as f_3 is concerned, f_2 is simply irrelevant information.

One way to avoid preemption on the basis of irrelevant information like this, which just happens to occur earlier on in an argument path, is first to introduce the notion of a minimal argument, as follows.

Definition 15 (Minimal argument) An argument of the form $\alpha(X, Y)$ is defined as a *minimal argument* for Z in the context $\langle\Gamma, X; \Phi\rangle$ if and only if (1) $\alpha(X, Y) \in \Phi$, (2) $Z \subseteq Y$, and (3) there is no $Y' \subset Y$ such that $Z \subseteq Y'$ for some argument $\alpha(X, Y') \in \Phi$.

Intuitively, the argument $\alpha(X, Y)$ is a minimal argument for a characterization Z of an original problem situation X if this argument establishes Z—that is, $Z \subset Y$—while establishing as little other information about X as possible. Minimal arguments avoid irrelevant information: if $\alpha(X, Y)$ is minimal, all of the other factors contained in Y must actually have been used in establishing Z.

It is now natural to refine the treatment of preemption suggested earlier by relying on the notion of minimal arguments to rule out preemption on the basis of irrelevant information.

Definition 16 (Preemption)
An argument of the form $\alpha(X, U) \underset{c}{\rightarrow} Y$ is *preempted* in the context $\langle \Gamma, X; \Phi \rangle$ if and only if there is a set V such that (i) either $V = X$ or there is a minimal argument $\alpha(X, V, U)$ for $U \cap Fac(c)$ in Φ, and (ii) there is a precedent constituent $d \in \Gamma$ such that (1) $V \cap Fac(d) \neq \emptyset$ and (2) $Out(d) = \overline{Out(c)}$.

If the argument $\alpha(X, U)$ can be extended through the constituent c to the characterization Y, it must be on the basis of $U \cap Fac(c)$, the factors that the characterization U shares with that constituent. By requiring that $\alpha(X, V, U)$ should be a minimal argument for $U \cap Fac(c)$, we require that V occurs as an essential step in some argument for $U \cap Fac(c)$, the very reason for advancing to the further conclusion Y. Clauses (1) and (2) then tell us that the $\alpha(X, U) \underset{c}{\rightarrow} Y$ should be preempted when there is a constituent d already applicable on the basis of V itself that provides a reason for a conflicting conclusion.

Returning to the previous example, it is easy to see that the argument

$$\{f_1\} \underset{c_1}{\rightarrow} \{f_1, f_2\} \underset{c_2}{\rightarrow} \{f_1, f_2, f_3\} \underset{c_4}{\rightarrow} \{f_1, f_2, f_3, f_4\}.$$

is no longer inappropriately preempted on the basis of the refined definition. The reason, of course, is that

$$\{f_1\} \underset{c_1}{\rightarrow} \{f_1, f_2\} \underset{c_2}{\rightarrow} \{f_1, f_2, f_3\},$$

although an argument for $\{f_1, f_2, f_3\} \cap Fac(c_4)$, is not a minimal argument for this characterization.

I believe that the notion of preemption set out here may have a useful role to play in the analysis of extended arguments in precedent-based reasoning—a role in determining preferences among competing arguments, similar to the notion of trumping. But of course, this idea of preemption needs to be explored more carefully and tested on concrete examples; in addition, the interplay between the preference criteria based on preemption and trumping would have to be examined. To illustrate some of the issues involved here, consider a final case base containing the precedent constituents $c_1 = \langle \{f_1\}, f_2 \rangle$, $c_2 = \langle \{f_2\}, f_3 \rangle$, $c_3 = \langle \{f_2\}, \overline{f_4} \rangle$, and $c_4 = \langle \{f_2, f_3\}, f_4 \rangle$, with original problem situation $X = \{f_1\}$; and let us suppose the interpretation has been developed far enough to contain the argument

$$\{f_1\} \underset{c_1}{\rightarrow} \{f_1, f_2\} \underset{c_2}{\rightarrow} \{f_1, f_2, f_3\}.$$

It seems that that this argument cannot be extended with c_3 to yield

$$\{f_1\} \underset{c_1}{\rightarrow} \{f_1, f_2\} \underset{c_2}{\rightarrow} \{f_1, f_2, f_3\} \underset{c_3}{\rightarrow} \{f_1, f_2, f_3, \overline{f_4}\},$$

since that argument is trumped by c_4. But it seems likewise that the argument cannot be extended with c_4 to yield

$$\{f_1\} \underset{c_1}{\rightarrow} \{f_1, f_2\} \underset{c_2}{\rightarrow} \{f_1, f_2, f_3\} \underset{c_4}{\rightarrow} \{f_1, f_2, f_3, f_4\},$$

since this argument is preempted by c_3. In such a case, should preemption take precedence over trumping, so that we would accept the first of these extended arguments; should trumping take precedence over preemption, so that we should accept the second; or should we accept neither argument, since the first is trumped and the second preempted?

References

[1] Kevin Ashley. Toward a computational theory of arguing with precedents: accomodating multiple interpretations of cases. In *Proceedings of the Second International Conference on Artificial Intelligence and Law (ICAIL-89)*, pages 93–110. The Association for Computing Machinery, 1989.

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

[3] L. Karl Branting. Reasoning with portions of precedents. In *Proceedings of the Third International Conference on Artificial Intelligence and Law (ICAIL-91)*, pages 145–154. ACM Press, 1991.

[4] L. Karl Branting. A computational model of *ratio decidendi*. *Artificial Intelligence and Law*, 2:1–31, 1994.

[5] Rupert Cross. *Precedent in English Law*. Oxford University Press, 1968. Second edition.

[6] Arthur Goodhart. Determining the *ratio decidendi* of a case. *Yale Law Journal*, 40:161–183, 1930.

[7] John Horty. Moral dilemmas and nonmonotonic logic. *Journal of Philosophical Logic*, 23:35–65, 1994.

[8] John Horty. Some direct theories of nonmonotonic inheritance. In D. Gabbay, C. J. Hogger, and J. A. Robinson, editors, *Handbook of Logic in Artificial Intelligence and Logic Programming, Volume 3: Nonmonotonic Reasoning and Uncertain Reasoning*, pages 111–187. Oxford University Press, 1994.

[9] John Horty. Nonmonotonic foundations for deontic logic. In Donald Nute, editor, *Defeasible Deontic Logic*. Kluwer Academic Publishers, 1997.

[10] David Makinson and Karl Schlechta. Floating conclusions and zombie paths: two deep difficulties in the "directly skeptical" approach to defeasible inheritance nets. *Artificial Intelligence*, 48:199–209, 1991.

[11] Henry Prakken and Giovanni Sartor. A dialectical model of assessing conflicting arguments in legal reasoning. *Artificial Intelligence and Law*, 4:331–368, 1996.

[12] Henry Prakken and Giovanni Sartor. Modelling reasoning with precedents in a formal dialogue game. *Artificial Intelligence and Law*, forthcoming.

[13] Jean Paul Sartre. *L'Existentialisme est un Humanisme*. Nagel, 1946. Translated as "Existentialism is a Humanism" in *Existentialism from Dostoevsky to Sartre*. W. Kaufmann (ed.), Meridian Press, 1975.

[14] Bas van Fraassen. Values and the heart's command. *The Journal of Philosophy*, 70:5–19, 1973.