Changing Legal Systems: Legal Abrogations and Annulments in Defeasible Logic

Guido Governatori[1] and Antonino Rotolo[2]

[1] NICTA, Queensland Research Laboratory, Australia,
guido.governatori@nicta.com.au
[2] CIRSFID and Law School, University of Bologna, Italy, antonino.rotolo@unibo.it

Abstract. In this paper we investigate how to represent and reason about legal abrogations and annulments in Defeasible Logic. We examine some options that embed in this setting, and in similar rule-based systems, ideas from belief and base revision. In both cases, our conclusion is negative, which suggests to adopt a different logical model. This model expresses temporal aspects of legal rules, and distinguishes between two main timelines, one internal to a given temporal version of the legal system, and another relative to how the legal system evolves over time. Accordingly, we propose a temporal extension of Defeasible Logic suitable to express this model and to capture abrogation and annulment. We show that the proposed framework overcomes the difficulties discussed in regard to belief and base revision, and is sufficiently flexible to represent many of the subtleties characterizing legal abrogations and annulments.

1 Introduction

Although formal investigations on norm change have drawn some attention in the deontic logic and AI & Law communities, they are still underdeveloped. This gap in the literature is not only obstructive to providing comprehensive models of normative reasoning, but has serious implications, for example, in developing multi-agent systems (MAS).

Indeed, it is widely acknowledged in the literature on MAS that normative concepts can play a crucial role in modeling agents' interaction (see [25]). In fact, while the main objective in MAS is to design systems of autonomous agents, it is likewise important that agent systems may exhibit global desirable properties. Like in human societies, such properties are ensured if the interaction of artificial agents, too, adopts institutional and organizational models whose goal is to regiment agents' behaviour through normative systems in supporting coordination, cooperation and decision-making. To keep agents autonomous it is often suggested that norms should not simply work as hard constraints, but rather as soft constraints [8]. In this sense, norms should not limit in advance agents' behaviour, but would instead provide standards which can be violated, even though any violations should result in sanctions or other normative effects applying to non-compliant agents.

This is the preprint version of the article.
The article has been accepted for publication in the *Logic Journal of the IGLP*

The original publication is available at Oxford University Press.

However, detecting agents' compliance can get into a very complex task if agents' interaction must be tested against dynamic normative systems, such as the law. Legal systems change over time, and so it is of paramount importance to provide efficient and reliable models of norm change in order to design MAS in realistic contexts.

Mainly inspired by [1], most formal models of norm change usually focus on the dynamics of obligations and permissions. However, as rightly noted on the occasion of a recent workshop on this topic[3], "these systems did not explicitly refer to possible changes in the underlying norms [...]". In fact, "new norms may be created and old norms may need to be retracted. In this dynamic setting, it is essential to distinguish norms from obligations and permissions as studied by deontic logic, to understand the formal properties specific for the dynamics of norms, and to describe how such objects can be manipulated [...]". Unfortunately, "a formal model that captures the relevant features of norm change is still lacking".

The aim of our work is to make some steps in this direction by investigating the notion of legal modification. Legal modifications are the ways through which the law implements norm dynamics [18]. Modifications can be either explicit or implicit. In the first case, the law introduces norms whose peculiar objective is to change the system by specifying what and how other existing norms should be modified. In the second case, the legal system is revised by introducing new norms which are not specifically meant to modify previous norms, but which change in fact the system because they are incompatible with such existing norms and prevail over them. (They prevail because, for example, have a higher ranking status in the hierarchy of the legal sources or because have been subsequently enacted.)

The most interesting case is when we deal with explicit modifications, which permit to classify a large number of modification types. In general, we have different types of modifying norms, as their effects (the resulting modifications) may concern, for example, the text of legal provisions, their scope, or their time of force, efficacy, or applicability [18,12,13]. Derogation is an example of scope change: a norm n supporting a conclusion P and holding at the national level may be derogated by a norm n' supporting a different conclusion P' within a regional context. Hence, derogation corresponds to introducing one or more exceptions to n. Temporal changes impact on the target norm in regard to its date of force (the time when the norm is "usable"), date of effectiveness (when the norm in fact produces its legal effects) or date of application (when conditions of norm applicability hold). An example of change impacting on time of force is when a norm n is originally in force in 2007 but a modification postpones n to 2008. Substitution is an example of textual modification, as it replaces some textual components of a provision with other components. For instance, some of its applicability conditions are replaced by other conditions.

We are interested here in studying the concepts of *abrogation* and *annulment*.

Annulment is usually seen as a kind of repeal, as it makes a norm invalid and removes it from the legal system. Its peculiar effect applies *ex tunc*: annuled

[3] `http://icr.uni.lu/normchange07/`

norms are prevented to produce all their legal effects, independently of when they are obtained.

The nature of *abrogation* is more controversial. In some cases, it is important to see whether the abrogation is the result of judicial review, legislation, or referenda. But again, despite domestic peculiarities, abrogations, too, work as if they were norm removals, even though they are different from annulments; the main point is usually that abrogations operate *ex nunc* and so do not cancel the effects that were obtained before the modification. If so, it seems that abrogations cannot operate retroactively[4].

Here, we have two main sources of complexities.

First of all, some (but not all) jurists argue that abrogations are not properly norm removals. In fact, if a norm n_1 is abrogated in 2007, its effects are no longer obtained after then. But, if a case should be decided in 2008 but the facts of the case are dated 2006, n_1, if applicable, will anyway produce its effects because the facts held in 2006, when n_1 was still in force (and abrogations are not retroactive). Accordingly, n_1 is still in the legal system, even though is no longer in force after 2007.

This point leads us to examine the core problem regarding abrogations: What are the legal effects these modifications should block? Suppose that a norm n_1 in force in 2006 states that, if your annual income is less than 5,000 euros, you are a needy person and norm n_2 says that a needy person has the right to live for free in a council house. If n_1 is retroactively *annulled* in 2007, this counts as n_1's removal since 2006, and all its effects are blocked. Imagine now that two norms n_3 and n_4 are added in 2007 stating that needy people's income is less than 3,000 euros and that needy people are eligible for medical aid. If n_1 is *abrogated* in 2007, jurists may argue that at least some of its indirect effect (in this case, obtained via n_2: right to house) should not be extinguished in 2007, whereas the propagation of the qualification "needy person" (with an income of less than 5,000 euros) should not be propagated from 2006 to 2007, since this would make n_4 applicable. In fact, jurists [18] say that abrogations can at most block some, but not *all*, past effects (otherwise, we would have annulments).

If so, we can indeed see abrogations as norm removals, but they operate as such only if properly parametrized to temporal constraints. In the examples mentioned above, norm n_1 works as if it were removed after 2007 and in regard to any case dated after 2007. If so, abrogations do remove norms but affect only some of the legal effects potentially derivable from the abrogated norms.

Hence, besides many legal complexities—which are often related to the peculiarities of actual legal systems—what we have to bear in mind is that the abrogations and annulments implement the following different reasoning patterns:

---

[4] However, this is not always true. Even where retroactive abrogations are prohibited (such as in the Italian system), the problem is open in some contexts. Suppose an ordinary court is called upon to decide a case in which a norm n applies, but the court argues that n violates some constitutional provisions and so it suspends the trial proceedings referring to the constitutional court to decide on the illegitimacy and abrogation of n. Constitutional court's decision and abrogation of n is necessarily posterior to the case.

– in one case norms are removed with all their effects, whereas
– in other case norms are somehow removed but some or all their effects propagate if obtained before the modification.

The problem of determining what legal effects should be blocked cannot be solved in general. Of course, the case of annulments is easy: all effects are removed. As far as abrogations are concerned, it is often the case that direct effects of a norm should be blocked but this does not necessarily hold for its indirect effects (as discussed above); however, we cannot exclude cases where indirect effects should propagate whereas the direct effect should be blocked, or all past effects should propagate, or, again, norm removals should apply at a certain time and some, but not all, past effects hold. Hence, any formal models for abrogations and annulments should be sufficiently flexible to capture all these options.

What is the most adequate formal method for modelling abrogations and annulments? Clearly, a temporal representation may help, but a preliminary point is whether we can abstract from this aspect and move to a general analysis where time is not considered.

We address these issues using Defeasible Logic (DL) [22,3], but analogous considerations can be extended to other nonmonotonic (sceptical) rule-based systems. Although other options are available, rule-based systems seem a natural way to represent legal systems: legal norms are usually viewed as rules specifying some applicability conditions and a legal effect. Hence, after providing in Section 2 a brief overview of DL, this paper proceeds by exploring two possible formal methods.

The first method is aimed at capturing at least some basic aspects of abrogations and annulments without resorting to temporal reasoning. If this research option were feasible, we could avoid a number of technical complexities arising within any temporal model for legal modifications. We first discuss whether it is possible to adjust theory revision in DL to capture abrogations and annulments. Section 3.1 considers an immediate method to adjust revision of belief sets in DL in order to capture annulment. Section 3.2 examines a possible alternative in which all operations, including contraction, are captured by only adding a suitable set of new rules. Even though this second option is better for modelling abrogation and annulment, some basic problems remain unsolved. Section 3.3 takes advantage of some ideas from the previous section and discusses how base revision in DL can be applied to capture norm removals. However, we argue that this approach, too, is not fully satisfactory, as it cannot express the notion of retroactivity, which is crucial in distinguishing annulments from abrogations.

In fact, since techniques from theory and base revision are not adequate, the second part of this paper proposes a different method and conceptual model, which is based on a temporal extension of DL able to capture many of the subtleties of legal modifications. Section 4 briefly describes the basic idea of this model. Section 5 develops such a temporal extension of DL: Section 5.1 describes the new formal language; Section 5.2 states the proof theory. Section 6 applies the framework to represent abrogation and annulment. Some conclusions end the paper.

2 Overview of Defeasible Logic

DL is based on a logic programming-like language and it is a simple, efficient but flexible non-monotonic formalism capable of dealing with many different intuitions of non-monotonic reasoning. DL is closely related to logic programming [4] and an argumentation semantics exists [11]. DL has a linear complexity [20] and also has several efficient implementations (e.g., [6]). In addition, some preliminary works on legal modifications in DL have been recently proposed [12,13].

A *defeasible theory* D is a structure $(F, R, >)$ where F is a finite set of facts, R a finite set of rules, and $>$ an acyclic superiority relation on R. *Facts* are represented as literals and are indisputable statements. A *rule* expresses a relationship between a set of premises and a conclusion. We have in DL three types of rules conveying the strength of the relationships: strict rules, defeasible rules and defeaters. A *strict* rule has the form $A_1, \ldots, A_n \rightarrow B$ and states the strongest kind of relationship since its conclusion always holds when the premises are indisputable. *Defeasible* rules have the form $A_1, \ldots, A_n \Rightarrow B$ and cover the case when the conclusion normally holds when the premises tentatively hold; *defeaters* have the form $A_1, \ldots, A_n \rightsquigarrow B$ and consider a situation where the premises do not warrant the conclusions: in defeaters the premises simply prevent another rule to support the opposite. Finally, the superiority relation ($>$) provides information about the relative strength of rules, i.e., about which rules can overrule which other rules.

Accordingly, a conclusion can be labelled either as definite or defeasible. A definite conclusion is an indisputable conclusion, while a defeasible conclusion can be retracted if additional premises become available. DL is based on a constructive proof theory for conclusions. Hence, we can say that a derivation for a conclusion exists and that it is not possible to give a derivation for a conclusion. Based on these two ideas conclusions will be tagged according to their strength and type of derivation:

- $+\Delta B$, meaning that we have a definite proof for B (a definite proof is a proof where we use only facts and strict rules);
- $-\Delta B$, meaning that it is not possible to build a definite proof for B;
- $+\partial B$, meaning that we have a defeasible proof for B;
- $-\partial B$, meaning that it is not possible to give a defeasible proof for B.

In what follows we will refer to $+\Delta$, $-\Delta$, $+\partial$ and $-\partial$ as *proof tags*, and we will give formal conditions under which we can label a conclusion with one of these proof tags.

Provability is based on the concept of a *derivation* (or proof) in $D = (F, R, >)$. A derivation is a finite sequence $P = (P(1), \ldots, P(n))$ of tagged literals satisfying four conditions (which correspond to inference rules for each of the four kinds of conclusion). $P(1..i)$ denotes the initial part of the sequence P of length i.

Some notational conventions before presenting proof conditions for DL derivations. Each rule is identified by a unique label. $A(r)$ denotes the set of antecedents of a rule r, while $C(r)$ denotes its consequent. If R is a set of rules, R_s is the set of all strict rules in R, R_{sd} the set in R of strict and defeasible rules, R_d the set of defeasible rules, and R_{dft} the set of defeaters. $R[B]$ denotes the set of rules in R with consequent B. If B is a literal, $\sim B$ denotes the complementary literal (if B is a positive literal C then $\sim B$ is $\neg C$; and if B is $\neg C$, then $\sim B$ is C).

Here are the proof conditions for strict derivations:

$+\Delta$: If $P(i+1) = +\Delta B$ then either
(1) $B \in F$, or
(2) $\exists r \in R_s[B]$: $\forall A \in A(r) : +\Delta A \in P(1..i)$

$-\Delta$: If $P(i+1) = -\Delta B$ then
(1) $B \notin F$ and
(2) $\forall r \in R_s[B]$: $\exists A \in A(r) : -\Delta A \in P(1..i)$

Strict proofs are just derivations based on detachment for strict rules. Given a strict rule $A_1, \dots, A_n \rightarrow B$, where we have definite proofs for all A_i's, we can deduce B ($+\Delta B$).

DL is a sceptical non-monotonic formalism: with a possible conflict between two conclusions (i.e., one is the negation of the other), DL refrains to take a decision and we deem both as not provable unless we have some more pieces of information that can be used to solve the conflict. One way to solve conflicts is to use a superiority relation over rules. The superiority relation gives us a preference over rules with conflicting conclusions. In case we have a conflict between two rules we prefer the conclusion of the strongest of the two rules. The superiority relation is applied in defeasible proofs.

