

Alma Mater Studiorum Università di Bologna
Archivio istituzionale della ricerca

A deontic argumentation framework based on deontic defeasible logic

This is the final peer-reviewed author's accepted manuscript (postprint) of the following publication:

Published Version:

Governatori, G., Rotolo, A., Riveret, R. (2018). A deontic argumentation framework based on deontic defeasible logic. Berlin : Springer Verlag [10.1007/978-3-030-03098-8_33].

Availability:

This version is available at: <https://hdl.handle.net/11585/660576> since: 2019-11-26

Published:

DOI: http://doi.org/10.1007/978-3-030-03098-8_33

Terms of use:

Some rights reserved. The terms and conditions for the reuse of this version of the manuscript are specified in the publishing policy. For all terms of use and more information see the publisher's website.

This item was downloaded from IRIS Università di Bologna (<https://cris.unibo.it/>).
When citing, please refer to the published version.

(Article begins on next page)

This is a post-peer-review, pre-copyedit version of the article

A Deontic Argumentation Framework Based on Deontic Defeasible Logic *published in* [[International Conference on Principles and Practice of Multi-Agent Systems](#)] PRIMA 2018: [PRIMA 2018: Principles and Practice of Multi-Agent Systems](#), pp.484-492.

Print ISBN 978-3-030-03097-1

Online ISBN 978-3-030-03098-8

The final authenticated version is available online at:

http://dx.doi.org/10.1007/978-3-030-03098-8_33

Springer's Self-archiving policy are available at:

<https://www.springer.com/gp/open-access/publication-policies/self-archiving-policy>

A Deontic Argumentation Framework based on Deontic Defeasible Logic

Guido Governatori¹, Antonino Rotolo², and Régis Riveret¹

¹ Data61, CSIRO, Australia

² CIRSFD, University of Bologna, Italy

Abstract. Deontic Defeasible Logic (DDL) 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. In this paper we present an argumentation system that corresponds to a variant of DDL. The resulting machinery is able to grasp in a natural way intuitions behind deontic reasoning with conditional norms featuring obligations, prohibitions, and (strong or weak) permissions.

Keywords: Argumentation · Deontic Argumentation · Deontic Defeasible Logic.

1 Introduction

Computational models of argument address defeasible claims raised on the basis of partial, uncertain and possibly conflicting pieces of information. Argumentation is pervasive in artificial intelligence, with many application domains [2].

Normative systems, and in particular legal systems, constitute a rich test bed and a major application domain for formal models of argument [13]. There, models of argument have applications ranging from case-based reasoning [14] to strategic studies in legal interactions [16, 17].

When representing and reasoning upon norms, deontic concepts such as obligation, prohibition and permission play a crucial role; and there exist some studies of deontic reasoning with formal models of argument, and a few argument-based models focus on (conditional) norms, deontic operators and their interplay [3, 4].

Besides these undertakings in deontic argumentation, many deontic formalisms have been previously designed [6]. Amongst formalisms with practical applications, DDL has been perhaps the most developed to represent and reason upon norms [8–10]. Moreover and interestingly, Defeasible Logic (DL) has possible interpretations in terms of arguments [7, 15, 12], but its deontic variants have received little or no consideration to construct argument-based frameworks for deontic reasoning. In this paper, we consider constructs from DDL to build a deontic argumentation system.

Contribution Following the approach in DDL, we offer a rich formalism able to express relevant aspects of deontic reasoning, such as contrary-to-duty obligations and preferences about permissions. The deontic argumentation framework is described in the remainder of the paper.

2 Deontic Argumentation System

This section presents the deontic argumentation system. We first specify its language, then arguments are constructed. Eventually, the justification and rejection of arguments are defined.

2.1 Language

The following definitions provide the building blocks of our formalism:

- literals and modalities ;
- preference operators for obligations and permissions.
- constitutive and deontic rules.

The attention is restricted to a simple propositional language with atomic negation and supplemented with a set of deontic operators $\{O, P\}$ where O indicates an obligation, and P a permission.

Definition 1. A literal is a **plain literal** iff it is an atomic proposition p or the negation of an atomic proposition, i.e., $\neg p$. A literal is a **deontic literal** iff it has either the form Ol or Pl or $\neg Ol$ or $\neg Pl$ where l is a plain literal. A literal is either a plain literal or a deontic literal.

