Book Title
Book Editors
IOS Press, 2003

Norm Modifications in Defeasible Logic

Guido Governatori [a,1], Monica Palmirani [b], Regis Riveret [b], Antonio Rotolo [b] and Giovanni Sartor [b]
[a] *School of ITEE, University of Queensland*
[b] *CIRSFID, University of Bologna*

Abstract. This paper proposes a framework based on Defeasible Logic (DL) to reason about normative modifications. We show how to express them in DL and how the logic deals with conflicts between temporalised normative modifications. Some comments will be given with regard to the phenomenon of retroactivity.

Keywords. Norm Modifications, Defeasible Logic

1. Introduction

This paper proposes a logical framework based on DL to deal with norm-modifications. We have different types of modifications according to how they affect the law. The impact might concern, e.g., the law text, its scope, or the time of its force, efficacy, or applicability [3]. So we can identify four main categories of modifications: (1) textual changes, (2) change of the norm scope, (3) temporal changes, (4) normative system changes. Textual changes intervene when a law is repealed, replaced, integrated or relocated. Changes of scope might be consequent on derogation or extension. Temporal changes impact on the date in force, date of efficacy or date of application of the destination norm. As a first step, in this paper we will focus on three kinds of modifications: substitution (which replaces some textual components of a provision with other textual components, or a provision with another provision), derogation (the derogating provision limits the effects of the derogated provision), and annulment (which cancels *ex tunc* a provision and prevents it to produce any normative effect).

In particular, we are interested in logically investigating the following issues: (A) Conditional modifications, which apply under, and are conditioned to, the occurrence of some uncertain events. (B) The notion of conflict between textual modifications and the logical strategies for solving them. (C) The concept of time-forking of modifications; indeed, there are modifications affecting the effects previously obtained by other modifications. This is the case, in particular, when we have retroactive modifications. Retroactive modifications lead to the forking (branching) of the versioning chain of modifications in order to keep trace both of the past modifications and of the new versioning chains (see [2]). The system is based on the following building blocks or assumptions, which are needed in DL to correctly represent the dynamics of normative systems. **Normative Conditionals:** It is possible to distinguish between different kinds of normative conditional. Here we will identify the following types of normative rules [1]: (1) *Rules for persistent*

[1] Correspondence to: Guido Governatori, School of ITEE, The University of Queensland, Brisbane, Australia, QLD 4072, guido@itee.uq.edu.au.

obligations, which, if applicable, permit to infer literals to be modalised by obligations that persist unless some other, subsequent, and incompatible events or states of affairs terminate them. For example, the obligation of paying the damages in a car crash will hold until such damages have been paid. (2) *Rules for co-occurrent obligations*, which allow for the inference of obligations which hold on the condition and only while the antecedents of these rules hold. For example, the obligation not speak loud in the church will hold only when the agent is in the church. (3) *Rules for counts-as links*, which express the idea of institutional power. For example, if i signs a document on behalf of her boss j, such a document is as it were signed by j only if i has been empowered to do this: i's signature counts as j's signature [1]. Here we will simply view the counts-as link as a normative conditional whose consequences are not necessarily deontic. In this context, modifying meta-norms, namely rules regulating norm-modifi cations, will be represented as counts-as rules. **Time:** Norm application and modifi cation take place along the axis of time. Since we will operate in a temporalised setting, we will not only impose that obligations be temporalised, but also that rules and any literal are labelled by time instants. In particular, a rule is represented as $(a : t \Rightarrow b : t') : t''$. t and t' indicate the time at which a and b hold, while t'' is the time of the rule being in force. We will assume here the time to be linear and discrete. **Normative provisions:** In general, complex normative provisions have an internal structure and can be decomposed into a number of nested units. Here we will assume that the rules of our logical theory correspond to the atomic normative provisions constituting complex provisions. This simplifi cation will keep the system manageable.

2. Basic Formal Language

Defeasible logic is a flexible sceptical non-monotonic formalism that has proven able to represent various aspects of normative reasoning. We adapt here the machinery developed in [1] to represent temporalised normative positions to reason both on the normative provisions in a legal systems and the meta-norms describing the modifi cations of legal texts.

