Temporalised Normative Positions in Defeasible Logic

Guido Governatori[*]
School of ITEE
University of Queensland, Australia
guido@itee.uq.edu.au

Antonino Rotolo
CIRSFID and Law Faculty
University of Bologna, Italy
rotolo@cirsfid.unibo.it

Giovanni Sartor
CIRSFID and Law Faculty
University of Bologna, Italy
sartor@cirsfid.unibo.it

ABSTRACT

We propose a computationally oriented non-monotonic multi-modal logic arising from the combination of temporalised agency and temporalised normative positions. We argue about the defeasible nature of these notions and then we show how to represent and reason with them in the setting of Defeasible Logic.

1. MOTIVATION AND LAYOUT

An increasing number of works on agents assume that in artificial societies normative concepts may play a decisive role, allowing for the flexible co-ordination of autonomous agents [5, 20, 15]. In particular, it seems crucial to model organisations of agents in terms of policy-based normative systems; accordingly an organisation should be characterised by specifying the normative positions relevant to design its structure. These positions include duties, permissions, but also powers, as for instance powers of creating further normative positions on the head of other agents. In this paper we will develop a formal machinery to account for several fundamental concepts that are required to model policy-based normative systems. These concepts will be embedded in a non-monotonic and computationally-oriented framework based on Defeasible Logic (DL).

From the conceptual standpoint, it is well known that the basic deontic qualifications (obligatory, forbidden, permitted and facultative) are not sufficient to capture all fundamental normative notions, such as the concepts of rights and power. For this reason, we will first provide an account in DL of the notion of other-directed obligation [14] to express, e.g., the first Hohfeldian set of fundamental concepts: duty, right, noright, and privilege. Second, we shall focus on different kinds of normative conditionals. This will enable us to characterise also the idea of normative power and articulate many potestative concepts such as the second Hohfeldian set of concepts: power, liability (or, subjection, to avoid confusion with the notion of liability, as used in tort law), disability, and immunity.

[*]This work was partially supported by Australia Research Council under Discovery Project No. DP0558854 on "A Formal Approach to Resource Allocation in Web Service Oriented Composition in Open Marketplaces".

ICAIL '05, June 6-11, 2005, Bologna, Italy.
Copyright 2005 ACM 1-59593-081-7/05/0006 ...$5.00.

In general, we shall see how the analysis of normative conditionality and normative positions has to include temporal aspects in order to capture a number of nuances in the perspective of framing a fine-grained classification of such concepts. In fact, normative determinations take place along the axis of time: Normative preconditions hold or happen in certain temporally characterised occasions, and consequently their effects too hold or happen within temporal bounds. In providing an analysis of normative conditionality, we cannot refrain from introducing some temporal notions, though avoiding as much as possible the complexities of temporal logics. In this regard, we will basically distinguish between "persistent" normative positions, which follow the so-called law of temporal inertia or temporal persistence, and those that are temporally co-occurent with regard to some events or states of affairs.

2. THE LEGAL FRAMEWORK

This section provides a systematic account of the basic legal concepts discussed in the paper. Rather than providing general definitions, we prefer to illustrate all notions by way of examples. We will follow the approach and notation developed in [23]. The notation adopted there is based on the use of modal operators and on formalisms taken from the Event Calculus. We prefer in this section to use this notation as it looks more intuitive, though it can be hardly embedded, as it is, within DL.

These are the contexts whose logical behaviour will be caracterised in the section.

2.1 Actions

With regard to actions, we distinguish two characterisations, a behavioural and productive characterisation. The first is concerned with specifying the behaviour performed by the agent, the second, the result the agent achieves.

$Does_{Tom}$[smoke]
(*Tom* is smoking)

$Brings_{Alex}$[the air is polluted]
(*Alex* brings it about that the air is polluted)

The productive action can also consist in making so that a conditional holds true.

2.2 Obligation and Permission

By applying to actions the usual deontic operators, we obtain obligations. Obligations can concern both behavioural and productive actions:

Obl $Does_j$[lecture]
(it is obligatory that j lectures)

Obl $Brings_j$[k's personal data are cancelled]
(it is obligatory that j brings it about that k's personal data are cancelled)

With regard to the obligations whose content is the negation of an action, we speak of prohibition.

Obl NON $Does_j$[smoke]
(it is obligatory that j does not smoke, or also, it is forbidden that j smokes)

In order to express the basic deontic modalities, besides obligations we also need permissions, which also concern both kinds of actions. Here are a couple of examples:

Perm $Does_j$[park in the University courtyard]
(it is permitted that j parks in the University courtyard)

Perm $Brings_j$[k's personal data are accessible to the police authority]
(it is permitted that j brings it about that k's personal data are accessible to the police authority)

We shall later develop a logical analysis of permissions. By now, let us just state some properties which appear to be important to us. First of all a permission must be incompatible with a prohibition. Secondly, the holding of a permission does not mean simply that there is no prohibition (we want a strong notion of a permission).

The simplest way to meet this requirement consists in viewing the permission of an action as the negation of its prohibition. This is the idea that was followed in [23].

Here we take a slightly different approach: We model permissive propositions as defeaters of the corresponding obligations. Accordingly, to state that proposition A is permitted is equivalent to say the normative proposition ***Obl*** $\neg A$ is not derived. This allows us to exploit the treatment of defeaters in defeasible logic to model permissions. As we shall see, this idea leads us to an unpleasant implication, so long as we remain within defeasible logic: We can never infer that something is permitted, since defeaters block inferences, but do not establish any conclusion. To remedy this fact, we identify the derivation of the permission of A with the situation where a defeater of ***Obl*** $\neg A$ exists, whose antecedent can be established and which is superior to the all rules it attempts to defeat (all rules for ***Obl*** A, whose antecedent can be established).

2.3 Directed Normative Propositions

An essential component of legal concepts, though this has been denied by some legal philosophers, and more notably by Hans Kelsen (see [23]), is their teleological dimension: normative propositions. Normative propositions are made legally binding (and are recognised as such) since their adoption and practice advance certain goals. These goals may consist in either:

- a collective one (collective health)
- and individual one (the health of any individual)

The idea of legal teleology, when the goal of a certain proposition consists in the benefit of a particular individual leads us to the idea of a *directed normative proposition*, that is a normative proposition which is aimed at (which has the goal of) satisfying he intererest of a particular person. A specification of this idea is provided by the notion of an other-directed obligation, as in the following examples.

Obl$^{Mary}Does_{Tom}$[pay € 1,000 to *Mary*]
(it is obligatory, toward Mary, that Tom pays € 1,000 to *Mary*)

Obl^{Tom}NON $Does_{Mary}$ [communicate to others *Tom*'s trade secrets]
(it is obligatory, toward Tom, that *Mary* does not communicate to others *Tom*'s trade secrets)

In general, by an *other-directed obligation* like

Obl$^{k}Does_jA$

we mean that is *obligatory, toward k*, that j does A, namely it is obligatory, that j does A, in order to advance the interest of k.

The idea of an other-directed obligation is complemented by the idea of an *other-directed permission*: this corresponds to the negation of an other-directed obligation, or rather to the statement that such an obligation is not be applied. For instance, the proposition

Perm$^{Tom}Does_{Mary}$[raise a building up to 15 meters high]
(it is permitted, toward *Tom*, that *Mary* raises a building up to 15 meters high)

expresses the idea that Mary is allowed to raise the building, as far as Tom's interest are concerns (this is intended to defeat the obligation, towards Tom, that Mary does not rise the building).