Notation 1 Given a set of literals \mathcal{L} , the set of plain literals in \mathcal{L} is denoted as $\text{Lit}_{\mathcal{L}}$ and the set of modal literals as $\text{ModLit}_{\mathcal{L}}$. However, in the remainder, the set of literals may be left implicit, and we may omit the subscript \mathcal{L} .

We introduce two preference operators, \otimes for obligations and \odot for permissions. These operators are used to build chains of preferences, called \otimes - and \odot -expressions. Intuitively, an \otimes -expression such as $l_1 \otimes l_2 \otimes \dots \otimes l_n$ indicates that the obligation that l_1 is preferred to the one that l_2 , which is preferred to l_3 etc.

Definition 2. Let $\odot \in \{\otimes, \odot\}$. An \odot -**expression** is defined as follows:

1. every literal $l \in \text{Lit}$ is an \odot -expression;
2. if A is an \odot -expression and $c_1, \dots, c_k \in \text{Lit}$, then $A \odot c_1 \odot \dots \odot c_k$ is an \odot -expression;
3. nothing else is an \odot -expression.

Notation 2 Given a set of literals \mathcal{L} , the set of \odot -expressions defined by \mathcal{L} is denoted $\text{Pref}_{\odot, \mathcal{L}}$ or simply Pref_{\odot} if \mathcal{L} is left implicit.

Definition 3. Let Lbl be a set of arbitrary labels. A set of rules Rul is a **well-formed set of rules** over a set of literals \mathcal{L} iff:

$$\text{Rul} \subseteq (\text{Rul}_{\text{d}}^{\text{O}} \cup \text{Rul}_{\text{d}}^{\text{P}} \cup \text{Rul}_{\text{d}}^{\text{c}}) \cup (\text{Rul}_{\text{dft}}^{\text{O}} \cup \text{Rul}_{\text{dft}}^{\text{c}})$$

such that

$$\begin{aligned} \text{Rul}_{\text{d}}^{\text{O}} &= \{r : a_1, \dots, a_n \Rightarrow_{\text{O}} b \mid r \in \text{Lbl}, \{a_1, \dots, a_n\} \subseteq \text{Lit} \cup \text{ModLit}, b \in \text{Pref}_{\otimes}\} \\ \text{Rul}_{\text{d}}^{\text{P}} &= \{r : a_1, \dots, a_n \Rightarrow_{\text{P}} b \mid r \in \text{Lbl}, \{a_1, \dots, a_n\} \subseteq \text{Lit} \cup \text{ModLit}, b \in \text{Pref}_{\odot}\} \\ \text{Rul}_{\text{d}}^{\text{c}} &= \{r : a_1, \dots, a_n \Rightarrow_{\text{c}} b \mid r \in \text{Lbl}, \{a_1, \dots, a_n\} \subseteq \text{Lit}, b \in \text{Lit}\} \\ \text{Rul}_{\text{dft}}^{\text{O}} &= \{r : a_1, \dots, a_n \rightsquigarrow_{\text{O}} b \mid r \in \text{Lbl}, \{a_1, \dots, a_n\} \subseteq \text{Lit} \cup \text{ModLit}, b \in \text{Lit}\} \\ \text{Rul}_{\text{dft}}^{\text{c}} &= \{r : a_1, \dots, a_n \rightsquigarrow_{\text{c}} b \mid r \in \text{Lbl}, \{a_1, \dots, a_n\} \subseteq \text{Lit}, b \in \text{Lit}\}. \end{aligned}$$

Rules with an arrow \Rightarrow are defeasible rules, while rules with an arrow \rightsquigarrow are so-called defeaters which are essentially used to specify exceptions to defeasible rules. A defeasible rule can be used to support its consequent, whereas a defeater does not support its consequent.

Notation 3 The set of antecedents of a rule r is denoted $A(r)$, and $C(r)$ denotes its consequent. Other abbreviations are, for example, $Rul^O = Rul_d^O \cup Rul_{dft}^O$, and $Rul[b]$ to denote the set of rules whose consequent is b , and $Rul_d[b]$ the set of defeasible rules whose consequent is b .

Consequents of rules can be incompatible, and such incompatibilities are captured through complementary literals.