Our language is based on a (numerable) set of atomic proposition $Prop = \{p, q, \dots\}$, a set of rule labels $\{r_1, r_2, \dots\}$, a discrete totally ordered set of instants of time $\mathscr{T} = \{t_1, t_2, \dots\}$, the modal operator ***Obl*** of obligation, and the negation sign $\neg$. A plain literal is either an atomic proposition or the negation of it. If l is a plain literal then ***Obl***l and $\neg$***Obl***l are modal literals. A literal is either a plain literal or a modal literal. Given a literal l with $\sim l$ we denote the complement of l, that is, if l is a positive literal p then $\sim l = \neg p$, and if $l = \neg p$ then $\sim l = p$. Finally we introduce the notion of temporal literals. A temporal literal is a pair $l : t$ where l is a literal and t is an instant of time. Intuitively the meaning of a temporal literal $l : t$ is that l holds at time t. Knowledge in defeasible logic can be represented in two ways: facts and rules. *Facts* are indisputable statements, represented either in form of states of affairs (literal and modal literal) and actions that have been performed. For example, "John is a minor". In the logic, this might be expressed as $Minor(John)$. A *rule* is a relation between a set of of premises (conditions of applicability of the rule) and a conclusion. In this paper the admissible conclusions are either normative provisions (obligations, permissions) or rules themselves, in addition the conclusions and the premises will be qualifi ed with the time when they hold. We consider two classes of rules: counts-as rules and deontic rules. Counts-as rules describe the inference mechanism of the institution on which norms are formalised and can be used to establish definitions as well as conditions for the creation and modification of other rules. Deontic rules on the other hand give the conditions under which normative qualifications (obligations and permissions) hold.

Beside the above classification rules can be partitioned according to their strength into *strict rules* (denoted by $\rightarrow$), *defeasible rules* (denoted by $\Rightarrow$) and defeaters (denoted by $\rightsquigarrow$). Strict rules are rules in the classical sense: they are monotonic and whenever the premises are indisputable so is the conclusion. Defeasible rules, on the other hand, are non-monotonic: they can be defeated by contrary evidence. Finally defeaters are the weakest rules: they do not support conclusions, but can be used to block the derivation of opposite conclusions. Thus we define the set of rule Rules using the following recursive definition:

- A rule is either a counts-as rule or a deontic rules or the empty rule $\perp$, where
- If r is a rule and $t \in \mathscr{T}$, then $r : t$ is a temporalised rule. (The meaning of a temporalised rule is that the rule is in force at time t).
- Let A be a finite set of temporal literals, C be a temporal literal and r a temporalised rule, then $A \hookrightarrow C$, $A \hookrightarrow r$ and $A \hookrightarrow {\sim}r$ are counts-as rules (henceforth we use $\hookrightarrow$ as a metavariable for either $\rightarrow$ when the rule is a strict rule, $\Rightarrow$ when the rule is a defeasible rule, and $\rightsquigarrow$ when the rule is a defeater).
- Let A be a finite set of temporal literals and C be a temporal plain literal. Then $A \hookrightarrow_O C$ is a deontic rule (henceforth we use $\hookrightarrow_O$ as a metavariable for either $\rightarrow_O$ when the rule is a strict rule, $\Rightarrow_O$ when the rule is a defeasible rule, and $\rightsquigarrow_O$ when the rule is a defeater).

For a rule r we will use $A(r)$ to indicate the body or antecedent of the rule and $C(r)$ for the head or consequent of the rule. The above inductive definition makes it possible to have nested rules, i.e., rules occurring inside other rules. However, it is not possible for a rule to occur inside itself. Thus for example, the following is a rule

$$p : t_p, \boldsymbol{Obl} q : t_q \Rightarrow (\boldsymbol{Obl} p : t_p \Rightarrow_O s : t_s) : t_r \qquad (1)$$

Eq. (1) means that if p is true at time t_p and q is obligatory at time t_q, then the deontic rule $\boldsymbol{Obl} p : t_p \Rightarrow_O s : t_s$ is in force at time t_r. The intuition we want to represent is that deontic rules are meant to introduce obligations. We do not admit deontic literals and rules as conclusions of deontic rules since the meaning of modalised rules and nested deontic modalities is not clear. Every temporalised rule is identified by it rule label and its time. Formally we can express this relationship by establishing that every rule label r is a function $r : \mathscr{T} \mapsto \text{Rules}$.

Thus a temporalised rule $r : t$ returns the value/content of the rule 'r' at time t. This construction allows us to uniquely identify rules by their labels, and to replace rules by their labels when rules occur inside other rules. In addition there is no risk that a rule includes its labels in itself. For example if we associate the temporal rule $(\boldsymbol{Obl} p : t_p \Rightarrow_O s : t_s) : t_r$ to the pair $r_1 : t_r$ then we can concisely rewrite eq. (1) as

$$p : t_p, \boldsymbol{Obl} q : t_q \Rightarrow r_1 : t_r \qquad (2)$$

We have to consider two temporals dimension for norms in a normative systems. The first dimension is when the norm is effective in the normative system, and the second when the norm is in force in the normative system. So far temporalised rule capture only one dimension, the effectiveness one. To cover the other dimension we introduce the notion of temporalised rule with viewpoint. A temporalised rule with viewpoint is an expression $s@t$ where s is a temporalised rule, and $t \in \mathscr{T}$. Thus the expression $r_1 : t_1@t_2$ represents a rule r_1 in force at time t_2 and effective from time t_1.

