i-lex. Scienze Giuridiche, Scienze Cognitive e Intelligenza artificiale
Rivista quadrimestrale on-line: www.i-lex.it
Novembre 2012, numero 17

NARROWING CONCEPTS IN LEGAL REASONING

Antonino Rotolo*

1. Introduction

An important research issue in AI and law is to design computer systems whose performance is constrained by suitable sets of legal norms: in this sense, norms establish what legality criteria should apply to their functioning. The law is an adaptive and flexible normative system, where legal norms provide standards which can be violated, even though any violations should result in sanctions or other normative effects applying to non-compliant agents. To do that, it is necessary to monitor the behaviour of agents and enforce the sanctions. Norms thus allow agents to optimize their performance by reasoning about the trade off between respecting the norm - thus incurring in the related compliance costs - and the risk of being sanctioned. However, this additional flexibility of norms is not enough, because it could lead the agent to respect the norm (or otherwise to be sanctioned) even in circumstances where the respect of the norm does not give any advantage to the system, thus wasting his resources while the whole system achieves only a suboptimal state.

Norms are like plans which aim at achieving the social goals shared by the members of a society.[1] The legislator could try to specify all the circumstances which a norm applies to and all the exceptional contexts where it does not apply, but, as well known in the planning community of AI, universal plans rarely are a practicable strategy. An agent should rather produce a partial plan and revise it when part of it becomes unfeasible. In the same way as replanning allows an agent to revise its plans while keeping fixed its original goals, legal interpretation allows norms to be adapted after their creation to the unforeseen situations in order to achieve the social goal they have been planned for.

Thus, the research question of this paper is: How to logically model the interpretation mechanism of law that adjust concepts to concrete cases, so to design more flexible systems regulated by norms?

We started investigating this research issue in some earlier work.[2] To an-

*CIRSFID and Department of Law, Alma Mater Studiorum – Università di Bologna, Italy.

[1] G.BOELLA,L.VANDERTORRE, *Norm Negotiation in Multiagent Systems*, in: *International Journal of Cooperative Information Systems*, 16, 1, 2007, 97–122.

[2] G. Boella, G. Governatori, A. Rotolo, L. van der Torre, *Lex minus dixit quam voluit, lex magis dixit quam voluit: A Formal Study on Legal Compliance and Interpretation*, in: P. Casanovas et al., eds.: *AI Approaches to the Complexity of Legal Systems*, Springer, Berlin, 2010, 162–183; G. Boella, G. Governatori, A. Rotolo, L. van der Torre, *A Logical Understanding of Legal Interpretation*, in: F. Lin et al., eds.: *Principles of Knowledge Representation and Reasoning: Proceedings of the Twelfth International Conference, KR 2010*, AAAI Press, Menlo Park, 2010, 563–565; A. Rotolo, C. Roversi, *Constitutive Rules and Coherence in Legal Argumentation: The Case of Extensive and Restrictive*

Suggested citation: Antonino Rotolo, Narrowing Concepts in Legal Reasonng
in: i-lex, 17, 2012, pp.249-266 (www.i-lex.it)
Invited paper

swer this question we worked with a modal extension of Defeasible Logic (DL, henceforth)[3] and we conceptually used the well-known distinction of constitutive (or counts-as) and regulative norms (henceforth, legal rules).The idea was that interpreting norms may require to revise theories of constitutive rules that characterize the concepts occurring in legal rules. Indeed, legal concepts can be inferentially characterized by arbitrarily large and connected theories of constitutive rules, and so, when we broaden or narrow the scope of legal concepts we are doing nothing but changing (revising or contracting) those theories. The advantages of this approach are thus that it allows us to make these interpretive arguments more transparent and to show interesting connections with techniques from the domain of revision theory.

However, our earlier analysis is defective in at least two respects:

- it is based on a quite cumbersome proof-theoretic machinery: here, we try to rethink the revision procedures in an argumentation setting;
- no real investigation of formal properties of the revision operations is presented in our earlier work: here, we discuss more technical options and present some first results suggesting that standard AGM model[4] can be sometimes unsuitable.

In this paper we confine our attention to only one type of revision: the case when the applicability scope of any legal rule is restricted.

The layout of the paper is as follows. Section 2 briefly recalls the main idea of we developed elsewhere and provides an informal background. Section 3 presents a variant of DL, which is the logical framework we use to revise theories: we first study the proof theory; this choice allows us to easily develop an argumentation semantics for that logic. With this done, we provide a first definition of applicability restriction (Section 4)Some technical discussion about possible alternatives and results are then offered in Section 5. Such alternatives determine the possibility that different revision outcomes are obtained: we briefly outline for future research some criteria that can be used to choose the best theory among these outcomes (Section 6).

2. The Background

As well known, norms have typically a conditional structure such as $b_1,\dots,b_n \Rightarrow_{\mathrm{OBL}} l$ (if $b_1,\dots,b_n$ hold, then l is obligatory); an agent is compliant with respect to this norm if l is obtained whenever $b_1,\dots,b_n$ is derived. Due to the complexities and dynamics of the world, norms cannot take into account all the possible conditions where they should or should not be applied, giving rise to the so called "penumbra".[5] After all, not only the world changes, giving

Interpretation, in C. Dahlman, E. Feteris, eds.: *Legal Argumentation Theory: Cross-Disciplinary Perspectives*, Springer, Dordrecht, forthcoming.

[3] G. Governatori, A. Rotolo, *BIO Logical Agents: Norms, Beliefs, Intentions in Defeasible Logic*, in: *Agents and Multi-Agent Systems*, 17, 1, 2008, 36–69.

[4] Carlos E. Alchourrón, Peter Gärdenfors, David Makinson, *On the Logic of Theory Change: Partial Meet Contraction and Revision Functions*, in: *Journal of Symbolic Logic*, 50, 2, 1985, 510–530.

[5] H. L. A. Hart, *The Concept of Law*, 2nd ed., Clarendon, Oxford, 1994, chap. 7; G. Sartor, *Legal Reasoning: A Cognitive Approach to the Law*, Springer, Dordrecht, 2005.

rise to circumstances unexpected to the legislator who introduced the norm, but even the ontology of reality can change with respect to the one constructed by the law to describe the applicability conditions of norm.

Normative systems regulating real societies have two mechanisms to cope with this problem. First they distinguish regulative legal rules (obligations, prohibitions and permissions; henceforth, legal rules) from constitutive rules. While the former, which are usually changed only by the legislative system, specify the ideal behaviour, the latter ones provide, by means of counts-as definitions, an ontology of institutional concepts. The applicability conditions of legal rules very often refer to these institutional concepts, rather than to so called brute facts.

Second, the judicial system, at the moment in which a case concerning a violation is discussed in court, can be empowered to interpret, i.e., to change norms, under some restrictions not to go beyond the purpose from which the legal rules stem. The distinction between legal and constitutive rules (ontology vs norms) suggests that legal interpretation does not amount to revising norms, but to interpreting legal concepts, i.e., to revising constitutive rules.

