Possible World Semantics for Defeasible Deontic Logic

Guido Governatori[1], Antonino Rotolo[2] and Erica Calardo[2]

[1] NICTA, Software Systems Research Group, Australia
[2] CIRSFID, University of Bologna, Italy

Abstract. Defeasible Deontic Logic is a simple and computationally efficient approach for the representation of normative reasoning. Traditionally defeasible logics are defined proof theoretically based on the proof conditions for the logic. While several logic programming, operational and argumentation semantics have been provided for defeasible logics, possible world semantics for (modal) defeasible logics remained elusive. In this paper we address this issue.

1 Introduction

Defeasible Logic (DL) [24, 3] is historically the first of a family of approaches based on the idea of logic programming without negation as failure. DL is a simple, efficient but flexible (skeptical) non-monotonic formalism capable of dealing with many intuitions of non-monotonic reasoning. The logic was designed to be easily implementable right from the beginning, unlike most other approaches, and has a linear complexity [22]. Recent implementations include DR-Prolog [1] and DR-DEVICE [5].

DL proved to be modular and flexible. In particular, propositional DL has been extended in various directions to study several aspects of normative and deontic reasoning [14–16]. A significant extension of DL was to embed in the logic different types of modal operators (capturing notions such as directed and undirected deontic statements, actions, counts-as, beliefs, and intentions) [14, 15]. The result was a number of logics having still linear complexity and being able, e.g., to model the deliberation of cognitive agents and their interplay with normative systems. Some implementations have been recently developed for modal extensions [18, 20].

An open research problem in this approach is how to semantically interpret the modal operators of the logic. Indeed, so far the main concerns were proof theory and the development of efficient computational methods to calculate the extension of any logical theory. Available semantic approaches to DL are, for example, the argumentation semantics proposed in [11]. However, this approach does not look much promising if the purpose is to characterize the modal operators added to DL: [11]'s argumentation semantics for DL simply provides a different, argument-based, and more intuitive representation of DL proof theory but does not add anything conceptually new to it. Hence, the research task we address in this paper is far from obvious, since it has to do with establishing significant connections between non-monotonic and modal logics. This work is a first and preliminary attempt in this direction as far as DL is concerned. We will show how to interpret any multi-modal extension of DL in neighbourhood semantics.

The layout of the paper is as follows. Section 2 provides an informal presentation of Defeasible Deontic Logic. Section 3 presents a general multi-modal logical framework

(Defeasible Multi-modal Logic), based on DL, which covers all existing variants of Defeasible Deontic Logic. Section 4 discusses how to interpret Defeasible Multi-modal Logic in neighbourhood semantics and identifies one open problem.

2 Defeasible Deontic Logic: An Informal Presentation

DL has three basic kinds of features: facts, rules, and a superiority relation among rules. Facts are indisputable statements. Rules are usually of three types: *strict rules*, marked by the arrow $\rightarrow$, correspond to the monotonic part of the logic and support indisputable conclusions whenever their antecedents, too, are indisputable[3]. *Defeasible rules*, marked by $\Rightarrow$, can be defeated by contrary evidence. *Defeaters*, marked by $\rightsquigarrow$, cannot lead to any conclusion but are used to defeat some defeasible rules by producing evidence to the contrary. The superiority relation ($>$) provides information about the relative strength of rules, i.e., about which rules can overrule which other rules.

Defeasible Deontic Logic is a family of logics that extend DL by adding deontic and other modal operators. The purpose is to study the interplay between deontic concepts (such as obligations and permission) and other modal components such as counts-as concepts and agents' actions [15], or agents' beliefs and intentions [14]. The resulting extended language is based on a family of different rules, where each type is labeled by a different modal operator $\Box_i$: the idea is that each rule, if parametrized by $\Box_i$, it is meant to govern the derivation of formulas modalized with $\Box_i$.

The approach we have elsewhere developed in Defeasible Deontic Logic is thus twofold. First, we take a constructive interpretation of any modal operator $\Box_i$: if we can build a derivation of p using rules for $\Box_i$, then we also have a derivation of $\Box_i p$. Second, derivability in classical logic is replaced with a practical and feasible notion like derivability in DL. Thus the intuition is that we are allowed to derive $\Box_i p$ if we can prove p with the mode $\Box_i$ in DL. For example, a rule like $p_1, \ldots, p_n \Rightarrow_{\mathrm{OBL}} q$ means that, if $p_1, \ldots, p_n$ are the case or proved, then the logical machinery allows us to derive q with mode OBL, and so OBLq. In general, for any $\Box_i$

$$\frac{\Gamma \quad \Gamma \Rightarrow_{\Box_i} q}{\Gamma \mid\!\sim \Box_i q}$$

Defeasible Deontic Logic defines some interaction patterns between modalities: in particular, one permits to use rules for a modality $\Box_i$ as they were for another modality $\Box_j$ (*rule conversions*), and one considers conflicts between rules.

Rule Conversions The notion of *rule conversion* allows us to model peculiar interactions between different modal operators (for an extensive conceptual discussion, see [14]). To give an example, suppose we have that $a \Rightarrow_{\mathrm{BEL}} b$ and that we derive a using a rule labeled by INT. Can we conclude INTb? If the answer is positive, conversions can be represented as follows:

$$\frac{\Gamma \mid\!\sim \mathrm{INT}\psi \quad \psi \Rightarrow_{\mathrm{BEL}} \phi}{\Gamma, \mathrm{INT}\psi \mid\!\sim \mathrm{INT}\phi} \text{ Conversion}$$

[3] For the sake of simplicity, we will not consider those rules in the logics discussed in this paper.

In many cases this is a reasonable conclusion to obtain. Indeed, if an agent believes to visit Italy if she visits Rome, and she has the intention to visit Rome, then it seems rational that she has the intention to visit Italy. When such a conversion is allowed in the logic, we will write that $\mathsf{Convert}(\mathrm{BEL},\mathrm{INT})$. A similar notation applies to any other pair of modalities for which we want to accept conversions.

Conflicts DL is a skeptical non-monotonic logic and thus is able to handle conflicts. Defeasible Deontic Logic behaves in the same way. In a multi-modal setting, we can establish what modalities can be incompatible with each other, and, also, we can impose various forms of consistency [14].