An issue we need to consider here is that we have two different types of normative conditionals: conditionals that initiate an action or a state of affairs which persists until an interrupting event occurs, and conditionals where the conclusion is co-occurrent with the premises. To represent this distinction we introduce a further distinction of rules, orthogonal to the previous one, where rules are partitioned in persistent and transient rules. A persistent rule is a rule whose conclusion holds at all instants of time after the conclusion has been derived, unless interrupting events occur; transient rules, on the other hand, establish the conclusion only for a specific instant of time. We use the following notation to differentiate the various types of rules: with $\hookrightarrow^t_O$ we represent a transient deontic rule, $\hookrightarrow^p_O$ a persistent deontic rule, $\hookrightarrow^t$ a transient counts-as rule, and $\hookrightarrow^p$ a persistent counts-as rule.

Given a set R of rules, we denote the set of strict rules in R by R_s, the set of strict and defeasible rules in R by R_{sd}, the set of defeasible rules in R by R_d, and the set of defeaters in R by R_{dft}. $R[q:t]$ denotes the set of rules in R with consequent $q:t$. We will use R^c for the set of counts-as rules, R^O to denote the set of deontic rules. The set of transient rules is denoted by R^t and the set of persistent rules by R^p. Finally we assume a set of rule modifiers. A rule modifier is a function $m : \mathscr{T} \times \text{Rules} \mapsto \mathscr{T} \times \text{Rules}$.

The combination of the above two constructions allows us to use rule modifiers on rule labels. Thus $m(r_1 : t_1) : t_2$ returns the rule obtained from r_1 as such as time t_1 after the application of the modification corresponding to the function m and the results refers to the content of the rule at time t_2. Given this basic notion of rule modifier, we can define some specific norm-modifications: Annulment, PSubstitution and TSubstitution (partial and total substitution), and Derogate (derogation). Let see their intuitive reading. Recall that $A(r)$ denotes the set of literals occurring in the antecedent of a rule r and $C(r)$ denotes the consequent of r. Suppose $r \in R$ is a generic defeasible rule (either counts-as or deontic) $(a_1 : y_1, \dots a_n : y_n \Rightarrow b : j) : t$. Then $\text{Annulment}(mr, r) : t'$ says that r is annulled at t' by the meta-rule mr.

Let $\mathscr{A}_r$ and $\mathscr{C}_r$ specify the substitutions of literals to be applied, within the rule r, respectively in its antecedent and in its consequent. Where $l \geq 1$ and $l + m \leq n$,

$$\text{PSubstitution}(mr, r, \mathscr{A}_r(a'_l : x_l/a_l : y_l, \dots, a'_{l+m} : x_{l+m}/a_{l+m} : y_{l+m})) : t'$$
$$\text{PSubstitution}(mr, r, \mathscr{C}_r(b' : k/b : j)) : t'$$

say, respectively, that we operate, at t' through mr, a substitution which replaces a subset or all literals in the antecedent of r with other literals $a'_l : x_l, \dots, a'_{l+m} : x_{l+m}$, and $b : j$ with $b' : k$ in the consequent of r. The new version of r will hold at t'.

$$\text{TSubstitution}(mr, r'/r, A(r') = \{d_1 : s_1, \dots, d_o : s_o\}, C(r') = e : z) : t'$$

indicates total substitution of r, i.e. that r is replaced, at t' through mr, by a different rule r' holding at t' and having the antecedent and consequent as specified within the predicate[1].

[1] To simplify the discussion, we will assume that r' will have the same strength of r.

Let $\mathscr{D}$ and $\mathscr{E}$ indicate the additional literals used, respectively in the antecedent of r' and r'', to specify how to derogate to r. Then

$$\text{Derogate}(mr, r, r', r'', \mathscr{D}(q_1 : z_1, \ldots, q_m : z_m), \mathscr{E}(b' : k)) : t'$$

indicates derogation to r. Derogation may affects "spatial" condition of application of a norm r or its conceptual range of application. In the first case, a norm r holding at the national level may be derogated by a norm r' in the event we operate, for example, within a regional context. In the second case, derogation may affect the conceptual range of r. In the first sense, for example, if r states that the tax rate in Italy corresponds to 30% of total income, a rule r' will state that, if one is resident in Emilia Romagna, under the same conditions, the tax rate will be 25% of the total income. In both cases, anyway, we have to identify logical exceptions of r. Thus the predicate says that r is derogated, at t' through mr, by a rule r' holding at t', which last includes in its antecedent the same conditions of r plus some additional conditions $q_1 : z_1, \ldots, q_m : z_m$. Also, such a rule r' produces the effect b' rather than b. However, under r's conditions and $q_1 : z_1, \ldots, q_m : z_m$ we should also block the derivation of the consequent of r: this is done by rule r'', a defeater, which holds as well at t' and that will have in the consequent the negation of the consequent of r[2]. Let us now characterise the modifiers that correspond to the above predicates. If $r : (a_1 : y_1, \ldots a_n : y_n \Rightarrow b : j) : t$