Defeasible proofs proceed in three phases: we first look for an argument supporting the conclusion we want to prove (an applicable rule for the conclusion). Second, we look for arguments for the opposite of what we want to prove. Third, we rebut the counterarguments. This can be done by showing that the counterargument is not founded (i.e., some of the premises do not hold), or by defeating the counterargument, i.e., the counterargument is weaker than an argument for the conclusion we want to prove. Formally,

$+\partial$: If $P(i+1) = +\partial B$ then either
(1) $+\Delta B \in P(1..i)$ or
(2.1) $\exists r \in R_{sd}[B] \forall A \in A(r) : +\partial A \in P(1..i)$ and
(2.2) $-\Delta \sim B \in P(1..i)$ and
(2.3) $\forall s \in R[\sim B]$ either
  (2.3.1) $\exists A \in A(s) : -\partial A \in P(1..i)$ or
  (2.3.2) $\exists t \in R_{sd}[B]$ such that
    $\forall A \in A(t) : +\partial A \in P(1..i)$ and $t > s$.

$-\partial$: If $P(i+1) = -\partial B$ then
(1) $-\Delta B \in P(1..i)$ and
(2.1) $\forall r \in R_{sd}[B]$ $\exists A \in A(r) : -\partial A \in P(1..i)$ or
(2.2) $+\Delta \sim B \in P(1..i)$ or
(2.3) $\exists s \in R[\sim B]$ such that
  (2.3.1) $\forall A \in A(s) : +\partial A \in P(1..i)$ and
  (2.3.2) $\forall t \in R_{sd}[B]$ either
    $\exists A \in A(t) : -\partial A \in P(1..i)$ or $t \not> s$.

3 Revising Defeasible Theories

We address in this section the problem of how to embed in DL some ideas from belief and base revision in order to capture annulment and abrogation.

We attack two different (but related) problems raised by these modifications:

- how to block either some or all norm effects;
- how to model norm removals in legal systems.

For the sake of clarity, we will be mostly proceeding by exploring separately these two issues. This choice can also meet expectations of those who think that the notion of norm removal is not logically necessary, since blocking all effects of a norm n is indeed a good method to "neutralize" n and so to capture its annulment, while blocking some of its effects precisely amounts to its abrogation (recall the discussion of Section 1).

As we argued, even though such modifications have a temporal flavour, we first move to a general analysis where time is not considered. Also, we assume that a defeasible theory can represent the basic logical structure of a legal system [12,13]. It is a general tenet in the literature that one reason why legal reasoning is defeasible depends on the fact that, in many cases, norm conclusions can be obtained only if we do not have stronger norms attacking them [24]. DL theories consist of a set of rules (which may be defeasible), a set of facts, and a set of priorities over rules (which establish their relative strength). In this perspective, rules naturally correspond to legal norms, while priorities represent the criteria used to solve legal conflicts. Hence, a general picture like this provides a standard for capturing the basics of legal systems [23]. With this said, let us begin with our discussion on annulment and abrogation.

3.1 Annulments are Contractions: Blocking Legal Effects by Revising Theory Extensions

Approaches based on AGM usually assume that a belief set B is a theory, i.e., a set of formulas closed under a logical consequence relation, thus $B = \mathrm{Cn}(B)$. Let us consider the equivalent of this notion in DL.

Given a Defeasible Theory T, the Herbrand Base HB_T of T is the set of all the atoms occurring in T. The extension of a Defeasible Theory T is defined as the 4-tuple [3]:

$$E(T) = (\Delta^+(T), \Delta^-(T), \partial^+(T), \partial^-(T)),$$

where $\#^{\pm}(T) = \{p | p \in HB_T, T \vdash \pm\#p\}$, $\# \in \{\Delta, \partial\}$.

Definition 1. *Let $T = (F, R, >)$ be a Defeasible Theory. We define another Defeasible Theory $T' = (\emptyset, R', \emptyset)$ such that R' is the smallest set satisfying the following conditions*

- *if $p \in \Delta^+(T)$, then $\to p \in R'$;*
- *if $p \in \partial^+(T)$, then $\Rightarrow p \in R'$;*
- *if $p \notin \Delta^+(T) \cup \Delta^-(T)$, then $p \to p \in R'$;*
- *if $p \in \Delta^-(T)$, then $R'_s[p] = \emptyset$;*
- *if $p \in \partial^-(T)$, then $R'_d[p] = \emptyset$;*
- *if $p \notin \partial^+(T) \cup \partial^-(T)$, then $p \Rightarrow p \in R'$.*

We will say that T' is the theory generated by the extension of T.

Proposition 1. *Let T be a defeasible theory and T' the theory generated from the extension of T. For every $p \in HB_T$, $T \vdash \# \pm p$ iff $T' \vdash \# \pm p$.*

Proof. See Appendix A for the proof.

The meaning of the result in Proposition 1 is that for every theory (and so every set of conclusions), we can generate a new equivalent theory without looking at the structure of the original theory: in fact, classically two theories are equivalent if they have the same extension (the same set of conclusions).

The above result gives us an immediate way to define contraction for revision based on belief sets. We define $T^{\ominus}_{c} = T'$ such that $E(T) = (\Delta^{+}(T), \Delta^{-}(T), \partial^{+}(T), \partial^{-}(T))$ and T' is the theory generated by the extension

$$(\Delta^{+}(T) - \{c\}, \Delta^{-}(T), \partial^{+}(T) - \{c\}, \partial^{-}(T)).$$

It is easy to verify that the above way to define contraction satisfies all AGM postulates[5].

Let us examine *annulment*. When we annul a norm in a legal system, this means that all (direct and indirect) legal effects deriving from it must be cancelled as well. For example, if we have a normative system T containing only the rules $A \Rightarrow B$ and $B \Rightarrow C$, then the annulment of the former rule (assuming the fact A) should block both B and C. Intuition seems to suggest that contraction is the right operation to capture annulment. Hence, the question is how to use contraction in this case. What one could do here is simply to remove the consequent of the rule. However, the (positive defeasible) extension of T (i.e., $\partial^{+}(T)$) is $\{A, B, C\}$,[6] and contracting B leaves C in the extension. Hence, this immediate use of contraction is not representative of legal annulment. As we said, we have to consider all consequences of the formula to be contracted. In the above example, C can only be derived if B does. Accordingly, annulment of any rule $A_1, \ldots, A_n \Rightarrow B$ could be defined as follows. Let $T = (F, R, \succ)$ be a Defeasible Theory. Then

$$T^{\ominus}_{A_1,\ldots,A_n \Rightarrow B} = \begin{cases} T & \text{if } A_1,\ldots,A_n \Rightarrow B \notin R \text{ or } \{A_1,\ldots A_n\} \not\subseteq \partial^{+} \\ (F', R', \succ') & \text{otherwise} \end{cases}$$
$$\text{such that} \qquad (1)$$
$$(F', R', \succ') \text{ is the theory generated by } E(T) - (\Delta^{+}(T'), \emptyset, \partial^{+}(T'), \emptyset)$$
$$\text{and } T' = (F = \{B\}, R, \succ).$$

The contraction operation reflecting annulment is defined by "removing" the consequent of the rule. In addition, the theory T' generates all consequences of B with respect to T. Then $T^{\ominus}_{A \Rightarrow B}$ is the theory generated by the extension $E(T) - (\Delta^{+}(T'), \emptyset, \partial^{+}(T'), \emptyset)$. However, let us consider another example.

Example 1. Assume to work with the following theory:

$$T = (F = \{A\}, R = \{A \Rightarrow B,\ B \Rightarrow C,\ A \Rightarrow C\}, \succ = \emptyset).$$

Thus,

$$T' = (F = \{B\}, R = \{A \Rightarrow B,\ B \Rightarrow C,\ A \Rightarrow C\}, \succ' = \emptyset).$$

---

[5] See Appendix B for a discussion of the type of contraction proposed here and its relationship with AGM postulates.

[6] Whenever clear from the context, we will use the term 'extension of a theory' as either the positive defeasible extension of it or the full extension of the theory (see Definition 1).

Hence, $(\partial^+(T) = \{A,B,C\}) - (\partial^+(T') = \{B,C\}) = \{A\}$, and this leads (by applying Definition 1) to obtain that $T^{\ominus}_{A\Rightarrow B}$ corresponds to

$$T'' = (\emptyset, R = \{\Rightarrow A\}, \emptyset).$$

This procedure is not satisfactory unless more sophisticated measures are added. Example 1 shows that the procedure does not properly work, as C has *multiple causes* (B and A): with T'' we exclude $A \Rightarrow B$ by dropping B (and its consequences), but this leads to drop, too, C and so to exclude $A \Rightarrow C$, which is too much.

In addition, the above procedure requires to change the set of facts, which seems to us meaningless. Why cannot we change the set of facts? The facts of a theory are only those pieces of evidence in a case used to *apply* rules (norms) and not to *change* them: hence they should not be considered when one modifies norms. Accordingly, if norms are represented as rules, then reasoning only on the consequences of a theory is not representative of norm change.

For example, the norm (i.e., the rule)

$$HighIncome \Rightarrow TopMarginalRate$$

says that if the income of a person is in excess of the threshold for high income, then the top marginal rate must be applied. If it is a fact that Nino exceeded the threshold (i.e., $HighIncome \in F$) then he has to pay the top marginal rate. Thus the extension is $\{HighIncome, TopMarginalRate\}$; contracting with *HighIncome* results in the theory just consisting in

$$\Rightarrow TopMarginalRate,$$

namely in a rule stating that, no matter what your income is, you will have to pay taxes at the top marginal rate. Thus, revising the evidence on which a case is based results in a change in the legislation, which seems a non-sense when applied to real legal systems.

The idea behind Definition 1 and (1) is that we have to generate a new normative system from the revised extension of corresponding source normative system. However, there are at least three reasons why Definition 1 and (1) do not seem satisfactory:

1. they may change the set of facts, and so do not differentiate between norms and instances of cases;
2. they revise theories regardless of the logical structure of the source theories;
3. they do not correctly account for *ex tunc* modifications, such as annulment.

Changing facts or generating new theories whose structure does not reflect the theories from which they have been obtained trivialise the concept of legal change. Indeed, it is crucial in the law to establish *what rules* generate *which effects*. Therefore, the contraction function defined in this section does not offer a suitable method for modelling annulment (and, in general, norm changes), even if it satisfies all AGM postulates.

3.2 Intermezzo: Revising Theories by Adding New Rules

The difficulties under points 1 and 2 above (at the end of Section 3.1) can be alleviated by adopting in DL the approach proposed in [7] to deal with belief revision of rule-based non-monotonic formalisms, where change operators are not applied to the set of facts and are all implemented by adding new rules and changing priorities. This permits to incrementally modify the legal system, taking into account the logical structure of the source theory. Let us briefly recall the basic features of this approach.

Let us examine *expansion*. Following [10], expansion adds a formula A to $\partial^+(T)$ *only if* $\neg A \not\in \partial^+(T)$. Hence, the case where $\neg A \in \partial^+(T)$ is irrelevant. However AGM decided to also add A in this case. In [7] T is kept unchanged, following [10] rather than [1]. Let $c = P_1, \ldots P_n$ be the formulas to be added. Expansion can be defined as follows:

$$T_c^+ = \begin{cases} T & \text{if } {\sim}P_i \in \partial^+(T) \text{ or } {\sim}P_i = P_j \text{ for some } i,j \in \{1,\ldots,n\} \\ (F, R', \succ') & \text{otherwise} \end{cases}$$
$$\text{where} \quad (2)$$
$$\begin{aligned} R' &= R \cup \{\Rightarrow P_1, \ldots, \Rightarrow P_n\} \\ \succ' &= (\succ \cup \{\Rightarrow P_i \succ r \mid i \in \{1,\ldots,n\}, r \in R[{\sim}P_i]\}) - \\ &\quad \{r \succ \Rightarrow P_i \mid i \in \{1,\ldots,n\}, r \in R[{\sim}P_i]\}. \end{aligned}$$

Thus, rules that prove each of the literals P_i are added, and it is ensured that these are strictly stronger than any possibly contradicting rules.

Let us examine *contraction*, which seems the right candidate to capture at least some aspects of annulments (and, also, of abrogations)[7]:

$$T_c^- = \begin{cases} T & \text{if } P_1, \ldots, P_n \not\in \partial^+(T) \\ (F, R', \succ') & \text{otherwise} \end{cases}$$
$$\text{where} \quad (3)$$
$$\begin{aligned} R' &= R \cup \{P_1, \ldots, P_{i-1}, P_{i+1}, \ldots, P_n \leadsto {\sim}P_i \mid i \in \{1,\ldots,n\}\} \\ \succ' &= \succ - \{s \succ r \mid r \in R' - R\}. \end{aligned}$$

Intuitively, (3) aims at preventing the proof of all the P_is. To achieve this it is ensured that at least one of the P_is will not be proven. The new rules in R' ensure that if all but one P_i have been proven, a defeater with head ${\sim}P_j$ will fire. Having made the defeaters not weaker than any other rules, the defeater cannot be "counterattacked" by another rule, and P_j will not be proven, as an inspection of the condition $+\partial$ in Section 2 shows.

This approach slightly deviates from the AGM postulates, in particular from those for contraction. The second AGM postulate states that we contract a formula only by deleting some formulas, but not by adding new ones. This postulate cannot be adopted here because it contradicts the sceptical nonmonotonic

[7] For space reasons, [7]'s treatment of revision is omitted.

nature of DL. To see this, suppose that we know A, and we have rules $\Rightarrow B$ and $A \Rightarrow \neg B$. Then A is sceptically provable and B is not. But if we decide to contract A, B becomes sceptically provable. Note that this behaviour is not confined to DL but holds in any sceptical nonmonotonic formalism [7]. Another peculiarity of this approach is the clear distinction between facts and rules and that facts are indisputable and cannot be changed. Thus, the negation of facts correspond to contradictions, and contracted facts are still included in the extension of the theory.

The advantages of [7]'s proposal are clear, as legal systems are changed by only adding new rules. In this sense, even though it works on theory extensions (suitable new rules ensure that some literals are included in extensions, or are excluded from them), this approach seems closer to base revision (see Section 3.3). But, independently of this question, one problem is still open: how to adjust this approach to account for legal modifications? A legal system T is modified by selecting, as a target, one or more norms of T, whereas [7]'s proposal parametrises operations to sets of literals. Let us bear in mind these points and proceed with our discussion.