The ontology of legal concepts is built via constitutive rules having the so-called counts-as form[6]: $r: a_1, \ldots, a_n \Rightarrow_c b$. For example, a bicycle is considered as a vehicle by the following constitutive rule:

$$r_0 : Bike \Rightarrow_c Vehicle.$$

Constitutive rules have a defeasible character, for example, a bicycle for children cannot be considered as a vehicle:

$$r_1 : Bike, ForChildren \rightsquigarrow_c \neg Vehicle,$$

where $r_1 \succ r_0$. As usual in DL, our language includes (1) a superiority relation $\succ$ that establishes the relative strength of rules and is used to solve conflicts, (2) special rules marked with $\rightsquigarrow$, called defeaters, which are not meant to derive conclusions, but to provide reasons against the opposite.

The set of legal rules is kept to be fixed: any judge during the interpretation process can argue about their applicability conditions but cannot either add new rules nor cancel them. Only legislators have the power to change legal rules.

Legal rules have the form $r: b_1, \ldots, b_n \Rightarrow_{\mathrm{OBL}} l$, for example:

$$r_2 : Vehicle, Park \Rightarrow_{\mathrm{OBL}} \neg Enter.$$

This rule reads as follows: if we have a vehicle and we are in a park, then it is (defeasibly) forbidden to enter.

Finally, as usually assumed by many legal theorists[7], we assign goals to legal rules. In the social delegation cycle[8] norms are planned starting from goals shared by the community of agents. However, such goals play also another role: they pose the limits within which the interpretation process of the judicial systems must stay when interpreting norms.

[6] J.R. Searle, *The Construction of Social Reality*, The Free Press, New York, 1996.

[7] For a discussion, see G. Sartor, *Legal Reasoning: A Cognitive Approach to the Law*, Springer, Dordrecht, 2005.

[8] G. Boella, L. van der Torre, *Norm Negotiation in Multiagent Systems*, in: *International Journal of Cooperative Information Systems*, 16, 1, 2007, 97–122.

Note that the goal alone is not sufficient to specify a norm, since there could be many ways to achieve that goal and some guidance should be given to the citizens. Thus, the norm works like a partial plan the legislator sets up in advance. The judicial system is left with the task of dynamically adapting the applicability of the regulative norm by revising the constitutive norms referring to its applicability conditions, in order to fulfil the goal of the norm also under unforeseen circumstances.

We define a set G of goals and a function $\mathcal{G}$ which maps legal rules into elements of G. For example, if $\mathcal{G}(r_2) = road\_safety$, this means that the goal of the rule prohibiting to enter into parks is to promote road safety.

Checking legal compliance requires to establish if a legal rule $r : b_1, \ldots, b_n \Rightarrow_{\mathrm{OBL}} l$ is violated by a fact or action l happened under some circumstances H. Let us assume that r states that $\neg l$ ought to be the case. However, l is not necessarily a violation, because we also have to check whether H, via the constitutive rules, matches with the applicability conditions $b_1, \ldots, b_n$ of r. In easy cases, these match and l directly amount to a violation. However, jurists argue that we have cases where this does not hold, as for example when there is a discrepancy between the literal meaning of $b_1, \ldots, b_n$ and the goal assigned to the rule r by the legislator. If so, even though H matches with $b_1, \ldots, b_n$, we do not have a violation because H should not match with $b_1, \ldots, b_n$. A non-literal interpretation of $b_1, \ldots, b_n$ would exclude H as a circumstance falling within the scope of r, since the goal of the norm would be achieved anyway: *lex magis dixit quam voluit*, the law said more than what the legislator was meaning to say. Analogously, not all cases in which H mismatches with $b_1, \ldots, b_n$ are not violations. We could have that *lex minus dixit quam voluit*, the law said less than what the legislator was meaning to say: here a non-literal, goal-based interpretation of r would lead to broaden its applicability scope to match H, thus making the agent a violator.

In this paper, we consider only the case when "the law said more than what the legislator was meaning to say". Here is an example illustrating the main idea.

Example 1. *Suppose Mary enters a park with her bike, thus apparently violating rule r_2 above about vehicles' circulation. Police stops her when she is still on her bike in the park and fines her. Mary thinks this is unreasonable and sues the municipality because she thinks that here the category "vehicle" should not cover bikes.*

The conceptual domain T of the normative system, corresponding to a set of constitutive rules, allows us to derive that any bike a is indeed a vehicle. The goal of the norm r_2 is reducing pollution $\mathcal{G}(r_2) = \neg pollution$. In court, the judge has to establish if Mary violated r_2 or not.

If T is the case, the judge could argue that r_2 clearly applies to Mary:

$$T = \{r_0 : Bike \Rightarrow_c Vehicle,\ r_3 : 2\_wheels, Transport, \neg Engine \Rightarrow_c Bike\}$$

But suppose that the judge can show that, if Mary's case fulfils the applicability conditions of r_2 (Mary's bike is a vehicle) then a goal which is incompatible with the goal assigned to r_2 would be promoted. Since $\mathcal{G}(r_2) = \neg pollution$, prohibiting to circulate with bikes in parks would encourage people to get around parks by *car and then walk. This would be against the goal of r_2 and so the judge has good reasons to exclude that bikes are vehicles when r_2 should be applied. Accordingly, when arguing in this way, the judge may interpret r_2 by reducing its applicability conditions as far as Mary's case is concerned. He thus contracts T in order to obtain in T that Mary's bike is not a vehicle in the context of the current situation, by adding a defeater r_4 blocking the Vehicle conclusion: r_4 : $Bike, Park \rightsquigarrow_c \neg Vehicle$ and by stating that r_4 is stronger than r_0: $\succ = \{r_4 \succ r_0\}$.*

3. The Logical Framework

3.1. Proof Theory

The following framework is an extension of DL, which slightly revises earlier works[9] While counts-as rules do not prove modalised literals, the system develops a constructive account of those modalities that rather correspond to obligations and goals: rules for these concepts are thus meant to devise suitable logical conditions for introducing modalities. For example, while a counts-as rule such as $a_1, \ldots, a_n \Rightarrow_c b$, if applicable, will basically support the conclusion of b, rules such as $a_1, \ldots, a_n \Rightarrow_{\mathsf{OBL}} b$ and $d_1, \ldots, d_n \Rightarrow_{\mathsf{Goal}} e$ if applicable, will allow for deriving $\mathsf{OBL}b$ and $\mathsf{Goal}e$, meaning the former that b is obligatory, the latter that e is a goal promoted by the facts used to derive it.

In our language, for $X \in \{c, \mathsf{OBL}, \mathsf{Goal}\}$, *strict rules* have the form $\phi_1, \ldots, \phi_n \rightarrow_X \psi$. *Defeasible rules* have the form $\phi_1, \ldots, \phi_n \Rightarrow_X \psi$. A rule of the form $\phi_1, \ldots, \phi_n \rightsquigarrow_X \psi$ is a *defeater*. Accordingly, the mode denoted by X determines the type of conclusion one can obtain, and the three types of rules establish the strength of the relationship. Strict rules provide the most stronger connection between a set of premises and their conclusion: whenever the premises are deemed as indisputable so is the conclusion; then we have defeasible rules: a defeasible rule, given the premises, allows us to derive the conclusion unless there is evidence for its contrary; finally we have defeaters. A defeater suggests that there is a connection between its premises and the conclusion, but this connection is not strong enough to warrant the conclusion on its own; on the other hand a defeater shows that there is some evidence for the conclusion, thus it can be used to defeat rules for the opposite conclusion.