- the modifier corresponding to $\text{Annulment}(mr, r) : t'$ assigns the empty rule $r : (\perp) : t'$ to r as holding at t. The rule r is thus dropped at t' from the system and so, at t', r is not in force.
- the modifier corresponding to

  $$\text{PSubstitution}(mr, r, A(a'_l : x_l / a_l : y_l, \ldots, a'_{l+m} : x_{l+m} / a_{l+m} : y_{l+m})) : t'$$

  assigns, to r at t, the following rule

  $$r : ((A(r) - \{a_l : y_l, \ldots, a_{l+m} : y_{l+m}\}) \cup \{a'_l : x_l, \ldots, a'_{l+m} : x_{l+m}\} \Rightarrow b : j) : t'$$

  while the modifier corresponding to $\text{PSubstitution}(mr, r, \mathscr{C}(b' : k / b : j)) : t'$ assigns, to r at t, the rule $r : (a_1 : y_1, \ldots a_n : y_n \Rightarrow b' : k) : t'$
- the modifier corresponding to

  $$\text{TSubstitution}(mr, r'/r, A(r') = \{d_1 : s_1, \ldots, d_o : s_o\}, C(r') = e : z) : t'$$

  assigns, to r at t, the rule $r' : (d_1 : s_1, \ldots, d_o : s_o \Rightarrow e : z) : t'$
- the expression $\text{Derogate}(mr, r, r', r'', \mathscr{E}(q_1 : z_1, \ldots, q_m : z_m), \mathscr{F}(b' : k) : t'$ corresponds to applying two distinct modifiers. The first modifier assigns, to r at t, the rule $r' : (A(r) \cup \{q_1 : z_1, \ldots, q_m : z_m\} \Rightarrow b' : k) : t'$ The second assigns, to r at t, the rule $r'' : (A(r) \cup \{q_1 : z_1, \ldots, q_m : z_m\} \rightsquigarrow \sim b : j) : t'$.

Modifications	Conditions
Annulment$(mr,r):t'$ PSubstitution$(mr',r,\mathscr{A}_r(a'_l:x_l/a_l:y_l,\ldots,a'_{l+m}:x_{l+m}/a_{l+m}:y_{l+m})):t''$	$t'=t''$
Annulment$(mr,r):t'$ PSubstitution$(mr',r,\mathscr{C}_r(b':k/b:j)):t''$	$t'=t''$
Annulment$(mr,r):t'$ TSubstitution$(mr',r'/r,A(r')=\{d_1:s_1,\ldots,d_o:s_o\},C(r')=e:z):t''$	$t'=t''$
Annulment$(mr,r'):t'$ TSubstitution$(mr',r'/r,A(r')=\{d_1:s_1,\ldots,d_o:s_o\},C(r')=e:z):t''$	$t'=t''$
Annulment$(mr,r):t'$ Derogate$(mr',r,r',r'',\mathscr{E}(q_1:z_1,\ldots,q_m:z_m),\mathscr{F}(b':k)):t''$	$t'=t''$
Annulment$(mr,r'):t'$ Derogate$(mr',r,r',r'',\mathscr{E}(q_1:z_1,\ldots,q_m:z_m),\mathscr{F}(b':k)):t''$	$t'=t''$
Annulment$(mr,r''):t'$ Derogate$(mr',r,r',r'',\mathscr{E}(q_1:z_1,\ldots,q_m:z_m),\mathscr{F}(b':k)):t''$	$t'=t''$
PSubstitution$(mr,r',\mathscr{A}_{r'}(a'_l:x_l/a_l:y_l,\ldots,a'_{l+m}:x_{l+m}/a_{l+m}:y_{l+m})):t'$ Derogate$(mr',r,r',r'',\mathscr{E}(q_1:z_1,\ldots,q_s:z_s),\mathscr{F}(b':k)):t''$	$t'=t''$ and $\exists u,v,\ l\leq u\leq l+m$ and $1\leq v\leq s$ such that $a_u=q_v$
PSubstitution$(mr,r'',\mathscr{A}_{r'}(a'_l:x_l/a_l:y_l,\ldots,a'_{l+m}:x_{l+m}/a_{l+m}:y_{l+m})):t'$ Derogate$(mr',r,r',r'',\mathscr{E}(q_1:z_1,\ldots,q_s:z_s),\mathscr{F}(b':k)):t''$	$t'=t''$ and $\exists u,v,\ l\leq u\leq l+m$ and $1\leq v\leq s$ such that $a_u=q_v$
TSubstitution$(mr,r'''/r,A(r''')=\{d_1:s_1,\ldots,d_o:s_o\},C(r''')=e:z):t''$ Derogate$(mr',r,r',r'',\mathscr{E}(q_1:z_1,\ldots,q_s:z_s),\mathscr{F}(b':k)):t''$	$t'=t''$
TSubstitution$(mr,r'''/r',A(r''')=\{d_1:s_1,\ldots,d_o:s_o\},C(r''')=e:z):t''$ Derogate$(mr',r,r',r'',\mathscr{E}(q_1:z_1,\ldots,q_s:z_s),\mathscr{F}(b':k)):t''$	$t'=t''$
TSubstitution$(mr,r'''/r'',A(r''')=\{d_1:s_1,\ldots,d_o:s_o\},C(r''')=e:z):t''$ Derogate$(mr',r,r',r'',\mathscr{E}(q_1:z_1,\ldots,q_s:z_s),\mathscr{F}(b':k)):t''$	$t'=t''$