Notation 4 The complementary of a literal q is denoted by $\sim q$; if q is a positive literal p , then $\sim q$ is $\neg p$, and if q is a negative literal $\neg p$, then $\sim q$ is p .

Definition 4. Let $l \in Lit \cup ModLit$ denote a literal. A set of literals is a set of **complementary literals** of l , denoted $Compl(l)$, iff:

- if $l = p \in Lit$ then $Compl(l) = \{\sim l\}$;
- $Compl(Ol) = \{\neg Ol, O\sim l, \neg Pl, P\sim l\}$, $Compl(\neg Ol) = \{Ol, \neg Pl\}$, $Compl(Pl) = \{O\sim l, \neg Pl\}$, $Compl(\neg Pl) = \{\neg O\sim l, Pl, \neg P\sim l\}$.

As usual, we can define a *superiority relation* between rules to determine their relative strength in case of conflict. As shown in [1], we can disregard the superiority relation in the discussion, since modular transformations exist that empty this relation while maintaining the same conclusions in the language [8]. This result holds both for ambiguity blocking and ambiguity propagating DL [7]. It also applies to deontic extensions of DL (including the one with \otimes and \odot operators), by means of the notion of inferiorly defeated rules [11].

2.2 Arguments and attacks

Defining the notion of argument in the current context is not obvious. The complexity mainly resides in the richness of the language (especially the presence of the operators \otimes and \odot) and in the constructive nature of the introduction of modalities. The derivation of a modal literal such as Ob depends on the constructive provability of b using rules such as $a_1, \dots, a_n \Rightarrow_O b$, and the derivation of $\neg Ob$ depends on showing that there is no proof for Ob . We propose thus the following definition of arguments.

Definition 5. An **argument** A for a conclusion p generated from a set of rules Rul is a (possibly infinite) tree where

1. the root node is labelled by literal p ;
2. any node is labelled by either a literal $h \in Lit \cup ModLit$ or no literals; and such that:
 1. if the node labelled by h has children h_1, \dots, h_n ($n > 0$), then all arcs connecting h_1, \dots, h_n to h are labelled by exactly one rule $r \in Rul$ with $A(r) = \{b_1, \dots, b_n\}$ such that $h_1 = b_1, \dots, h_n = b_n$, and either

- (a) if $r \in \text{Rul}_d^O$ and $C(r) = c_1 \otimes \dots \otimes c_m$ then $h = Oc_k$ ($1 \leq k \leq m$);
 - (b) if $r \in \text{Rul}_d^P$ and $C(r) = c_1 \odot \dots \odot c_m$ then $h = Pc_k$ ($1 \leq k \leq m$);
 - (c) if $r \in \text{Rul}_d^C$ then $h = C(r)$;
 - (d) if $r \in \text{Rul}_{dft}^O$ then $h = p = \neg OC(r)$ is the root of the argument;
 - (e) if $r \in \text{Rul}_{dft}^C$ then $h = p = C(r)$ is the root of the argument;
2. if the node labelled by h has no children (i.e. h is a leaf node), then either
- (a) h is labelled by no literals;
 - (b) $h = \neg \Box l$ ($\Box \in \text{Mod}$);
 - (c) $h = Pl$.

The interpretation of item (d) is as follows. First of all, notice that we have to do with a case where a defeater is considered. A defeater for O with head p does not positively prove anything, but it can attack any obligation rule $a_1, \dots, a_n \Rightarrow_O \neg p$ supporting $\neg p$ (i.e., proving $O\neg p$). Conceptually, this means that the defeater can be a reason for stating that p is not obligatory, i.e., that $\neg Op$ is the case. Notice, also, that such a defeater—as any defeater here, and as done in standard argumentation semantics for DL [7]—can only label arcs leading to a root. In the modal case, this makes the interaction among arguments simple, as the concept of derivation for a negative modal literal depends on the relation between the argument considered and other arguments attacking the former one. Hence—as we will see—that $\neg Op$ is justified depends on the absence of successful arguments whose conclusion is $O\neg p$.