Definition 1 (Language)**.** *Let* PROP *be a set of propositional atoms,* $\mathsf{MOD} = \{c, \mathsf{OBL}, \mathsf{Goal}\}$*, and* Lbl *be a set of labels. The sets defined below are the smallest sets closed under the given construction conditions:*

Literals

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

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

Modal literals

$$\mathsf{ModLit} = \{Xl, \neg Xl | l \in \mathsf{Lit}, X \in \{\mathsf{OBL}, \mathsf{Goal}\}\}$$

[9] G. Governatori, A. Rotolo, *BIO Logical Agents: Norms, Beliefs, Intentions in Defeasible Logic*, in: *Agents and Multi-Agent Systems*, 17, 1, 2008, 36–69; G. Governatori, A. Rotolo, *A Computational Framework for Institutional Agency*, in: *Artificial Intelligence and Law*, 16, 1, 2008, 25–52.

Rules Rul $= (\mathrm{Rul}_{\mathrm{s}}^{X} \cup \mathrm{Rul}_{\mathrm{s}}^{c}) \cup (\mathrm{Rul}_{\mathrm{d}}^{X} \cup \mathrm{Rul}_{\mathrm{d}}^{c}) \cup (\mathrm{Rul}_{dft}^{X} \cup \mathrm{Rul}_{dft}^{c})$, *where* $X \in \{\mathrm{OBL}, \mathrm{Goal}\}$*, such that*

$$\begin{aligned}
\mathrm{Rul}_{\mathrm{s}}^{X} &= \{r : a_1,\dots,a_n \rightarrow_X b | r \in \mathrm{Lbl}, \{a_1,\dots,a_n\} \subseteq \mathrm{Lit} \cup \mathrm{ModLit}, b \in \mathrm{Lit}\}\\
\mathrm{Rul}_{\mathrm{s}}^{c} &= \{r : a_1,\dots,a_n \rightarrow_c b | r \in \mathrm{Lbl}, \{a_1,\dots,a_n\} \subseteq \mathrm{Lit}, b \in \mathrm{Lit}\}\\
\mathrm{Rul}_{\mathrm{d}}^{X} &= \{r : a_1,\dots,a_n \Rightarrow_X b | r \in \mathrm{Lbl}, \{a_1,\dots,a_n\} \subseteq \mathrm{Lit} \cup \mathrm{ModLit}, b \in \mathrm{Lit}\}\\
\mathrm{Rul}_{\mathrm{d}}^{c} &= \{r : a_1,\dots,a_n \Rightarrow_c b | r \in \mathrm{Lbl}, \{a_1,\dots,a_n\} \subseteq \mathrm{Lit}, b \in \mathrm{Lit}\}\\
\mathrm{Rul}_{dft}^{X} &= \{r : a_1,\dots,a_n \rightsquigarrow_X b | r \in \mathrm{Lbl}, \{a_1,\dots,a_n\} \subseteq \mathrm{Lit} \cup \mathrm{ModLit}, b \in \mathrm{Lit}\}\\
\mathrm{Rul}_{dft}^{c} &= \{r : a_1,\dots,a_n \rightsquigarrow_X b | r \in \mathrm{Lbl}, \{a_1,\dots,a_n\} \subseteq \mathrm{Lit}, b \in \mathrm{Lit}\}
\end{aligned}$$

We use $A(r)$ *to denote the set of* antecedents *of the rule r, and* $C(r)$ *to denote its* consequent*. Other obvious abbreviations are, for example,* $\mathrm{Rul}[b]$ *to denote the set of rules whose consequent is b and*

$$\begin{aligned}
\mathrm{Rul}^{c} &= \{r : a_1,\dots,a_n \hookrightarrow_c b | \hookrightarrow \in \{\rightarrow, \Rightarrow, \rightsquigarrow\}\}\\
\mathrm{Rul}_{\mathrm{sd}} &= \{r : a_1,\dots,a_n \hookrightarrow_X b | X \in \mathrm{MOD}, \hookrightarrow \in \{\rightarrow, \Rightarrow\}\}\\
\mathrm{Rul}_{\mathrm{s}}[b] &= \{r : a_1,\dots,a_n \rightarrow_X b | X \in \mathrm{MOD}\}
\end{aligned}$$

A normative theory is the knowledge base which is used to reason about the applicability of legal rules included in the theory itself.

Definition 2 (Normative Theory)**.** *A normative theory is a structure* $D = (F, R^c, R^{\mathrm{OBL}}, R^{\mathrm{Goal}}, \succ)$*, where*

- $F \subseteq \mathrm{Lit} \cup \mathrm{ModLit}$ *is a finite set of facts;*
- $R^c \subseteq \mathrm{Rul}^c$ *is a finite (non-empty) set of counts-as rules,*
- $R^{\mathrm{OBL}} \subseteq \mathrm{Rul}^{\mathrm{OBL}}$ *is a finite set of obligation rules,*
- $R^{\mathrm{Goal}} \subseteq \mathrm{Rul}^{\mathrm{Goal}}$ *is a finite set of goal rules,*
- $\succ$ *is an acyclic relation (called superiority relation) defined over* $(R^c \times R^c) \cup (R^{\mathrm{OBL}} \times R^{\mathrm{OBL}}) \cup (R^{\mathrm{Goal}} \times R^{\mathrm{Goal}})$.

Proofs are sequences of literals and modal literals together with the so-called proof tags $+\Delta$, $-\Delta$, $+\partial$ and $-\partial$. If $X \in \{c, \mathrm{OBL}, \mathrm{Goal}\}$, given a normative theory D, $+\Delta^X q$ means that literal q is provable in D with the mode X using only facts and strict rules, $-\Delta^X q$ means that it has been proved in D that q is not definitely provable in D with the mode X, $+\partial^X q$ means that q is defeasibly provable in D with the mode X, and $-\partial^X q$ means that it has been proved in D that q is not defeasibly provable in D with the mode X.

Definition 3. *Given a normative theory D, a proof in D is a linear derivation, i.e, a sequence of labelled formulas of the type* $+\Delta^X q$*,* $-\Delta^X q$*,* $+\partial^X q$ *and* $-\partial^X q$*, where the proof conditions defined in the rest of this section hold.*

Definition 4. *Let D be a normative theory. Let* $\# \in \{\Delta, \partial\}$ *and* $X \in \{\mathrm{OBL}, \mathrm{Goal}\}$*, and* $P = (P(1),\dots,P(n))$ *be a proof in D. A literal/modal literal q is* #-provable *in P if there is a line* $P(m)$*,* $1 \leq m \leq n$*, of P such that either*

1. *q is a literal and $P(m) = +\#^c q$ or*
2. *q is a modal literal Xp and $P(m) = +\#^X p$ or*
3. *q is a modal literal $\neg Xp$ and $P(m) = -\#^X p$.*

*A literal q is $\#$-*rejected *in P if there is a line $P(m)$ of P such that*

1. *q is a literal and $P(m) = -\#^c q$ or*
2. *q is a modal literal Xp and $P(m) = -\#^X p$ or*
3. *q is a modal literal $\neg Xp$ and $P(m) = +\#^X p$.*

The definition of Δ^X describes just forward (monotonic) chaining of strict rules[10]:

$+\Delta^X$: If $P(n+1) = +\Delta^X q$ then
(1) $q \in F$ if $X = c$ or $Xq \in F$ or
(2) $\exists r \in R_s^X[q]:\ \forall a \in A(r)\ a$ is Δ-provable.