Table 1. Conflicts.

3. Conflicts between Norm-modifications

Table 1 summarises all cases of basic conflicts between different norm modifications. Notice that in all cases a conflict obtains only if the conflicting modifications apply to the same time instant. Annulment of r is obviously incompatible with any partial substitution in r (first and second rows from the top). The same applies to a total substitution which replaces r with r' when we also have that r or r' is annulled (third and fourth row). A similar intuition holds for the three subsequent rows: it is impossible to derogate to r if this last rule, or r' or r'', are dropped from the system. Exactly for the same reasons, derogation to r is incompatible with total substitution of r, or of r' or r'' (the first three rows from the bottom). Finally, the cases in the fourth and fifth rows from the bottom state that a partial substitution in the antecedent of a rule is incompatible with a derogation if at least one literal used in r' or r'' to derogate to r is replaced in r' or r''.

4. The Inference Machinery

A defeasible theory D is a structure $D=(\mathscr{T},F,R^c,R^O,\prec)$ where $\mathscr{T}$ is discrete totally ordered set of instants of time, F is a finite set of temporalised literals, R^c is a finite set of temporal counts-as rules with viewpoint, R^O is a finite set of temporalised deontic rules with viewpoint, $\prec$, the superiority relation, is defined as $(R^c\times R^c\times\mathscr{T})\cup(R^O\times R^O\times\mathscr{T})$. A conclusion in Defeasible Logic can have one of the following four forms:

$+\Delta@t\ q:t'$ meaning that q is definitely provable, at time t' with viewpoint t, in D (i.e., using only facts and strict rules).

[2] Again, to simplify the discussion, we will assume that r is defeasible and that r' will have the same strength of r. Of course, that r be defeasible is a necessary requirement, otherwise r'' could not block r.

$-\Delta@t\ q : t'$ meaning that we have proved that q is not definitely provable, at time t' with viewpoint t, in D.

$+\partial@t\ q : t'$ meaning that q is defeasibly provable, at time t' with viewpoint t, in D

$-\partial@t\ q : t'$ meaning that we have proved that q is not defeasibly provable, at time t' with viewpoint t, in D.

For example, $+\partial_O@t_1\ q : t_0$ means that we have a defeasible proof for ***Obl***q at t_0, or, in other words, that ***Obl***q holds at time t_0 when we use the rules in force in the normative system at time t_1. To specify whether a conclusion $q : t$ has been obtained via transient rules or via persistent rules we will introduce auxiliary proof tags indicating persistency or transiency. The proof tags are labelled with the mode used to derive the rule, according to their appropriate proof conditions. It is not possible to give the complete set of proof condition in this paper. Here we concentrate only on the proof conditions to derive defeasible persistence of both rules, and literals with both counts-as and obligation mode. The proof conditions given here are extensions of those given in [1]; the omitted proof conditions can be analogously obtained.

Provability is based on the concept of a *derivation* (or proof) in D. A derivation is a finite sequence $P = (P(1),\dots,P(n))$ of tagged literals satisfying the proof conditions (which correspond to inference rules for each of the kinds of conclusion). $P(1..n)$ denotes the initial part of the sequence P of length n. A strict derivation (i.e., a conclusion tagged with Δ) is just a monotonic derivation using forward chaining of rules, that is, modus ponens. In Defeasible Logic a defeasible derivation, on the other hand, has three phases. In the first phase we propose an argument in favour of the conclusion we want to prove. In the simplest case this consists of an applicable rule for the conclusion (a rule is applicable if the antecedent of it has already been proved). Then in the second phase we examine all possible counter-arguments (rules for the opposite conclusion). Finally we have to rebut the count-arguments. Thus we have to provide evidence against the counter-argument. Accordingly we can demonstrate that the argument is not as such (i.e., some of its premises are not provable), or we can show that the counter-argument is weaker than an argument for the conclusion. For persistent conclusions we have another method. We can use a derivation of the conclusion at a previous time provided that no terminating event occurred in between.