3.3 Revising Normative Bases: Blocking Effects vs. Removing Rules

The main problem with revision based on belief sets is that this approach does not mimic how the law implements norm changes, since "new" rules are generated to reflect the changes. Legal effects of rules can be used to guide how norms should be changed, but they should not determine *what* and *how* rules are changed. Therefore the alternative to revision based on belief sets is base revision [19,21]. As is well-known, base revision does not operate on the extension of a theory, but rather applies to the theory "generators" (i.e., the non-logical axioms of the theory). This idea can be naturally coupled with partitioning the elements of a theory into "facts" and "rules", where the former cannot be revised (unless update is used), while the latter may be subject to revision.

Usually, belief revision operations are defined as contraction followed by expansion (according to Levi's Identity). Therefore, revision often results in some rules to be removed from the base of a theory. Base revision allows us to adopt different strategies, namely, to modify rules. As we discussed, in the law modifications sometimes require to remove existing rules, sometimes to introduce them, or, finally, to partially modify them. In this perspective, assuming a rule-based representation of norms, revision on bases using modification techniques seems in general more flexible and closer to the legal practice, since it allows for the conceptual distinction of these types of changes. In addition, as argued e.g. in [9], base revision results in theories that are closer to the structure of the theories to be revised.

In the following subsections we examine some alternative options to model abrogations and annulments by resorting to base revision techniques.

Annulments: Revising Normative Bases by Adding Exceptions Which Block Legal Effects Let us consider an example to introduce the idea of modification of bases. Suppose we want to revise a theory containing a rule $r_1 : A \Rightarrow B$ and contract B when C is the case (let us say that C implies $\neg B$). The revision of the rule is $r'_1 : A, \neg C \Rightarrow B$. This means that we have modified the original rule taking into account the exception provided by C. This intution was proposed, for example, in [9]. DL has an elegant mechanism to deal with exceptions. An exception is simply implemented by a rule capturing the connection between the exceptional antecedent and the conclusion to be blocked. Thus, in the example above, instead of changing r_1 into r'_1, we may simply add a new rule such as $r_2 : C \Rightarrow \neg B$ or $r_2 : C \leadsto \neg B$, and state that $r_2 \succ r_1$. As we have seen in Section 3.2, this idea has been originally proposed for DL in [7], but there were still the open problems of adapting this method for modeling legal modifications and of setting change operations in such a way as to parametrise them with respect to the proper target of legal modifications, namely, legal rules.

Let us see how to adjust [7]'s definitions for modeling *annulment*. Let T be a theory and $r : A_1, \ldots, A_n \Rightarrow B$ be the rule to be annuled. The simplest solution is to frame this modification in terms of [7]'s contraction of the head of r:

$$T^{annul1}_{A_1,\ldots,A_n \Rightarrow B} = T^-_B \tag{4}$$

This solution directly applies (3). However, (4) is too strong since it forces the removal of B from the extension (unless B is a fact). If we have two different (and independent) rules applicable at the same time and with the same head, and we just annul one of them, the other should still be able to produce its effect. But, unfortunately, (4) affects the second rule as well.

Thus, we have to give an alternative annulment operation based on a variant of the contraction operation.

$$\begin{aligned} T^{annul2}_{r:A_1,\ldots,A_n \Rightarrow B} = &\begin{cases} T & \text{if } B \notin \partial^+(T) \\ (F, R', \succ') & \text{otherwise} \end{cases} \\ &\text{where} \\ &R' = R \cup \{r' : \leadsto \sim B\} \\ &\succ' = \succ \cup \{(r', r)\} \cup \{(s, r') | s \in R[B] - \{r\}\} \end{aligned} \tag{5}$$

Consider the following examples.

Example 2. Let us consider the following theory:

$$T = (F = \{A\}, R = \{r_1 : A \Rightarrow B,\ r_2 : B \Rightarrow C\}, \succ = \emptyset).$$

Clearly, $\partial^+(T) = \{B, C\}$. Hence,

$$T^{annul2}_{r_1:A \Rightarrow B} = (F, R \cup \{r'_1 : \leadsto \neg B\}, \succ' = \emptyset).$$

In the resulting theory we prove $-\partial B$, which makes r_2 inapplicable, thus preventing the positive conclusion of C.

Example 3. Let us consider again the theory in Example 1:

$$T = (F = \{A\}, R = \{r_1 : A \Rightarrow B,\ r_2 : B \Rightarrow C,\ r_3 : A \Rightarrow C\}, \emptyset).$$

The annulment of r_1 still amounts to adding $r'_1 : \rightsquigarrow \neg B$ to R, which prevents the conclusion of all literals depending *only* on B. Accordingly, C will be in the extension, as it is obtained through r_3. In addition, if $r_4 : \Rightarrow B$ were in R, r_4 would be stronger than r'_1, thus obtaining B.

Annulments and Abrogations: Blocking Legal Effects vs. Removing Rules
An account of annulment like (5) is closer to the legal practice, as it precisely focuses on modifications of norms and does not merely work on the modification of the effects of norms, a result which could be obtained by any suitable but arbitrary combination of facts and rules. However, things can be viewed from a different perspective. Even though this approach can simulate *ex tunc* modifications like annulments (since it allows us to block all norm effects), (5) fails to remove norms. One may argue that "neutralizing" all effects of a norm is what we actually need, because this operation works as if we were removing it. However, when a norm is legally annulled, it is indeed "removed" from the legal system, and this is sometimes crucial, because, if not removed, a "neutralized" norm still exists in the legal system and this fact can be used as a premise for applying other rules [18].

Accordingly, it seems that we should remove the rule to be annulled from the set of rules:

$$T_r^{annul3} = (F, R - \{r\}, \succ) \tag{6}$$

This solution is simple and effective: it removes the rule r and removes, in the resulting theory, all consequences which can be derived from r.

But, then, we have another problem: how to deal with *ex nunc* modifications, such as abrogations? In this case, the modification of a rule should not necessarily prevent the derivation of its conclusions.

Let us consider Example 2 and assume that the abrogation of r_1 does not prevent the derivation of B and C. This means that, if B and C were derivable before the modification, then they should remain in the extension of the revised theory. Here, we have two options.

- First, we can argue, as done above with annulment, that when a norm is abrogated, it is "removed" from the legal system. But, if r_1 is removed following a similar procedure to that stated in (6), the extension of the revised theory will *not* contain B as well as C, whereas abrogations can also admit of cases where both conclusions should be maintained.
- Thus, a second option does not remove the rule, but adds a suitable set of new rules which allows us to derive what should not be blocked.

Let us suppose to work on the second option. However, what can we do in this case if both B and C should not be dropped? It seems hard to adjust (5) in order to maintain both B and C. At most, what we can do is preventing the derivation of B and maintaining C. Only in this case, if $T = (F, R, >)$ is a defeasible theory, then the *abrogation* of a norm $r : A_1, \ldots, A_n \Rightarrow B$ runs as follows:

$$T^{abr}_{r:A_1,\ldots,A_n \Rightarrow B} = \begin{cases} T & \text{if } r \notin R \\ (F, R', >') & \text{otherwise} \end{cases}$$

$$\begin{aligned} &\text{where} \\ &R' = R \cup \{r^- : \rightsquigarrow \neg B,\ r' : \Rightarrow B'\} \\ &\qquad \cup \{t' : (A(t) - \{B\}) \cup \{B'\} \to C(t) | t \in R_s \text{ and } B \in A(t)\} \\ &\qquad \cup \{t' : (A(t) - \{B\}) \cup \{B'\} \Rightarrow C(t) | t \in R_d \text{ and } B \in A(t)\} \\ &\qquad \cup \{t' : (A(t) - \{B\}) \cup \{B'\} \rightsquigarrow C(t) | t \in R_{dft} \text{ and } B \in A(t)\} \\ &>' = > \cup \{(w, r^-) | w \in R[B] - \{r\}\} \cup \{(t', s) | (t, s) \in >\} \cup \{(s, t') | (s, t) \in >\} \end{aligned} \tag{7}$$

where B' is a new literal not appearing in T.

Proposition 2. *Given a theory T and a rule $r : A_1, \ldots, A_n \Rightarrow B$ such that $T \vdash +\partial B$, then for every $C \in HB_T - \{B\}$, $T \vdash C$ iff $T^{abr}_r \vdash C$.*

Proof. See Appendix A for the proof.

Example 4. Consider the following theory:

$$T = (F = \{A, D\}, R = \{r : A \Rightarrow B,\ t : B \Rightarrow C,\ s : D \Rightarrow \neg C,\ w : E \Rightarrow B\}, (t, s) \in >)$$

Hence, according to (7), $T^{abr}_{r:A_1,\ldots,A_n \Rightarrow B}$ is as follows:

$$\begin{aligned} T^{abr}_{r:A_1,\ldots,A_n \Rightarrow B} = (F = \{A, D\} \\ R = \{r : A \Rightarrow B,\ t : B \Rightarrow C,\ s : D \Rightarrow \neg C,\ w : E \Rightarrow B \\ r^- : \rightsquigarrow \neg B,\ r' : \Rightarrow B',\ t' : B' \Rightarrow C\} \\ > = \{(t, s), (t', s), (w, r^-)\}) \end{aligned}$$

The fact A makes r applicable, but the introduction of r^- blocks the derivation of B using r. However, C is derived via r' and t' (which is stronger than s). Note that (7) is such that the defeater r^- attacks only r (we are abrogating rule r only): hence, if E were in F, B would be obtained from w.

In sum, we have the following possibilities:

- We omit to model annulments and abrogations as corresponding to rule removals. Hence, we represent them working only on rule conclusions and so adopt (5) and (7). However, (7) is partially satisfactory, as it blocks the derivation of the head of the abrogated rule, which does not necessarily hold for all cases of abrogations.
- We address the issue that annulments and abrogations correspond to rule removals. Thus, (6) works for annulments, but it seems quite hard to find an adequate counterpart for abrogation.

- We do not care whether annulments and abrogations correspond to rule removals and are free to adopt, together with (7), either (5) or (6). But, as we said, (7) is problematic.

Of course, we do not exclude that the above problems can be settled. For example, some limits of (7) can be avoided by combining the introduction of exceptions and the removal of the abrogated rule. This can be done by applying the idea in (6) and subsequently reinstate the conclusions that should not be blocked. This can be done by simply using expansion $^+$ as defined in (2). More precisely, suppose $c = C_1, \ldots, C_n$ are the consequences of the rule to be abrogated which we want to maintain.

Definition 2. *Let $T = (F, R, >)$ be a Defeasible Theory such that $r : A_1, \ldots, A_n \Rightarrow B \in R$. Then*

$$T^{abr'}_{r:A_1,\ldots,A_n \Rightarrow B} = (T')^+_c$$

such that $T' = (F, R - \{r\}, >)$ and $c = C_1, \ldots, C_n \in \partial^+(T'')$, where

- *$T'' = (F = \{B\}, R, >)$;*
- *for every C_k, $1 \leq k \leq n$, $C_k \notin \Delta^+(T')$.*

But, even in that case, another difficulty arises when we have to deal with *retroactive* modifications: as we already mentioned, retroactivity is a typical feature of legal modifications. This problem is discussed in the following section.

3.4 Revision and Retroactivity

A norm modification is an operation such that a normative system (consisting of norms and the consequences of cases) is transformed into a different normative system. Accordingly, dynamics of a normative system are described by a sequence of operations.

Suppose we have a system, let us call it T_0, in which we introduce a new rule r and subsequently we remove another rule, let us say s. The system obtained from the first operation is T_1, while the final system is T_2. Thus

$$T_2 = ((T_0)^+_r)^-_s.$$

So far so good. But let us suppose that the removal of s is retroactive. How can we model this case? The idea is that every time we have a retroactive modification we should reconstruct the normative system at the time when the retroactive modification is effective. For example, if the modification is effective since yesterday, we have to recover the system as it was yesterday by undoing the operations leading to the normative system we have today, then we have to apply the retroactive modification and finally redo the other modifications. So, if in the example above s is a retroactive modification effective from T_0, the sequence of modifications still adds r and removes s, but the sequence of theories is as follows:

$$T'_1 = (T_0)^-_s \text{ and so } T_2 = ((T_0)^-_s)^+_r.$$

T_1' is meant to represent the retroactive removal of s in the system as it was yesterday (before adding r). Accordingly, T_2 corresponds to the system in which s is retroactively removed and r is added.

Is this procedure in agreement with the intuition behind retroactive legal modifications? Our answer is negative. The point is that it is possible to define transformations moving from one normative system T_i to T_{i+1} where the transformation is effective at T_i itself, and so the system to be changed is not the target of the modification but the source of it. Let us consider the following example. The normative system T_0 just consists of the fact A. T_1 is obtained from T_0 by retroactively adding two rules $A \Rightarrow B$ and $B \Rightarrow C$ and these rules are effective in T_0. Then the next transformation, leading to T_2 is the removal of $A \Rightarrow B$ from T_0. But then we have two different versions of T_0. Analogous considerations apply when we work on rule consequences and model modifications adding defeaters.

The reason why we have multiple versions of a normative system is that norms have different temporal dimensions: the time of validity of a norm (when the norm enters in the normative system) and the time of effectiveness (when the norm can produce legal effects). Thus, if one wants to model norm modifications, then normative systems must be modeled by more complicated structures. In particular, a normative system is not just the set of norms valid in it, but it should also consider the normative systems where the norms are effective. Accordingly, a normative system is a structure

$$N_i = (T_i, \langle T_0^i, T_1^i, \ldots \rangle)$$

where

- T_i is the theory modeling the set of norms/rules and facts valid in the normative system N_i, and
- $\langle T_0^i, T_1^i, \ldots \rangle$ is the sequence of theories encoding the effective norms for all "versions" of the normative system.