An other directed permission towards k that j does A does not entail that k is prohibited from preventing j from doing A: it only consists in rejecting that k is forbidden, for the benefit of k, from doing A . Thus k, notwithstanding the permission can still attempt to prevent A in all legal ways (for instance, Tom can threaten Mary the he shall leave his house and sell his land to a constructor that will build a 20 storeys skyscraper, if she raises her building).

2.4 Obligational Rights

On the basis of the notion of directed obligation, we can construct the notion of an *obligational right*, which is the other face of a directed obligation. For instance

OblRight$_{Tom}Does_{Mary}$[pay € 1,000 to *John*]
(*Tom* has the obligational right that *Mary* pays € 1,000 to *John*)

means that:

Obl^{Tom}($Does_{Mary}$[pay € 1,000 to John]
(it is obligatory, toward *Tom*, that *Mary* pays € 1,000 to *John*)

Thus we obtain the idea of a right as consisting in the protection of an interest (through the obligation upon somebody else), an idea which was defended in particular by Bentham and Jhering (see [23]; for a discussion of directed obligations and rights, see, e.g., [14]).

We can also introduce the notion of permissive right, which is a directed permission aimed at satisfying the interest of the permitted person. Note that the holder of a permissive right, in the sense of its beneficiary, is the author of the permitted action, rather than the person toward which this permission exists: In a permissive right ***PermRight***$^{k}Does_jA$, j is the beneficiary of the right, while k, the addressee of the permission, bears the burden of it (in the sense that k's interest is not protected by means of a prohibition that j does A). For instance, to express that Mary has the permissive right toward Tom to raise a building up to 15 meters, we write:

PermRight$^{Tom}Does_{Mary}$[raise a building up to 15 meters]

Similarly, to express that Ali, a Muslim worker, has the permissive right to abstain from work on Fridays, against his employer Mary, we write:

PermRight^{Mary}(NON $Does_{Ali}$[work on Friday])

Note that the distinction between being one's permission to do A and another's prohibition to prevent A also applies to permissive rights: ***PermRight***$^{k}Does_jA$ does not entail that k is prohibited from preventing A.

2.5 Hohfeldian Notions

Further combinations of these notions lead us to the ideas of permissive rights and absolute rights, for which we refer to [23], and to the Hohfeldian notions of a privilege and a noright, which we can simply define as follows.

The Hohfeldian notion of a *privilege* can be expressed through our notion of other-directed permission: To say that j has a privilege toward k with regard to A means that j is permitted, toward k, to omit A. Equivalently, we may use the notion of *noright* to express that one does not have the obligational right that another does a certain action, that is, to denote the situation when the latter is permitted toward the first to omit that action. The ideas we have just defined lead us to first square of Hohfeldian concepts, which we call the Hohfeldian obligational set, where:

$$\begin{tikzcd} \boldsymbol{OblRight}_k Does_j A \arrow[r, dash, "\textbf{equivalent}"] \arrow[d, dash, "\textbf{incompatible}"'] & \boldsymbol{Obl}^k Does_j A \arrow[d, dash, "\textbf{incompatible}"] \\ \boldsymbol{NoRight}^k Does_j A \arrow[r, dash, "\textbf{equivalent}"'] & \boldsymbol{Privilege}^k Does_j A \end{tikzcd}$$

2.6 Normative Conditionals

Besides obligations and permissions, the main structure of legal knowledge is provided by normative conditionals. Normative conditionals establish a connection according to which and antecedent normatively determines a consequent, which we represent as:

IF A THEN^n B

where the superscript n indicates the specifically normative nature of the connection. Here is an example, which uses the notions we introduced above.

FORANY (x)
IF [x is below 18 years]
THEN^n ***Forb*** *Does*$_x$[buy alcoholic drinks]
(for any x, if x is below 18 years, then it is forbidden that x buys alcoholic drinks)

We distinguish different types of normative conditional, according to the nature of their consequent:

Deontic initiation
IF [Tom does not deliver the merchandise in time]
THEN^n [Tom becomes obliged to pay a penalty of € 1,000]

Deontic termination
IF [Mary renounces to the payment of the penalty]
THEN^n [Tom ceases to be obliged to pay it]

Deontic emergence
IF [Tom is inside a mosque]
THEN^n [Tom is forbidden to wear shoes]

Qualificatory initiation
IF [Tom is born in Italy]
THEN^n [Tom becomes an Italian citizen]

Qualificatory initiation
IF [Tom acquires another citizenship]
THEN^n [Tom ceases to be an Italian citizen]

Qualificatory state-emergence
IF [an object is permanently attached to the soil]
THEN^n [the object is an immovable good]

Qualificatory event-emergence
IF [Tom drives while being drunk]
THEN^n [Tom commits a criminal offence]

Conditionals may, on the one hand, be concerned with the a deontic or a non deontic property (state of affairs) or, on the other hand, they may be concerned with an event determining the initiation of the property (which continues after the event has taken place), with a property determining the emergence of another property (which only exists so long as the determining property exists), or with an event determining the emergence of another event. Here are an example of non deontic supervenience and initiation.

FORANY (x,t)
IF [x is a piece of land, water source, a tree, a building, or is anyway permanently attached to the land] *holds at time t*
THEN^n [x is an immovable good] *holds at time t*

FORANY (x,t)
IF [x makes an offer to the public, containing all terms of the contract it concerns] *happens at time t*
THEN^n [x makes a contractual offer] *happens at time t*

2.7 Notions of Power

The idea of a normative conditional leads us to the notion of a power. In a first sense we say that a person has a power (an action power) when, according to a normative conditional, an action of that person determines a normative effect. Thus we view the two expressions below as equivalent:

FORANY (x,y)
WHEN [animal y does not belong to anybody]
THEN^n IF *Does*$_x$ [capture y]
THEN^n ([x is the owner of y] *initiates*)

FORANY (x,y)
WHEN [animal y does not belong to anybody]
THEN^n ***ActionPower***$_x$ [x is the owner of y] *initiates*
VIA
[capturing y]
(for any person x and animal y, if y does not belong to anybody, then x has the action-power of initiating x's ownership of the animal, by capturing y)

In a more specific sense, we may speak of a power only under a teleological perspective, that is when the normative connection between an antecedent action and a consequent normative effect has the function of enabling and promoting the achievement of the effect through the action. This excludes from this restricted notion of a power the case when an action determines a penalty against its author (we do not want to say that a person has the power of achieving the subjection to punishment by committing a crime). When such a normative connection also has the further purpose of promoting the interest of the author of the action, we speak of a potestative-right. Thus, we have the following distinction:

- An enabling-power is an action-power intended to enable the normative result of the action
- A potestative-right is an enabling power intended to further the interests of the power holder

This is the notion of a right to which those authors which have defended a power-based theory of rights, like

FORANY (x, y)
WHEN [animal y does not belong to anybody]
THEN^n ***PotestativeRight***$_x$
[x is the owner of y] *initiates*
VIA [capturing y]
(for any person x and animal y, if y does not belong to anybody, then x has the potestative-right of initiating x's ownership of the animal, by capturing y)

When an agent j has the power of creating a normative position *Pos* concerning another person k, we can say that k is subject to this power, and when there is no such power we can say that j has a disability to create that position and that k has an immunity with regard to it. This leads us to the Hohfeldian potestative set.