The consistency between modalities require to define incompatibility relations between them as well as specific methods to solve conflicts between the corresponding types of rule. Many complex conflict patterns can be identified [14]. For the purpose of this paper, we will introduce a binary and asymmetric relation $\mathsf{Conflict}$ over the set of modalities that defines which types of rules are in conflict and which rule types prevail. For example, if we have $\mathsf{Conflict}(\mathrm{OBL},\mathrm{INT})$, this means that any rule of the form $p_1,\ldots,p_n \Rightarrow_{\mathrm{OBL}} q$ is potentially in conflict with any rule of the form $d_1,\ldots,d_n \Rightarrow_{\mathrm{INT}} \neg q$ and that, in case of an actual conflict (i.e., when both rules fire), the obligation prevails over the intention.

3 Defeasible Multi-modal Logic

In this section we present a general multi-modal logical framework, called Defeasible Multi-modal Logic, which covers all existing variants of Defeasible Deontic Logic. Hence, we abstract from any specific interpretation of the modal operators and assume to work with a language based on any arbitrary number set of modal operators. The limitation is that each modal operator can logically behave as only one of those introduced in [14, 15].

3.1 The Language

A defeasible theory consists of a set of *facts* or indisputable (non modal) statements, n sets of rules for the modalities $\Box_1,\ldots,\Box_n$, a set of *conversions* saying when a rule of one type can be used also as another type, a set of *conflict relations* saying when two rule types can be in conflict and which rule type prevails, and a *superiority relation* $>$ among rules saying when a single rule may override the conclusion of another rule. For any $\Box_i$, $1 \leq i \leq n$, $a_1,\ldots,a_n \Rightarrow_{\Box_i} b$ is a *defeasible rule* that can be defeated by contrary evidence; $a_1,\ldots,a_n \rightsquigarrow_{\Box_i} b$ is a *defeater* that is used to defeat some defeasible rules by producing evidence to the contrary. It is worth noting that modalised literals can occur only in the antecedent of rules: the reason of this is that the rules are precisely used to derive modalised conclusions.[4]

The language of Defeasible Multi-modal Logic is only built from a set of propositional constants, which are denoted using lowercase letters such as $a,b,c,\ldots p,q,s$.

[4] Clearly, this is a simplification aimed at keeping the logic manageable. For a version of Defeasible Multi-modal Logic that admits iterations, see, e.g., [15].

Propositional variables are rather denoted in this paper using uppercase letters such as $A,B,C,\dots P,Q,S$.

Definition 1 (Language). *Let* PROP *be a set of propositional atoms (propositional constants),* $\mathrm{MOD} = \{\Box_1,\dots,\Box_n\}$ *be the set of modal operators, and* Lab *be a set of labels. The sets below are the smallest sets closed under the following rules:*

Literals

$$\mathrm{Lit} = \mathrm{PROP} \cup \{\neg p | p \in \mathrm{PROP}\}$$

If q is a literal, $\sim q$ denotes the complementary literal (if q is a positive literal p then $\sim q$ is $\neg p$; and if q is $\neg p$, then $\sim q$ is p);

Modal literals

$$\mathrm{ModLit} = \{\Box_i l, \neg\Box_i l | l \in \mathrm{Lit}, \Box_i \in \mathrm{MOD}\};$$

Rules $\mathrm{Rule} = \mathrm{Rule}_d \cup \mathrm{Rule}_{dft}$, *where for* $\Box_i \in \mathrm{MOD}$

$$\mathrm{Rule}_d = \{r : a_1,\dots,a_n \Rightarrow_{\Box_i} b | r \in \mathrm{Lab}, A(r) \subseteq \mathrm{Lit} \cup \mathrm{ModLit}, b \in \mathrm{Lit}\}$$
$$\mathrm{Rule}_{dft} = \{r : a_1,\dots,a_n \rightsquigarrow_{\Box_i} b | r \in \mathrm{Lab}, A(r) \subseteq \mathrm{Lit} \cup \mathrm{ModLit}, b \in \mathrm{Lit}\}$$

where $A(r)$ denotes the set $\{a_1,\dots,a_n\}$ of antecedents *of the rule r and $C(r)$ its* consequent *b. We use some abbreviations, such as superscript for modal operators, subscript for rule type, and* $\mathrm{Rule}[b]$ *for rules whose consequent is b, for example:*

$$\mathrm{Rule}^i = \{r : a_1,\dots,a_n \rhd_{\Box_i} b | (r : a_1,\dots,a_n \rhd_{\Box_i} b) \in \mathrm{Rule}, \rhd \in \{\Rightarrow, \rightsquigarrow\}\}$$
$$\mathrm{Rule}_d[b] = \{r \in \mathrm{Rule}_d | A(r) = b\}$$

Definition 2 (Conversion and Conflict Relations; Reflexive Predicate). *Let* $\mathsf{Convert} \subseteq \mathrm{MOD} \times \mathrm{MOD}$ *denote the conversion relation.*

The conflict relation $\mathsf{Conflict} \subseteq \mathrm{MOD} \times \mathrm{MOD}$ *is such that*

$$\forall \Box_i, \Box_j \in \mathrm{MOD},\ \mathsf{Conflict}(\Box_i,\Box_j) \Rightarrow \neg(\mathsf{Conflict}(\Box_j,\Box_i)) \quad \textit{(asymmetry)}$$

With $\mathsf{Reflexive}(\Box_i)$ *we will qualify modal operators with the property that they can derive a for $\Box_i a$. When clear from the context, in using* Convert, Conflict *and* Reflexive *we will denote the modalities by only referring to their identifying subscripts.*

3.2 Proof Theory

Definition 3. *A* Defeasible Theory *is a structure*

$$(F,R,>)$$

where F is a set of literals, R is a set of defeasible rules and defaters, and the superiority relation $>$ is such that $> = >^{sm} \cup >^{\mathrm{Conflict}}$, where $\forall \Box_i, \Box_j \in \{\Box_1,\dots,\Box_n\}$, $>^{sm} \subseteq R^i \times R^i$ such that if $r > s$, then if $r \in \mathrm{Rule}^i[p]$ then $s \in \mathrm{Rule}^i[\sim p]$ and $>$ is acyclic; and $>^{\mathrm{Conflict}}$ is such that

$$\forall r \in \mathrm{Rule}^i[p], \forall s \in \mathrm{Rule}^j[\sim p],\ \textit{if}\ \mathsf{Conflict}(i,j),\ \textit{then}\ r >^{\mathrm{Conflict}} s$$

Definition 4. *A* conclusion *of a defeasible theory D is a tagged literal that can have one of the following forms:*

$+\partial q$ *meaning that q is defeasibly provable in D with 'factual' mode;*
$+\partial_{\Box_i} q$ *meaning that q is defeasibly provable in D with mode $\Box_i$;*
$-\partial q$ *meaning that q is defeasibly refutable in D with 'factual' mode;*
$-\partial_{\Box_i} q$ *meaning that q is defeasibly refutable in D with mode $\Box_i$.*

The intuition is that if we derive $+\partial p$ (with 'factual' mode), then p holds, while when we prove $+\partial_{\Box_i} p$, this means that $\Box_i p$ holds.