A revision of a legal system is an operation that transforms a normative system into another normative systems by 'changing' the rules in it. In particular, the operation should specify what rules are to be changed, when they are changed, and when the changes are effective. Thus a norm change can be seen as a "transaction" from a normative system

$$N_i = (T_i, \langle T_0^i, T_1^i, \ldots \rangle)$$

into a normative system

$$N_{i+1} = (T_{i+1}, \langle T_0^{i+1}, T_1^{i+1}, \ldots \rangle),$$

where there exists some j such that $T_j^{i+1} = change(T_j^i)$ for some $change$ operation. For example, the abrogation of a rule r may be modeled as

$$T_{i+1}^{i+1} = (T_i^i)_r^{abr},$$

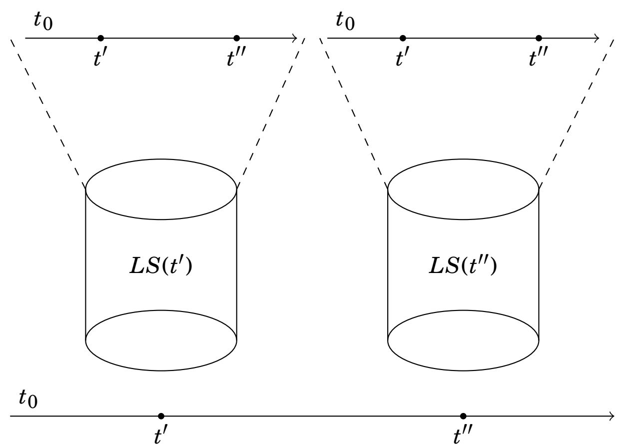


Fig. 1. Legal System at t' and t''

and the retroactive annulment of r, as

$$T_j^{i+1} = (T_j^i)_r^{annul} \qquad \text{(for } j < i\text{)}.$$

In addition, once a norm has been introduced in a normative system the norm continues in general to be in the normative system unless it is explicitly removed. This means that the norm must be included in all theories succeeding the theory in which it has been first introduced. Accordingly, it could be very cumbersome to keep track of the changes and where the changes have to been applied. In real normative systems norms are introduced at a particular time, they are effective at a particular time, and so are changes –changes are norms themselves. Thus, to obviate the issue of keeping track of the changes, and at the same time to offer a conceptual model of norm changes, we have proposed in [12,13] an extension of DL with time, where we consider the two temporal dimensions of relevance for norm change (effectiveness and validity). This is done by labelling rules with two time values, one for the validity time of the norms, and the other for their effectiveness time; furthermore, the labels indicate whether these 'changes' persist or not. The idea that changes are norms themselves is captured by the notion of meta-rule, i.e., a rule whose elements can be rules themselves and not only literals. The next section offers the conceptual background of the proposal presented in [12,13].

4 A Temporal Model for Legal Systems and Norm Change: Outline

The discussion of Section 3.4 suggests that the dynamics of a legal system LS are more correctly captured by a time-series

$$LS(t_1), LS(t_2), \ldots, LS(t_j)$$

of its versions. Each version of LS is called a *norm repository* [12,13].

The passage from one repository to another is effected by legal modifications or simply by persistence [13]. But dynamics of norm change and retroactivity need to introduce another time-line within each version of LS (see Figure 1). Clearly, retroactivity does not imply that we can really change the past: this is "physically" impossible. Rather, we need to set a mechanism through which we are able to reason on the legal system from the viewpoint of its current version but as if it were revised in the past: when we change some $LS(i)$ retroactively, this does not mean that we modify some $LS(k)$, $k < i$, but that we move back from the perspective of $LS(i)$. Hence, we can "travel" to the past along this inner time-line, i.e., from the viewpoint of the current version of LS where we modify norms.

Elements contained in, or derived from, theories can propagate across these time-lines. Hence, propagation concerns the derived conclusions of rules (when some consequent P holds), the rules themselves, and also derivations (i.e., queries: $+\partial P$). This introduces several options regarding how modifications affect a legal system over time:

- conclusions may persist within a certain repository or across different repositories;
- derivations may persist within a certain repository or across different repositories;
- rules may persist within a certain repository or across different repositories.

For example, a rule r, enacted at time t' may persist until t'', but this could mean two things:

1. r propagates, for example, within $LS(t')$ across the inner time-line of that version of the legal system;
2. r carries over from the legal system $LS(t')$ to the legal system $LS(t'')$ (outer time-line), where it is still in force at t'' (inner time-line).

Analogous considerations apply to conclusions and derivations. This duplication of time-lines is necessary when we need to reason about retroactive modifications. These modifications are typically effective in the current legal system but they are retroactive, and so we need to reason from the viewpoint of this version of legal system (outer time-line) and modify past effects (inner time-line within the current version). As we will see, while in abrogations legal effects can carry over from the past into the present (outer time-line), with annulments we have to do two operations: changing virtually the past (inner time-line within the version where the modification is effective) and blocking legal effects coming from past versions of the legal system (outer time-line).

The precise implementation in DL of these intutions is far from obvious and only a partial solution was offered in [13]. The development of a complete DL temporal model for abrogation and annulment is a problem addressed in the remainder of this paper.

5 Temporal Defeasible Logic

Temporal Defeasible Logic (TDL) is an umbrella expression to designate extensions of DL to capture time. TDL has proved useful in modelling temporal aspects of normative reasoning, such as temporalised normative positions [16]; in addition, it was suggested that the notion of a temporal viewpoint may solve the problem of retroactive modifications [12,13]. In this section we present some variants that deal with temporal dimensions as recalled above and presented in [14]. Dynamic aspects of legal reasoning are captured by two means: by first introducing temporal coordinates and, second, normative modifications.

[16] extended DL with temporalised literals, i.e., every occurrence of a literal in the logic has associated to it a timestamp. Thus we have expressions of the type a^t, meaning that a holds at time t. This means that we have to give the condition to prove a literal at time t. So besides the straightforward extension of the conditions given in Section 2, we have to consider whether a conclusion is transient (holding at precisely one instant of time) or whether it is persistent. To prove that a holds at t, we can prove that a held at a previous instant t' and then for all instant in between t and t', it is not possible to terminate a. We will refer to this property as *persistence of a conclusion*.

However, the other components of our knowledge, too, have their temporal validity: we can speak of the time of force of a rule, i.e., the time when a rule can be used to derive a conclusion given a set of premises. In this perspective we can have expressions like

$$(r : a^{t_a} \rightarrow b^{t_b})^{t_r} \tag{8}$$

meaning that the rule r is in force at time t_r, or in other words, we can use the rule to derive the conclusion at time t_r. The full semantics of this expression is that at time t_r we can derive that b holds at time t_b if we can prove that a holds at time t_a. But now we are doing a derivation at time t_r, so the conclusion b^{t_b} is derived at time t_r and the premise a^{t_a} must be derived at time t_r as well. In the same way a conclusion can persist, this applies as well to rules and then to derivations.

Let us consider the following example from a hypothetical taxation law. If the taxable income of a person at January 31, for the previous year is in excess on 100,000$, then the top marginal rate computed at February 28 is 50% of the total taxable income. And this provision is in force from January 1. This rule can be written as follows:

$$(\mathit{Threshold}^{31Jan} \Rightarrow HighMarginalRate^{28Feb})^{1Jan}$$

Let us suppose that the last instalment for the salary was paid to an employee on January 4, and that it makes the total taxable income greater than the threshold stated above. We use $\mathit{Threshold}^{4Jan}$ to signal that the threshold of 100,000$ has been certified on January 4. Clearly $\mathit{Threshold}^{4Jan}$ is a persistent property, thus in this case we can derive that the threshold was reached by January 31. So let us ask what the top marginal rate for the employee is if she lodges a tax return on January 20. What we have to do is to see whether the rule is still in force on

January 20. Given that the norm was valid from January 1, and no changes were made to the legislation in between, the rule persists. Thus from the point of view of January 20, the top marginal rate is 50%. Suppose now that there is a change in the legislation and that the above norm is changed on February 15, and the change is that the top marginal rate is 30%.

$$(Threshold^{31Jan} \Rightarrow MediumMarginalRate^{28Feb})^{15Feb}$$

In this case if the employee lodges her tax return after February 15, the top marginal rate is 30% instead of 50%.

From the above example it is clear that what we derive depends on what rules are valid, and on the normative content of rules, at the time when we do the derivation. In addition, the above example illustrates the case that the content of a rule can be changed. Thus we have to devise a mechanism to capture this phenomenon. To this end we introduce meta-rules, i.e., rules where the consequent is itself a rule and not only a simple proposition. In addition, to keep track of the norm changes, i.e., to represent the different versions of a legal system, we use the notion of repository, i.e., a snap-shot of rules and literals known to exist at a specific time instant. In the rest of the section we will give a formal presentation of the notions discussed so far.

5.1 Language

The language of TDL is based on a (numerable) set of atomic proposition $Prop = \{p, q, \ldots\}$, a set of rule labels $\{r_1, r_2, \ldots\}$, a discrete totally ordered set of instants of time $\mathcal{T} = \{t_1, t_2, \ldots\}$, the negation sign $\neg$, and the rule signs $\rightarrow$ (for strict rules), $\Rightarrow$ (for defeasible rules) and $\leadsto$ (for defeaters). A *plain literal* is either an atomic proposition or the negation of it. Given a literal l with $\sim l$ we denote the *complement of* l, that is, if l is a positive literal p then $\sim l = \neg p$, and if $l = \neg p$ then $\sim l = p$. If l is a literal and t is an instant of time, i.e., $t \in \mathcal{T}$, then l^t is a *temporalised literal*. If l^t is a temporalised literal and $x \in \{tran, pers\}$,[8] then $l^{(t,x)}$ is a *duration literal*. If $l^{(t,x)}$ is a duration literal, $y \in \{tran, pers\}$ $t' \in \mathcal{T}$, then $l^{(t,x)}@(t', y)$ is a *fully temporalised literal*.

A *rule* is a relation between a set of premises (conditions of applicability of the rule) and a conclusion. In this paper the admissible conclusions are either literals or rules themselves; in addition the conclusions and the premises will be qualified with the time when they hold. We consider two classes of rules: *meta-rules* and *proper rules*. Meta-rules describe the inference mechanism of the institution on which norms are formalised and can be used to establish conditions for the creation and modification of other rules or norms, while proper rules correspond to norms in a normative system. In what follows we will use Rules to denote the set of rules, and MetaRules for the set of meta-rules, i.e., rules whose consequent is a rule.

[8] Intuitively *tran* and *pers* are labels attached to the various elements of the language to specify whether these elements persist over time (*pers*) or not (*tran*). See the rest of the section for the full details.

A *temporalised rule* is either an expression $(r:\bot)^{(t,x)}$ (the void rule) or $(r:\emptyset)^{(t,x)}$ (the empty rule) or $(r:A\hookrightarrow B)^{(t,x)}$, where r is a rule label, A is a (possibly empty) set of temporalised literals, $\hookrightarrow$ is a rule sign, B is a duration literal, $t\in\mathcal{T}$ and $x\in\{tran,pers\}$.

We have to consider two temporal dimensions for norms in a normative system. The first dimension is when the norm is in force in a normative system, and the second is when the norm exists in the normative system from a certain viewpoint. So far temporalised rules capture only one dimension, the time of force. To cover the other dimension we introduce the notion of temporalised rule with viewpoint. A *temporalised rule with viewpoint* is an expression

$$(r:A\hookrightarrow B)^{(t,x)}@(t',y),$$

where $(r:A\hookrightarrow B)^{(t,x)}$ is a temporalised rule, $t'\in\mathcal{T}$ and $y\in\{tran,pers\}$.

Finally, we introduce meta-rules, that is, rules where the conclusion is not a simple duration literal but a temporalised rule. Thus a *meta-rule* is an expression

$$(s:A\hookrightarrow(r:B\hookrightarrow C)^{(t',x)})@(t,y),$$

where $(r:B\hookrightarrow C)^{(t',x)}$ is a temporalised rule, $r\neq s$, $t\in\mathcal{T}$ and $y\in\{tran,pers\}$. Notice that meta-rules carry only the viewpoint time (the validity time) but not the "in force" time. The intuition behind this is that meta-rules yield the conditions to modify a legal system. Thus they specify what rules (norms) are in a normative system, at what time the rules are valid, and the content of the rules. Accordingly, these rules must have an indication when they have been inserted in a normative system, but then they are universal (i.e., apply to all instants) within a particular instance of a normative system.

Every temporalised rule is identified by its rule label and its time. Formally we can express this relationship by establishing that every rule label r is a function

$$r:\mathcal{T}\mapsto \text{Rules}.$$

Thus a temporalised rule r^t returns the value/content of the rule 'r' at time t. This construction allows us to uniquely identify rules by their labels[9], and to replace rules by their labels when rules occur inside other rules. In addition there is no risk that a rule includes its label in itself. In the same way a temporalised rule is a function from $\mathcal{T}$ to Rules, we will understand a temporalised rule with viewpoint as a function with the following signature:

$$\mathcal{T}\mapsto(\mathcal{T}\mapsto \text{Rules}).$$

As we have seen in Section 4, a legal system LS is a sequence of versions $LS(t_0), LS(t_1),\dots$. The temporal dimension of viewpoint corresponds to a version while the temporal dimension temporalising a rule corresponds to the time-line

[9] We do not need to impose that the function is an injective: while each label should have only one content at any given time, we may have that different labels (rules) have the same content.

inside a version. Thus the meaning of an expression $r^{t_v}@t_r$ is that we take the value of the temporalised rule r^{t_v} in $LS(t_r)$. Accordingly, a version of LS is just a repository (set) of norms (implemented as temporal functions).

Accordingly, given a rule r, the expression $r^t@t'$ gives the value of the rule (set of premises and conclusion of the rule) at time t in the repository t'. The content of a void rule, e.g., $(r:\bot)^t@t'$ is $\bot$, while for the empty rule the value is the empty set. This means that the void rule has a value for the combination of the temporal parameters, while for the empty rule, the content of the rule does not exist for the given temporal parameters. Another way to look at the difference between the empty rule and the void rule is to consider that a rule is a relationship between a set of premises and a conclusion. For the void rule this relationship is between the empty set of premises and the empty conclusion; thus the rule exists but it does not produce any conclusion. For the empty rule, the relationship is empty, thus there is no rule. Alternatively, we can think of the function corresponding to temporalised rules as a partial function, and the empty rule identifies instants when the rule is not defined.