$$\begin{array}{ccc}
\textbf{\textit{EnablingPower}}_j\textbf{\textit{Pos}}_k & \text{— equivalent —} & \textbf{\textit{Subjection}}^j_k\textbf{\textit{Pos}} \\
| & & | \\
\textbf{incompatible} & & \textbf{incompatible} \\
| & & | \\
\textbf{\textit{Disability}}^k_j\textbf{\textit{Pos}} & \text{— equivalent —} & \textbf{\textit{Immunity}}^j_k\textbf{\textit{Pos}}
\end{array}$$

2.8 Proclamation and Proclamative Power

A special instance of the idea of a potestative right concerns the case when one person has normative ability to create a normative position (in general, to realise a normative proposition) by stating this intention. The act of stating one's intention to produce a certain normative result is what we call proclamation, and a proclamation is effective (produces its intended result) when there is a rule like the following:

IF $Procl_x \varphi$ THEN^n $Brings_x\ \varphi$
(for any person x, and proposition φ, if x proclaims φ, then φ is realised)

When such a rule exists, we say that the author of the proclamation has a corresponding proclamative power. For instance, if a rule like the following holds:

IF $Procl_{Tom}$[Tom's contract with Mary is terminated]
THEN^n $Brings_{Tom}$[Tom's contract with Mary is terminated]

we can say that Tom has the declarative power of realising the termination of his contract:

ProclPow$_{Tom}$[Tom's contract with Mary is terminated]

According to our classification above, proclamation rules can be seen as rules concerned with an the emergence of an event. According to the content of a proclamation, namely, its intended result, we can have further distinctions:

1. $Procl_j$([***Obl*** $Does_j A$] *initiates*): j proclaims that her own obligation to do A initiates (promise).
2. $Procl_j$([***Obl*** $Does_k A$] *initiates*): j proclaims that k's obligation to do A initiates. (command)
3. $Procl_j$([***Obl*** $Does_j A$] *terminates*): j proclaims that her own obligation to do A terminates (withdrawal of a promise).
4. $Procl_j$([***Obl*** $Does_k A$] *terminates*): j proclaims that k's obligation to do A terminates (withdrawal of a command).

When a proclamation is effective (according to a proclamation rule), its intended effect follows, according to logical inference. Here is an example. Given:

P_1: $Procl_{John}$
[*Brothers Karamazov* belongs to Mary]
initiates concerns making a small gift
P_2: $Procl_{John}$
[*Brothers Karamazov* belongs to Mary]
initiates

R: FORANY (φ, x)
WHEN [$Procl_x \varphi$ concerns making a small gift]
THEN^n IF $Procl_x \varphi$ THEN^n $Brings_x \varphi$

we can conclude that:

[*Brothers Karamazov* belongs to Mary] *initiates*

On the basis of the idea of a proclamation, we can develop further interesting legal concepts, such as the concept of the notion of representation. On the other hand, we can also build the idea of a source of law.

3. THE LOGICAL FRAMEWORK

3.1 Introduction

Our approach is motivated by the inherent computational complexity of multimodal logics (see, e.g., [12]). In addition, very often the notion of modality adopted for multiagent systems is by its own nature non-monotonic and so does not lend itself to necessitation [10]. In general, the addition of (normal) modal operators to the classical propositional base leads to the increase of complexity of the logic. This is mainly due to: (1) the rules to introduce modalities, such as the necessitation rule, (2) the axioms governing the behaviour of modalities and their mutual interaction. These problems are even more crucial in multiagent systems where the combination of a number of modalities is usually required and the need of efficient reasoning mechanisms is compelling (see, e.g., [7]).

This paper thus proposes a general solution to the mentioned problems by exploiting the nice computational features of DL and by setting specific rules for introducing modal operators: Rules are primarily meant to introduce modalities in terms of provability of literals. This solution keeps the system manageable. A literal will be modalised with, say, X if it is deduced via rules specifically devised to express concept X. Given the conceptual framework previously described, this methodology is applied to the following building blocks, which are needed to correctly represent the normative aspects of policy-based systems of agents[1].

[1]Notice that in DL we will not deal explicitly with quantifiers. We may have that rules with free variables are interpreted as rule schemas, that is, as the set of all ground instances; in such cases we assume that the Herbrand universe is finite.

3.1.1 Actions

The formalism aims to capture the idea of modal agency [6] but also to reason on explicit actions. Modal logics of agency are very general since actions are simply taken to be relationships between agents and states of affairs. In this regard, we will focus on he idea of personal and direct action to realise a state of affairs, formalised by the modal operator *Brings*: $Brings_i A$ means that the agent i brings it about that A. Different axiomatisations have been provided for it but almost all include $Brings_i A \rightarrow A$, $\neg Brings_i \top$, $(Brings_i A \wedge Brings_i B) \rightarrow Brings_i (A \wedge B)$, and are closed under logical equivalence. If these are some general properties for *Brings*, a specific axiom advanced in [22] for this operator is $Brings_i Brings_j A \rightarrow \neg Brings_i A$. It corresponds to the idea that the bring-about operator expresses actions performed directly and personally: It is counter-intuitive that the same agent brings it about that A and brings it about that somebody else achieves A. Accordingly, we will devise a set of rules to encode the action transitions occurring, under certain circumstances, as the results of actions. The applicability of these rules will allow for the introduction of the operator *Brings*. In addition, since we want to be also able to reason about explicit actions, we will extend the language of DL by adding a set of action symbols. These symbols will occur only in the antecedent of action rules, and this precisely because such rules are meant to derive the states of affairs resulting from the performance of actions.

3.1.2 Normative Conditionals

It is possible to distinguish different kinds of normative conditional. As was argued, all of them indeed can be conceptually reduced to a single normative conditional "IF A THEN^n B": what differentiates each kind of link is the type of antecedent and/or consequent that occur in the conditional link. Thus a normative conditional is meant to simply obtain any kind of normative consequence. However, to make things clear, we will distinguish, in the syntax of DL, different kinds of rules, and we simply establish in the proof conditions for literals the appropriate criteria to account for this reduction. We will identify the following types of normative rules:

Rules for persistent obligations. These rules, if applicable, permit to infer literals to be modalised by obligations that persist unless some other, subsequent, and incompatible states of affairs, actions, or obligations terminate them. For example, the obligation towards an agent of paying the damages she caused in a car crash will hold until the agent has not paid such damages.

Rules for co-occurrent obligations. These rules 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.

Rules for the count-as link. The count-as relation is meant to express the idea of institutional power [24]. 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. In principle, this kind of ability should be distinguished from the practical capacity to obtain a certain state of affairs. The exercise of a power may not be successful: its being successful, within the institutional context, depends on whether that institution makes it effective. In [13, 7], for example, different formalisations of this notion have been developed. We will not commit ourselves in choosing one of the these approaches. Both approaches characterise the count-as link by also introducing the a modality D_s, which is meant, though according to different intepretations, to qualify the facts and propositions that hold within a given institution s. As we shall see, here it not required to use D_s, which would be necessary in our framework only in the case in which we should deal with facts and propositions that hold within different institutions. From the conceptual point of view, we will view the count-as link as a generic normative conditional whose consequences are not necessarily deontic. Deontic rules will be then a considered as generic normative rules which specifically allow for the deduction of deontic literals. In DL, we will syntactically distinguish the two types of rules, but we will impose, according to this view, that the provability of deontic consequences is a sub-type of the provability via count-as rules.