In [1] the rules are given, but the formalism we have introduced in the previous sections allows us to have rules in the head of counts-as rules, thus we have to admit the possibility that rules are not only given but can be derived. Thus in the proof conditions we have to cater for this option. Accordingly we have to give conditions that allow us to derive rules instead of literals. For the sake of simplicity we will assume that all rules in R can be overruled/modified. Then we have to extend the notation $R[x : t]$ to the case where x is a rule label (and norm-modifications). Given a set of (counts-as) rules R and a set of rule modifiers $M = \{m_1,\dots,m_n\}$, then

$$R[r : t_r] = \{s \in R : C(s) = m_i(v : t_v) \text{ and } m_i(v : t_v) = r : t_r\}$$

$R[r : t_r]$ gives the set of nested rules whose head results in the rule $r : t_r$ after the application of the rule modifier; and

$$R[\sim r : t_r] = \{s \in R : C(s) = m_i(r : t_r) \text{ and } m_i(r : t_r) \text{ is in conflict with } r : t_r\}$$

The set $R[\sim r : t_r]$ gives the set of rules that modify $r : t_r$ and the modifi cation is in conflict with the $r : t_r$, see Table 1 for such conflicts. We can now give the proof condition for $+\partial^p$ to derive a rule.

If $P(n+1) = +\partial^p @t\ r : t_r$ then
1a) $r : t_r @t \in R^c$ or
1b) $\exists s : t_s \in R^c[r : t_r]$ such that $+\partial @t\ s : t_s \in P(1..n)$ and
$\quad\forall a : t' \in A(s), +\partial @t\ a : t' \in P(1..n)$; and
2) $\forall v : t_v \in R^c[\sim r : t_r]$ if $+\partial @t\ v : t_v \in P(1..n)$, then either
$\quad$2.1) $\exists b : t'' \in A(v)$ such that $-\partial @t\ b : t'' \in P(1..n)$ or
$\quad$2.2a) $v : t_v \prec_t r : t_r$ if 1a obtain or
$\quad$2.2b) $v : t_v \prec_t s : t_s$ if 1b obtain; or
3) $+\partial^p @t'\ r : t_r \in P(1..n), t' < t$ and
$\quad$3.1) $\forall t'', t' \leq t'' < t, \forall s : t_s \in R[\sim r : t_r]$ if $+\partial @t''\ s : t_s \in P(1..n)$, then either
$\quad\quad$3.1.1) $\exists a : t_a \in A(s), -\partial @t''\ a : t_a \in P(1..n)$ or $t_s < t_r$; and
4) $+\partial^p @t\ r : t'_r \in P(1..n), t'_r < t_r$ and
$\quad$4.1) $\forall t', t'_r \leq t'' < t_r, \forall s : t_s \in R[\sim r : t_r]$ if $+\partial @t'\ s : t_s \in P(1..n)$, then either
$\quad\quad$4.1.1) $\exists a : t_a \in A(s), -\partial @t'\ a : t_a \in P(1..n)$ or $t_s < t'_r$.

Let us briefly examine the above proof conditions. To prove a rule at time t, the rule must be in force at time t, i.e., the rule must be one of the given rules (condition 1a). There is a second possibility the rule is derived from another rule. The second rule must be provable and applicable at t (condition 1b). However, this is not enough since there could have been modifi cations to the rule effective at t. Thus we have to show that either all eventual modifi cations are not applicable (2.1) or the modifi cations are not successful since they are defeated (2.2a and 2.2b). Finally the rule could be provable because it was persistent, i.e., it was persistently in force before (3), and no modifi cation occurred in between. The possible modifi cations in force after the rule was in force are not applicable to the rule. Or (4) the rule was persistently effective before, and its effectiveness was not revoked. The inference condition for positive persistent defeasible proofs is as follows.