Remark 1. Item (b) of the last condition where the modality \Box is a permission P captures the case where the conclusion l as strong permission—i.e. a permission derived from a rule with \Rightarrow_P —is defeated (as we will see later). If the modality \Box is an obligation O , then it captures the case where a weak permission is put forward. Similarly, the last item (c) of condition (2) captures the case where a weak permission is assumed. In both cases, such a weak permission is not directly expressed in a specific deontic rule and cannot be constructively reflected in the tree-structure. In this sense, such nodes can only be leaf nodes.

Example 1. Suppose we have the following rule set:

$$\{r_1 : \Rightarrow_O c_1 \otimes c_2, \quad r_2 : Oc_1, \neg Op \Rightarrow_P q\}.$$

Then, we can build an argument as in Figure 1.

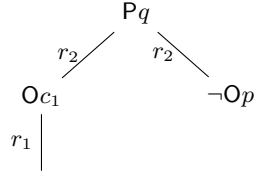


Fig. 1. An argument.

□

We may employ some auxiliary terminology.

- A *supportive argument* is a finite argument in which no defeaters are used.
- An argument is *positive* iff no defeater is used in it.
- A *constitutive argument* is an argument where all rules are constitutive rules.
- Any literal/modal literal labelling any node of an argument A is a *conclusion* of A .

Definition 6. Let A denote any argument with height $j \geq 1$ for any literal p .³ The **top subargument** of A , denoted A^t , is the top subargument of A with height 1. Let us use $R(A^t)$ to denote the rule associated with the arcs arriving at the root of A^t .

On the basis of arguments, we provide the core notion of the approach —*argument agglomeration set*— which gathers all arguments that are strictly needed to accept an argument. Such agglomeration set caters for two cases:

- when nodes are labelled by rule conclusions, in case of \otimes - or \odot -expressions in the head of rules such as $a \otimes b$ and $p \odot q$, the fact that Ob or Pq label nodes means that a and $O\neg p$ have been concluded;
- when leaf nodes are labelled modal literals such as $\neg Op$ we fall in the case discussed in Remark 1; even here, conditions external to the single argument at stake must be checked, which is required to verify that such an argument is justified.

Definition 7. Let $A \in \text{Args}$ be an argument such that for every node h of A labelled by a modal literal $\neg Ol$, $\neg Pl$, Ol or Pl , the arcs leading to h are labelled by

- (for literals such as Ol) $r : b_1, \dots, b_n \Rightarrow c_1 \otimes \dots \otimes c_m$ and $l = c_k$ ($1 \leq k \leq m$), and
- (for literals such as Pl) $r : b_1, \dots, b_n \Rightarrow d_1 \odot \dots \odot d_m$ and $l = d_k$ ($1 \leq k \leq m$).

An **argument agglomeration set** $\text{Aggl}(A) \subseteq \text{Args}$ w.r.t. A is a smallest set of arguments such that $A \in \text{Aggl}(A)$ and:

- for each (leaf) node labelled $\neg Ol$ or $\neg Pl$,
 - there is an argument $B \in \text{Aggl}(A)$ whose conclusion is $p \in \text{Compl}(Ol)$;
 - $\text{Aggl}(B) \subseteq \text{Aggl}(A)$;
- for each node labelled c_j ($1 \leq j < k$),
 - there is an argument $C \in \text{Aggl}(A)$ whose conclusion is $\neg c_j \in \text{Compl}(c_j)$;
 - $\text{Aggl}(C) \subseteq \text{Aggl}(A)$;
- for each node labelled d_j ($1 \leq j < k$),
 - there is an argument $D \in \text{Aggl}(A)$ whose conclusion is $q \in \text{Compl}(Pd_j)$;
 - $\text{Aggl}(D) \subseteq \text{Aggl}(A)$.

Remark 2. The agglomeration set of any argument A gathers all arguments that are strictly needed to accept the construction of A . Thus, the agglomeration set of A can be viewed as a single argument where special arcs connect nodes in A labelled by modal literals obtained by rules supporting \odot -expressions.

Example 2. Suppose the following rules:

$$\left\{ \begin{array}{ll} r_1 : \Rightarrow_c \neg c_1, & r_3 : \Rightarrow_O c_1 \otimes c_2, \\ r_2 : \rightsquigarrow_O \neg p, & r_4 : Oc_2, \neg Op \Rightarrow_P d_1 \odot d_2 \end{array} \right\}.$$

Then, we can build the following arguments⁴:

³ As usual, the height of an argument is the number of edges on the longest path between the root and a leaf node.