3.1.3 Time

Since we will operate in a temporalised setting, we will not only impose that obligations be temporalised, but also that any literal may be labelled by time instants. For the sake of simplicity, here will assume the time to be linear and discrete. Notice that Event Calculus' temporal notions and predicates "*initiates*" and "*terminates*" will be rendered in DL by reframing the idea of DL superiority relation between different types of temporalised rules.

3.1.4 Proclamations

The last logic component we will use is the modal operator *Procl*, which will capture the idea of proclamation [7]. It expresses any speech act (proclaiming) that is intended to modify the institutional world. The action of proclamation is not necessarily successful: $Procl_i A$ is just an attempt of i to achieve A. Whether it is successful or not, will depend on whether an institution, for example, will make it effective by means of appropriate count-as rules.

3.2 The System

DL [18, 1] is a simple, efficient but flexible non-monotonic formalism that can deal with many different intuitions of non-monotonic reasoning [2], and efficient and powerful implementations have been proposed [17, 4]. In the last few years DL has been applied in many fields; in addition DL encompasses other existing formalisms proposed in the AI & Law field such as Prakken and Sartor's [21] and Loui and Simari's [25] (see, [9]), and recent work shows that DL is suitable for extensions with modal and deontic operators [10] and violations which in turn allows DL to be used to represent and reason with business contracts [11, 8]. Here we propose a non-monotonic logic of institutional agency based on the framework for DL developed in [1].

As usual with non-monotonic reasoning, we have to specify 1) how to represent a knowledge base and 2) the inference mechanism used to reason with the knowledge base. The language of Normative Defeasible Logic consists of a finite set of agents $\mathscr{A}$, a (numerable) set of atomic propositions $Prop = \{p, q, \dots\}$, a set of action symbols $Act = \{\alpha_i, \beta_i, \dots\}_{i \in \mathscr{A}}$, a discrete totally ordered set of instants of time $\mathscr{T} = \{t_1, t_2, \dots\}$, the modal operator $\boldsymbol{Obl}$ of obligation, and the parametrised modal operators $\boldsymbol{Obl}^i$, $Brings_i$, and $Procl_i$ (where $i \in \mathscr{A}$), and the negation sign $\neg$. We define the permission operator $\boldsymbol{Perm}^i$ as the non-derivation via defeaters of $\boldsymbol{Obl}^i\neg$.

We supplement the usual definition of literal (an atomic proposition or the negation of it), with the following clauses

1. an action symbol is a literal and so is the negation of an action symbol;

2. if l is a literal then $Procl_i l$ is a literal;

3. if l is a literal then $Brings_i l$, and $\neg Brings_i l$, are literals if l is different from $Brings_i m$, $\neg Brings_i m$, for some literal m; and

4. if l is a literal then, $\boldsymbol{Obl}^i l$ is a literal, if l is different from $\boldsymbol{Obl}^i m$, for some literal m.

We will call literals obtained from clause i) above action literals. Literals obeying conditions ii)–iv) are called modal literals. 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*, on the other hand, describes the relationship between a set of literals (premises) and a literal (conclusion), and we can specify how strong the relationship is and the mode the rule connects the antecedent and the conclusion. As usual rules allow us to derive new conclusions given a set of premises. Since rules have a mode, the conclusions will be modal literals. As far as the strength of rules is concerned we distinguish between *strict rules*, *defeasible rules* and *defeaters*; for the mode we have: count-as rules, describing the basic inference mechanism internal to an institution; deontic rules or rules for conditional obligations, determining the conditions under which an obligation holds; and rules for agency, or results-in rules, i.e., rules that encode the transitions from state to state occurring as the result of actions performed by the agents within the organisation. As we will see, the idea of conditional obligations and results-in rules is to introduce modalised conclusions. Accordingly if we have a results-in rule for p for an agent i, then this means that the rule allows for the derivation of $Brings_i p$, and if we have an applicable conditional obligation for an agent i whose conclusion is p, then the rule can be used to support the derivation of $\boldsymbol{Obl}^i p$.

Formally a *rule* r consists of its *antecedent* (or *body*) $A(r)$ ($A(r)$ may be omitted if it is the empty set) which is a finite set of temporal literals, an arrow ($\rightarrow$ for strict rules, $\Rightarrow$ for defeasible rules, and $\leadsto$ for defeaters), and its *consequent* (or *head*) $C(r)$ which is a temporal literal. 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 count-as rules, R^{O^i} to denote the set of rules for obligation for agent i, and R^i for the set of results-in rules for agent i. Given the intended use and meaning of the rules, and the constraints we imposed on the definition of literals, we have to impose some constraints on the literals that can appear in the head of a rule. Thus

1. literals modalised by $\boldsymbol{Obl}^i$ or $\boldsymbol{Perm}^i$ are not permitted in the head of count-as rules,

2. if $r \in R^{O^i}$, then modal literals whose main operator is $\boldsymbol{Obl}^i$ and $\boldsymbol{Perm}^i$ are not permitted in the head of r,

3. if $r \in R^i$, then modal literals whose main operator is $Brings_i$ are not permitted in the head of r and so are action symbols parametrised with agent i.

As we have seen in the previous sections 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: Let $\hookrightarrow$ stand for any of $\rightarrow$, $\Rightarrow$ and $\leadsto$; then with $\hookrightarrow^t_{O^i}$ we represents a transient obligation rule, $\hookrightarrow^p_{O^i}$ a persistent obligation rule, $\hookrightarrow^t_i$ a transient results-in rule, $\hookrightarrow^p_i$ a persistent results-in rule, $\hookrightarrow^t_c$ a transient basic rule, and $\hookrightarrow^p_c$ a persistent basic rule. The set of transient rules is denoted by R^t and the set of persistent rules by R^p.

Strict rules are rules in the classical sense: whenever the premises are indisputable (e.g., facts) then so is the conclusion. An example of a strict rule is "every minor is a person". Written formally:

$$minor(X) : t \rightarrow^t_c person(X) : t.$$

Defeasible rules are rules that can be defeated by contrary evidence. An example of such a rule is "every person has the capacity to perform legal acts to the extent that the law does not provide otherwise"; written formally:

$$person(X) : t \Rightarrow^p_c hasLegalCapacity(X) : t.$$

The idea is that if we know that someone is a person, then we may conclude that he/she has legal capacity, *unless there is other evidence suggesting that he/she has not.*

Defeaters are rules that cannot be used to draw any conclusions. Their only use is to prevent some conclusions. In other words, they are used to defeat some defeasible rules by producing evidence to the contrary. For example the defeater

$$WeakEvidence : t \leadsto^p_c \neg guilty : t$$

states that if pieces of evidence are assessed as weak, then they can prevent the derivation of a "guilty" verdict; on the other hand they cannot be used to support a "not guilty" conclusion.

The *superiority relation* among rules is used to define priorities among rules, that is, where one rule may override the conclusion of another rule. For example, given the defeasible rules

$$\begin{array}{rl} r: & person(X) : t \Rightarrow^p_c hasLegalCapacity(X) : t \\ r': & minor(X) : t \Rightarrow^p_c \neg hasLegalCapacity(X) : t \end{array}$$

which contradict one another, no conclusive decision can be made about whether a minor has legal capacity. But if we introduce a superiority relation $\succ$ with $r' \succ r$, then we can indeed conclude that the minor does not have legal capacity. It turns out that we only need to define the superiority relation over rules with contradictory conclusions. Also notice that a cycle in the superiority relation is counter-intuitive from the knowledge representation perspective. In the above example, it makes no sense to have both $r \succ r'$ and $r' \succ r$. Consequently, the defeasible logic we discuss requires an acyclic superiority relation.