Definition 5. *A* derivation *P is a sequence $P(1),\dots,P(n)$ of tagged modal literals satisfying the proof conditions below (for $0 \leq m < n$)*[5]

If $P(m+1) = +\partial_{\Box_i} l$, then
(1) $\sim l \notin F$, $\mathsf{Reflexive}(i)$; and
(2) $\exists r \in R_d[l]$ such that
  (1) $\forall \Box_k a \in A(r) \cap \mathrm{ModLit}$, $+\partial_{\Box_k} a \in P(1..m)$,
    $\forall \neg\Box_k a \in A(r) \cap \mathrm{ModLit}$, $-\partial_{\Box_k} a \in P(1..m)$,
    $\forall a \in A(r) \cap \mathrm{Lit}$, $+\partial a \in P(1..m)$, if $r \in R^i$; or
  (2) $\forall a \in A(r)$, $+\partial_{\Box_k} a \in P(1..m)$, if $r \in R^j$, $A(r) \neq \emptyset$ and $\mathsf{Convert}(j,i)$; and
(3) $\forall s \in R[\sim l]$ either
  (1) $s \in R^j$ and $\neg\mathsf{Conflict}(j,i)$; or
  (2) $s \in R^i \cup R^j$, $\mathsf{Conflict}(j,i)$ and either
    (1) $\exists a \in A(s) \cap \mathrm{Lit}$, $-\partial a \in P(1..m)$ or
    (2) $\exists \Box_k a \in A(s) \cap \mathrm{ModLit}$, $-\partial_{\Box_k} a \in P(1..m)$ or
    (3) $\exists \neg\Box_k a \in A(s) \cap \mathrm{ModLit}$, $+\partial_{\Box_k} a \in P(1..m)$; or
  (3) $s \in R^j$, $\mathsf{Convert}(j,k)$, $\mathsf{Conflict}(k,i)$ and
    (1) $A(s) = \emptyset$ or
    (2) $A(s) \cap \mathrm{ModLit} \neq \emptyset$ or
    (3) $\exists a \in A(r)$ $-\partial_{\Box_k} a \in P(1..m)$; or
  (4) $\exists t \in R[l]$ such that
    (1) $\forall \Box_k a \in A(r) \cap \mathrm{ModLit}$, $+\partial_{\Box_k} a \in P(1..m)$,
      $\forall \neg\Box_k a \in A(t) \cap \mathrm{ModLit}$, $-\partial_{\Box_k} a \in P(1..m)$,
      $\forall a \in A(r) \cap \mathrm{Lit}$, $+\partial a \in P(1..m)$,
      and $t > s$; or
    (2) $A(t) \neq \emptyset$, $A(t) \cap \mathrm{ModLit} = \emptyset$,
      $\forall a \in A(t)$ $+\partial_{\Box_k} a \in P(1..m)$, if $t \in R^k$, $s \in R^m$, $\mathsf{Convert}(k,j)$ and
      $\mathsf{Conflict}(j,m)$.

If $P(m+1) = +\partial l$ then
(1) $l \in F$ or
(2) $+\partial_{\Box_i} l \in P(1..m)$ for some $\Box_i \in \mathrm{MOD}$ such that $\mathsf{Reflexive}(i)$.

[5] For space reasons we give only the proof conditions for the positive proof tags. The conditions for the negative proof tags can be obtained from the positive ones by the *Principle of Stronge Negation* [2, 12]. The strong negation of a formula is closely related to the function that simplifies a formula by moving all negations to an innermost position in the resulting formula and replaces the positive tags with the respective negative tags and vice-versa.

To prove a literal with 'factual' mode we have two possibilities: the literal is given as a fact, or the literal is derived with a reflexive modality. The derivation of $+\partial_{\Box_i}$ has three phases (Clauses 2, 3 and 3.4). Clause (1) is to ensure consistency of what we derive using reflexive modalities. According to clause (2) we have to have an applicable rule for the conclusion we want to prove and the rule should be appropriate for the modality of the conclusions. Here we have two cases: we use a rule for the same modality, or we use a conversion. For the same modality (clause 2.1), each element of the antecedent of the rule must be proved with its modality ($+\partial$ if factual, $+\partial_{\Box_k} l$ for a modal literal $\Box_k l$, and $-\partial_{\Box_k} l$ for $\neg\Box_k l$). For a conversion (clause 2.2), the body of the rule must not be empty, all literals in the body are not modal literal, and all them must be be provable with the modality the rule converts to. In the second phase (clause 3) we have to consider the possible attacks to the conclusion, and we have to discard them. There are several options to discard a rule. (1) the rule is not really attacking the conclusion. This is the case when the rule for the opposite conclusion is not in a conflict relation with the modality of the conclusion. (2) we can discard a conflicting rule, when the rule is not applicable: this means that one of the elements of the body of the rule is not provable with the appropriate modality. (3) if the attack is from a rule potentially using conversion, then we have to show that the conversion does not hold. The final case (4) is where we rebut the attacking rule. To do so, we have to show that there is a stronger and applicable rule (4.1) or that there is a rule that converts into a modality that conflicts with (and defeats) the modality of the attacking rule.

Definition 6. *The* extension *of a Defeasible Theory D is the structure* $(+\partial_{\Box_i}, -\partial_{\Box_i})$, *where* $\pm\# = \{p : D \vdash \pm\# p\}$.

In this paper we are concerned with theories corresponding to standard models of modal logics. One problem is that theories containing rules like $\Box_i a \Rightarrow_{\Box_i} a$ might not be able to produce a conclusion, since they determine loops. To obviate this problem we propose a syntactic criterion to avoid loops based on the concept of dependency graph of the literals in a Defeasible Theory.

Definition 7 (Dependency graph). *Let* $\mathrm{Lit}(D)$ *be the set of literals occurring in a Defeasible Theory D. The* dependency graph *of D is the directed graph* (N, E) *where:*

- $N = \{p, \Box_i p : p \in \mathrm{PROP}, \{p, \neg p, \Box_j p, \neg\Box_j p, \Box_j\neg p, \neg\Box_j\neg p\} \cap \mathrm{Lit}(D) \neq \emptyset\}$;
- $(n, m) \in E$ *iff*
  - $n = \Box_i m$, $\mathsf{Reflexive}(i)$ *and* $\exists r \in R^i[m] \cup R^i[\sim m]$;
  - $m = \Box_i l$ *and* $\exists r \in R^i[l] \cup R^i[\sim l]$ *such that* $\{n, \sim n\} \cap A(r) \neq \emptyset$.