If $P(n+1) = +\partial^p @t\ q : t'$ then
1) $+\Delta^p @t\ q : t' \in P(1..n)$, or
2) $-\Delta @t \sim q : t' \in P(1..n)$, and
$\quad$2.1) $\exists r : t_r \in R^p_{sd}[q : t']$: $+\partial @tr : t_r \in P(1..n), \forall a : t_a \in A(r : t_r), +\partial @t\ a : t_a \in P(1..n)$ and
$\quad$2.2) $\forall s : t_s \in R[\sim q : t]$: if $+\partial @t\ s : t_s$, then either $\exists a : t_a \in A(s : t_s), -\partial @t\ a : t_a \in P(1..n)$; or
$\quad\quad$2.2.1) $\exists w : t_w \in R[q : t]$: $+\partial @t\ w : t_w \in P(1,,n)$ and
$\quad\quad\quad\forall a \in A(w : t_w), +\partial @t\ a : t_w \in P(1..n)$ and $w \succ s$; or
3) $\exists t'' \in \mathscr{T} : t'' < t$ and $+\partial^p @t''\ q : t' \in P(1..m)$ and
$\quad$3.1) $\forall t'''\ t'' < t''' \leq t\ \forall s : t_s \in R[\sim q : t']$: if $+\partial @t'''\ s : t_s \in P(1..n)$, then either
$\quad\quad$3.1.1) $\exists a : t_a \in A(s : t_s), -\partial @t'''\ a : t_a \in P(1..n)$ or
$\quad\quad$3.1.2) $\exists v : t_v \in R[q : t'], +\partial @t'''\ v : t_v \in P(1..n)$ and
$\quad\quad\quad\forall b : t_b \in A(v : t_v) + \partial @t'''\ b : t_b \in P(1..n)$ and $s : t_s \prec_{t'''} v : t_v$; or
4) $\exists t'' \in \mathscr{T} : t'' < t'$ and $+\partial^p @t\ q : t'' \in P(1..m)$ and
$\quad$4.1) $\forall t'''\ t'' < t''' \leq t'\ \forall s : t_s \in R[\sim q : t''']$: if $+\partial^p @t\ s : t_s \in P(1..n)$, then either
$\quad\quad$4.1.1) $\exists a : t_a \in A(s : t_s), -\partial @t\ a : t_a \in P(1..n)$ or
$\quad\quad$4.1.2) $\exists v : t_v \in R[q : t'''] + \partial @t\ v : t_v \in P(1..n)$ and
$\quad\quad\quad\forall b : t_b \in A(v : t_v) + \partial @t\ b : t_b \in P(1..n)$ and $s : t_s \prec_{t'''} v : t_v$.

The rationale of above proof conditions is the same of those in [1]. The main difference is that here every time we use a rule we have to verify that the rule is provable in the system. In addition we have to cater for both the persistence of the effectiveness of the rule (4) and that the rule was previously persistently in force (3).

5. Time-forking and Norm-modification

Retroactivity may be applied to basic rules as well as to norm-modifications. Cases of retroactivity of basic rules occur when we have rules such as $r : (a : x \Rightarrow^p_O p : y) : z$. r is a permanent (defeasible) deontic rule in force since z, persisting through subsequent instants $z_* \geq z$, stating that if a occurs at x then p is obligatory since time instant y. Suppose we have that $y = x + 1$ and $z \geq 5$, and in particular that $z = 5$. If the theory contains $a : 2$, this allows for the derivation of $+\partial^p_O@5\ p : 3$, namely $\boldsymbol{Obl}p$ at time 3 from the viewpoint of 5, which is the time-instantiation of r being in force. At 5, in fact, r is in force; on the other hand, there is no time-constraint over x and y, and so a occurring at 2 permits to derive retroactively (namely, with respect to 5) p at 3. Suppose that the theory contains also

$$s : (b : j \Rightarrow^p_O \neg p : j) : k$$

such that $k = 2$. If $b : 3$ we would derive $+\partial^p_O@k\ \neg p : 3$. Since $k = 2$, then $+\partial^p_O@5\ \neg p : 3$ thus getting a conflict. If we want to have r prevailing over s the theory should say that $r \succ_t s$, $t = 5$. Notice, however, that we will still derive $+\partial^p_O@k_*\ \neg p : 3$ where $2 \leq k_* \leq 4$. This means that, if we have another rule

$$s' : (\boldsymbol{Obl}\neg p : i \Rightarrow^p_O q : i) : m$$

such that $m = 2$, then we will derive $+\partial^p_O@m_*\ \neg q : i$ where $2 \leq m_* \leq 4$ but $-\partial^p_O@m_*\ \neg q : i$ where $m_* > 4$. Let us see now cases of retroactivity of norm-modifications. Suppose r is defined as at the beginning of this section. Again, if we have $a : 2$, then $+\partial^p_O@5\ p : 3$. Imagine the system contains the following meta-rule, stating the annulment of r:

$$mr : (c : x' \Rightarrow^p \text{Annulment}(mr, r) : y') : z'$$

such that $z' = 4$ while $z = 3$ (3 is the time-instantiation of the being-in-force of r). This meta-rule assumes for its application that r is in force, which is true from time instant 3 on (rules persist over time unless cancelled or modified). Let c be a future but uncertain event under which the norm-modification will apply. This means that $x' > z$. However, we state that mr works retroactively, e.g., that $y' = 3$. Suppose we have $c : 7$. Accordingly, we will derive $+\partial^p@z'_*$ Annulment$(mr, r) : 3$, where, e.g., $z'_* = 5$. As we have stated, a modification is nothing but a function which assigns, to the modified rule, another rule which holds at the time of the modification and so at the time-instantiation of the modification and at every subsequent time instant. In the case of annulment, the value of the the function is the empty rule, since the rule associated to the name r is dropped from the system. Thus, after the application of mr, we will also get $+\partial^p@z'_*\ (r : \bot) : y'$. Thus, from the viewpoint of $z'_* = 5$, if $a : 2$, this will no longer permit to derive the modalised consequent of r, i.e., $\boldsymbol{Obl}p : 3$ because, at that perspective, r does not exist anymore: $-\partial^p_O@z'_*\ p : 3$. Finally, notice that if mr were a transient rule, we would be able, under appropriate conditions, to model the notion of temporary annulment: in fact, in transient rules the effects are co-occurrent with the conditions. Logically, this is just a variation of the intuition just discussed. Let us see two cases, one of retroactive partial substitution and one retroactive derogation. Even here, dealing with these cases is a matter of variation of the previous logic intuitions. So we will only provide a brief discussion. Given r

as above, suppose we have the following meta-rule, stating the substitution, in the body of r, of $a:x$ with $d:x$:

$$mr' : (c : x' \Rightarrow^p \text{PSubstitution}(mr', r, \mathscr{A}_r(d : x/a : x)) : y') : z'$$

such that $z' = 4$ while $z = 3$. If we assume the same conditions of the case of the annulment just discussed, the fact of having $a : 2$ permits to derive $\boldsymbol{Obl}p : 3$ only if the viewpoint is such that $z_* < 4$. When mr' comes to be in force, the conditions for obtaining $\boldsymbol{Obl}p$ from r will be no longer a but d. Let us finally focus on the notion of retroactive derogation. Given r as above, suppose we have the following meta-rule, stating the derogation of r when the additional condition q obtains and such that the different effect p' should follow under this additional condition:

$$mr'' : (c : x' \Rightarrow^p_c \text{Derogate}(mr'', r, r', r'', \mathscr{D}(q : x), \mathscr{E}(p' : y) : y') : z'$$

Again, we assume the same conditions of the case of the annulment just discussed. Under those conditions we can reiterate the same argument and so we can derive $+\partial^p_O@z_* \; p : 3$ given $a : 2$. However, if mr'' is made applicable, we will derive $+\partial^p@z'_* \; \text{Derogate}(mr, r, r', r'', \mathscr{E}(q : x), \mathscr{F}(p' : y)) : y'$, namely that this holds from the viewpoint of z'_* where $z'_* \geq 4$. According to the definition of Derogate, this conclusion is associated to adding in the theory the following two rules:

$$r' : (a : x, q : x \Rightarrow^p_O p' : y) : y' \qquad r'' : (a : x, q : x \rightsquigarrow^p_O \neg p : y) : y'$$

So, after the application of mr'', we will have r' and r'' in force since time instant 3. Thus, if we have $a : 2$ but also $q : 2$, then we will no longer derive $\boldsymbol{Obl}p$. In particular, at time 3, this conclusion will be blocked and we will get $+\partial^p_O@3 \; p'$.

6. Summary

We extended the defeasible logic presented in [1] by allowing nested rules in the head of counts-as rules. This extension increases the expressive power of the logic and it allows us to represent meta-norms describing norm-modifications. We outlined the inferential mechanism needed to deal with the derivation of rules. Then we described some issues related to norm versioning and we illustrated the techniques with some relevant norm-modifications such as annulment, partial and total substitution and derogation. We showed that the formalism introduced here is able to deal with complex scenarios such as retroactivity and time-forking.

References

[1] G. Governatori, A. Rotolo, and G. Sartor. Temporalised normative positions in defeasible logic. In *Proceedings of the Tenth International Conference on Artificial Intelligence and Law*, pages 25-34, New York, ACM, 2005.

[2] J. Martinek and J. Cybulka. Dynamics of legal provisions and its representation. In *Proceedings of the Tenth International Conference on Artificial Intelligence and Law*, pages 20-24, New York, ACM, 2005.

[3] M. Palmirani. Time model in normative information systems. In *Post-proceedings of the ICAIL Workshop on the Role of Legal Knowledge in e-Governement*, Wolf, Tilburg, 2005.