⁴ Arrows indicate the type of rule used.

$$A : \quad \Rightarrow Oc_2, \neg Op \Rightarrow Pd1, \quad B : \quad \Rightarrow \neg c_1, \quad C : \quad \rightsquigarrow_O \neg p$$

which agglomerate as illustrated in Figure 2.

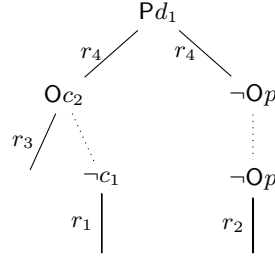


Fig. 2. An argument agglomeration set, where the dotted arc represent the relation between arguments in the agglomeration. □

Eventually, arguments supporting complementary literals attack each other.

Definition 8. An argument A **attacks** an argument B iff

1. there exists a node of B labelled by m , and
2. there exists a node of A labelled by $l \in \text{Compl}(m)$.

A set of arguments S attacks an argument B iff there is an argument A in S that attacks B .

2.3 Justified and rejected arguments

The justification of arguments has been thoroughly studied in the literature, and multiple semantics have been proposed. As we are dealing with DL, we resort to the argumentation semantics for variants of DL as presented in [7].

The usual definition of accepted arguments is slightly adapted to embrace argument agglomeration sets.

Definition 9. An argument A is an **accepted argument** w.r.t a set of arguments S iff A is finite, and every argument attacking any argument in any $\text{Aggl}(A)$ is attacked by S .

From accepted arguments, and similarly as [5], we can define justified arguments using a ‘characteristic function’.

Definition 10. Let Args be a set of arguments. The **deontic justification characteristic function** of Args is a function $J_i : \text{pow}(\text{Args}) \rightarrow \text{pow}(\text{Args})$ such that:

- $J_0 = \emptyset$, and
- $J_{i+1} = \{A \in \text{Args} \mid A \text{ is accepted w.r.t. } J_i\}$.

Definition 11. Let $Args$ be a set of arguments. The **set of justified arguments** of $Args$ is $JArgs = \bigcup_{i=1}^{\infty} J_i$.

Definition 12. Let $Args$ be a set of arguments. A literal is a **justified literal** if it is a conclusion of a supportive argument in $JArgs$.

Once justified arguments and literals are established, rejected arguments and literals can be determined. We first define rejected arguments with respect to a generic set of arguments which is then instantiated as a set of justified arguments.

Definition 13. An argument A is a **rejected argument** by a set of arguments S iff either

1. a proper subargument B of A is in S , or
2. B is attacked by a finite argument.

Definition 14. Let T be a set of arguments. The **deontic rejection characteristic function** of $Args$ is a function $R_i(T) : \text{pow}(Args) \rightarrow \text{pow}(Args)$ such that

- $R_0(T) = \emptyset$, and
- $R_{i+1}(T) = \{A \in Args \mid A \text{ is rejected by } R_i(T) \text{ and } T\}$.

Definition 15. The **set of rejected arguments** w.r.t. T is $RArgs = \bigcup_{i=1}^{\infty} R_i(T)$. An argument is rejected if it is rejected w.r.t. $JArgs$.

From justified and rejected arguments with respect to justified arguments, we define rejected literals.

Definition 16. A literal l is a **rejected literal** by T iff there is no supportive argument for l in $Args - RArgs(T)$. A literal l is rejected if it is rejected by $JArgs$.

In reference to the above definition, we can note that the set of justified arguments $JArgs$ is included in $Args - RArgs(JArgs)$. Furthermore, some arguments in $Args - RArgs(T)$ may be neither justified nor rejected. Consequently, a literal may be neither justified nor rejected. In this case, we may say that the status of the literal is undetermined.

Example 3. Let us suppose two arguments A and B attacking each other. Argument A supports literal a , while argument B supports literal $\neg a$. The set of justified arguments is empty, and thus the set of rejected arguments is empty. Consequently, literals a and $\neg a$ are neither justified nor rejected. Their status is undetermined. \square

3 Conclusion

A deontic rule-based argumentation framework has been devised to capture normative knowledge and reasoning upon it. To do so, we have been inspired by works in DDL and extended the argumentation machinery developed in [7].