4 Neighborhood Semantics for Defeasible Multi-modal Logic

4.1 The Background

Despite some difficulties (see [10]) and some alternatives [13], neighbourhood models are still the main semantics for non-normal modal logics [8].

As we will see, Defeasible Multi-modal Logic can be in fact interpreted as a non-normal multi-modal logic. Hence, this sub-section recalls some standard notions and proposes a couple of new results that are needed for the remainder of the paper. Assume our multi-modal language is defined as follows, where PROP is a set of atomic sentences:

$$p|\neg l|l \wedge l|\Box_1 l|\Diamond_1 l|\ldots|\Box_n l|\Diamond_n l$$

such that $p \in$ PROP.

Definition 8. *A multi-modal neighbourhood frame $\mathscr{F}$ is a structure $\langle W, \mathscr{N}^{\Box_1}, \ldots, \mathscr{N}^{\Box_n} \rangle$ where*

- *W is a non-empty set of possible worlds;*
- *$\mathscr{N}^{\Box_1}, \ldots, \mathscr{N}^{\Box_n}$ are functions $W \mapsto 2^{2^W}$.*

Definition 9. *A multi-modal neighbourhood model $\mathscr{N}$ is a structure $\langle W, \mathscr{N}^{\Box_1}, \ldots, \mathscr{N}^{\Box_n}, v \rangle$ where $\langle W, \mathscr{N}^{\Box_1}, \ldots, \mathscr{N}^{\Box_n} \rangle$ is a multi-modal neighbourhood frame and v is an evaluation function* PROP $\mapsto 2^W$.

In the reminder of the paper, for simplicity we will call “frame” a multi-modal neighbourhood frame and “model” a multi-modal neighbourhood model.

Definition 10 (Truth in a model). *Let $\mathscr{M}$ be a model $\langle W, \mathscr{N}^{\Box_1}, \ldots, \mathscr{N}^{\Box_n}, v \rangle$ and $w \in W$. The truth of any formula A in $\mathscr{M}$ is defined inductively as follows:*

1. *standard valuation conditions for the boolean connectives;*
2. *$\mathscr{M}, w \models \Box_i A$, $1 \leq i \leq n$, iff $||A|| \in \mathscr{N}_w^{\Box_i}$,*
3. *$\mathscr{M}, w \models \Diamond_i A$, $1 \leq i \leq n$, iff $W - ||A|| \notin \mathscr{N}_w^{\Box_i}$,*

where $||A||_{\mathscr{M}}$, the truth set of A wrt to $\mathscr{M}$, is thus defined:[6]

$$||A||_{\mathscr{M}} = \{w \in W : \mathscr{M}, w \models A\}.$$

A formula A is *true at a world* in a model iff $\mathscr{M}, w \models A$; *true in a model $\mathscr{M}$*, written $\mathscr{M} \models A$ iff for all worlds $w \in W$, $\mathscr{M}, w \models A$; *valid in a frame $\mathscr{F}$*, written $\mathscr{F} \models A$ iff it is true in all models based on that frame; *valid in a class of frames* iff it is valid in all frames in the class.

As usual, we can characterize many different classes of neighbourhood frames. Let us consider below a few of them that are relevant for our purposes:

Definition 11. *A frame $\mathscr{F} = \langle W, \mathscr{N}^{\Box_1}, \ldots, \mathscr{N}^{\Box_n} \rangle$ is*

- $\Box_i$-reflexive *iff, for any $X \subseteq W$, if $X \in \mathscr{N}_w^{\Box_i}$, then $w \in X$;*
- $\Box_i$-coherent *iff for any $w \in W$ and $X \subseteq W$, $X \in \mathscr{N}_w^{\Box_i} \Rightarrow W - X \notin \mathscr{N}_w^{\Box_i}$;*
- $\Box_i$-$\Box_j$-coherent *iff for any $w \in W$ and $X \subseteq W$, $X \in \mathscr{N}_w^{\Box_i} \Rightarrow W - X \notin \mathscr{N}_w^{\Box_j}$.*

Let us consider the few inference rules and schemata that we will use or discuss in the remainder[7]:

[6] Whenever clear from the context we will drop the references to the model.
[7] With standard schemata and inference rules, we adopt [8]’s terminology.

Axiom Schemata and Inference Rules

D:= $\Box_i A \rightarrow \neg\Box_i \neg A$

T:= $\Box_i A \rightarrow A$

bimodal-**D**:= $\Box_i A \rightarrow \neg\Box_j \neg A$

bimodal conversion axiom:= $(A_1 \wedge \ldots \wedge A_n \rightarrow \Box_i B) \rightarrow (\Box_j A_1 \wedge \ldots \wedge \Box_j A_n \rightarrow \Box_j B)$, where $n \geq 1$.[8]

RE:= $\dfrac{\vdash A \equiv B}{\vdash \Box_i A \equiv \Box_i B}$

The following are standard results (see [8, 13, 9]) about propositional non-normal modal logics and neighbourhood frames.

Theorem 1. *For any modal operator $\Box_i$,*

- **D** *is valid in the class of $\Box_i$-coherent frames;*
- **T** *is valid in the class of $\Box_i$-reflexive frames.*

Let us know consider the bridge axiom schema bimodal **D**. Its semantic characterization is almost a straightforward result:

Theorem 2. *For any couple of modal operators $\Box_i$ and $\Box_j$, bi-modal* **D** *is valid in the class of $\Box_i$-$\Box_j$-coherent frames.*

Less standard is the semantic characterization of bimodal conversion axiom, which seems a good approximation of the idea of conversion in Defeasible Multi-modal Logic. Let us define the following:[9]

Definition 12. *A frame $\mathscr{F} = \langle W, \mathscr{N}^{\Box_1}, \ldots, \mathscr{N}^{\Box_n}\rangle$ is $\Box_i$-$\Box_j$-*convertible *iff*

$$\forall w \in W,\ \forall X_1, \ldots, X_n, Y \subseteq W$$
$$[(w \in \bigcap_{1\leq k\leq n} X_k \Rightarrow Y \in \mathscr{N}_w^{\Box_i})\ \&\ (X_1, \ldots, X_n \in \mathscr{N}_w^{\Box_j})] \Rightarrow Y \in \mathscr{N}_w^{\Box_j}. \quad (1)$$

Hence,

Theorem 3. *For any couple of modal operators $\Box_i$ and $\Box_j$, bimodal conversion axiom is valid in the class of $\Box_i$-$\Box_j$-convertible frames.*

Proof. ($\Rightarrow$) Suppose that the bimodal conversion axiom is not valid. We show that the corresponding property does not hold too. If the schema is false, then there a model and a world $w \in W$ in it such that, for some literals $p_1, \ldots, p_n, q$

[8] If $A_1 \wedge \ldots \wedge A_n$ is inconsistent and $\Box_j$ does not obey the **D** axiom, the antecedent of $\Box_j A_1 \wedge \ldots \wedge \Box_j A_n \rightarrow \Box_j B$ may be consistent and we can infer $\Box_j B$, for any B. This depends on combining classical propositional calculus with a logic for $\Box_j$ which does not contain $(\Box_j A \wedge \Box_j B) \rightarrow \Box_j (A \wedge B)$. Fortunately, we assume **D** for any $\Box_j$, otherwise, to avoid this problem we should impose that $A_1 \wedge \ldots \wedge A_n$ is consistent.

