A semantic view on reasoning about priorities (extended abstract)

Henry Prakken*
Computer/Law Institute
Faculty of Law, Free University Amsterdam

February 24, 1995

Abstract

This paper gives a logical analysis in semantic terms of reasoning about preference relations. A method is proposed for extending logics for reasoning with prioritised information with means to express information about priorities as premises, and for defining the logical consequences of such premises. The method is first presented in general terms and then applied to two existing systems, prioritised default logic and a model-theoretic formulation of pointwise circumscription; also its application to argument-based systems is discussed. Finally, the analysis is applied to a realistic example from legal reasoning.

1 Introduction

In the logical study of common-sense reasoning the need for logics that can handle prioritised information is widely acknowledged. Accordingly, many prioritised logics (e.g. [1, 3, 10, 11]) have been developed. A severe restriction of them, however, is that they only formalise the *use* of priorities in resolving contradictions. Yet in many domains preference relations only hold in certain circumstances, and whether those circumstances hold, must often be derived from further information. Hence, priorities should themselves be derivable from the premises. Moreover, information about the priorities is often itself incomplete or inconsistent, and priorities must therefore often themselves be derived as consequences of a prioritised logic.

One domain in which these phenomena occur is legal reasoning. Every legal system has many conflict resolution principles. As an illustration of how they work, consider Section 5 of the Dutch act on general provisions (AB), that expresses the temporal principle, 'the later rule prevails over the earlier one', and Section 1637c of the Dutch civil code (BW), stating that statute rules concerning

*Supported by a fellowships grant of the Royal Netherlands Academy of Arts and Sciences, and by Esprit WG 8319 'Modelage'. I thank Gerhard Brewka for suggesting some improvements.

labour contracts have priority over conflicting statute rules about any other type of contract. Clearly, these rules only apply if certain other domain information holds: for instance, for applying 1637c it must be known of a certain thing that it is a contract or that it is a labour contract. These things are real legal issues, about which further information exists, and which need to be reasoned about. Moreover, the conflict rules can themselves be in conflict: imagine, for example, two conflicting statute rules, a later one concerning any type of contract and an earlier one concerning labour contracts: then 1637c BW and 5 AB disagree on which rule takes precedence. To these conflicts other conflict rules apply, based, for instance, on the general hierarchical structure of the legal system or on specificity.

About specificity a special remark is in order. In legal reasoning, and presumably also in many other domains, this criterion is only one of several conflict resolution rules, and not even the most important one. Therefore, a logic with a built-in specificity comparator (e.g. [3]) is for our purposes inadequate.

The aim of this paper is to formalise these aspects of reasoning with prioritised information. First the general idea will be sketched, after which it is instantiated for two prioritised logics, and applied to the just given legal example.

2 The general setting

In this section the main idea of the paper will be explained and some general notions will be formally defined. I will assume a very general and abstract picture of what a logic for reasoning with prioritized information (LPI) is. An LPI assumes input information in the form of an 'input pair' $(P, (O, \leq))$. P contains the premises in the usual sense and O is a set ordered by an ordering relation $\leq$[1] and uniquelly determined by P. The ordering is intended to resolve conflicts between the premises. Usually it is defined on the elements of P, i.e. $O = P$. Moreover, usually the user has to specify the ordering, which means that the same set O can be ordered in more than one way. An exception is when the LPI assigns the priorities on the basis of a logical specificity definition. Finally, I assume the language of an LPI to be expressive enough for expressing ordering information.

While in standard logic each set of premises has only one set of consequences, in an LPI there are usually more such sets, resulting from the fact that often not all conflicts between the premises can be solved with the priorities. One may think, for example, of the extensions of prioritised default logics [1] or the formulas true in the preferred models of a logic for preferential entailment [3, 7]. In this paper I will call the output sets of an LPI 'extensions'; moreover, I will assume them to be closed under deductive consequence. Note that in many LPI's there is a special category of 'strict' premises, i.e. premises that are in every extension.