A defeasible theory D is a structure $(F, R, \succ)$ where F is a finite set of facts, R a finite set of rules (comprising strict rules, defeasible rules and defeater), and $\succ$ a superiority relation over R.

Let X range over the modes of rules. A *conclusion* of D is a tagged literal and can have one of the following four forms:

$+\Delta_X q : t$ meaning that q is definitely provable, at time t, in D (i.e., using only facts and strict rules of mode X).

$-\Delta_X q : t$ meaning that we have proved that q is not definitely provable, at time t, in D.

$+\partial_X q : t$ meaning that q is defeasibly provable, at time t, in D.

$-\partial_X q : t$ meaning that we have proved that q is not defeasibly provable, at time t, in D.

For example, $+\partial^t_{O^i} q : t_0$ means that we a defeasible proof for $\boldsymbol{Obl}^i q$ at t_0, or, in other words, that $\boldsymbol{Obl}^i q$ holds at time t_0. However, these tags do not take care whether a conclusion $q : t$ has been obtain via transient rules (that is, q holds only at time t) or via persistent rules, in such a case for every t' such that $t < t'$, the property q persists at time t', unless we have other evidence on the contrary, i.e., a piece of evidence that terminates the property q. To reflect these issues we will introduce auxiliary proof tags that indicate whether a conclusion is persistent or transient. The proof tags are labelled with the mode used to derive the rule, according to the proof conditions given below.

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

Before introducing the proof conditions for the proof tags relevant to this paper we provide some auxiliary notions.

Given a temporal literal q and a proof $P = (P(1), \dots, P(n))$ in D we will say that q is Δ-*provable* in P, or simply Δ-provable, if there is a line $P(m)$ of the derivation such that either:

1. if $q = l : t$, then either:
   - $P(m) = +\Delta l : t$, or
   - $P(m) = +\Delta_i l : t$, or
   - $Brings_i l : t$ is Δ-provable in $P(1..m-1)$.
2. if $q = Brings_i l : t$, then either:
   - $P(m) = +\Delta_i l : t$, or
   - $P(m) = +\Delta_j Brings_i l : t$, $i \neq j$, or
   - $Brings_j Brings_i l : t$ is Δ-provable in $P(1..m-1)$, $i \neq j$.
3. if $q = \neg Brings_i l : t$, then either:
   - $P(m) = -\Delta_i l : t$, or
   - $P(m) = +\Delta_j \neg Brings_i l : t$, $i \neq j$, or
   - $Brings_j \neg Brings_i l : t$ is Δ-provable in $P(1..m-1)$, $i \neq j$.
4. if $q = \boldsymbol{Obl}^i l : t$, then either:
   - $P(m) = +\Delta_{O^i} l : t$, or
   - $P(m) = +\Delta_j \boldsymbol{Obl}^i l : t$, or
   - $Brings_i l : t$ is Δ-provable in $P(1..m-1)$.
5. if $q = \boldsymbol{Perm}^i l : t$, then either:
   - $P(m) = +\Delta_{O^i} l : t$, or
   - $P(m) = +\Delta_j \boldsymbol{Perm}^i l : t$, or
   - $Brings_j \boldsymbol{Obl}^i l : t$ is Δ-provable in $P(1..m-1)$, or
   - $Brings_j \boldsymbol{Perm}^i l : t$ is Δ-provable in $P(1..m-1)$.

The definition of ∂-provable has the same first four clauses where Δ is replaced with ∂; however, the last clause, the clause for $\boldsymbol{Perm}^i$, has the following additional condition

$$P(m) = +\pi_{O^i} l : t$$

As we will see π is a special proof tag designed explicitly to handle this case.

In a similar way we can define a literal to be Δ- and ∂-rejected by taking, respectively, the definition of Δ-provable and ∂-provable and changing all positive proof tags into negative proof tags, adding a negation in front of the literal when the literal is prefixed by a modal operator $Brings_j$, and replacing all the *or*s by *and*s. Thus, for example, we can say that a literal $Brings_i l$ is ∂-rejected if, in a derivation, we have a line $-\partial_i l$, and the literal $\neg Brings_i \neg l$ is ∂-rejected if we have $+\partial_i \neg l$ and so on.

Let X be a modal operator and # is either Δ or ∂. A literal l is $\#_X$-*provable* if the modal literal Xl is #-provable; l is $\#_X$-rejected if the literal Xl is #-rejected.

Let X range over the set $\{c, O^i, i\}$. Given a strict rule r we will say that the rule is Δ_X-*applicable* iff

1. $r \in R^X$ and $\forall a_k : t_k \in A(r)$, $a_k : t_k$ is Δ-provable; or
2. if $X \neq c$ and $r \in R^c_s$, i.e., r is a count-as rule, then $\forall a_k : t_k$, $a_k : t_k$ is Δ_X-provable.

The conditions for a rule r to be ∂_X-*applicable* are the same as those for Δ_X-applicable, but where we replace Δ with ∂.

Let X range over the set $\{c, O^i, i\}$. Given a strict rule r we will say that the rule is Δ_X-*discarded* iff

1. $r \in R^X$ and $\exists a_k : t_k \in A(r)$, $a_k : t_k$ is Δ-rejected; or
2. if $X \neq c$ and $r \in R^c_s$, i.e., r is a count-as rule, then $\exists a_k : t_k$, $a_k : t_k$ is Δ_X-rejected.

The conditions for a rule r to be ∂_X-*discarded* are the same as those for Δ_X-discarded, but where we replace Δ with ∂.

We are now ready to define the proof theory of defeasible logic, that is, the inference conditions to derive tagged literals from a given theory D.

We begin with the proof conditions to determine whether a literal is a definite transient conclusion of a theory D.

$+\Delta^t_X$: If $P(n+1) = +\Delta^t_X q : t$, then
1) $q : t \in F$, or
2) $\exists r \in R^t_s[q : t]$: r is Δ_X-applicable.

The above condition is the normal condition for definite (positive) proofs in defeasible logic, that is, monotonic derivations using forward chaining, or modus ponens. The only thing to notice is the case of $+\Delta^t_X$ were X is a modal operator. In this case a strict count-as rule can be understood as a modal rule, either as a deontic rule or a results-in rule, if all the literals in the body of the rule are modalised with the appropriate operator. For example, given the count-as rule

$$A, B \rightarrow^t_c C$$

we can derive $+\Delta_{O^i} C$ if both $\boldsymbol{Obl}^i A$ and $\boldsymbol{Obl}^i B$ are Δ-provable. However, if the rule were

$$A, \boldsymbol{Obl}^i B \rightarrow^t_c C$$

then $+\Delta_{O^i} C$ would not provable. The rule would not Δ_{O^i}-applicable since, in the body of the rule, we have $\boldsymbol{Obl}^i B$, and this requires that $\boldsymbol{Obl}^i \boldsymbol{Obl}^i B$ instead of $\boldsymbol{Obl}^i B$ is Δ_{O^i}-provable.