For a rule $(r:A \hookrightarrow B)^{(t,x)}@(t',y)$ or a meta-rule $(r:A \hookrightarrow B)@(t,x)$ we will use $A(r)$ to indicate the body or antecedent of the rule, i.e., A, and $C(r)$ for the head or consequent of the rule, i.e., B. Given a temporalised rule $(r:A \hookrightarrow B)^{(t,x)}$, its complement is defined as follows:

$$R[\sim r^t] = \{(r:\bot)^{(t,x)}\} \cup \{(r:\emptyset)^{(t,x)}\} \cup \{(r:A' \hookrightarrow B')^{(t,x)} | A' \neq A \text{ or } B' \neq B\}$$

Finally, for every literal, rule, and every temporal dimension, we have the specification whether the element is persistent or transient for that temporal dimension. The interpretation of transient and persistent elements is as follows: a transient temporalised literal $l^{(t,tran)}$, means that l holds at time t, while a persistent temporal literal $l^{(t,pers)}$ signals that l holds for all instants of time after t (t included), for the time-line of the legal system in which the literal is found. For a transient fully temporalised literal $l^{(t,x)}@(t',tran)$ the reading is that the validity of l at t is specific to the legal system corresponding to repository associated to t', while $l^{(t,x)}@(t',pers)$ indicates that the validity of l at t is preserved when we move to legal systems after the legal system identified by t'. An expression $r^{(t,tran)}$ sets the value of r at time t and just at that time, while $r^{(t,pers)}$ sets the values of r to a particular instance for all times after t (t included).

We will often identify rules with their labels, and, when unnecessary, we will drop the labels of rules inside meta-rules. Similarly, to simplify the presentation and when possible, we will only include the specification whether an element is persistent or transient only for the elements for which it is relevant for the discussion at hand.

Meta-rules describe the inference mechanism of the institution on which norms are formalised and can be used to establish conditions for the creation and modification of other rules or norms, while proper rules correspond to norms in a normative system. Thus a temporalised rule r^t gives the 'content' of the rule 'r' at time t; in legal terms it tells us that norm r is in force at time t. The expression

$$(p^{t_p}, q^{t_q} \Rightarrow (p^{t_p} \Rightarrow s^{(t_s,pers)})^{(t_r,pers)})@(t,tran)$$

means that, for the repository at t, if p is true at time t_p and q at time t_q, then $p^{t_p} \Rightarrow s^{(t_s,pers)}$ is in force from time t_r onwards.

A legal system is represented by a temporalised defeasible theory, i.e., a structure

$$(\mathcal{T}, F, R^{\text{nm}}, R^{\text{meta}}, R^{\text{mod}}, \prec)$$

where $\mathcal{T}$ is a totally ordered discrete set of time points, F is a finite set of facts (i.e., fully temporalised literals), R^{nm} is a finite set of unmodifiable rules, R^{meta} is a finite set of meta rules, R^{mod} is a finite set of proper rules, and $\prec$, the superiority relation over rules is formally defined as

$$\mathcal{T} \mapsto (\mathcal{T} \mapsto \text{Rules} \times \text{Rules}).$$

An unmodifiable rule is a rule such that $\forall t, t', t'', t'''\ r^t@t' = r^{t''}@t'''$.[10] This means that the content/value of the rule is the same across all repositories for all instants. The superiority relation $\prec$ determines the relative strength for rules for every instant in every version of the legal system. Thus it is possible that a rule r is both stronger and weaker than another rule s in two versions of the legal system, and then that two rules in different repositories have opposite relative strengths. To illustrate this case, consider a simplified version of the water restrictions in force in South East Queensland in January 2007,[11] where it is permitted to water gardens in residential properties on Tuesday, Thursday and Saturday for odd number properties and on Wednesday, Friday and Sunday for even number properties; and watering is otherwise forbidden. This regulation can be represented as follows

$$r : \Rightarrow \neg watering, \quad o : OddNumber \Rightarrow watering, \quad e : EvenNumber \Rightarrow watering$$

where the superiority contains, among others:

$$o \prec^{2007}_{Monday} r, \quad r \prec^{2007}_{Tuesdays} o, \quad e \prec^{2007}_{Monday} r, \quad r \prec^{2007}_{Wednesdays} e$$

This means that, according to the regulation in force in 2007, on Tuesday rule o is stronger than rule r, but on Monday r is stronger than o.

5.2 Proof Conditions

We are now ready to define how conclusions can be obtained in TDL. Notice that the main difference between the proof conditions given here and those of basic DL (of course besides the presence of the temporal dimensions) is that, in basic DL,

---

[10] Unmodifiable rules cannot be simulated in general by strict meta-rules with an empty antecedent. While we do not consider it in the present paper, it is possible to extend meta-rules so that the conclusion of a meta-rule is a meta-rule as well. Therefore, if unmodifiable rules were to be implemented as meta-rules, it would not be possible to ensure that they could not be changed.

[11] Given the worsening of the drought more stringent restrictions were first introduced to replace those presented here and then lifted when the rain eased the drought.

rules are always given as elements of the theory, while here every time we have to use a rule, we have to ensure that the rule is derivable from the theory. Given the structure of a theory and the types of the rules we have, the proof conditions for rules are slightly different from those for literals (though they follow the same intuition). Accordingly, we will give separate proof conditions for deriving literals and for deriving rules.

The main notion at hand is the notion of derivation (or proof). A *proof* P is a finite sequence of tagged expressions such that:

1. Each expression is either a temporalised rule or a temporalised literal;
2. Each tag is one of the following: $+\Delta t@t'$, $-\Delta t@t'$, $+\partial t@t'$, $-\partial t@t'$;
3. The proof conditions "Strict Rule Provability", "Defeasible Rule Provability", "Strict Literal Provability" and "Defeasible Literal Provability" given below are satisfied by the sequence P.

Given a proof P we use $P(n)$ to denote the n-th element of the sequence, and $P[1..n]$ denotes the first n elements of P.

A proof tag has four components: (1) sign, (2) tag, (3) derivation time and (4) repository time. Accordingly, the meaning of the proof tags is a follows:

- $+\Delta t@t'\ x^{t_x}$: we have a definite derivation of x^{t_x} at time t using the elements in the repository at time t';
- $-\Delta t@t'\ x^{t_x}$: we can show that it is not possible to have a definite derivation of x^{t_x} at time t using the elements in the repository at time t';
- $+\partial t@t'\ x^{t_x}$: we have a defeasible derivation of x^{t_x} at time t using the elements in the repository at time t';
- $-\partial t@t'\ x^{t_x}$: we can show that it is not possible to have a defeasible derivation of x^{t_x} at time t using the elements in the repository at time t'.

In the presentation of the proof conditions we will adopt the following convention for the various times involved: t_d is the time with respect to which we do the derivation and it refers to the time-line within a repository, t_r is the repository time, thus it is the time-line of the legal system as a whole. Finally, the last temporal dimension is the object time, which in the case of a rule is the time of force t_v, for a literal a it is the time when the literal holds; we use a^{t_a} for a temporal literal. The derivation and the repository times are parameters of the proof tags.

The general mechanism for a derivation in the present framework is as follows. First of all, a derivation corresponds to a query, and the query is parametrised by two temporal values: the repository time and the derivation time. The repository time is used to time-slice the information relevant for the query using the time-line of the legal system. This means that we retrieve all elements of the theory where the repository time is equal to the repository time of the query and all elements whose repository time is less than the repository time of the query but the elements carry over due to persistence over repositories. After this step we have the legal system in force at the repository time. At this stage the derivation time kicks in. Similarly to what we have done in the previous step, we use the value of the derivation time to time-slice the legal system under analysis. In particular we consider all rules whose time of force is equal to the derivation time, or rules whose time of force precedes the current derivation time but carries over to it because such rules are marked as persistent. Finally, we consider the temporalised literals in the rules resulting from the two previous steps, and we check whether the literals are provable with the time with which they appear in the rules.

Strict Rule Provability

If $P(n+1) = +\Delta t_d @ t_r\ r^{t_v}$ then
1) $r^{t'_v} @ t'_r \in R^{\text{nm}}$ or
2) $\exists s @ t'_r \in R_s^{\text{meta}} : \forall a^{t_a} \in A(s), +\Delta t_d @ t_r\ a^{t_a} \in P[1..n]$, or
3) $+\Delta t'_d @ t''_r\ r^{t'_v}$.

where:

1. if r is persistent, then $t'_v \le t_v$; if r is transient, then $t_v = t'_v$;
2. if facts, rules and meta-rules are persistent across repositories, then $t'_r < t_r$, otherwise $t'_r = t_r$;
3. $t'_d < t_d$ if conclusions are persistent within a repository; otherwise $t'_d = t_d$;
4. $t''_r < t_r$ if conclusions are persistent across repositories; otherwise $t''_r = t_r$.

Notice that for clause (2) we must be able to prove the antecedent of the meta-rule s with exactly the same reference point, i.e., the combination of derivation time t_d and repository time t_r as the reference point of the conclusion we prove, i.e., r^{t_v}; whether the literals used to apply s are obtained by persistence or by a direct derivation with the appropriate time reference depends on the proof conditions for literals and the variant of TDL at hand. Finally clause (3) is the persistence clause for strict derivation of rules.

Defeasible Rule Provability

If $P(n+1) = +\partial t_d @ t_r\ r^{t_v}$, then
1) $+\Delta t_d @ t_r\ r^{t_v}$ or
2) $-\Delta t_d @ t_r\ {\sim} r^{t_v}$ and
2.1) $r^{t'_v} @ t'_r \in R^{\text{mod}}$ or $\exists s^{t_s} \in R_{sd}^{\text{meta}}[r^{t'_v}] : \forall a^{t_a} \in A(s), +\partial t'_d @ t''_r\ a^{t_a} \in P[1..n]$ and
2.2) $\forall m^{t_m} \in R[{\sim} r^{t_v}]$ either
.1) $\exists b^{t_b} \in A(m) : -\partial t''_d @ t'''_r\ b^{t_b} \in P[1..n]$ or
.2) $m^{t_m} \prec_{t_d}^{t_r} r^{t_r}$, if $r^{t'_v} @ t'_r \in R^{\text{mod}}$ or $m^{t_m} \prec_{t_d}^{t_r} s^{t_s}$, if $r^{t'_v} @ t'_r \notin R^{\text{mod}}$ or
.3) $\exists w^{t_w} \in R[r^{t''_v}] : \forall c^{t_c} \in A(w), +\partial t'''_d @ t''''_r\ c^{t_c} \in P[1..n]$ and $m^{t_m} \prec_{t_d}^{t_r} w^{t_w}$

where

1. if r is persistent, then $t'_v \le t_v$; if r is transient, then $t_v = t'_v$;
2. if a^{t_a}, (resp. b^{t_b}, c^{t_c}) is persistent within the repository at t_r, then $t'_d \le t_d$ (resp. $t''_d \le t_d$, $t'''_d \le t_d$); if a^{t_a} (resp. b^{t_b}, c^{t_c}) is transient within the repository at t_r, then $t'_d = t_d$) (resp. $t''_d = t_d$, $t'''_d = t_d$);
3. if a^{t_a}'s, b^{t_b}'s and c^{t_c}'s are persistent with respect to repositories (i.e., conclusions are persistent), then $t''_r, t'''_r, t''''_r \le t_r$; otherwise $t''_r, t'''_r, t''''_r = t_r$;

4. if $r^{t'_v}$ and s (i.e., facts, rules, and meta-rules) are persistent with respect to repositories, then $t'_r \leq t_r$; otherwise $t'_r = t_r$.

A rule r is defeasibly provable at time t_d, given the information available in a repository t_r, if (1) the rule is strictly provable with the same parameters, or (2) we have definitely rejected that the content of the rule is different from what we want to prove, and then we have some justification to the claim. This means that (2.1) the rule is given in the theory, i.e., $r \in R^{\text{mod}}$. In this case, the rule can be given with a previous validity time (t'_v, $t'_v < t_v$) if that parameter is labelled as persistent. Similarly for the enactment time (or in-force time) t'_r. Notice that for a given rule, there are no constraints for the derivation time (t_d): given rules are understood as universally valid for that temporal dimension. For the second part of (2.1) we have that there is a meta-rule having r as its conclusion. In this case we have to check that the antecedent of the rule has been available to make the rule applicable at the derivation time t_d. This aspect depends on the particular variant of TDL one wants to adopt. The antecedent could have been derived at a previous time t_a, $t_a < t_d$, in a variant where conclusions persist within a repository; or with the same derivation time $t_a = t_d$, but in a previous version of the repository $t'_r < t_r$ if conclusions persist over repositories. We will fully explain these concepts when we present the proof conditions for literals. Clause (2.2) ensures that there are no justified reasons to claim the content of the rule is different from what we want to prove. Remember the given interpretation of rules (rule labels) as function from the temporal dimensions to the content of a rule (i.e., the relationship between the antecedent, a set of premises and the conclusion). Thus for every combination of temporal parameters for a rule, there is only a single value for the content of the rule. So, if we want to prove $+\partial 10@1\ (r : a^{t_a} \Rightarrow b^{t_b})$ we have to ensure that there is no way in which the content of rule r at time 10 in repository 1 is different from $a^{t_a} \Rightarrow b^{t_b}$.

Strict Literal Provability

If $P(n+1) = +\Delta t_d@t_r\ p^{t_p}$, then
1) $p^{t'_p}@t'_r \in F$; or
2) $\exists r^{t_v} \in R_s[p^{t'_p}]$, $+\Delta t_d@t_r\ r^{t'_v} \in P[1..n]$, $t'_v = t_d$ and
$\forall a^{t_a} \in A(r): +\Delta t_d@t_r\ a^{t_a} \in P[1..n]$; or
3) $+\Delta t'_d@t'_r\ p^{t'_p} \in P[1..n]$

where:

1. if p is persistent, then $t'_p \leq t_p$; if p is transient, then $t'_p = t_p$;
2. if r is persistent, then $t'_v \leq t_v$; if r is transient, then $t_v = t'_v$;
3. if facts, rules and meta-rules are persistent across repositories, then $t'_r < t_r$, otherwise $t'_r = t_r$;
4. if conclusions are persistent within a repository, then $t'_d < t_d$; otherwise $t'_d = t_d$;
5. if conclusions are persistent across repositories, then $t'_r < t_r$; otherwise $t'_r = t_r$.

What we want to point out for strictly literal provability is the mechanism governing persistence of conclusions. While persistence of rules and facts (within or across repositories) is a property of the single instances, persistence of conclusions is a property characterising variants of TDL. In this case, a conclusion is persistent within a repositories if it is possible to carry over this derivation from one instant to a successive instant while keeping the time reference relative to the repository unchanged. This means that, for example, if one is able to prove $+\Delta t@1\ p$, for $t = 10$ then for all $t' > 10$, $+\Delta t'@1\ p$ can be proved. Notice that in this case all we have to do is to provide a strict proof for p at time t using the information in the repository at time 1. For persistence of conclusions across repositories, on the other hand, we keep fixed the derivation time, but once a conclusion has been proved in a repository, it can be used in all repositories succeeding it. Thus, for example, if we prove $+\Delta 10@t\ p$, with $t = 2$, then for all $t' > 2$, we have $+\Delta 10@t'\ p$. Notice that, in this case, it is possible, as it often happens with abrogation (see Section 6 for details), that the reason for proving p in repository t no longer subsists in repositories after t. The two types of persistence of conclusions can be combined.