Generally speaking, there are two ways of adapting an LPI to reasoning about priorities. One way is to change the proof theory of the LPI, to the effect that the priorities required by the proof theory are provable within the LPI itself. In the present paper, however, I will employ a semantic perspective. My aim

[1] In this paper $x \leq y$ will stand for 'y is preferred over x'. Also, I will usually write $(P, \leq)$ for $(P, (O, \leq))$ and let O be clear from the context.

is to develop a criterion for testing whether extending the proof theory of an LPI 'works', whether the consequences of a certain input indeed contain the appropriate ordering. The idea is simple: given a premise set P, an extension is *faithful* with respect to an ordering $\leq$ iff what it says about $\leq$ is indeed true; and what an extension says about $\leq$ is true iff the extension is indeed 'created' with $\leq$, i.e. if it is an extension of $(P, \leq)$.

In the logical study of metalevel reasoning similar notions have been proposed before. For example, (cf.[5]) use the term 'introspective truthfulness' for the question whether theories stating that they have certain properties indeed have these properties: for example, they call a consistent theory that says of itself that it is inconsistent, introspectively untruthful. And Genesereth and Nilsson [4, p.253], when discussing metalevel formalisations of inferences at the object level, call such a formalisation 'introspectively faithful' iff a change in an object level database is always made with rules 'prescribed' by the metalevel formalisation and never with rules 'forbidden' by that formalisation.

Below I assume that the object language of an LPI contains special predicate symbols for the ordering relations and terms for the objects that are ordered. $\leq$ will be denoted by the predicate symbol $\preceq$ and related predicate symbols such as $\prec$ will be defined in the usual sense. Furthermore, in examples where the ordering is on premises I will usually name defeasible premises by $d_1, \ldots, d_n$ and strict (or 'factual') premises by $f_1, \ldots, f_n$.

Before the notion of faithfulness can be defined, two subtleties must be explained. Firstly, the relevant axioms making the ordering of the appropriate type should somehow be made part of every extension, as well as the definitions of predicates like $\prec$ and $\not\prec$. This can be done, for example, by including them in the strict premises. In the rest of this paper I will assume that this has been done.

Consider next the premises

$$\begin{array}{lll} d_1 : a & d_2 : \neg a & d_3 : \neg a \wedge b \\ f_1 : d_1 \prec d_2 \vee d_1 \prec d_3 & & \end{array}$$

and assume that d_1, d_2, d_3 are ordered by a partial preorder. Note that since f_1 is a strict premise, it will be in every extension of these premises. Assume now that $\leq$ is the identity relation; then our definitions should account for the fact that the formula $d_1 \not\prec d_2 \wedge d_1 \not\prec d_3$ is true of $\leq$ and that this formula is inconsistent with any extension containing f_1.

Accordingly, the definition of a faithful extension is based on the notion of the 'description' of an ordering: this is the set of all atoms and negations of atoms that are true of the ordering.

Definition 2.1 *Let ℓ be the language of any LPI, O a set of objects $o_1, \ldots, o_n$, which in ℓ are named $t_1, \ldots, t_n$, and $\leq$ any ordering of O. Then the function $d_O : Pow(O \times O) \longrightarrow Pow(\ell)$ is defined as follows. For all $\leq$:*

$$d_O(\leq) = \{t_i \preceq t_j \mid (o_i, o_j) \in \leq\} \cup \{(t_k \not\preceq t_l) \mid (o_k, o_l) \notin \leq\}$$

We can now define the notion of a faithful extension: an extension is faithful iff it is consistent with the description of an ordering with which it can be constructed

(Below 'E is an L-extension of ...' means 'E is according to the definitions of L an extension of ...').

Definition 2.2 *Let L be an LPI. E is a* faithful L-extension *of a set P of well-formed formulas of L iff for some ordering $\leq$:*

1. *E is an L-extension of $(P, \leq)$; and*
2. *$\leq$ is compatible with E, i.e. $d(\leq)^2 \cup E$ is consistent.*

3 Some prioritised logics extended

In this section I will discuss the application of the above ideas to some existing LPI's: a consistency-based system, a preferred-model semantics and argument-based systems.

3.1 Brewka's prioritised default logic (PDL)

PDL [1] is a prioritised version of normal default logic. Brewka himself has also extended PDL for dealing with reasoning about priorities. The basic idea is the same as in the present paper, but the formalisation is slightly different. Apart from this, the present paper can be regarded as a generalisation of Brewka's ideas.