$-\Delta_X^t$: If $P(n+1) = -\Delta_X^t q : t$, then
1) $q : t \notin F$, and
2) $\forall r \in R_s^t[q : t]$: r is Δ_X-discarded.

To prove that a definite (transient) conclusion is not possible we have to show that all attempts to give a definite proof of the conclusion fail.

We can now move to persistent definite conclusions.

$+\Delta_X^p$: If $P(n+1) = +\Delta_X^p q : t$, then
1) $q \in F$; or
2) $\exists r \in R_s^p[q : t]$: r is Δ_X-applicable; or
3) $\exists t' \in \mathscr{T}$: $t' < t$ and $+\Delta_X^p q : t' \in P(1..n)$.

The first two clauses are the same as the corresponding clauses in the definition of definite transient conclusion and what the condition adds is the persistence condition (3). The persistence condition allows us to reiterate literals definitely proved at previous times, or, using the terminology of Section 2, that a persistent property q initiated at a time before the current time. For example, given the the theory

$$(F = \{DropGlass_{Tom} : t_0\},$$
$$R = \{DropGlass_{Tom} : t \rightarrow_{Tom} BrokenGlass : t\},$$
$$\succ = \emptyset)$$

we can derive $+\Delta_i^p BrokenGlass : t_0$, and $+\Delta_i^p BrokenGlass : t_1$, $t_0 < t_1$. Thus, if Tom drops a glass breaking it at time t_0, then the glass remains broken afterwards; thus it is broken at time t_1.

$-\Delta^p$: If $P(n+1) = -\Delta_X^p q : t$, then
1) $q \notin F$; and
2) $\forall r \in R_s^p[q : t]$: r is Δ_X-discarded; and
3) $\forall t' \in \mathscr{T}$: $t' < t$ and $-\Delta_X^p q : t' \in P(1..n)$.

In addition to the conditions we have for the transient case, here we have to check that for all instants of time before now the persistent property has not been proved.

According to the above conditions to prove that q is a definite conclusion of D at time t we have to consider whether q is a definite transient conclusion of D at time t or if it is a definite persistent conclusion at t. Thus

$+\Delta_X$: If $P(n+1) = +\Delta_X q : t$ then
1) $+\Delta_X^t q : t \in P(1..n)$ or
2) $+\Delta_X^p q : t \in P(1..n)$.

A definite conclusion is not provable at time t if it is not possible to prove it persistently nor transiently.

$-\Delta_X$: If $P(n+1) = -\Delta_X q : t$ then
1) $-\Delta_X^t q : t \in P(1..n)$ and
2) $-\Delta_X^p q : t \in P(1..n)$.

Defeasible derivations have an argumentation like structure divided in three phases. In the first phase we put forward a supported reason (rule) for the conclusion we want to prove. Then in the second phase we consider all possible (actual and not) reasons against the desired conclusion. Finally in the last phase, we have to rebut all the counterarguments. This can be done in two ways: we can show that some of the premises of a counterargument do not obtain, or we can show that the argument is weaker than an argument in favour of the conclusion. This is formalised by the following (constructive) proof conditions.

Again we start with the conditions for transient defeasible conclusions.

$+\partial_X^t$: If $P(n+1) = +\partial_X^t q : t$ then
1) $+\Delta_X q : t \in P(1..n)$, or
2) $-\Delta_X \sim q : t \in P(1..n)$ and
2.1) $\exists r \in R_{sd}[q : t]$: r is ∂_X-applicable and
2.2) $\forall s \in R[\sim q : t]$ either s is ∂_X-discarded or
$\exists w \in R[q : t]$: w is ∂_X-applicable and $w \succ s$.

$-\partial_X^t$: If $P(n+1) = -\partial_X^t q : t$ then
1) $-\Delta_X^t q : t \in P(1..n)$, and
2) $+\Delta_X \sim q : t \in P(1..n)$ or
2.1) $\forall r \in R_{sd}^t[q : t]$: either r is ∂_X-discarded or
2.2) $\exists s \in R[\sim q : t]$: s is ∂_X-applicable and
$\forall w \in R[q : t]$: w is either ∂_X-discarded or $w \not\succ s$.

The above conditions are, essentially, the usual conditions for defeasible derivations in defeasible logic, we refer the reader to [18, 1, 10] for more thorough treatments. The only points we want to highlight here are:

- clause 2 requires that the complement of the literal we want to prove is not definitely provable (or definitely provable for $-\partial$), but it does not specify whether it is persistent or transient: remember that what we want to achieve is to see whether the literal or its complement are provable at t but not both; in the same way, and for the same reason, q can be attacked by any compatible rule for the complement of q.
- count-as rules, as in the case of definite derivations, can play the role of deontic rules and results-in rules when all the literals in the body are ∂_X-derivable, and iii) that current derivations take precedences over persistent literals. We will return on this issue after we have introduced the conditions for persistent derivations.

The inference conditions for persistent defeasible proofs are as follows.

$+\partial_X^p$: If $P(n+1) = +\partial_X^p q : t$ then
1) $+\Delta_X^p q : t \in P(1..n)$, or
2) $-\Delta_X \sim q : t \in P(1..n)$, and
2.1) $\exists r \in R_{sd}^p[q : t]$: r is ∂_X-applicable, and
2.2) $\forall s \in R[\sim q : t]$: either s is ∂_X-discarded or
$\exists w \in R[q : t]$: w is ∂_X-applicable and $w \succ s$; or
3) $\exists t' \in \mathscr{T}$: $t' < t$ and $+\partial_X^p q : t' \in P(1..m)$ and
3.1) $\forall s \in R[\sim q : t'']$: $t' < t'' \leq t$, s is ∂_X-discarded, or
$\exists w \in R[q : t'']$: w is ∂_X-applicable and $w \succ s$

$-\partial_X^p$: If $P(n+1) = -\partial_X^p q : t$ then
1) $-\Delta_X^p q : t$, and
2) $+\Delta_X \sim q : t$, or
2.1) $\forall r \in R^p[q : t]$: either r is δ_X-discarded or
2.2) $\exists s \in R[\sim q : t]$: r is ∂_X-applicable and
$\forall w \in R[q : t]$ either w is ∂_X-discarded or $w \not\succ t$; and
3) $\forall t' \in \mathscr{T}$: $t' < t$, if $+\partial_X^p q : t' \in P(1..m)$, then
3.1) $\exists s \in R[\sim q : t'']$: $t' < t'' \leq t$, s is ∂_X-applicable, and
$\forall w \in R[q : t'']$: either w is ∂_X-discarded or $w \not\succ s$.

Clauses 1 and 2 of the above proof conditions are the same as the corresponding clause for transient defeasible derivations, the only difference is in clause 3. In the same way we have a persistence clause in definite derivation we have a persistence clause to defeasibly prove/disprove a literal. Thus to show that a literal holds defeasibly at t we can check that the literal has been proved at a time t', $t' < t$, and that for every instant of time between the two the property has not been terminated. This amounts to show that all possible termination events were not triggered or they are weaker than some reasons in favour of the persistence of the property.