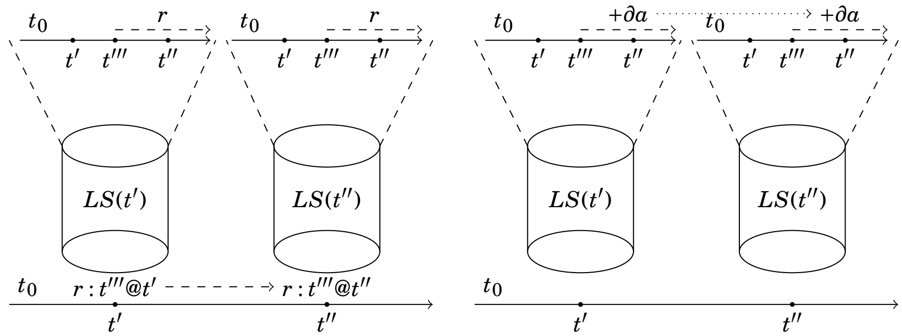
Defeasible Literal Provability

If $P(n+1) = +\partial t_d@t_r\ p^{t_p}$, then
1) $+\Delta t_d@t_r\ p^{t_p} \in P[1..n]$ or
2) $-\Delta t_d@t_r\ {\sim}p^{t_p} \in P[1..n]$ and
2.1) $r^{t'_v}@t_r \neq \emptyset$, $r^{t'_v}@t'_r \in R_{sd}[p^{t'_p}]$, $+\partial t'_d@t'_r\ r^{t'_v} \in P[1..n]$ and
$\forall a^{t_a} \in A(r)$, $+\partial t'_d@t'_r\ a^{t_a} \in P[1..n]$, and
2.2) $\forall s^{t_s} \in R[{\sim}p^{t_{\sim p}}]$ if $+\partial t''_d@t''_r s^{t'_s} \in P[1..n]$, then either
.1) $\exists b^{t_b} \in A(s), -\partial t''_d@t''_r\ b^{t_b} \in P[1..n]$ or
.2) $\exists w^{t_w} \in R[p^{t''_p}]$ such that $+\partial t''_d@t''_r\ w^{t_w} \in P[1..n]$ and
$\forall c^{t_c} \in A(w), +\partial t''_d@t''_r\ c^{t_c} \in P[1..n]$ and $s^{t_s} \prec^{t_r}_{t_d} w^{t_w}$

where

1. if p is persistent, $t'_p \leq t_{\sim p} \leq t_p$, otherwise $t'_p = t_{\sim p} = t_p$;
2. $t'_s \leq t_v$, if s is persistent, otherwise $t_s = t'_s = t_v$;
3. $t_d \leq t'_s$, if s is persistent, otherwise $t_s = t'_s = t_d$;
4. if conclusions are persistent over derivations (i.e., $+\partial t'_d@t_r p^{p_t}$ implies $+\partial t_d@t_r p^{p_t}$ where $t'_d < t_d$), $t'_d \leq t''_d \leq t_d$; otherwise $t'_d = t''_d = t_d$;
5. if conclusions are persistent over repositories, $t'_r \leq t''_r \leq t_r$; otherwise $t'_r = t''_r = t_r$.

The conditions to defeasibly derive literals inherit intuitions from standard DL and the features described for the other types of derivation. The mechanism of persistence of conclusions over derivation (or within a repository) is essentially the same as that of strict conclusions. The main difference regards the way conclusions persist across repositories. In this case it is not enough that a defeasible derivation existed in previous repository. What is required is that the rule used to prove the conclusion still exists in the current repository w.r.t. the time it was



(a) Rule Persistence. A persistent rule r enacted at time t' and in force at t''' carries over from the legal system $LS(t')$ to the legal system $LS(t'')$, where it is still in force at t'''.

(b) Causal Conclusion Persistence. A conclusion is causal if it persists from a legal system $LS(t')$ to a legal system $LS(t'')$ even if the rules used to derive it are no longer effective in $LS(t'')$.

Fig. 2. Rules and Conclusions Persistence

valid in the previous repository, plus the conclusions that are to be proved can carry over from one repository to successive ones. Thus, for example, if we are able to prove $+\partial 10@1\ p^{t_p}$ because we can prove $+\partial 10@1\ (r : a^{t_a} \Rightarrow p^{t_p})$ and $+\partial 10@1\ a^{t_a}$, then we must be able to prove $+\partial 10@2\ a^{t_a}$ and that r has not been revoked after 1.

The proof conditions given above produce classes of variants of TDL, according to conditions on the temporal parameters. In particular, it is possible to define variants capturing different types of persistence. Of particular relevance to norm modifications we mention *rule persistence* and *causal conclusion persistence*.

Generally once a norm has been introduced in a legal system, or better in a specific version of it, the norm continues to be in the system unless it is explicitly removed. This means that the norm must be included in all versions succeeding the one in which it has been first introduced (see Figure 2(a) for a graphical representation of this phenomenon). This effect is achieved by specifying that the derivation of rules is persistent over repositories. On the other hand, if we can prove a conclusion with respect to a specific version of the legal system in some cases we have to propagate it to successive versions. In particular, this is the case when we have causal conclusions. However, for some type of norm modifications, namely annulment, we have to block the persistence of conclusions over repositories when the reasons for deriving these conclusions are no longer in the system. See Figure 2(b) for a graphical representation of causal conclusion persistence. This effect depends on whether derivations of conclusions are persistent over repositories, and it is in function of the particular type of modification we want to implement.

To illustrate these ideas consider the following theory:

$$(r : a^{10} \Rightarrow b^{(20,pers)})^{10}@(1, tran) \qquad (s : b^{30} \Rightarrow c^{(30,pers)})^{15}@(1, pers)$$

Since r is marked as transient, the rule can be used only in repository 1, while s can be used in all repositories after repository 1.[12] Given $a^{10}@1$ we can first derive $+\partial 10@1\ b^{(20,pers)}$. Since b is persistent we have $+\partial 10@1\ b^{30}$, since b persists from 20 to 30. But the second rule cannot be applicable, since its validity time is 15. Thus, to apply it we have to assume that derivations are persistent within a repository. If this is the case then we obtain $+\partial 15@1\ b^{30}$, which makes rule s applicable, and from which we get $+\partial 15@1\ c^{30}$. If we have that conclusions are persistent across repositories, then we can conclude $+\partial 15@2\ c^{30}$. Notice that we can conclude $+\partial 15@2\ c^{30}$ even if the reasons for deriving it (i.e., rule r) do not persist across repositories. The point to note for conclusion persistence is that, if we have a derivation in a preceding repository and the derivation is not 'killed' in successive repositories, we can carry over the conclusion from the repository where the conclusion has been proved to successive repositories.

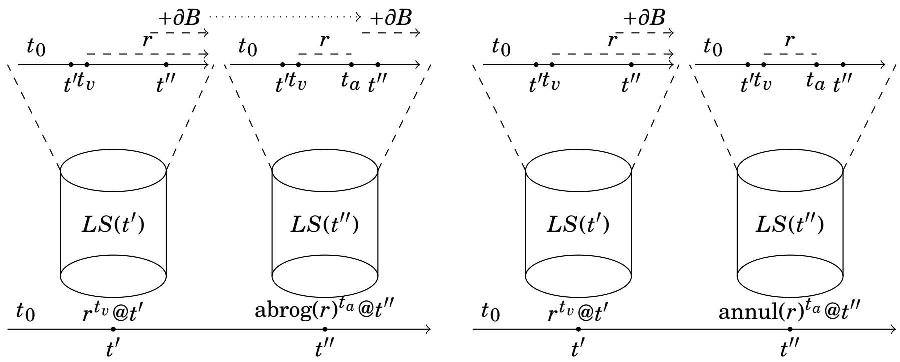
6 Abrogation and Annulment in TDL

Let us apply TDL to abrogation and annulment. Both modifications cancel norms from a legal system. Let $LS(t')$ be the repository containing a modifiable rule $(r : A \hookrightarrow B)^{(t_v,pers)}$ such that $t' < t_v < t''$, and $LS(t'')$ be the subsequent repository where we apply the modification of r, which is effective from a certain time t_a. Then $LS(t'')$ will contain $(r : \emptyset)^{(t_a,pers)}$. This makes the previous version of r inapplicable in $LS(t'')$ from t_a, and so, there, we no longer obtain B using r:[13] condition (2.1) for Defeasible Literal Provability states that $(r : A \hookrightarrow B)$ is applicable if it is provable with that content, but this does not hold after the modification (see condition (2.2) for Defeasible Rule Provability).

Does this solution solve all problems? Suppose the modification is retroactive, such that $t_v < t_m < t''$. This means that $LS(t'')$ contains an applicable meta-rule such as $(mr : A' \hookrightarrow (r : \emptyset)^{(t_a,pers)})@t''$. Note that the effect of mr is persistent to guarantee that r is null from t_a onwards. With some examples of *abrogation*, this measure may work, as we block the derivation of B, based on r in $LS(t'')$, from t_a onwards. Accordingly, if B was derived (with its own appropriate time) in $LS(t')$, it can carry over from $LS(t')$ to $LS(t'')$ (see Figure 3(a)). But this does not apply to annulments, for which B can carry over only through the inner time-line of $LS(t')$. In [13] we suggested that the solution is that annulment is obtained by blocking persistency of derivations across repositories. In other words, conclusions of the annulled rule are only derived in the repository in which the modification does not occur (see Figure 3(b)). However, no technical solution was offered. Our solution is as follows. Let the positive defeasible extension of a theory T be the set $E^{+\partial}(T) = \{p^{t_p}@t | T \vdash +\partial\#t'@tp^{t_p}\}$.

[12] To make the example simpler we have used two different scales for the derivation time and the repository time. Anyway in legal reasoning these will be on the same time scale.

[13] For simplicity, let us explicitly reason for the moment on the repository time and time of force only. For example, B will also have a temporal parameter and, if persistent, will hold from then onwards. Let us assume that B is persistent and its time is slightly after t_a.



(a) Abrogation. In $LS(t')$ rule r produces a persistent effect B. B carries over by persistence to $LS(t'')$ even if r is no longer in force.

(b) Annulment. In $LS(t')$ rule r is applied and produces a persistent effect B. Since r is annulled in $LS(t'')$, B must be undone as well.

Fig. 3. Abrogation and Annulment

Abrogation Given a rule $(r : A \hookrightarrow b^{t_b})^{t_r}@t$, the abrogation of r at t_a in repository $t'$is defined as follows:

$$T_r^{abr(t_a,t')} = \begin{cases} T & \text{if } r \notin E^{+\partial}(T) \\ (F, R', \prec) & \text{otherwise} \end{cases} \tag{9}$$

where $R' = R \cup \{(abr_r : \Rightarrow (r : \bot)^{(t_a,pers)})@(t',pers)\}$, where $t' > t$. The abrogation simply terminates the applicability of the rule. More precisely this operation sets the rule to the void rule. The rule is not removed from the system, but it has now a form where no longer can produce effects. This is in contrast to what we do for annulment where the rule to be annulled is set to the empty rule. This essentially amounts to removing the rule from the repository. From the time of the annulment the rule has no longer any value.

Annulment The definition of a modification function for annulment depends on the underlying variants of TDL, in particular whether conclusions persist across repositories. In a variant where conclusions do not persist over repositories, the operation can be simply defined by the introduction of a meta-rule setting the rule to be annulled to $\emptyset$, with the time when the rule is annulled and the time when the meta-rule is inserted in the legal system.

$$T_r^{ann\text{-}tran(t_a,t')} = \begin{cases} T & \text{if } r \notin E^{+\partial}(T) \\ (F, R', \prec) & \text{otherwise} \end{cases} \tag{10}$$

where $R' = R \cup \{(mr : \Rightarrow (r : \emptyset)^{(t_a,pers)})@(t',pers)\}$ Thus to annul the rule $r^{t_r}@t$ at t_a we introduce a meta-rule whose conclusion is the empty rule.

In a variant where conclusions persist over repositories we need some additional technical machinery: given a set of duration literals D, a set of temporalised literals T and a total discrete ordered $(\mathscr{T}, <)$, we define

$$D \cap_{(\mathscr{T},<)} T = \left\{ l^t \in T | \exists l^{(t',x)} \in D : t' = t \text{ if } x = tran, \text{ and } t \leq t' \text{ otherwise} \right\}$$

Given a duration literal $b^{(t,x)}$ and a theory T, we define the *dependence set*, i.e., the set of literals (called *critical literals*) potentially depending on it, as follows:

$$Dep(b^{(t,x)}) = \left\{ b^{(t,x)} \right\} \cup \left\{ c^{(t_c,x_c)} | \exists r \in R : C(s) = c^{(t_c,x_c)} \wedge A(r) \cap_{(\mathscr{T},<)} Dep(b^{(t,x)}) \neq \emptyset \right\}$$

Then, if the annulment applies at t_a in repository t'

$$T^{annul(t_a,t')}_{(r:a_1,\ldots,a_n \hookrightarrow b^{(t_b,x)})^{t_r}@t} = \begin{cases} T & \text{if } r \notin E^{+\partial}(T) \\ (F, R', \prec') & \text{otherwise} \end{cases} \tag{11}$$

where[14]

$$\begin{aligned} R' = R \cup \{ & (r : \emptyset)^{(t_a,pers)}@(t',pers), \\ & (r^{\sim} : \rightsquigarrow \sim b^{(t_b,x)})^{(t_a,pers)}@(t',tran), \\ & (r^{ann} : \Rightarrow ann(b)^{(t_b,x)})^{(t_a,pers)}@(t',tran)\} \\ \cup \{ & (s^{rep} : A(s) - Dep(C(r)) \cup \{ann(a)^{t_a} | a \in A(s) \cap_{(\mathscr{T},<)} Dep(C(r))\} \hookrightarrow \\ & \qquad ann(C(s))^{(t_a,pers)})^{(t_a,tran)}@(t',tran), \\ & (s^{ann} : ann(C(s))^{t_a} \rightsquigarrow \sim C(r))^{(t_a,pers)}@(t',tran) | \\ & \qquad \text{if } A(s) \cap_{(\mathscr{T},<)} Dep(C(r)) \neq \emptyset \} \\ \cup \{ & (s^{nan} : A(s) \Rightarrow \sim ann(C(r))^{(t_a,tran)})^{(t_a,tran)}@(t',tran) | \\ & \qquad \text{if } C(s) \in Dep(C(r)) \wedge A(r) \cap_{(\mathscr{T},<)} Dep(C(r)) = \emptyset \} \\ \prec' = \prec \cup & \{(t', t_a, r^{rep}, r) | r \in R\} \cup \{(t', t_a, r^{nan}, s^{ann}) | r, s \in R\} \end{aligned}$$

The idea behind this construction is to introduce new (auxiliary) literals to signal whether literals are eventually revoked (declared null) as a consequence of an annulment. Then, for every rule where literals depending on the conclusion of the rule to be annulled occur in the antecedent, we create a copy of the rule where all critical literals are replaced by auxiliary literals. Moreover, for each critical literal its auxiliary literal is the body of a defeater for the complement of the critical literal. Finally, for each rule for a critical literal different from the conclusion of the rule to be annulled where no critical literal appears in the antecedent, we create a defeasible rule with the same body and as conclusion the complement of the critical literal.