To show that a literal q is defeasibly provable with the mode X we have two choices: (1) We show that q is already definitely provable; or (2) we need to argue using the defeasible part of a normative theory D. For this second case, some (sub)conditions must be satisfied. First, we need to consider possible reasoning chains in support of $\sim q$ with the mode X, and show that $\sim q$ is not definitely provable with that mode (2.1 below). Second, we require that there must be a strict or defeasible rule with mode X for q which can be applied (2.2 below). Third, we must consider the set of all rules which are not known to be inapplicable and which permit to get $\sim q$ with the mode X (2.3 below). Essentially, each such a rule s attacks the conclusion q. For q to be provable, s must be counterattacked by a rule t for q with the following properties: (i) t must be applicable, and (ii) t must prevail over s. Thus each attack on the conclusion q must be counterattacked by a stronger rule. In other words, r and the rules t form a team (for q) that defeats the rules s.

$+\partial^X$: If $P(n+1) = +\partial^X q$ then
(1)$+\Delta^X q \in P(1..n)$ or
(2) (2.1) $-\Delta^X \sim q \in P(1..n)$ and
(2.2) $\exists r \in R_{sd}^X[q]$ such that $\forall a \in A(r)\ a$ is ∂-provable, and
(2.3) $\forall s \in R^X[\sim q]$ either $\exists a \in A(s)$ such that a is ∂-rejected, or
(2.3.1) $\exists t \in R^X[q]$ such that $\forall a \in A(r)\ a$ is ∂-provable and $t \succ s$

Definition 5. *Given a normative theory D, $D \vdash \pm\#^X l$, where $\# \in \{\Delta, \partial\}$ and $X \in \{c, \mathrm{OBL}, \mathrm{Goal}\}$, iff there is a proof $P = (P(1), \ldots, P(n))$ in D such that $P(n) = \pm\#^X l$.*

[10] For space reasons, in the remainder we present only the proof conditions for $+\Delta$ and $+\partial$. Conditions for the negative tags can be easily obtained using the notion of *#-rejected*.

Definition 6. *Given a normative theory D, the* universe *of D (U^D) is the set of all the atoms occurring in D. The* extension *of D (E^D), is defined as follows*[11]*:*

$$E^D = (\Delta_D^+, \Delta_D^-, \partial_D^+, \partial_D^-)$$

where for $X \in \{\mathrm{OBL}, \mathrm{Goal}\}$

$$\Delta_D^+ = \{Xl : D \vdash +\Delta^X l\} \cup \{l : D \vdash +\Delta^c l\}; \qquad \partial_D^+ = \{Xl : D \vdash +\partial^X l\} \cup \{l : D \vdash +\partial^c l\};$$
$$\Delta_D^- = \{Xl : D \vdash -\Delta^X l\} \cup \{l : D \vdash -\Delta^c l\}; \qquad \partial_D^- = \{Xl : D \vdash -\partial^X l\} \cup \{l : D \vdash -\partial^c l\}.$$

It is worth noting that the logic enjoys the following properties.

Proposition 1. *Let D be a normative theory where the transitive closure of* $\succ$ *is acyclic. For every* $\# \in \{\Delta, \partial\}$*,* $X \in \{c, \mathrm{OBL}, \mathrm{Goal}\}$*:*

- *it is not possible that both* $D \vdash +\#^X p$ *and* $D \vdash -\#^X p$*;*
- *if* $D \vdash +\partial^X p$ *and* $D \vdash +\partial^X {\sim}p$*, then* $D \vdash +\Delta^X p$ *and* $D \vdash +\Delta^X {\sim}p$*.*

Theorem 1. *For every normative theory D, the extension* E^D *can be computed in time linear to the size of the theory, i.e.,* $O(|U^D| * |R|)$*.*

3.2. The Argumentation Semantics

Let us now provide an argumentation semantics for the logic presented in the previous section.[12]

Definition 7 (Argument and Supportive Argument)**.** *An* argument $\mathcal{A}$ *for a conclusion p generated from a normative theory* $D = (F, R^c, R^{\mathrm{OBL}}, R^{\mathrm{Goal}}, \succ)$ *is a (possibly infinite) tree with the following structure:*

1. *all nodes are labeled by elements* $h \in \mathrm{Lit} \cup \mathrm{ModLit}$*;*
2. *the root is labeled by p;*
3. *if* $h_1, \ldots, h_n$ *label the children of h then, either*
   (a) *if* $h \in \mathrm{Lit}$ *then there is a rule r in* R^c *with body* $b_1, \ldots, b_n$ *(where* $h_1 = b_1, \ldots, h_n = b_n$*) and head q (where* $h = q$*), or*
   (b) *if* $h = Xl \in \mathrm{ModLit}$ *then there is a rule r in* R^X *with body* $b_1, \ldots, b_n$ *(where* $h_1 = b_1, \ldots, h_n = b_n$*) and head q (where* $h = q$*),*

   such that all arcs connecting $h_1, \ldots h_n$ *to h are labeled by the rule r;*
4. *if a rule r labeling an arc is a defeater, then* $h = p$ *is the root of the argument;*
5. *the leaf nodes are labeled by either elements of F or by modal literals having the form* $\neg Xl$*.*

Some auxiliary terminology:

[11] When clear from the context, we will omit the subscript *D* in $\Delta_D^+, \Delta_D^-, \partial_D^+, \partial_D^-$.
[12] Cf. G. Governatori, M.J. Maher, D. Billington, G. Antoniou, *Argumentation Semantics for Defeasible Logics*, in: *Journal of Logic and Computation*, 14, 5, 2004, 675–702.

- *The nodes labeled by modal literals having the form* $\neg Xl$ *are called* open nodes*.*
- *A* supportive argument *is a finite argument in which no defeater is used.*
- *An argument is* positive *iff no defeater is used in it.*
- *A* strict argument *is an argument in which only strict rules are used.*
- *An argument that is not strict is called* defeasible*.*

Definition 8 (Top subargument of an argument)**.** *Let* $\mathcal{A}$ *any argument with height* $j \geq 1$ *for any literal* p*. The top subargument* $\mathcal{A}^t$ *of* $\mathcal{A}$ *is the top subargument of* $\mathcal{A}$ *with height* 1*. Let us use* $R(\mathcal{A}^t)$ *to denote the rule associated with the arcs arriving at the root of* $\mathcal{A}^t$*.*

In characterizing the argumentation semantics for our logic we disregard the superiority relation to keep the discussion and the technicalities simple. There are ways according to which this restriction does not affect the generality of the approach. For example, it is possible to give a modular transformation that empties the superiority relation while maintaining the same conclusions in the language of the original theory.[13]