Let us illustrate how the above conditions work with the help of the following theory.

$$
\begin{aligned}
(F = \{&A : t_0,\, B : t_2,\, C : t_2,\, D : t_3\},\\
R = \{&r_1 :\ A : t \Rightarrow^p_c E : t,\\
&r_2 :\ B : t \Rightarrow^p_c \neg E : t,\\
&r_3 :\ C : t \rightsquigarrow^p_c E : t,\\
&r_4 :\ D : t \Rightarrow^t_c \neg E : t\},\\
\succ = \{&r_3 \succ r_2,\, r_1 \succ r_4\})
\end{aligned}
$$

At time t_0, r_1 is the only applicable rule; accordingly we derive $+\partial^p_c E : t_0$, or, using the terminology of Section 2, E initiates at t_0. At time t_1 no rule is applicable, and the only derivation permitted is the derivation of $+\partial^p_c E : t_1$ using the persistence condition; so E holds at t_1. At time t_2 both r_2 and r_3 are applicable, but r_4 is not. If r_2 prevailed, then it would terminates E. However, it is rebutted by r_3, thus, in this case we derive $+\partial^p_c E : t_2$. Finally at time t_3, rule r_4 is applicable, thus we derive $+\partial^t_c \neg E$ and $-\partial^p_c E : t_3$, which means that r_4 terminates E. Notice that, even if r_4 is weaker than r_1, the rule that has initiated E at t_0, the latter is not applicable at t_3, thus it does not offer any support to maintain E.

In the definition of a literal to be ∂-provable, we stated that $\boldsymbol{Perm}^i l$ is provable if there is a line of a derivation containing $+\pi_{O^i} l$. Here we give the proof conditions for such a proof tag.

$+\pi^t_{O^i}$: If $P(1..n) = +\pi^t_{O^i} q : t$, then
1) $-\Delta{\sim}p \in P(1..n)$; and
2) $\exists r \in R^t_{dft}[q : t]$: r is ∂_{O^i}-applicable, and
3) $\forall s \in R[{\sim}q : t]$: either s is ∂_{O^i}-discarded or $r \succ s$.

$+\pi^p_{O^i}$: If $P(1..n) = +\pi^p_{O^i} q : t$, then
1) $-\Delta{\sim}p \in P(1..n)$; and
2) $\exists r \in R^p_{dft}[q : t]$: r is ∂_{O^i}-applicable, and
$\forall s \in R[{\sim}q : t]$: either s is ∂_{O^i}-discarded or $r \succ s$; or
3) $\exists t' \mathscr{T}$: $t' < t$, $+\pi^p_{O^i} q : t' \in P(1..n)$, and
$\forall s \in R[{\sim}q : t'']$: $t' < t'' \leq t$ either s is ∂_{O^i}-discarded or $r \succ s$.

While it is possible to define $-\pi$, we will refrain from it since it is not clear how we can make use of it.

In standard deontic logic ***Perm*** is defined as the dual of ***Obl***. Something similar to this has been also adopted here. Given this assumption, it is redundant to introduce in our system an additional type of rule devised to capture the idea of permission. However, since we do not have rules that embed in the consequent obligations, it could seem hard to represent permissions, which correspond to $\neg\boldsymbol{Obl}\neg$. But, as it is possible to see in the proof conditions, this is not a real problem: It suffices to understand a literal p to be permitted if there is a defeater with the head p such that this defeater overrides all obligation rules that allow to infer $\neg p$.

To derive a permission, let us say $\boldsymbol{Perm}^i p$ we have to show that the derivation of the corresponding prohibition ($\boldsymbol{Obl}^i \neg p$) fails. In the logic we have developed, obligation rules are meant to introduce a positive modal conclusion, thus it is not possible to prove directly $\neg\boldsymbol{Obl}^i$. There could be two reasons why $\boldsymbol{Obl}^i \neg p$ fails to be provable: the theory does not have enough resources to prove it, i.e., there are no applicable obligation rules for p, or the obligation rules for $\neg p$ are all defeated by obligation rules for p. As we have seen, only strict and defeasible rules can be used to support a conclusion, but any rule can be used to prevent the derivation of a conclusion. Thus if we do not want to obtain the stronger conclusion that p is permitted because it is obligatory we have to use a defeater to defeat a rule for $\neg p$. For example, suppose there is a norm that prohibits to U-turn at traffic lights unless there is a "U-turn permitted" sign. This scenario can be represented as follows:

$$
\begin{aligned}
r_1 :&\ \Rightarrow^t_{O^i} \neg Uturn\\
r_2 :&\ UturnSign \rightsquigarrow^t_{O^i} Uturn.
\end{aligned}
$$

where $r_2 \succ r_1$. If both rules are applicable then we have $-\partial_{O^i} \neg Uturn$ and $+\pi_{O^i} Uturn$, from which we obtain $\boldsymbol{Perm}^i Uturn$, i.e., U-turn is permitted. If r_2 were a defeasible rule, then we would get $+\partial_{O^i} Uturn$, which means $\boldsymbol{Obl}^i Uturn$, that is, we have the obligation to U-turn.

3.3 Example

Let us illustrate our system with the help of a toy, but more concrete, example. Suppose the agent i wants to buy on-line a software from company X. To do this, i should contact X by using the on-line service provided by X. Any agent has to process her purchase order within 8 time instants since she connected to the on-line service; otherwise a timeout feature will automatically log the agent off. As soon as the purchase order is sent to X and processed, the system will log the agent off. If i sends a purchase order which matches X's conditions, and provided that X made an on-line offer for the software, this counts as i's proclamation that the software is purchased. Once that the act of purchasing is performed some obligations follow:

- X is obliged towards i to deliver an invoice; this obligation will not longer hold when the invoice is sent;
- after receiving the invoice, i in turn will have to pay X the price.

The scenario is represented as follows (bold type expressions denote action symbols, the italicised ones state of affairs; to save space, *DInvoice* abbreviates "Deliver the Invoice"):

$$
\begin{aligned}
F = \{&\mathbf{Advertising_X} : t_0, \mathbf{ProvideService_X} : t_0, \mathbf{ConnectService_i} : t_1,\\
&\mathbf{SendOrder_i} : t_2, Brings_X DInvoice : t_4, Brings_i Pay : t_7\}\\
R = \{&r_1 : \mathbf{ConnectService_i} : t_x \Rightarrow^p_i UseService : t_x,\\
&r_2 : TimeOut : t_{x+8} \Rightarrow^p_X \neg UseService : t_{x+8},\\
&r_3 : Brings_i Purchase : t_y \Rightarrow^p_X \neg UseService : t_{y+1},\\
&r_4 : Brings_i UseService : t_z, \mathbf{SendOrder_i} : t_z \Rightarrow^t_c Procl_i Purchase : t_z,\\
&r_5 : \mathbf{Advertising_X} : t_k, \mathbf{ProvideService_X} : t_k \Rightarrow^p_c Brings_X Offer : t_k,\\
&r_6 : Brings_X Offer : t_z, Procl_i Purchase : t_z \Rightarrow^t_c Brings_i Purchase : t_z,\\
&r_7 : Brings_i Purchase : t_z \Rightarrow^p_{O^i} Brings_X DInvoice : t_{z+1},\\
&r_8 : Brings_X DInvoice : t_h \rightsquigarrow^p_{O^i} \neg Brings_X DInvoice : t_{h+1},\\
&r_9 : Brings_i Purchase : t_q, Brings_X DInvoice : t_q \Rightarrow^p_{O^X} Brings_i Pay : t_{q+1},\\
&r_{10} : Brings_i Pay : t_v \rightsquigarrow^p_{O^X} \neg Brings_i Pay : t_{v+1}\}\\
\succ = \{&r_{10} \succ r_9,\, r_8 \succ r_7,\, r_3 \succ r_1,\, r_2 \succ r_1\}
\end{aligned}
$$

Let us comment this scenario by focusing on the relevant time instants. At t_0 we have $\mathbf{Advertising_X}$, $\mathbf{ProvideService_X}$ and these count as $Brings_X Offer$. At t_1 i connects to on-line service provided by X and so we get $Brings_i UseService$. At t_2 agent i sends her purchase order to X: again, we have $Brings_i UseService$, but we also get $Procl_i Purchase$; since $Brings_X Offer$ still holds, we obtain via r_6 $Brings_i Purchase$. At t_3 we will infer $Brings_X \neg UseService$; since the agent has already made the purchase, X is obliged to deliver the invoice. At t_4 this obligation holds, but we also have $Brings_X DInvoice$. Thus, at point t_5, the obligation to deliver the invoice will no longer apply to X. At t_6 we get $\boldsymbol{Obl}^X Brings_i Pay$. At t_7, i pays X; so, by virtue of r_{10}, at any subsequent instant of t_7, i is no longer obliged to pay.