[9] For the same reason we mentioned in footnote 8 above, if $\Box_j$ does not obey **D**, we should impose that $\bigcap_{1\leq k\leq n} X_k \neq \emptyset$. We thank one of DEON anonymous reviewers for pointing out these aspects of the bimodal conversion schema.

(a) $\mathscr{M}, w \models p_1 \wedge \ldots \wedge p_n \rightarrow \Box_i q$,
(b) $\mathscr{M}, w \models \Box_j p_1 \wedge \ldots \wedge \Box_j p_n$, but
(c) $\mathscr{M}, w \not\models \Box_j q$.

Point (a) means that (i) $w \in \bigcap_{1 \leq k \leq n} ||p_k|| \Rightarrow ||q|| \in \mathscr{N}_w^{\Box_i}$; point (b) means that (ii) $||p_k|| \in \mathscr{N}_w^{\Box_j}$ where $1 \leq k \leq n$; point (c) means that (iii) $||q|| \notin \mathscr{N}_w^{\Box_j}$. Hence, $\exists w \in W$ such that

$$(w \in \bigcap_{1 \leq k \leq n} ||p_k|| \Rightarrow Y \in \mathscr{N}_w^{\Box_i}) \ \& \ (||p_1||, \ldots, ||p_n|| \in \mathscr{N}_w^{\Box_j}) \ \& \ ||q|| \notin \mathscr{N}_w^{\Box_j}$$

which falsifies (1).

($\Leftarrow$) Suppose property does not hold and let us show that the bimodal conversion axiom is not valid too. Hence, there is a model $\mathscr{M}$ where $\exists w \in W$, $\exists \alpha_1, \ldots, \alpha_n$, $\exists \gamma$ such that

$$\begin{array}{ll} (i) & (w \in \bigcap_{1 \leq k \leq n} \alpha_k \Rightarrow \gamma \in \mathscr{N}_w^{\Box_i}) \ \& \\ (ii) & (\alpha_1, \ldots, \alpha_n \in \mathscr{N}_w^{\Box_j}) \ \& \\ (iii) & \gamma \notin \mathscr{N}_w^{\Box_j} \end{array}$$

Let us define a valuation for some propositions $p_1, \ldots, p_n, q$ and establish that $v(p_1) = \alpha_1, \ldots, v(p_n) = \alpha_n$ and $v(q) = \gamma$. Hence, we have the following:

$$\begin{array}{ll} (i') & w \in ||p_1 \wedge \cdots \wedge p_n|| \Rightarrow ||q|| \in \mathscr{N}_w^{\Box_i} \\ & \mathscr{M}, w \models \bigwedge_{1 \leq k \leq n} p_k \Rightarrow \mathscr{M}, w \models \Box_i q \\ & \mathscr{M}, w \models \bigwedge_{1 \leq k \leq n} p_k \rightarrow \Box_i q \\ (ii') & ||p_1||, \ldots, ||p_n|| \in \mathscr{N}_w^{\Box_j} \\ & \mathscr{M}, w \models \Box_j p_1, \ldots, \mathscr{M}, w \models \Box_j p_n \\ & \mathscr{M}, w \models \bigwedge_{1 \leq k \leq n} \Box_j p_k \\ (iii') & \mathscr{M}, w \not\models \Box_j q. \end{array}$$

Hence, bimodal conversion axiom is falsified.

4.2 From Defeasible Logic to Neighbourhood Semantics

Definition 13. *The D-extension E of a Defeasible Theory D is the smallest set of literals and modal literals such that:*

(a) $l \in E$ *iff* $l \in +\partial$*;*
(b) $\Box_i l \in E$ *iff* $l \in +\partial_{\Box_i}$*;*
(c) $\neg\Box_i l \in E$ *iff* $l \in -\partial_{\Box_i}$*;*

where l ranges on the set of literals.

Definition 14. *A* Defeasible Rule Theory *is a structure* $(R, >)$*, where R is a set of defeasible rules and defeaters, and* $>$ *is as in Definition 3.*

Definition 15. *A set of literals is* consistent *iff it does not contain p and* $\neg p$ *for any literal p.*

Definition 16. *Let L be a consistent set of literals, and $D = (R, >)$ be a defeasible rule theory. The D-*extension* of L is the extension of the defeasible theory $(L, R, >)$.*

Proposition 1. *Let D be a defeasible rule theory such that the transitive closure of $>$ is acyclic. Then, the D-extension of L is consistent iff L is consistent.*

Proof. The proof is based on those of Proposition 3.3 of [3], and Theorem 2 of [15].

Definition 17. *A D-extension is $\Box_i$-*complete *iff for every atom p, p and $\neg p$ are in $+\partial_{\Box_i} \cup -\partial_{\Box_i}$.*

Proposition 2. *Let D be defeasible rule theory such that the dependency graph is acyclic. Then D is $\Box_i$-complete for any $\Box_i \in \mathrm{MOD}$.*

Proof. The proof is an extension of the proof of the analogous result for defeasible logic, Theorem 2.1 of [4].