In the following presentation of PDL I will represent normal defaults $a : b/b$ as $a \Rightarrow b$. Furthermore, I will adapt Brewka's notation of orderings to the one of the present paper (Brewka writes 'd_1 has priority over d_2' as $d_1 < d_2$). The input of PDL is a *prioritised default theory*: this is a triple $(D, W, <)$, where D is a finite set of normal defaults, W is a set of first-order formulas and $<$ is a strict partial order on D. Given a set E of formulas, a default $a \Rightarrow b$ is *active* in E iff $a \in E$, $b \notin E$ and $\neg b \notin E$. Then:

Definition 3.1 ([1]) *Let $\Delta = (D, W, <)$ be a prioritised default theory and $\ll$ a strict total order containing $<$: E is the prioritised extension of Δ* generated by *$\ll$ iff $E = \cup E_i$, where $E_0 := Th(W)$, and*

$$E_{i+1} = \begin{cases} E_i & \textit{if no default is active in } E_i, \\ Th(E_i \cup \{b\}) & \textit{otherwise, where } b \textit{ is the consequent of the} \\ & \ll\textit{-maximal default that is active in } E_i. \end{cases}$$

E is a PDL extension of $\Delta = (D, W, <)$ iff there is a strict total order containing $<$ that generates E.

In extending PDL to $EPDL$, recall that W is assumed to contain the axioms of a strict partial order. Now a set of formulas is an $EPDL$-extension of (D, W) iff for some ordering $<$ it is a *faithful* PDL-extension of $(D, W, <)$.

As Brewka [2] shows, PDL's guarantee of the existence of extensions is lost in $EPDL$. A counterexample is $\Delta = (W, D), W = \emptyset; D = \{d_1, d_2\}$ and

[2] Here and below O will be left implicit.

$$d_1: \quad \top \Rightarrow d_1 \prec d_2 \qquad\qquad d_2: \quad \top \Rightarrow d_2 \prec d_1$$

The following example shows a difference between Brewka's and my way of extending PDL. Consider $\Delta = (D, W)$ where $W = \emptyset$, $D = \{d_1, d_2, d_3\}$ and

$$\begin{array}{llll} d_1: & \top \Rightarrow a & d_3: & \top \Rightarrow d_1 \not\prec d_2 \\ d_2: & \top \Rightarrow \neg a & d_4: & \top \Rightarrow d_2 \not\prec d_1 \end{array}$$

Brewka requires consistency of an extension with $d(\ll)$ instead of with $d(<)$[3]. Then this theory has no faithful extensions, while with the present definitions it has two such extensions, one containing a and the other $\neg a$, and both containing $d_1 \not\prec d_2$ and $d_2 \not\prec d_1$. Which outcome is acceptable depends on how a partial ordering is intuitively interpreted: Brewka regards it as a partial description of a total ordering, while I regard it as genuinely partial.

Knowing that the present approach regards more PDL extensions as faithful than Brewka's, it is easy to show that every PDL-extension which is faithful according to Brewka's definitions, is also faithful according to definition 2.1 (proof omitted in extended abstract).

In the last section $EPDL$ will be applied to the legal-reasoning example from the introduction.

3.2 Preferential entailment

Applications of preferred-model semantics to nonmonotonic reasoning typically use (ab)normality predicates. The preferred models are those containing as few statements of abnormality as possible (where the comparison is with respect to set inclusion). If the set of these statements is ordered (as e.g. in [3]) then this ordering can be used to obtain a further ranking between multiple minimal models.

One formalisation of these ideas is the next definition. It is a model-theoretic version of prioritised pointwise circumscription [7], mixing ideas of [3] with familiar notions ($I_M(P)$ denotes the interpretation of the predicate letter P in the model M, and abnormality atoms are formulas of the form $ab(t)$).

Definition 3.2 *Let T be a first-order theory, A the set of abnormality atoms of the language of T ordered by a partial order $\leq$, and let Q and Z be disjoint sets of predicate letters such that their union is the set of all predicate letters of the language of T, minus $\{ab\}$. Let M and M' be two models of T. We say that $M \geq_{A,Q,Z} M'$ (M is preferred over M') iff:*

1. *M and M' have the same domain; and*
2. *For all $Q_i \in Q : I_M(Q_i) = I_{M'}(Q_i)$; and*
3. *For all $a \in A$ such that $M \models a$ and $M' \not\models a$ there is an $a' \in A$ such that $M' \models a'$, $M \not\models a'$ and $a < a'$.*

[3] In fact, he defines d in such a way that it only contains the atoms true of $\ll$.