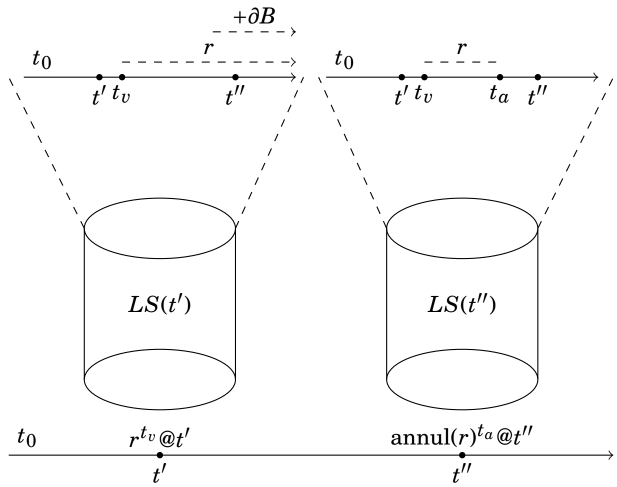
[14] To simplify the notation in the rest of the definition the rules are the conclusion of a meta-rule (each with a unique name), thus the expression $(r : a_1, \ldots, a_n \hookrightarrow b^{t_b})^{t_r}@t$ must be understood as the abbreviation of the meta-rule (with empty body) $(mr : \Rightarrow (r : a_1, \ldots, a_n \hookrightarrow b^{t_b})^{t_r})@t$. For a literal l, $ann(l)$ is a new literal not occurring in the theory.

Note that the above construction guarantees that for every pair $(l, ann(l))$ at most one of them is defeasibly provable, and that, if the strict part of the theory is consistent, then if $+\partial t@t'\ ann(l)$, then $-\partial t@t'\ l$. The intuition here is that the introduction of the meta-rule setting the rule to be annulled to $\perp$ determines that we no longer carry over the conclusion of the rule from one repository to the next one. However, this does not prevent conclusions depending on it to pass over (after all, at the time they were derived we had valid reasons to derive them, and unless some preventing reasons occurred after, we have no reasons to stop them to pass from one repository to next one). Hence, we need specific reasons to stop them. Thus the idea to refute them is to add the explanation that they were derived from 'causes' declared null in a successive step, and thus they must be null as well.

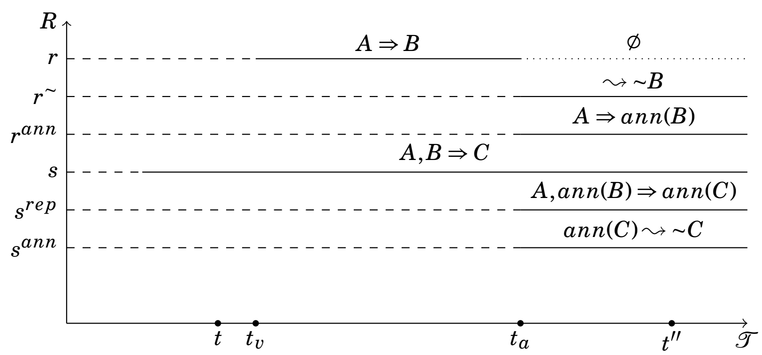
6.1 Discussion

Although the abrogation of a norm r at time t may raise a number of legal perplexities, its logical treatment is relatively straightforward, at least if it is taken to be a modification which, when effective, only blocks the normative effects of r holding after t. It should be noted that we model this modification as a quite peculiar type of norm removal. Indeed, we do not properly remove r, but only its content; this happens only from the time t when the abrogation applies. Accordingly, r is made entirely inapplicable after t, but its previous version, in force before t, is still in the legal repository, and so, if a case to be settled is dated before t, the abrogated norm can still be applied (see Section 1).

Annulments, though conceptually simple, require instead a number of technical devices. They are simple, as they properly correspond to norm removals and so are meant to block all past, present and future effects. (Unless, of course, the annulment is subsequently undone.) As we noted, if the conclusions of a norm r do not persist over repositories, annulments, too, are trivial. Things get difficult when r's conclusions persist over repositories. Here, the difficulty is to detect and block all effects of r even though they are obtained within the previous repositories. Figure 4 graphically shows how the machinery works. From an external point of view—from $LS(t')$ to $LS(t'')$—the annulment of a norm $r : A \Rightarrow B$ simply corresponds to the intuition of Figure 3(b). But, from the internal viewpoint of $LS(t'')$, we can see the machinery at work. The annulment applies from time t_a. Then, from t_a onwards r is set to be the empty rule and no longer exists in the repository (first line from the top). However, we should block all r's effects. All temporal instances of its persistent direct effect B are blocked by a defeater $r^{\sim}$ in force from t_a (second line from top). All temporal instances of its persistent indirect effect C, which can be obtained via the rule $A, B \Rightarrow C$ (marked on the third line from bottom), are undone as well by means of some auxiliary rules. Rule r^{ann} works in such a way as to mark B as a revoked consequence (third line from top); rule s^{rep} marks C as revoked, too, and captures the dependence of this revoked literal on other literals involved in s; finally, r^{ann}, in the case C is derived as revoked, defeats the derivation of C, and so blocks all temporal instances of the indirect effects of r.



(a) External view of the annulment of rule $r : A \Rightarrow B$ from $LS(t')$ to $LS(t'')$



(b) Internal view of annulment of $r : A \Rightarrow B$ in $LS(t'')$.

Fig. 4. Annulment: Removing and Blocking

Finally, let us see an example illustrating our approach.

Example 5. Consider a legal system at time 1 encoded in the following theory T:

$$
\begin{aligned}
F &= \{a^{10}@(1,pers), c^{10}@(1,pers), f^{10}@(1,pers)\},\\
R &= \{(r_1 : a^{10} \Rightarrow b^{(10,pers)})^{(10,pers)}@(1,pers),\\
&\quad (r_2 : b^{10}, c^{10} \Rightarrow d^{(10,pers)})^{(10,pers)}@(1,pers),\\
&\quad (r_3 : d^{10} \Rightarrow e^{(10,pers)})^{(10,pers)}@(1,pers),\\
&\quad (r_4 : f^{10} \Rightarrow e^{(10,tran)})^{(10,pers)}@(1,pers)\}\\
&<= \emptyset.
\end{aligned}
$$

Clearly we can prove $+\partial 10@1\ X$ for $X \in \{a,b,c,d,e,f\}$. Now suppose that the legal system is changed by revoking rule r_1, and that the change is valid from 10. The resulting legal system is the legal system at time 2.

If the change is an abrogation, then, the resulting legal system at time 2 is obtained by the addition of the rule $(r_1^{abr} : \bot)^{(10,pers)}@(2,pers)$. Accordingly, if a (the antecedent of r before the abrogation) were holding before 10 and within the repository at 2, we could still derive the consequent b.

The legal system at 2 is obtained by adding the rules (meta-rules) implementing the annulment function $T_{r_1}^{ann(10,2)}$. Namely we revise T by introducing the rules

$$
\begin{gathered}
(r_1 : \emptyset)^{(10,pers)}@(2,pers),\\
(r_1^{\sim} : \rightsquigarrow b^{(10,pers)})^{(10,pers)}@(2,tran),\\
(r_1^{ann} : \Rightarrow ann(b)^{(10,pers)})^{(10,pers)}@(2,tran),\\
(r_2^{rep} : c^{10}, ann(b)^{10} \Rightarrow ann(d)^{(10,pers)})^{(10,pers)}@(2,tran),\\
(r_2^{ann} : ann(d)^{10} \rightsquigarrow \neg d^{(10,pers)})^{(10,tran)}@(2,tran),\\
(r_3^{rep} : ann(d)^{10} \Rightarrow ann(e)^{(10,pers)})^{(10,pers)}@(2,tran),\\
(r_3^{ann} : ann(e)^{10} \rightsquigarrow \neg e^{(10,pers)})^{(10,tran)}@(2,tran),\\
(r_4^{nan} : f^{10} \Rightarrow \neg ann(e)^{(10,tran)})^{(10,tran)}@(2,tran).
\end{gathered}
$$

From the point of view of the legal system at 1 we have a derivation of b^{10} at 10 ($+\partial 10@1\ b^{10}$). Thus, allowing conclusions to persist over repositories, would mean that we can carry over the derivation of it to the repository at 2 (Clause 2.1 of Defeasible Literal Provability, plus condition on conclusion persistence). But setting rule r_1 to $\emptyset$ in 2 produces the effect that now the rule no longer exists and thus it cannot longer be used. Hence we block the derivation b^{10}, more precisely, $-\partial 10@2\ b^{10}$. When we look at d, without rules r_2^{rep} and r_2^{ann}, d^{10} ($+\partial 10@1\ d^{10}$) was obtained from the viewpoint of 1. Rule r_2 has not been revoked, and at the time the conclusion was derived, the rule was applicable (i.e., the antecedent was provable). Thus, the conclusion passes from 1 to 2, that is $+\partial 10@2\ d^{10}$ would be derivable. However, this conclusion was the result of an act declared null by the (retroactive) annulment. Finally, for e we have that there are two rules for it. In the first rule (r_3) where e depends on some annulled literal, but this is not the case for the second rule (r_4). In the annulment r_3 generates r_3^{rep}, and r_4 generates r_4^{nan}. These two rules are in conflict which each other, but $r_3^{rep} \prec_{10}^{2} r_4^{nan}$, thus r_4^{ann} prevails, and we are able to prove $ann(e)$, $-\partial 10@2\ ann(e)$, so r_3^{ann} is not applicable at 10 w.r.t. repository 2. Hence we can use r_4 to continue to derive $+\partial 10@2\ e$.

7 Conclusions

In this paper we investigated how to model legal abrogations and annulments in Defeasible Logic. Terminology may vary from one legal system to another, but, despite this, it is possible to identify in general two different reasoning patterns: in one case norms are removed with all their effects, whereas in other cases norms are removed but all or some of their effects propagate if obtained before the modification. We examined some ways to capture these intuitions in DL using techniques from revision based on belief sets and from base revision. We concluded that abrogation and annulment can only be partially represented in these settings. In addition, we argued that it is hard, if not impossible, to simulate retroactivity, which clearly refers to the temporal dimension. Hence, we illustrated a different conceptual starting point from which the problem can be addressed.

Hence, we extended the logic presented in [12] to capture different temporal aspects of abrogations and annulments. This extension increases the expressive power of the logic and it allows us to represent meta-norms describing norm-modifications by referring to a variety of possible time-lines through which conclusions, rules and derivations can persist over time.

We outlined the inferential mechanism needed to deal with the derivation of rules and literals. In particular, for each proof condition we identified several temporal constraints that permit to allow for, or block, persistency with respect to specific time-lines. This virtually leads to define different variants of TDL according to whether a condition is adopted or not. Then we described some issues related to norm modifications and we illustrated the techniques with respect to annulment and abrogation. We showed that the temporal formalism introduced here is able to deal with complex scenarios such as retroactivity. In particular, we solved the problem of how legal effects of *ex-tunc* modifications, such as annulment, can be blocked after the modification is applied. The idea we suggested is to block persistency of derivations across repositories. In other words, the conclusions of the annulled rule will only be derived in the repository in which the modification does not occur.

Typically there are two mainstream approaches to reasoning with and about time. A point based approach, as in the present paper, and an interval based approach [2]. Notice that the current approach is able to deal with constituents holding in an interval of time, thus an expression $\Rightarrow a^{[t_1,t_2]}$ meaning that a holds between t_1 and t_2 can just be seen as a shorthand of the pair of rules $\Rightarrow a^{(t_1,pers)}$ and $\leadsto \neg a^{(t_2,tran)}$. Currently it is not clear what benefits would result from an interval based temporalised defeasible logic for the intended application. Anyway we would like to point out that interval and duration based temporal defeasible logic have been developed [5,17]. [17] focuses on duration and periodicity and relationships with various forms of causality. [5] proposed a sophisticated interaction of defeasible reasoning and standard temporal reasoning (i.e., mutual relationships of intervals and constraints on the combination of intervals). In both cases it is not clear whether the techniques employed there are relevant to the application to norm modifications, and such works consider only a single temporal dimension, and do not have meta-rules.

Acknowledgements 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. Antonino Rotolo was supported by the EU project ESTRELLA (IST-2004-027655).

This paper is an extended and revised version of [14] and [15].