Definition 18. *Let D be a defeasible rule theory. The* canonical neighbourhood model *of D, $\mathscr{M}_D$ is the structure:*

$$(W, \{\mathscr{N}^{\Box_i}\}_{\Box_i \in \mathrm{MOD}}, v)$$

where

- *W is the set of all consistent D-extensions.*
- *each $\mathscr{N}^{\Box_i}$ is a function with signature $W \mapsto 2^{2^W}$ defined as follows:*
  - *$xS^i_j y$ iff $\exists r_j \in R$ such that*
    1. *$\sim C(r) \notin x$ if $\mathsf{Reflexive}(i)$;*
    2. *(a) $A(r) \subseteq x$ and $C(r) \in y$ if $r \in R^i$ or*
       (b) $A(r) \neq \emptyset$, $\Box_i A(r) \subseteq x$ and $C(r) \in y$ if $r \in R^j$ and $\mathsf{Convert}(j,i)$; and
    3. *$\forall s \in R[\sim C(r)]$ either*
       (a) $s \in R^j$ and $\neg\mathsf{Conflict}(j,i)$, or
       (b) $s \in R^i \cup R^j$, $\mathsf{Conflict}(j,i)$ and $\exists a \in A(s), a \notin x$; or
       (c) $s \in R^j$, $\mathsf{Convert}(j,i)$, and either $A(s) = \emptyset$ or $A(s) \cap \mathrm{ModLit} \neq \emptyset$ or $\exists a \in A(s)$ such that $\Box_i a \notin x$; or
       (d) $\exists t \in R[C(r)]$ such that
          i. $t > s$, $A(t) \subseteq x$ or
          ii. $t \in R^k$, $s \in R^m$ $\mathsf{Convert}(k,j)$, $\mathsf{Conflict}(j,m)$, $A(t) \neq \emptyset$, $A(t) \cap \mathrm{ModLit} = \emptyset$ and $\Box_k A(t) \in x$.
  - *$S^i_j(w) = \{x \in W : wS^i_j x\}$*
  - *$\mathscr{S}^i_j(w) = \bigcup_{C(r_k)=C(r_j)} S^i_k(w)$*
  - *$\mathscr{N}^{\Box_i}_w = \{\mathscr{S}^i_j(w)\}_{r_j \in R}$*
- *v is a (partial) valuation function such that $v(p,w) = 1$ iff $p \in w$, and $v(p,w) = 0$ iff $\neg p \in w$.*

The idea of the construction of the canonical model for a rule defeasible theory is to consider all possible theories/extensions that can be generated from it. This means that first we take all possible consistent sets of facts and we use them as input for the set of rules in R, and we take the maximal $\Box_i$-complete extensions. This is the parallel of the saturation in a standard Lindenbaum-Henkin construction to obtain the set of all maximal consistent sets. The second step is the construction of the neighbourhoods. This step is divided in two phases. In the first phase, we build relationships between possible worlds based on information in the rules. Remember that, given a rule $a_1,\dots,a_n \Rightarrow_{\Box_i} b$ the intuition is that if $a_1,\dots,a_n$ hold, then we conclude $\Box_i b$. This means that, if $a_1,\dots,a_n$ are in a possible world (extension), then we can saturate the extension by including $\Box_i b$ in it. At the same time we use one of the other standard construction of canonical models for modal logic: two maximally consistent sets x and y are related iff $\{\Box_i a \in x\} \subseteq \{a \in y\}$. The difference is that we build one of such relations for each rule in the defeasible rule theory (condition 2, part (a) is to take care of the case of conversion). The remaining conditions are to ensure that a rule really produces the conclusions according to the proof conditions for the logic at hand. The second phase of the construction of the neighbourhoods is to put together all relations obtained from the rule for the same combination of modal operator and literal.

Notice that in the codomain of each S^i_j we have all possible worlds where the conclusion of the rule holds. Thus the codomain corresponds to the truth set for that literal. In addition such a truth set is not empty, since each of it is a literal, and since we take the set of all maximal consistent and $\Box_i$-complete sets, there is at least one possible world where the literal holds. Thus we have the following result.

Lemma 1. *Let D be a defeasible rule theory, $\mathscr{M}$ the canonical model of D, and $\Box_i \in$ MOD.*

$$\forall w \in W, \forall X \in \mathscr{N}_w^{\Box_i}, \exists l : \Box_i l \in \mathrm{ModLit} :\ X = ||l|| \neq \emptyset$$

Theorem 4. *Let D be a defeasible rule theory, $\mathscr{M}$ the canonical model of D, and $w \in W$:*

1. $\mathscr{M}, w \models \Box_i p$ *iff* $w \vdash +\partial_{\Box_i} p$
2. $\mathscr{M}, w \models \neg\Box_i p$ *iff* $w \vdash -\partial_{\Box_i} p$

Proof (Sketch). The proof, by induction of the length of derivation on one side and the iterative construction of the extension, is based on the proof of Theorem 2.2 of [4] (Theorem 1 of [21]), which shows the equivalence of the proof conditions of DL [3] and the construction of the extension of a theory. The difference is that in [4] the heads of applicable rules (leading to conclusions) are added to the extension being constructed, while here, the elements expanding the extension are modal literals, and, in addition, we create instances of the relationships between the current extension and the extensions where the literal occurs unmodalised.

Theorem 5. *Let D be a defeasible rule theory and $\mathscr{M}$ the canonical model generated by D:*

1. *$\mathscr{M}$ is $\Box_i$-coherent for all $\Box_i \in$ MOD;*
2. *$\mathscr{M}$ is $\Box_i$-$\Box_j$-coherent for all $\Box_i, \Box_j \in$ MOD such that* $\mathsf{Conflict}(i,j)$*;*
3. *$\mathscr{M}$ is $\Box_i$-reflexive for all $\Box_i \in$ MOD such that* $\mathsf{Reflexive}(i)$.

Proof (Sketch). Part 1. By construction of the canonical model, every w is a consistent D-extension, thus for no fact p, p and $\neg p$ are in w, this means that it holds, that for w it is not that case that $w \vdash +\partial_{\Box_i} p$ and $w \vdash +\partial_{\Box_i} \neg p$. Suppose that $w \vdash +\partial_{\Box_i} p$ for some arbitrary literal p. By Theorem 4 $w \models \Box_i p$, thus $||p|| \in \mathcal{N}_w^{\Box_i}$. Since the conditions for $-\partial_{\Box_i}$ are the strong negation of that of $+\partial_{\Box_i}$, we have that $w \vdash -\partial_{\Box_i}\neg p$; again, by Theorem 4, $w \models \neg\Box_i\neg p$, thus $-||p|| \notin \mathcal{N}_w^{\Box_i}$.

Part 2. $\mathsf{Conflict}(i,j)$ means that every rule for i is superior to any rule for j. This means that every time the proof conditions for $+\partial_{\Box_i}$ are satisfied for p, then the proof conditions for $-\partial_{\Box_j}$ are satisfied for $\sim p$. Thus we can repeat the reasoning as in the previous case.