In extending this definition to reasoning about priorities each $a \in A$ of the form $ab(t_1, \ldots, t_n)$ is denoted by the function expression $AB(t_1, \ldots, t_n)$. It is important to note that the set of ordered objects is A and not the entire set of well-formed formulas. Although I could redefine the output of a preferential semantics based LPI in terms of extensions, it is perhaps more elegant to restate the definition of a faithful extension in terms of models.

Definition 3.3 *Let M be a model of a theory T. Then M* is compatible with $\leq$ *iff $I_M(\preceq) = \leq$.*

Definition 3.4 *Let T, A, Q and Z be as before. A model M is a* faithful model *of (T, Q, Z) iff for some ordering $\leq$ of A M is a $\geq_{A,Q,Z}$-maximal model of T and M is compatible with $\leq$.*

Consider the following example.

$$\begin{array}{ll} d_1: & \forall x((Quaker(x) \wedge \neg ab(d_1, x)) \rightarrow Pacifist(x)) \\ d_2: & \forall x(Republican(x) \wedge \neg ab(d_2, x)) \rightarrow \neg Pacifist(x)) \\ d_3: & \forall x((Politician(x) \wedge \neg ab(d_3, x)) \rightarrow AB(d_1, x) \prec AB(d_2, x)) \\ d_3: & \forall x((Inchurchchoir(x) \wedge \neg ab(d_4, x)) \rightarrow AB(d_2, x) \prec AB(d_1, x)) \\ f_1: & Quaker(n) \wedge Inchurchchoir(n) \wedge Republican(n) \wedge Politician(n) \end{array}$$

$$\begin{array}{l} Q = \{Quaker, Republican, Politician, Inchurchchoir\} \\ Z = \{Pacifist, \prec\} \end{array}$$

This results in two faithful models: one satisfying

$$\begin{array}{l} \{ab(d_1, n), \neg ab(d_2, n), AB(d_1, n) \prec AB(d_2, n), \\ \neg ab(d_3, n), ab(d_4, n), \neg Pacifist(n)\} \end{array}$$

and the other satisfying

$$\begin{array}{l} \{ab(d_2, n), \neg ab(d_1, n), AB(d_2, n) \prec AB(d_1, n), \\ \neg ab(d_4, n), ab(d_3, n), Pacifist(n)\}. \end{array}$$

3.3 Argument-based systems

Argument-based systems (e.g. [10, 11]) give a logical analysis of the phenomenon of constructing and comparing arguments for incompatible conclusions. Their output is an assessment of arguments, often in terms of three classes: arguments with which a dispute can be 'won', respectively, 'lost' and arguments which leave the dispute undecided.

In applying the above ideas to argumentation frameworks the extensions are the sets of conclusions of maximal compatible sets of winning arguments (in some systems there is only one such set (e.g. in [10]), but in other systems there can be more (e.g. [11]), analoguously to alternative extensions in PDL). Then the above ideas can be applied straightforwardly.

4 An application to legal reasoning

I will now formalise the legal example of the introduction in $EPDL$, thereby ignoring specificity considerations. I will use the following naming method: every default with free variables $x_1, \ldots, x_n$ is named with a function expression $r(x_1, \ldots, x_n)$, where r is the informal name of the default, and each instantiation with terms $t_1, \ldots, t_n$ is named by $r(t_1, \ldots, t_n)$. Below 1637c BW and 5 AB are instantiated for a conflict between a rule r_1 about any contract and an earlier rule r_2 about labour contracts. I will leave the facts of the case implicit and, to keep the default names readable, I will reflect the 'depth' of instantiations of a scheme r by using the shorthands r, r' and r''.

1637c: $ctr(r_1) \wedge labour(r_2) \Rightarrow r_1 \prec r_2$
5: $later(r_1, r_2) \Rightarrow r_2 \prec r_1$

If we next assume that 5 AB and 1637c BW are themselves regulations about (labour) contracts, we can also instantiate them for their own conflict, assuming in addition that 5 AB is later than 1637c BW.