Definition 9. *Let* $l \in \mathrm{Lit} \cup \mathrm{ModLit}$*. The set* $\mathrm{Compl}(l)$ *of complementary literals of* l *is defined as follows:*

- *if* $l = p \in \mathrm{Lit}$ *then* $\mathrm{Compl}(l) = \{\sim l\}$*;*
- *if* $l = Xp \in \mathrm{ModLit}$ *then* $\mathrm{Compl}(l) = \{\neg Xp, X{\sim}p\}$*.*

Definition 10 (Attack)**.** *An argument* $\mathcal{A}_i$ *attacks an argument* $\mathcal{A}_j$ *iff there is a literal/modal literal* l *such that*

1. $h = l$*, where* h *is a node in* $\mathcal{A}_i$*;*
2. $l \in \mathrm{Compl}(m)$*, where* $h' = m$ *such that* h' *is a node in* $\mathcal{A}_j$*.*

A set of arguments S *attacks an argument* $\mathcal{A}_j$ *if there is an argument* $\mathcal{A}_i$ *in* S *that attacks* $\mathcal{A}_j$*.*

Definition 11 (Support and Undercut)**.** *A set of arguments* S supports *an argument* $\mathcal{A}$ *if every proper subargument of* $\mathcal{A}$ *is in* S*.*

An argument $\mathcal{A}_i$ *is* undercut *by a set of arguments* S *if* S *supports an argument* $\mathcal{A}_j$ *attacking a proper subargument of* $\mathcal{A}_i$*.*

Definition 12 (Acceptable and Rejected Arguments)**.** *An argument* $\mathcal{A}$ *for* l *is* acceptable *w.r.t a set of arguments* S *if* $\mathcal{A}$ *is finite and every argument attacking* $\mathcal{A}$ *is undercut by* S*.*

An argument $\mathcal{A}$ *is* rejected *by sets of arguments* S *and* T *when*

1. *a proper subargument* $\mathcal{A}_s$ *of* $\mathcal{A}$ *is in* S*, or*

[13] Cf. G. Governatori, A. Rotolo, *BIO Logical Agents: Norms, Beliefs, Intentions in Defeasible Logic*, in: *Agents and Multi-Agent Systems*, 17, 1, 2008.

2. *$\mathcal{A}_s$ is attacked by an argument supported by T.*

Definition 13 (Rejected Arguments and Rejected Literal/Modal Literal)**.** *Let D be a normative theory and T a set of arguments.*

We define $R_i^D(T)$ as follows.

- $R_0^D(T) = \emptyset$;
- $R_{i+1}^D(T) = \{\mathcal{A} \in Args_D | \mathcal{A}$ *is rejected by* $R_i^D(T)$ *and* $T\}$.

The set of rejected arguments *in a normative theory D w.r.t. T is $RArgs^D = \bigcup_{i=1}^{\infty} R_i^D$. An argument is rejected if it is rejected w.r.t. $JArgs^D$.*

A literal/modal literal l is rejected by T if there is no supportive argument for l in $Args_D - RArgs^D(T)$. A literal/modal literal l is rejected if it is rejected by $JArgs^D$.

Definition 14 (Grounded Arguments)**.** *Let $D = (F, R^c, R^{\mathrm{OBL}}, R^{\mathrm{Goal}}, \succ)$ be a normative theory. An argument $\mathcal{A} \in Args_D$ is* grounded *iff, for each leaf node h in it, either $h \in F$ or, if $h = \neg Xp$ then Xp is rejected.*

Definition 15 (Justified Argument and Justified Literal/Modal Literal)**.** *Given any normative theory D, let $Args_D$ be the set of arguments that can be generated from D.*

We define J_i^D as follows.

- $J_0^D = \emptyset$;
- $J_{i+1}^D = \{\mathcal{A} \in Args_D | \mathcal{A}$ *is acceptable w.r.t.* J_i^D *and is grounded*$\}$.

The set of justified arguments *in a normative theory D is $JArgs^D = \bigcup_{i=1}^{\infty} J_i^D$.*

A literal/modal literal l is justified if it is a conclusion of a supportive argument in $JArgs^D$.

Theorem 2 (Characterization of the Logic in Argumentation Semantics)**.** *Let D be a normative theory, $X \in \{\mathrm{OBL}, \mathrm{Goal}\}$, and $l \in \mathrm{Lit} \cup \mathrm{ModLit}$.*

1. *$D \vdash +\Delta^X p$ iff there is a strict supportive argument for $l = Xp$ in $Args_D$.*
2. *$D \vdash +\Delta^c p$ iff there is a strict supportive argument for $l = p$ in $Args_D$.*
3. *$D \vdash -\Delta^X p$ iff there is no (finite or infinite) strict argument for $l = Xp$ in $Args_D$.*
4. *$D \vdash -\Delta^c p$ iff there is no (finite or infinite) strict argument for $l = p$ in $Args_D$.*
5. *$D \vdash +\partial^X p$ iff $l = Xp$ is justified.*
6. *$D \vdash +\partial^c p$ iff $l = p$ is justified.*
7. *$D \vdash -\partial^X p$ iff $= Xp$ is rejected by $JArgs^D$.*
8. *$D \vdash -\partial^c p$ iff $l = p$ is rejected by $JArgs^D$.*

Sketch. The proof can be obtained by easily extending earlier results.[14] Indeed, the addition of modal operators does not technically make things harder, but only requires to also consider when modal literals are obtained with rules for $X \in \{\mathrm{OBL}, \mathrm{Goal}\}$. The only significant difference is that supportive arguments that make use of rules with negative modal literals like $\neg Xl$ in their antecedents require that Xp is rejected, a case which is covered in Definition 15 by requiring that all justified arguments be grounded. □

4. The Contraction of Norm Applicability

In Section we have recalled the basic ideas presented elsewhere and have informally illustrated how to revise legal concepts in order to make legal rules applicable or to block their applicability. Here, we focus on the operation of contraction. We first show how to frame this operation by making use of the argumentation semantics presented in Section . This requires to preliminarily introduce some notions which we will use in the remainder (extended normative theory, goal demotion and promotion).

Definition 16 (Extended Normative Theory)**.** *An* extended normative theory *is a structure* $\langle (F, R^c, R^{\mathrm{OBL}}, R^{\mathrm{Goal}}, \succ), G, \mathcal{G}, > \rangle$*, where*

- $(F, R^c, R^{\mathrm{OBL}}, R^{\mathrm{Goal}}, \succ)$ *is a normative theory;*
- $G \subseteq \mathrm{Lit}$ *is a set of goals;*
- $\mathcal{G} : R^{\mathrm{OBL}} \mapsto G$ *is a function assigning a goal to each obligation rule;*
- $>$ *is a partial order over* G *defining the relative importance of the rule goals.*

Definition 17 (Goal Demotion and Promotion)**.** *Let* $\langle (F, R^c, R^{\mathrm{OBL}}, R^{\mathrm{Goal}}, \succ), G, \mathcal{G}, > \rangle$ *be an extended normative theory. Let* $H = \{f_1, \ldots, f_m\} \subseteq \mathrm{Lit} \cup \mathrm{ModLit}$ *be a set of facts,* $r : b_1, \ldots, b_n \hookrightarrow_{\mathrm{OBL}} l \in R^{\mathrm{OBL}}_{\mathrm{sd}}$ *be a regulative legal rule, and g the goal of r.*

H demotes the goal g *iff*

- *there is no justified argument for* $\mathrm{Goal}\neg g$ *if* $H \cap F = \emptyset$*;*
- *there is at least one justified argument for* $\mathrm{Goal}\neg g$ *if* $H \subseteq F$*.*

H promotes the goal g *iff*

- *there is no justified argument for* $\mathrm{Goal}\, g$ *if* $H \cap F = \emptyset$*;*
- *there is at least one justified argument for* $\mathrm{Goal}\, g$ *if* $H \subseteq F$*.*

The procedure for contracting the theories that characterize legal concepts can be framed in general as follows.