Part 3. If $\Box_i$ is reflexive, i.e., $\mathsf{Reflexive}(i)$ holds, then every time we have $w \vdash +\partial_{\Box_i} p$, then we have $w \vdash +\partial p$, and thus p is the corresponding maximal $\Box_i$-complete D-extension, i.e., $p \in w$. However, to have $w \vdash +\partial_{\Box_i} p$, we need to have a rule that is applicable in w, and then $wS^i_j x$ such that $p \in x$, thus $wS^i_j w$, which means that $\forall X \in \mathcal{N}_w^{\Box_i}$, $w \in X$.

4.3 Characterizing Conversions in Canonical Models

We have semantically characterized the bimodal conversion axiom. Apparently, this schema looks as a good approximation in non-normal multi-modal logics of the notion of conversion in Defeasible Multi-modal Logic. However, if we have $\mathsf{Convert}(i,j)$, the semantic property corresponding to the schema for $\Box_i$ and $\Box_j$ does not generally hold in the canonical model generated by any theory, canonical model that is defined in Definition 18 in order to capture, by construction, the notion of conversion. Indeed, what we have is the following:

Proposition 3. *For every Defeasible Rule Theory $D = (R, >)$*

(i) $\mathsf{Convert}(i,j) \Rightarrow$ *the canonical neighbourhood model of D is $\Box_i$-$\Box_j$-convertible.*
(ii) The canonical neighbourhood model of D is $\Box_i$-$\Box_j$-convertible $\not\Rightarrow \mathsf{Convert}(i,j)$.

Proof (Sketch). Consider Case (i) and suppose it does not hold. Hence, there is at least one Defeasible Rule Theory D such that $\mathsf{Convert}(i,j)$ and the canonical model of D is not $\Box_i$-$\Box_j$-convertible.

Hence, $\exists w \in W$ in the canonical model such that $\exists a_1,\ldots,a_n, b \in \mathrm{Lit}(D)$ (where $n \geq 1$ and $\mathrm{Lit}(D)$ is the set of literals occurring in D)

$$
\begin{aligned}
&(a)\ (w \in \textstyle\bigcap_{1\leq k\leq n} ||a_k|| \Rightarrow ||b|| \in \mathcal{N}_w^{\Box_i})\ \&\\
&(b)\ (||a_1||,\ldots,||a_n|| \in \mathcal{N}_w^{\Box_j})\ \&\\
&(c)\ ||b|| \notin \mathcal{N}_w^{\Box_j}
\end{aligned}
$$

Condition (a) can be satisfied because $w \notin \bigcap_{1\leq k\leq n} ||a_k||$ or because $||b|| \in \mathcal{N}_w^{\Box_i}$. By construction (Definition 18), the first case means either that $\{a_1,\ldots,a_n\}$ is not consistent—but this is excluded by Definition 18 —or that there exists an $a_k \in \mathrm{Lit}(D)$ for which no rule r in R can prove it. However, condition (b) requires that every a_k can be proved in D, so condition (a) only means in the canonical model that $||b|| \in \mathcal{N}_w^{\Box_i}$. From this last conclusion, it follows that either

1. there is a rule $r_i \in R^i[b]$, or
2. $\mathsf{Convert}(k,i)$ and there is a rule $r_k \in R^k[b]$ such that $A(r_j) \neq \emptyset$ and $\forall d \in A(r_j)$: $\Box_i d \in w$.

Since condition (b) guarantees that $\Box_j a_1 \dots \Box_j a_n \in w$, case 1. and condition (c) above jointly imply that $\mathsf{Convert}(i,j)$ does not hold. Case 2 implies that $||b|| \in \mathcal{N}_w^{\Box_j}$, contrary to condition (c).

Consider now Case (ii). The fact that the property of being $\Box_i$-$\Box_j$-convertible does not in general guarantee $\mathsf{Convert}(i,j)$ can be easily shown by considering the following Defeasible Rule Theory:

$$R = \{r_1 : b \Rightarrow_{\Box_i} c,\; r_2 : \Rightarrow_{\Box_i} c,\; r_3 : \Rightarrow_{\Box_j} a\}$$
$$> = \emptyset$$

Suppose that $\mathsf{Convert}(i,j)$ holds. Consider a world w where the facts contain a and $\neg b$. The conclusions of the resulting theory contain: $+\partial_{\Box_i} c$, $+\partial_{\Box_j} a$, $-\partial_{\Box_j} b$ and $-\partial_{\Box_j} c$. According to Definition 13, a, $\Box_i c$ and $\Box_j a$ are true in w; thus, given that the model is $\Box_i$-$\Box_j$-convertible, $\Box_j c$ holds in w, but as we have seen, $(\{a, \neg b\}, R, \emptyset) \vdash -\partial_j c$, Hence a contradiction.

The property characterizing the bimodal conversion axiom thus does not fully guarantee that conversions are captured in the canonical model: what we can know is only that when the conversion holds, the property holds, too.

The reason why the bimodal conversion axiom does not fully capture conversions depends on the fact that we cannot keep track of the way in which a modal literal is obtained. Hence, the translation in standard multi-modal logic of conversions does not work because given $a_1, \dots, a_n \to \Box_i b$ and $\Box_j a_1, \dots \Box_j a_n$, if we obtain $\Box_j b$ we do not really know if this last formula was obtained by applying the conversion to a rule for $\Box_i$ or it is rather obtained using other rules in the theory.

We leave this problem to our future work. Let us just outline here some technical details for a possible solution. In a nutshell the idea is the following:

- duplicate each literals occurring in the theory in order to keep track of the rule where the literal occurs;
- build canonical model from the defeasible rule theory translated into the new language;
- make a filtration of the generated canonical model in order to guarantee that the duplicates of the same original literal are logically equivalent, in such a way that the translation procedure is safe with respect to the theory extension.

We expect that in this new generated model the property for bimodal conversion holds and characterizes conversions.

The basic starting point would be the following definition.

Definition 19. *Let $D = (R, >)$ be any Defeasible Rule Theory. The* expansion $\mathrm{Exp}(D) = (\mathrm{Exp}(R), \mathrm{Exp}(>))$ *of D is defined as follows:*

- $\mathrm{Exp}(R) = \{r : a_1^{r,i}, \dots, a_n^{r,i} \rhd_{\Box_i} b^{r,i} | \forall r : a_1, \dots, a_n \rhd_{\Box_i} b,\ \rhd \in \{\Rightarrow, \leadsto\}\}$*;*
- $\mathrm{Exp}(>) = >$.

The language of the resulting theory $\mathrm{Exp}(D)$ would not require to substantially change the proof theory: it would be sufficient to make it possible, for any X and Y, to use the derivation of $a_1^X, \ldots, a_n^X$ to trigger any rule like $a_1^Y, \ldots, a_n^Y \rhd_{\Box_i} b^Y$. We should also take this into account to handle conflicts, since a^X and $\sim a^Y$ could be incompatible.