1637c': $ctr(5) \wedge labour(1637c) \Rightarrow 5 \prec 1637c$
5': $later(5, 1637c) \Rightarrow 1637c \prec 5$

But now we can do the same again, with one more prime added to every term and the rule name, and so on, ad infinitum. Logically, this is not a case of self reference (recall the naming method) since it is one instance of a rule that refers to another of its instances; moreover, there are two faithful extensions, one giving priority to all instances of 1637c BW and the other doing the same for all instances of 5 AB.

Let us now revise our assumptions: the two conflict rules are now not themselves about (labour) contracts, which prevents the infinite instantiation; moreover, 1637c BW is now later than 5 AB. This yields

1637c: $ctr(r_1) \wedge labour(r_2) \Rightarrow r_1 \prec r_2$
5: $later(r_1, r_2) \Rightarrow r_2 \prec r_1$
5': $later(1637c, 5) \Rightarrow 5 \prec 1637c$

Now 5' is the only rule that applies to the conflict between 5 and $1637c$; thus there is a unique faithful extension, containing $5 \prec 1637c$ and $r_3 \prec r_4$. Interestingly, this means that one instance of 5 AB makes another of its instances inferior to a competing rule. Logically as well as legally this seems perfectly acceptable.

5 Conclusion and related research

This paper has presented a general semantical notion for making prioritised logics suitable for formalising reasoning about priorities. Technically, the idea is simple; its main value lies in the tools it provides for formalising a wide range of examples that cannot be formalised in existing logics.

As for related research, I confine myself because of space limitations to just mentioning some relevant publications. In addition to [2] they are [6, 9, 8]. [8]

investigates an alternative method for making argument-based systems selfprioritising, while [6] and [9], unlike Brewka and the present paper, do not present their ideas as a new logical system: [9] addresses the issue mainly at the implementational level in the context of logic metaprogramming, while [6] describes a systematic way of representing priority information in an existing LPI, viz. [3].

A technical topic for further research is extending also the proof theory of prioritised logics. Another, more philosophical issue is whether the notion of introspective faithfulness has other interesting applications besides to reasoning about priorities. One might think, for example, of legal rules stating that other rules are 'invalid' or in certain circumstances 'inapplicable'.

References

[1] G. Brewka, Adding priorities and specificity to default logic. *Proceedings JELIA-94.*

[2] G. Brewka, Reasoning about priorities in default logic. *Proceedings AAAI-94.*

[3] H. Geffner and J. Pearl, Conditional entailment: bridging two approaches to default reasoning. *Artificial Intelligence* 53 (1992), 209-244.

[4] M.R. Genesereth and N.J. Nilsson, *Logical Foundations of Artificial Intelligence.* Morgan Kaufmann, 1987.

[5] F. Giunchiglia and A. Smaill, Reflection in constructive and non-constructive automated resoning. In H. Abramson and M. Rogers (eds): *Meta-Programming in Logic Programming*, Bristol. MIT Press, 123-140.

[6] T.F. Gordon, *The pleadings game. An Artificial-Intelligence model of procedural justice.* Doctoral dissertation Technische Hochschule Darmstadt, 1993.

[7] V.L. Lifschitz, Pointwise circumscription. In M.L. Ginsberg (ed.), *Readings in nonmonotonic reasoning.* Morgan Kaufman, 1987, 179-193.

[8] H. Prakken and G. Sartor, On the relation between legal language and legal argument: assumptions, applicability and dynamic priorities. Accepted for the Fifth International Conference on Artificial Intelligence and Law, 1995.

[9] G. Sartor, A simple computational model for nonmonotonic and adversarial legal reasoning. *Proceedings of the Fourth International Conference on Artificial Intelligence and Law*, ACM Press, 1993, 192-201.

[10] G.R. Simari and R.P. Loui, A mathematical treatment of defeasible argumentation and its implementation. *Artificial Intelligence* 53 (1992), 125-157.

[11] G. Vreeswijk, *Studies in defeasible argumentation.* Doctoral dissertation Free University Amsterdam, 1993.