[14] G. Governatori, M.J. Maher, D. Billington, G. Antoniou, *Argumentation Semantics for Defeasible Logics*, in: *Journal of Logic and Computation*, 14, 5, 2004: Proposition 5 and Theorem 17.

Definition 18 (Applicability Restriction)**.** *Let $\langle (F, R^c, R^{\mathrm{OBL}}, R^{\mathrm{Goal}}, \succ), G, \mathcal{G}, > \rangle$ be an extended normative theory, $r : b_1, \dots, b_n \hookrightarrow_{\mathrm{OBL}} l \in R^{\mathrm{OBL}}_{sd}$, g be the goal of r, and $H = \{f_1, \dots, f_m\} \subseteq F$. If*

1. *there exists at least one justified argument $\mathcal{A}$ whose top subargument $\mathcal{A}^t$ is such that $R(\mathcal{A}^t) = r$; and*
2. *either*
   - *(a) $H \cup \{\neg l\}$ and $H \cup \{l\}$* demote *g and there exists any b_k, $1 \leq k \leq n$, such that b_k is a node in each justified argument $\mathcal{C}$ for* $\mathrm{Goal}\neg g$ *when either $H \cup \{\neg l\} \subseteq F$ or $H \cup \{l\} \subseteq F$; or*
   - *(b) $H \cup \{\neg l\}$* promotes *g and $H \cup \{l\}$* demotes *g, and there exists any b_k, $1 \leq k \leq n$, such that b_k is a node in each justified argument $\mathcal{C}$ for* $\mathrm{Goal}g$ *when $H \cup \{\neg l\} \subseteq F$, or b_k is a node in each justified argument $\mathcal{C}'$ for* $\mathrm{Goal}\neg g$ *when $H \cup \{l\} \subseteq F$; or*
   - *(c) $H \cup \{\neg l\}$ and $H \cup \{l\}$* promote *g and there exists a b_k, $1 \leq k \leq n$, such that such that b_k is a node in each justified argument $\mathcal{C}$ for* $\mathrm{Goal}g$ *when either $H \cup \{\neg l\} \subseteq F$ or $H \cup \{l\} \subseteq F$;*

then the restriction of the applicability conditions of r with respect to the case H amounts to revising R^c and $\succ$ in $\langle (F, R^c, R^{\mathrm{OBL}}, R^{\mathrm{Goal}}, \succ), G, \mathcal{G}, > \rangle$ as follows:

$$R'^c = R^c \cup \{r : f_1, \dots, f_m \leadsto_c \sim b_k\} \qquad \succ' = \succ \cup \{r \succ t | \forall t \in R^c[b_k]\}$$

The normative theory $(F, R^c, R^{\mathrm{OBL}}, R^{\mathrm{Goal}}, \succ)$ in the resulting extended normative theory is denoted by D' and is such that

3. *for any other goal g', there is no argument $\mathcal{A} \in JArgs^{D'}$ for* $\mathrm{Goal}\neg g'$ *when $H \subseteq F$, where*
   - *$g \not> g'$, and*
   - *$g' = \mathcal{G}(z)$ where $z \in R^{\mathrm{OBL}}_{\mathrm{sd}}$ and there exists at least one argument $\mathcal{B} \in JArgs^{D'}$ whose top subargument $\mathcal{B}^t$ is such that $R(\mathcal{B}^t) = z$ when $H \subseteq F$.*

Remark 1. *We have three situations in which the applicability conditions of a regulative legal rule should be intuitively restricted (sub-conditions (a), (b), and (c) under point 2 above).*

First, we have that the goal of r is demoted not only by $\neg l$ but also by complying with the regulative legal rule; moreover, there exists at least one of the antecedents of the legal rule which is used in all goal arguments, in which either l or $\neg l$ occur, to prove $\neg g$ (the goal violation of the legal rule). Thus, we have reasons to block the constitutive rules supporting this antecedent.

Second, we have that the violation of r ($\neg l$) unexpectedly promotes the goal g of this rue, while compliance (l) doesn't; moreover, there exists at least one of the antecedents of the legal rule which is used in all goal arguments, in which l occurs, to support g (the goal promotion of the legal rule), while an antecedent is used in all goal arguments, in which $\neg l$ occurs, to support $\neg g$ (the *goal demotion of the legal rule). Thus, we have reasons to block the constitutive rules supporting such antecedents.*

Finally, we have that the goal of the norm is promoted independently from the fulfillment or violation of the norm.

The resulting revision is subject to the final constraint that no other goal g' of any applicable regulative rule would be demoted if g' is at least as important as g.

Notice that the operation of contraction is technically done by changing R^c and $\succ$. This revision is supposed to remove any b_k from the extension of the normative theory D' when $f_1, \ldots f_m$ are the case.

5. Refining the Argumentation Model

In comparison to other approaches[15], Definition 18 has the advantage of being much less cumbersome, since it makes use of a more intuitive framework that looks also more natural in an argumentative domain like legal reasoning. However, the model has still some limits:

(i) Definition 18 only works when any literal b to be contracted occurs in ∂_D^+; when $b \in \Delta_D^+$, no contraction is possible;

(ii) Definition 18 does not cover the case where the contraction of the literal is not made by directly attacking the rules supporting it.

5.1. Contraction Operation: Refinements

Let us consider the first problem: what to do when there is a strict supportive argument for the literal b? Since strict rules cannot be defeated, the only solution is rule removal.

Definition 19 (Rule Removal)**.** *Let $D = (F, R^c, R^{\mathrm{OBL}}, R^{\mathrm{Goal}}, \succ)$ be a normative theory. Let $\mathcal{A}_1, \ldots \mathcal{A}_n$ the strict arguments in $Args_D$ supporting p. The normative theory D^p_{-X} is equal to D except for R^c_{-X}, which is defined as follows:*

- *$R^c_{-X} = R^c - X$ and*
- *$X = \{w_1, \ldots, w_m\}$ is the smallest set of strict rules in R^c where, for each $k \in \{1, \ldots, n\}$, there is at least a $w_j \in X$ such that w_j occurs in $\mathcal{A}_k$.*

[15]Cf. G. Boella, G. Governatori, A. Rotolo, L. van der Torre, *Lex minus dixit quam voluit, lex magis dixit quam voluit: A Formal Study on Legal Compliance and Interpretation*, in: P. Casanovas et al., eds.: *AI Approaches to the Complexity of Legal Systems*, Springer, Berlin, 2010, 162–183; G. Boella, G. Governatori, A. Rotolo, L. van der Torre, *A Logical Understanding of Legal Interpretation*, in: F. Lin et al., eds.: *Principles of Knowledge Representation and Reasoning: Proceedings of the Twelfth International Conference, KR 2010*, AAAI Press, Menlo Park, 2010, 563–565.