With this done, the procedure to generate the canonical model of any expanded defeasible rule theory would not need any significant revision: the advantage would be that now we can keep track—in building the neighbourhoods—of the rules that are used to derive modalized expressions. Hence, we could evaluate $\Box_j p^{r,\Box_i}$, which means that p was proved using a rule r for $\Box_i$ and thanks to $\mathsf{Convert}(i, j)$. Notice that this does not change the language in a strict sense, since $p^{r,\Box_i}$ is still a propositional constant.

The last step would be to define a filtration [8] of the generated canonical model by stating that, for each each world w in the original canonical model $\mathscr{M}_D$:

$$[w] = \{v | w, v \in W \text{ and } \forall X, Y : \mathscr{M}_D, w \models p^X \text{ iff } \mathscr{M}_D, w \models p^Y\}$$

We thus ensure that, in the new model, we have $||p^X|| = ||p^Y||$. We expect that the new model preserves all properties of the original canonical model.

5 Summary and Related Work

The relation between non-monotonic and modal logics is a complex question [6, 7]: one of the first attempts to investigate the issue was in Autoepistemic Logic [23]. In the present setting of DL, a rather direct approach to the problem could be to proceed from [19]'s semantics, since the basic consequence relation of DL is cumulative reasoning. Another possibility would be to define an extension of the argumentation semantics of [11] and take the route of [17]; however, this approach would not really establish connections between DL and modal logic, but would simply import techniques from the latter into the former domain.

In this paper we thus addressed the problem by discussing the meaning of modal provability of DL in neighbourhood semantics. We presented a general multi-modal logical framework called Defeasible Multi-modal Logic, which is able to embed all existing variants of Defeasible Deontic Logic in the literature. We discussed how to interpret the logic in neighbourhood semantics and introduced a specific class of canonical models for that. We discussed one critical aspect of Defeasible Multi-modal Logic, the notion of conversion, and proved that the proposed semantical construction only partially characterize it. An open problem is to determine in general what classes of defeasible theories are are sound and complete with respect to which classes of neighborhood frames. Addressing this problem would precisely clarify the "modal meaning" of Defeasible Multi-modal Logic. We leave this issue to future research.

Acknowledement

NICTA is funded by the Australian Government as represented by the Department of Broadband, Communications and the Digital Economy and the Australian Research Council through the ICT Centre of Excellence program.

References

1. G. Antoniou and A. Bikakis. Dr-prolog: A system for defeasible reasoning with rules and ontologies on the semantic web. *IEEE Trans. Know. and Data Eng.*, 19(2):233–245, 2007.
2. G. Antoniou, David Billington, G. Governatori, and M. J. Maher. A flexible framework for defeasible logics. In *AAAI-2000*, pages 401–405, Menlo Park, CA, 2000. AAAI/MIT Press.
3. G. Antoniou, David Billington, G. Governatori, and M. J. Maher. Representation results for defeasible logic. *ACM Transactions on Computational Logic*, 2(2):255–287, 2001.
4. G. Antoniou, David Billington, G. Governatori, and M. J. Maher. Embedding defeasible logic into logic programming. *Theory and Practice of Logic Programming*, 6(6):703–735, 2006.
5. N. Bassiliades, G. Antoniou, and I. Vlahavas. A defeasible logic reasoner for the semantic web. *International Journal on Semantic Web and Information Systems*, 2(1):1–41, 2006.
6. A¿ Bochman. *A logical theory of nonmonotonic inference and belief change — numerical methods*. Springer, 2001.
7. A. Bochman. Default theory of defeasible entailment. In *KR 2008*, pages 466–475, 2008.
8. B. F. Chellas. *Modal Logic, An Introduction*. Cambridge University Press, Cambridge, 1980.
9. R. Fagin, J. Y. Halpern, Y. Moses, and M. Y. Vardi. *Reasoning about Knowledge*. MIT Press, 1995.
10. M. Gerson. The Inadequacy of the Neighbourhood Semantics for Modal Logic. *Journal of Symbolic Logic*, 40(2):141–148, 1975.
11. G. Governatori, M. J. Maher, G. Antoniou, and David Billington. Argumentation semantics for defeasible logics. *Journal of Logic and Computation*, 14(5):675–702, 2004.
12. G. Governatori, V. Padmanabhan, A. Rotolo, and A. Sattar. A defeasible logic for modelling policy-based intentions and motivational attitudes. *Logic Journal of the IGPL*, 17(3):227–265, 2009.
13. G. Governatori and A. Rotolo. On the Axiomatization of Elgesem's Logic of Agency and Ability. *Journal of Philosophical Logic*, 34(4):403–431, 2005.
14. G. Governatori and A. Rotolo. BIO logical agents: Norms, beliefs, intentions in defeasible logic. *Journal of Autonomous Agents and Multi Agent Systems*, 17(1):36–69, 2008.
15. G. Governatori and A. Rotolo. A computational framework for institutional agency. *Artificial Intelligence and Law*, 16(1):25–52, 2008.
16. G. Governatori and A. Rotolo. Changing legal systems: legal abrogations and annulments in defeasible logic. *Logic Journal of IGPL*, 18(1):157–194, 2010.
17. D. Grossi. On the logic of argumentation theory. In *AAMAS'10*, pages 409–416, Richland, SC, 2010. International Foundation for Autonomous Agents and Multiagent Systems.
18. E. Kontopoulos, N. Bassiliades, and G. Antoniou. A modal defeasible reasoner of deontic logic for the semantic web. *International Journal of Semantic Web and Information Systems*, 7(1):18–43, 2011.
19. S. Kraus, D. Lehmann, and M. Magidor. Nonmonotonic reasoning, preferential models and cumulative logics. *Artificial Intelligence*, 44:167–207, 1990.
20. H.P. Lam and G. Governatori. The making of SPINdle. In *RuleML 2009*, LNCS. Springer, 2009.
21. M. J. Maher and G. Governatori. A semantic decomposition of defeasible logic. In *AAAI-99*, pages 299–305, Menlo Park, CA, 1999. AAAI Press.
22. M.J. Maher. Propositional defeasible logic has linear complexity. *Theory and Practice of Logic Programming*, (6):691–711, 2001.
23. V. W. Marek and M. Truszczynski. Autoepistemic logic. *Journal of ACM*, 38(3):588–619, 1991.
24. D. Nute. Defeasible logic. In *Handbook of Logic in Artificial Intelligence and Logic Programming*. Oxford University Press, 1993.