The main source of difficulties resided in the introduction of modal and deontic-preference operators. In particular, the introduction of modalities required to significantly modify the concept of argument and the basic system of [7]. Indeed, the derivation of a modal literal such as Ob depends on the constructive provability of b using rules such as $a_1, \dots, a_n \Rightarrow_O b$, and the derivation of $\neg Ob$ requires that there is no proof for Ob . We have thus devised argument agglomeration sets which, to the best of our knowledge, have no counterparts in the argumentation literature.

References

1. Antoniou, G., Billington, D., Governatori, G., Maher, M.J.: Representation results for De-feasible Logic. *ACM Trans. Comput. Logic* **2**, 255–287 (April 2001)
2. Atkinson, K., Baroni, P., Giacomin, M., Hunter, A., Prakken, H., Reed, C., Simari, G.R., Thimm, M., Villata, S.: Towards artificial argumentation. *AI Magazine* **38**(3), 25–36 (2017)
3. Beirlaen, M., Heyninck, J., Straßer, C.: Structured argumentation with prioritized conditional obligations and permissions. *Journal of Logic and Computation* p. exy005 (2018)
4. Beirlaen, M., Straßer, C.: A structured argumentation framework for detaching conditional obligations. *CoRR* **abs/1606.00339** (2016)
5. Dung, P.M.: On the acceptability of arguments and its fundamental role in nonmonotonic reasoning, logic programming and n-person games. *Artificial Intelligence* **77**(2), 321–358 (1995)
6. Gabbay, D., Horty, J., Parent, X., van der Meyden, R., van der Torre, L. (eds.): *Handbook of Deontic Logic and Normative Systems*. College Publications (2013)
7. Governatori, G., Maher, M.J., Antoniou, G., Billington, D.: Argumentation semantics for Defeasible Logic. *J. Log. Comput.* **14**(5), 675–702 (2004)
8. Governatori, G., Rotolo, A.: BIO logical agents: Norms, beliefs, intentions in Defeasible Logic. *Autonomous Agents and Multi-Agent Systems* **17**(1), 36–69 (2008)
9. Governatori, G., Rotolo, A., Calardo, E.: Possible world semantics for defeasible deontic logic. In: *Deontic Logic in Computer Science*. Springer (2012)
10. Governatori, G., Rotolo, A., Sartor, G.: Temporalised normative positions in defeasible logic. In: *Proc. of the 10th Int. Conf. on Artificial Intelligence and Law*. pp. 25–34. ACM (2005)
11. Lam, H., Governatori, G.: What are the necessity rules in defeasible reasoning? In: *Proc. of the 11th Int. Conf. on Logic Programming and Nonmonotonic Reasoning*. pp. 187–192 (2011)
12. Lam, H.P., Governatori, G., Riveret, R.: On ASPIC⁺ and Defeasible Logic. In: Baroni, P. (ed.) *Proc. of 6th Int. Conf. on Computational Models of Argument*. IOS Press, Amsterdam (2016)
13. Prakken, H., Sartor, G.: Law and logic: A review from an argumentation perspective. *Artif. Intell.* **227**, 214–245 (2015)
14. Prakken, H., Wyner, A.Z., Bench-Capon, T.J.M., Atkinson, K.: A formalization of argumentation schemes for legal case-based reasoning in ASPIC⁺. *J. Log. Comput.* **25**(5), 1141–1166 (2015)
15. Riveret, R., Governatori, G., Rotolo, A.: Argumentation semantics for temporal defeasible logic. In: *Proc. of the 3rd Starting AI Researchers’ Symposium*. pp. 267–268. IOS Press (2006)
16. Riveret, R., Prakken, H., Rotolo, A., Sartor, G.: Heuristics in argumentation: A game-theoretical investigation. In: *Proc. of the 2nd Int. Conf. on Computational Models of Argument*. pp. 324–335. IOS Press (2008)
17. Sartor, G., Rudnianski, M., Rotolo, A., Riveret, R., Mayor, E.: Why lawyers are nice (or nasty): a game-theoretical argumentation exercise. In: *Proc. of the 12th Int. Conf. on Artificial Intelligence and Law*. pp. 108–117. ACM (2009)