4. CONCLUSIONS

Nute [19] proposed a Deontic Defeasible Logic which, in some respect, is similar to the framework presented here. Beside minor differences in the way rules are handled at the propositional level, the main difference is that Nute develops a system able to deal with non-temporalised and non-directed obligations. To do this, he uses one type of rule and adds explicit operators in the head of rules. Traditionally, in proof-theory, rules to introduce operators give the meaning of them. Thus using one and the same type of rule for obligation and factual conclusion does not show the real meaning of the operators involved. Moreover, it is not clear to us whether and how conversions, such as those that permit to obtain $+\Delta_{O^i}C$ from $A, B \rightarrow^t_c C$ and the fact that $\boldsymbol{Obl}^i A$ and $\boldsymbol{Obl}^i B$ are Δ-provable, can be dealt with only using a single rule type.

This paper originates from two lines of research. The first concerns the logical treatment of directed obligations. In this regard, we started, though with some differences, from the analysis of [14]. As regards other aspects such as the research into legal positions, and particularly into the notion of power, we are particularly indebted to the seminal work [13]. The second line of research concerns the temporal and dynamic treatment of obligations. In this regard, a large piece of recent literature is available. We are particularly indebted to the works by Sergot and colleagues on the use of Event Calculus, such as [3].

Let us see shortly some issues for future research. In this paper violations are viewed as neutral with regard to obligations. If an obligation holds at a certain instant, the violation will not be considered and such an obligation will simply continue to hold until it is complied with. In [11, 8] we showed how the non-classical (substructural) connective $\otimes$, whose interpretation is such that $p \otimes q$ is read as "q is the reparation of the violation of p", can be embedded within DL. This work has to be extended to cover temporalised normative concepts. [16] has proved, for the propositional case, that the set of tagged literals can be derived from the theory in linear time in the number of rules in it. We expect not to be hard to extend this result to the modal case. The distinction of different kinds of rules does not seem to affect the complexity of the theory. The case for basic rules is the same adopted in standard DL while, for the other components, we convert relevant rules into the appropriate "extended" modal literals. At this point, the inference mechanism is the same as the standard one.

5. REFERENCES

[1] G. Antoniou, D. Billington, G. Governatori, and M. Maher. Representation results for defeasible logic. *ACM Transactions on Computational Logic*, 2(2):255–287, 2001.

[2] G. Antoniou, D. Billington, G. Governatori, M. Maher, and A. Rock. A family of defeasible reasoning logics and its implementation. In W. Horn, editor, *Proc. ECAI 2000*, Amsterdam, 2000. IOS.

[3] A. Artikis, J. Pitt, and M. Sergot. Animated specifications of computational societies. In *Proc. AAMAS'02*. ACM, New York, 2002.

[4] N. Bassiliades, G. Antoniou, and I. Vlahavas. DR-DEVICE: A defeasible logic system for the Semantic Web. In H.J. Ohlbach and S. Schaffert, editors, *2nd Workshop on Principles and Practice of Semantic Web Reasoning*. Springer, 2004.

[5] R. Conte and C. Dellarocas. *Social Order in Multiagent Systems*. Kluwer, Boston, 2001.

[6] D. Elgesem. The modal logic of agency. *Nordic Journal of Philosophical Logic*, 2:1–48, 1997.

[7] J. Gelati, G. Governatori, A. Rotolo, and G. Sartor. Normative autonomy and normative co-ordination: Declarative power, representation, and mandate. *AI and Law*, 12(1-2):53–81, 2004.

[8] G. Governatori. Representing Business Contracts in RuleML. *Int. Journal of Coop. Inf. Syst.*, 14(2-3), 2005.

[9] G. Governatori, M. Maher, D. Billington, and G. Antoniou. Argumentation semantics for defeasible logics. *Journal of Logic and Computation*, 14(5):675–702, 2004.

[10] G. Governatori and A. Rotolo. Defeasible logic: Agency, intention and obligation. In A. Lomuscio and D. Nute, editors, *Proc. DEON 2004*, Berlin, 2004. Springer.

[11] G. Governatori and A. Rotolo. Modelling contracts using RuleML. In T. Gordon, editor, *Proc. Jurix 2004*, Amsterdam, 2004. IOS.

[12] J.Y. Halpern and Y. Moses. A guide to completeness and complexity for modal logics of knowledge and belief. *Artificial Intelligence*, 54(3):319–380, 1992.

[13] A.J.I. Jones and M. Sergot. A formal characterisation of institutionalised power. *Journal of IGPL*, 3:427–443, 1996.

[14] Ch. Krogh and H. Herrestad. Getting personal: Some notes on the relationship between personal and impersonal obligation. In M. Brown and J. Carmo, editors, *Proc. DEON 1996*, Berlin, 1996. Springer.

[15] A. Lomuscio and D. Nute, editors. *Proc. DEON 2004*, Berlin, 2004. Springer.

[16] M. Maher. Propositional defeasible logic has linear complexity. *Theory and Practice of Logic Programming*, (6), 2001.

[17] M.J. Maher, A. Rock, G. Antoniou, D. Billignton, and T. Miller. Efficient defeasible reasoning systems. *International Journal of Artificial Intelligence Tools*, 10(4):483–501, 2001.

[18] D. Nute. Defeasible logic. In *Handbook of Logic in Artificial Intelligence and Logic Programming*, volume 3. OUP, 1987.

[19] D. Nute. Norms, priorities, and defeasible logic. In P. McNamara and H. Prakken, editors, *Proc. DEON 1998*. IOS, Amsterdam, 1998.

[20] J. Pitt, editor. *Open Agent Societies: Normative Specifications in MAS*. Wiley, Chichester, 2003.

[21] H. Prakken and G. Sartor. A dialectical model of assessing conflicting arguments in legal reasoning. *AI and Law*, 4:331–368, 1996.

[22] F. Santos, A.J.I. Jones, and J. Carmo. Action concepts for describing organised interaction. In *Hawaii International Conference on System Sciences*. IEEE, Los Alamitos, 1997.

[23] G. Sartor. *Legal Reasoning*. Springer, Dordrecht, 2005.

[24] J.R. Searle. *The Construction of Social Reality*. Penguin, Harmondsworth, 1995.

[25] G. Simari and R. Loui. A mathematical treatment of argumentation and its implementation. *Artificial Intelligence*, 53:125–157, 1992.