Example 2. *Consider the following normative theory D:*

$$
\begin{aligned}
F &= \{a, b\}\\
R^c &= \{r_1 : a \rightarrow_c \neg c,\ r_2 : d \Rightarrow_c c,\ r_3 : b \rightarrow_c d, r_4 : d, a \rightarrow_c \neg c\}\\
R^{\mathrm{OBL}} &= \{r_4 : \neg c \Rightarrow_{\mathrm{OBL}} l\}\\
R^{\mathrm{Goal}} &= \{ \Rightarrow_{\mathrm{Goal}} g\}\\
\succ &= \emptyset
\end{aligned}
$$

Suppose we want to contract $\neg c$, which is in Δ^+ of the extension of D and triggers r_4. In this normative theory we have only two strict arguments for $\neg c$:

$$\mathcal{A}_1 = a \rightarrow_c \neg c \qquad \mathcal{A}_2 = b \rightarrow_c d, a \rightarrow_c \neg c$$

Since the two arguments do not share any rule, we have necessarily to remove at least two rules. An option that satisfies Definition 19 is the removal of r_1 and r_4, but also the removal of r_3 would be acceptable.

Notice that Definition 19 allows us to change arguments in a very flexible way, as it permits to contract a literal by removing rules that do not directly support it. This may be needed in the law. For instance, suppose we have the following rules

$$r_1 : Embryo \rightarrow_c Alive \qquad r_2 : Alive \rightarrow_c Person \qquad r_3 : Person, Kill \rightarrow_c Homicide$$

and we want that that killing an embryo does not amount to a homicide. Perhaps, it would be better to remove r_1 or r_2 rather than denying that killing a person is not a homicide (removing r_3).

The cost for having this flexibility is that the contraction of strict conclusions no longer satisfies AGM postulates for minimal change. In particular, let us define the contraction D^-_p of a strict conclusion p as D^p_{-X} and consider the following two postulates that reframe in our setting two well-known AGM postulates[16]: let $\# \in \{\Delta, \Gamma\}$

$$\text{If } p \notin \#^+_D \text{ then } E^{D^-_p} = E^D \qquad (\#\text{-Vacuity})$$

$$\text{If } p \in \#^+_D \text{ then } \Delta^+_D \subseteq \Delta^+_{(D^-_p)^+_p} \text{ and } \partial^+_D \subseteq \partial^+_{(D^-_p)^+_p} \qquad (\#\text{-Recovery})$$

Assume that expansion D^+_p is in general defined as follows: if $\# \in \{\Delta, \partial\}$

$$D^+_p = \begin{cases} D & \text{if } {\sim}p \in \#^+_D \\ (F, R'^c, R^{\mathrm{OBL}}, R^{\mathrm{Goal}}, \succ') & \text{otherwise} \end{cases}$$

where the following conditions hold:

when $\# = \partial$, then either $R'^c = R^c \cup \{w : \Rightarrow_c p\}$

$$\succ' = (\succ \cup\ \{w \succ r \mid \forall r \in R^c[{\sim}p]\}) - \{r \succ w \mid r \in R^c[{\sim}p]\} \text{ or}$$

$$R'^c = R^c \cup \{w : \rightarrow_c p\} \qquad (1a)$$

[16]Cf. D. Billington, G. Antoniou, M. Maher, G. Governatori, *Revising Nonmonotonic Belief Sets: The Case of Defeasible Logic*, in: *KI-99: Advances in Artificial Intelligence*, Springer, Berlin, 1999.

$$\text{when } \sharp = \Delta, \text{ then } R'^c = R^c \cup \{w :\rightarrow_c p\}. \tag{1b}$$

When we apply Definition 19, the only case of expansion we are interested in is the one at condition (1b) above. If so, it is easy to show that the following result holds:

Theorem 3. *For any literal p and any normative theory D, let $D^-_p = D^p_{-X}$. Then, the contraction operation does not satisfy the postulates of Δ-Vacuity and Δ-Recovery, unless $\forall w \in X : w \in R^c[p]$.*

In other words, we should impose that the rule removal only affects those rules that directly prove *p*.

Notice that this procedure in Definition 19 successfully removes any *b* from the set of strict conclusions of a normative theory:

Theorem 4 (Rule Removal Success)**.** *Let $D = (F, R^c, R^{\mathrm{OBL}}, R^{\mathrm{Goal}}, \succ)$ be any normative theory. Then $p \notin \Delta^+_{D-X}$.*

Sketch. By construction, Definition 19 guarantees that at least one strict constitutive rule is removed in every argument based on *D* supporting *p*. So *p* is not strictly provable in *D*. □

Similar considerations can be applied to the contraction of defeasible conclusions, which is introduced in Definition 18. The contraction operation described there works in such a way that any *p* is removed from the defeasible extension of the normative theory when some facts hold. Even in this case, the operation directly affects the defeasible rules that prove *p*, but a more flexible contraction operation can be introduced.

Definition 20 (Defeasible Contraction)**.** *Let $D = (F, R^c, R^{\mathrm{OBL}}, R^{\mathrm{Goal}}, \succ)$ be a normative theory. Let $\mathcal{A}_1, \dots \mathcal{A}_n$ be the arguments in $JArgs^D$ for p. The theory $D_{\flat p} = (F, R'^c, R^{\mathrm{OBL}}, R^{\mathrm{Goal}}, \succ')$ is such that*

(i) $R'^c = R^c \cup \{s :\rightsquigarrow_c \sim q\} \cup \{t :\rightsquigarrow_c \sim x\}$,

(ii) $\succ' = \succ -\ [\{r_k \succ s | r_k \in R^c[\sim C(s)], r_k$ *occurs in* $\mathcal{A}_k\ \forall k \in \{1, \dots, n\}\} \cup \{w \succ t \mid$ *for each argument* $\mathcal{B}$ *where all the subarguments of* $\mathcal{B}$ *are in* $JArgs^D$ *except the top subargument* $\mathcal{B}^t$ *and* $C(R(\mathcal{B}^t)) = x$*, either* $x = p$ *or* $R(\mathcal{B}^t)$ *occurs in* $\mathcal{A}_k\ \forall k \in \{1, \dots, n\}\}]$.

Remark 2. *The apparent complexity of Definition 20 depends on the fact that, the removal of p from ∂^+_D without affecting the rules in $R^c_{\mathrm{d}}[p]$ must take into account that blocking some other literals that are used to argue in favor of p may trigger in turn some rules proving p, like in the following set of rules:*

$$\begin{aligned} R^c = \{ & r_1 :\Rightarrow_c a, \\ & r_2 : a \Rightarrow_c b, \\ & r_3 : b \Rightarrow_c p, \\ & r_4 : a \Rightarrow_c c, \\ & r_5 :\Rightarrow_c \neg c, \\ & r_6 : \neg c \Rightarrow_c p\} \end{aligned}$$

Here, if we only add a defeater like $s :\rightsquigarrow_c \neg a$, we remove a from ∂^+_D, which occurred in all justified arguments for p, but in doing so, rule r_4 is no longer applicable and thus we trigger r_5 and r_6, once again having p in the defeasible extension. Hence, we have also to add a defeater like $t :\rightsquigarrow_c c$, which blocks rule r_5.

However, a result similar to the one in Theorem 3 can be proved also in this case:

Theorem 5. *For any literal p and any normative theory D, let $D^-_p = D_{\flat p}$. Then, the contraction operation does not satisfy the postulates of ∂-Vacuity and ∂-Recovery, unless, condition (i) and (ii) in Definition 20 are reframed as follows:*

(i') $R'^c = R^c \cup \{s :\rightsquigarrow_c {\sim}p\}$

(ii') $\succ' = \succ - \{r_k \succ s | r_k \in R^c_{\mathrm{d}}[p]\}$.

The following proposition shows that the proposed operations for defeasible contraction are successful:

Theorem 6 (Success of Defeasible Contraction)**.** *Let $D = (F, R^c, R^{\mathrm{OBL}}, R^{\mathrm{Goal}}, \succ)$ be any normative theory. Then $p \notin \partial^+_{D_{\flat p}}$ unless $p \in \Delta^+_D$.*

Sketch. On the basis of Theorem 2, we can look at the proof theory: an inspection of the proof conditions for ∂ directly shows that Definition 20 successfully blocks the derivation of p, unless it is derived using only strict rules. Notice that condition (ii) in Definition 20 ensures that, in case the attacks made by the added defeaters activate other (previously defeated) rules supporting p, these last potential derivations are made unsuccessful. □

6. Future Work: Arguments to the Best Theory

While Definition 18 guarantees to get one resulting normative theory, this does not apply to Definitions 19 and 20: here we may have cases where more results are available for revising the given theories of constitutive rules in order to align concepts with the goals of regulative legal rules:

Corollary 1. *The restricting procedures based on Definitions 19 and 20 lead to more than one outcome.*

This fact raises the question of how to choose among different theory revisions. This problem can take the form of a dialectical process where an exchange of arguments and counter-arguments are meant to establish which is the best theory supporting or excluding the contraction of the applicability conditions of a given legal rule.[17] In general, some criteria can be proposed, such as inclusiveness, simplicity, minimal change, and theory connectedness.[18] Let us just consider two of them.

[17] H. Prakken, G. Sartor, *The Three Faces of Defeasibility in the Law*, in: *Ratio Juris*, 17, 2, 2004, 118–139.

[18] Cf. A. Rotolo, C. Roversi, *Constitutive Rules and Coherence in Legal Argumentation: The Case of Extensive and Restrictive Interpretation*, in C. Dahlman, E. Feteris, eds.: *Legal Argumentation Theory: Cross-Disciplinary Perspectives*, Springer, Dordrecht, forthcoming.

Minimal change We have already discussed it in regard to some contraction procedures. In the previous sections we have shown how some well-known AGM criteria of minimal change can be re-framed in our logic setting. We have seen that a more liberal (and flexible, at least in the legal reasoning perspective) version of the contraction procedures requires to abandon such criteria. In our framework, minimal change still amounts to minimizing the changes of the theory extension. However, since the contraction is made by adding defeaters and by removing strict rules, one may state that the minimal change should be rather obtained by keeping the set of rules as close as possible to the original one. This idea is well-known in revision theory but is overlooked in rule-based defeasible systems. Notice that the second option (minimal rule change) is independent of the facts of the normative theory, whereas the other (minimal extension change) is context-dependent, since different facts may fire different rules. Obviously, these two options do not always lead to the same results: suppose the facts includes a and consider

$$R^c = \{r : a \Rightarrow b,\ s : b \Rightarrow e,\ t : b \Rightarrow d,\ z : b \Rightarrow e\}$$

If we want to block e (we want to restrict the interpretation of a concept), we have two options: add either two defeaters to override respectively s and z, or only one defeater that overrides r. The first option is better in terms of minimizing the change of conclusions (we only drop e), while the second one is better, as only one defeater is added (but two conclusions are dropped: b and e).

Theory connectedness This can be roughly measured by the number of links (based on sharing common literals) that connect each concept with the others. Intuitively, this is a desirable property, since it reflects the systematic character of legal systems. Consider, for instance, the following two sets of rules:

$$\begin{aligned} R^c = \{ & r_1 : Embryo \rightarrow_c Alive, \\ & r_2 : Alive, Kill \rightarrow_c Homicide, \\ & r_3 : Person \rightarrow_c Alive\} \\ R^{\mathrm{OBL}} = \{ & r_4 : Homicide \Rightarrow_{\mathrm{OBL}} Punished\} \end{aligned}$$

Since all counts-as rules are strict, what we can do is only to remove rules. Suppose we want to prevent the situation where one is punished because she "killed" an embryo. We have two options here: either (a) remove r_1 or (b) remove r_2. Option (b) is less satisfactory in this specific example: the impact on the theory is greater, as rule $_2$ is more connected to the other rules than r_1.

How to handle these criteria in an argumentation framework is left to future research. We could just notice that such a framework can be based on the following intuitions:

- An argument $\mathcal{A}$ is any normative theory $D = (F, R^c, R^O, R^G, \succ)$;

- Given an extended normative theory $Ex(D) = \langle (F, R^c, R^{\mathrm{OBL}}, R^{\mathrm{Goal}}, \succ), G, \mathcal{G}, > \rangle$ an attack $\mathcal{B}$ on $\mathcal{A}$ is another theory obtained from $\mathcal{A}$ using the contraction procedures discussed above, such that, either
  - (a) $\mathcal{A}$ is such that some p is not in the extension of $\mathcal{A}$ and this determines that a goal g is demoted; (b) $\mathcal{B}$ is such that some p is in the extension of $\mathcal{B}$ and this determines that the goal g is promoted;
  - (i) $\mathcal{A}$ is such that some p is not in the extension of $\mathcal{A}$ and this determines that a goal g is promoted; (ii) $\mathcal{B}$ is such that some p is in the extension of $\mathcal{B}$ and this determines that the goal g is promoted;
  - (1) $\mathcal{A}$ is such that some p is not in the extension of $\mathcal{A}$ and this determines that a goal g is promoted; (2) $\mathcal{B}$ is such that some p is not in the extension of $\mathcal{B}$ and this determines that the goal g is promoted.
- Criteria such as the ones mentioned above can be used to establish the set of justified and rejected arguments.

7. Summary

In this paper we proposed a framework for reconstructing the arguments supporting the restrictive interpretations of legal provisions. The contribution is based on the idea that the interpretation of legal concepts may require to change the constitutive rules defining them. Indeed, if our ontology classifies, for example, a bike as a vehicle, but we have reasons that this is not the case, then this implicitly leads to conclude that the ontology must be revised and that a bike, at least in the contexts under consideration, is not a vehicle. The revision procedure presented in this paper is driven and constrained by considering the goal of the regulative legal rules in which these concepts occur. Some interesting connections with revision theory techniques are considered.