References

1. C.E. Alchourrón, P. Gärdenfors, and D. Makinson. On the logic of theory change: Partial meet contraction and revision functions. *Journal of Symbolic Logic*, 50:510–530, 1985.
2. J.F. Allen. Towards a general theory of action and time. *Artificial Intelligence*, 23:123–154, 1984.
3. G. Antoniou, D. 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, D. 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. J.C. Augusto and G.R. Simari. Temporal defeasible reasoning. *Knowledge and Information Systems*, 3(3):287–318, 2001.
6. 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.
7. D. Billington, G. Antoniou, G. Governatori, and M.J. Maher. Revising nonmonotonic belief sets: The case of defeasible logic. In Wolfram Burgard, Thomas Christaller, and Armin B. Cremers, editors, *KI-99: Advances in Artificial Intelligence*, pages 101–112. Springer, 1999.
8. G. Boella and L. van der Torre. Fulfilling or violating obligations in multiagent systems. In *2004 IEEE/WIC/ACM International Conference on Intelligent Agent Technology (IAT 2004)*, pages 483–486. IEEE Computer Society, 2004.
9. P. Di Giusto and G. Governatori. A new approach to base revision. In P. Barahona and J.J. Alferes, editors, *9th Portuguese Conference on Artificial Intelligence (EPIA 1999)*, pages 327–341. Springer, 1999.
10. P. Gärdenfors. *Knowledge in Flux: Modeling the Dynamics of Epistemic States*. MIT, 1988.
11. G. Governatori, M.J. Maher, D. Billington, and G. Antoniou. Argumentation semantics for defeasible logics. *Journal of Logic and Computation*, 14(5):675–702, 2004.
12. G. Governatori, M. Palmirani, R. Riveret, A. Rotolo, and G. Sartor. Norm modifications in defeasible logic. In *Legal Knowledge and Information Systems (JURIX 2005)*, pages 13–22. IOS Press, 2005.
13. G. Governatori, M. Palmirani, R. Riveret, A. Rotolo, and G. Sartor. Variants of temporal defeasible logic for modelling norm modifications. In *11th International Conference on Artificial Intelligence and Law (ICAIL 2007)*, pages 155–159. ACM Press, 2007.
14. G. Governatori and A. Rotolo. Changing legal systems: Abrogation and annulment. Part I: Revision of defeasible theories. In R. van der Meyden and L. van der Torre, editors, *9th International Conference on Deontic Logic in Computer Science (DEON 2008)*, pages 3–18, 2008.
15. G. Governatori and A. Rotolo. Changing legal systems: Abrogation and annulment. Part II: Temporalised defeasible logic. In *Proceedings NorMAS 2008*, 2008.
16. G. Governatori, A. Rotolo, and G. Sartor. Temporalised normative positions in defeasible logic. In *10th International Conference on Artificial Intelligence and Law (ICAIL 2005)*, pages 25–34. ACM Press, 2005.

17. G. Governatori and P. Terenziani. Temporal extensions to defeasible logic. In M.A. Orgun and J. Thornton, editors, *20th Australian Joint Conference on Artificial Intelligence*, pages 476–485, 2007.
18. R. Guastini. *Teoria e dogmatica delle fonti*. Giuffré, Milan, 1998.
19. S.O. Hansson. *Belief Base Dynamics*. Phd thesis, Uppsala University, 1991.
20. M.J. Maher. Propositional defeasible logic has linear complexity. *Theory and Practice of Logic Programming*, 6(1):691–711, 2001.
21. B. Nebel. A knowledge level analysis of belief revision. In R. Brachman, H. Levesque, and R. Reiter, editors, *First International Conference on Principles of Knowledge Representation and Reasoning – KR'89*, pages 301–311. Morgan Kaufmann, 1989.
22. D. Nute. Defeasible reasoning. In *Proceedings of 20th HICSS*. IEEE press, 1987.
23. H. Prakken. *Logical Tools for Modelling Legal Argument*. Kluwer, Dordrecht, 1997.
24. G. Sartor. *Legal Reasoning*. Springer, Dordrecht, 2005.
25. L. van der Torre, G. Boella, and H. Verhagen, editors. *Normative Multi-agent Systems*, Special Issue of *JAAMAS*, vol. 17(1), 2008.

A Proofs of the Propositions

In [3] a thorough investigation of the proof theory of Defeasible Logic was carried out, and the conditions between the various proof tags for a literal (and its complement) were established. Given that we have four proof tags, we could have sixteen (2^4) possible combinations. However, only six are possible based on the proof conditions. These are listed in Table 5(a).[15]

	$+\Delta$	$-\Delta$	$+\partial$	$-\partial$
A	p		p	
B		p	p	
C			p	
D		p		
E		p		p
F				

(a) Possible combinations of tags for an atom

p \ $\neg p$	A	B	C	D	E	F
A	✓				✓	✓
B					✓	
C					✓	
D				✓	✓	✓
E	✓	✓	✓	✓	✓	✓
F	✓			✓	✓	✓

(b) Combinations of tags for a literal and its complement

Fig. 5. Relationships between proof tags

For a literal and its complement we have $6 \times 6 = 36$ combinations. However, only 17 are possible. See Table 5(b) for the list of the possible combinations.

For example, for an atom q, the combination DE means that $q \in \Delta^-$, $\neg q \in \Delta^-$ and $\neg q \in \partial^-$ corresponding to the theory where the only rule for $p/\neg p$ is $p \Rightarrow p$.

[15] The meaning of the figure is that if p is in a column then p is in the set corresponding to the column, otherwise p is not in the set. For row F this means that p is not in any of the sets. This is only possible if the theory contains $p \to p$, $\neg p \to \neg p$ and no other strict rules for p, $\neg p$.

Definition 3. *Let S be a set of atomic propositions and E be a 4-tuple of sets of literals where the sets of literals are subsets of $L = S \cup \{\neg p | p \in S\}$. E is a* defeasible logic extension *iff for all pairs of literals p, $\neg p$ in L one of the conditions listed in Table 5(b) is satisfied.*

If the extension of a theory satisfies property X listed in Table 5(a) for a literal l we will say that the theory is of type X for literal l (or simply of type X). Similarly for Table 5(b) and an atom q.

Proposition 1. *Let T be a defeasible theory and T' the theory generated from the extension of T. For every $p \in HB_T$, $T \vdash \# \pm p$ iff $T' \vdash \# \pm p$.*

Proof. We prove first the case, if $T \vdash \pm\# p$ then $T' \vdash \pm\# p$.

Case $+\Delta$. If $T \vdash +\Delta p$, then $p \in \Delta^+(T)$. Hence in T' we have the rule $\to p$, for which clause $+\Delta 1$ is satisfied. So $T' \vdash +\Delta p$.

Case $-\Delta$. If $T \vdash -\Delta p$, then $p \in \Delta^-(T)$. Hence there is no strict rule with head p in T', i.e., $R'_s[p] = \emptyset$. Therefore the clauses of proof condition for $-\Delta$ are vacuously satisfied, and $T' \vdash -\Delta p$.

Case $+\partial$. If $T \vdash +\partial p$, then $p \in \partial^+(T)$. This means that we have one the following two cases: (i) $R'[p] = \{\Rightarrow p\}$ or (ii) $R'[p] = \{\Rightarrow p, \to p\}$. Let us consider (ii) first. We can use $\to p$ to derive $T' \vdash +\Delta p$ as we have seen in the first case, and then $T' \vdash +\partial p$ by clause $+\partial 1$. For (ii) we use the result of [3] that shows that when $p \notin \Delta^+(T)$ but $p \in \partial^+(T)$, then $\neg p \notin \Delta^-(T) \cup \partial^-(T)$. This means that there are no rules for $\neg p$ in T', hence $+\partial 2.3$ is vacuously satisfied, and so $T' \vdash +\partial p$.

Case $-\partial$. If $T \vdash -\partial p$, then $p \in \partial^-(T)$, this means that the extension is of type E (Table 5(a)). Accordingly, $p \in \Delta^-(T)$ and in addition we have $\Delta^-(T) \cup \Delta^+(T) = \emptyset$ and $\partial^-(T) \cup \partial^+(T) = \emptyset$ (see [3]). Thus $R_{sd}[p] = \emptyset$, and $-\partial 1$ is vacuously satisfied. Therefore $T' \vdash -\partial p$.

We can now prove the other direction, i.e., if $T' \vdash \pm\# p$ then, $T \vdash \pm\# p$.

Case $T' \vdash +\Delta p$. This means $\to p \in R'_s[p]$, thus $p \in \Delta^+(T)$, thus, by definition, $T \vdash +\Delta p$.

Case $T' \vdash -\Delta p$. This means $R'_s[p] = \emptyset$, thus, $p \notin \Delta^-(T)$, and by definition $T \vdash -\Delta p$.

Case $T' \vdash +\partial p$. This means that we have either $\to p$ or $\Rightarrow p$ in $R[p]$. In both cases we have that $p \in \partial^+(T)$, thus by definition $T \vdash +\partial p$.

Case $T' \vdash -\partial p$. Since the superiority relation in T' is empty, we can consider three cases. (1) $T' \vdash +\Delta{\sim}p$ (2) $R_{sd}[p] = \emptyset$ or (3) $R[{\sim}p] \neq \emptyset$. For (1) we have that $R_s[{\sim}p] = \{\to {\sim}p\}$, thus ${\sim}p \in \Delta^-(T)$, and, by definition $T \vdash +\Delta{\sim}p$, which then implies $T \vdash -\partial p$. For (2) we have that $p \in \partial^-(T)$, thus $T \vdash -\partial p$. For (3) by construction of T' we have two possible cases $R[{\sim}p] = \{\to {\sim}p, \Rightarrow {\sim}p\}$ or $R[{\sim}p] = \{\Rightarrow p\}$. In the first case we have that $p \in \Delta^+(T)$. Accordingly T could be a theory of type AA, EA or FA. If it were AA, then $p \in \Delta^+(T)$ and so $\to p \in R'$ and so $T' \nvdash -\partial p$. Similarly if T were of type FA, then $p \notin \Delta^-(T) \cup \partial^-(T)$, and so $p \to p \in R'$, and again $t' \nvdash -\partial p$. Hence T is of type EA, which means that $p \in \partial^-(T)$, and so $T \vdash -\partial p$. For the second case, we have that T is either of type EB or EC, and again we can conclude that $T \vdash -\partial p$.

Proposition 2. *Given a theory* T *and a rule* $r : A_1, \dots, A_n \Rightarrow B$ *such that* $T \vdash +\partial B$*, then for every* $C \in HB_T - \{B\}$*,* $T \vdash C$ *iff* $T_r^{abr} \vdash C$*.*

Proof. What we have to show here is how to transform a proof in T into a proof in T' and the other way around. First of all we have that $+\partial B$ is provable in T, but the presence of r^- prevents us to use proofs in T as they are in T'. In case the abrogated rule was not the only rule needed to prove B in T, then we can use these proofs unchanged as proofs in T'; otherwise we are no longer able to prove $+\partial B$ in T'. By construction every rule where B occurs in the antecedent is supplemented by a rule (related, in the superiority relation, with the same rules as the rules it supplements) where B' replaces B. Therefore we have to show how to prove $+\partial B'$ in T'. There are no rules for $\neg B'$ in T', and the only rule for B' is r' whose antecedent is empty, thus the sequence $-\Delta\neg B', +\partial B'$ satisfies the conditions to be a proof in T'.

Now for every proof in T where we have $+\partial B$, we replace such an entry with the sequence $-\Delta\neg B'$, $+\partial B'$. It is immediate to verify that the resulting sequence is a proof in T'.

In T' we have that there is a proof P of $+\partial B$. Thus we can take a proof in T' and we can replace every entry of $+\partial B'$ with P. Again, it is immediate to verify that the resulting sequence is a proof in T.

B On AGM Postulates and Literal Contraction

In this section we analyse whether the definition of contraction given in Definition 1 conforms with the AGM postulates for contraction. To permit this analysis we discuss how to reframe the postulates within our approach based on Defeasible Logic. In what follows B is meant to be a belief set.

$B_c^{\ominus}$ *is a belief set.* A theory in classical logic is just a set of formulas closed under logical consequence, and it essentially corresponds to the notion of extension (maximal consistent set of formulas). The meaning of this postulate is to guarantee closure under the belief revision operations and consequently that the result of a contraction makes sense in the given logic.

As we have seen, in Defeasible Logic we have to separate the notions of extension and theory. Each theory generate a (defeasible) extension, and then we operate the contraction on the extension and finally we take the theory generated by the contracted extension as the theory corresponding to the contracted theory. A 4-tuple of sets of literals has some properties to satisfy in order to be considered an extension. Thus, we have to rephrase the postulate as follows:

> The contraction of an extension is an extension, i.e., $E(T)_c^{\ominus}$ is an extension.

We can easily verify that this properties is satisfied. The contraction removes a literal from the sets of literals that are positively provable, i.e., from $\Delta^+(T)$ and $\partial^+(T)$. Thus, the only theories that can be affected by the operation are those of type A* (resp. *A), BB and CC. The removal of a positively provable literal from a theory of type A* creates an extension of type F*. The set of combinations allowed in theories of type A* is a subset of that of type F*. Hence, the result of contraction is guaranteed to satisfy one of the conditions in Table 5(b) (the same argument holds for extensions of type *A). The removal of a positively provable literal from an extension of type BB or type CC will create, respectively, an extension whose type is either EB or BE and EC or CE. Therefore the resulting 4-tuple satisfies one of the conditions of Table 5(b).

$B^{\ominus}_{c} \subseteq B$. In classical logic an extension contains only the formulas that logically follow from the theory. In Defeasible Logic we have four possible types of conclusions. Thus in this case what we have to do is to check that the pointwise subset relationships holds. This means that the postulate can be rewritten as

$$E(T)^{\ominus}_{c} \subseteq E(T)$$

which corresponds to

$$\#^{\pm}(T)^{\ominus}_{c} \subseteq \#^{\pm}(T).$$

Now for $\#^{+}(T)^{\ominus}_{c} = \#^{+}(T) - \{c\} \subseteq \#^{+}(T)$ and $\#^{-}(T)^{\ominus}_{c} = \#^{-}(T)$.

If $c \notin B$ *then* $B = B^{\ominus}_{c}$. As we discussed for the previous postulates, AGM focuses on the classical logic notion of consequence relation. Thus $c \notin B$ means that c is not a consequence of the theory. In Defeasible Logic this corresponds to not being able to prove something, thus it is encoded by a negative proof tag (i.e., $-\Delta$ and $-\partial$). Accordingly $c \notin B$ can be understood as $c \in \partial^{-}(T)$ (remember that $\partial^{-}(T) \subseteq \Delta^{-}(T)$). Thus we have the following version of the postulate

if $c \in \partial^{-}(T)$ then $E(T) = E(T)^{\ominus}_{c}$

Clearly the revised postulate is satisfied since contraction removes the literal c for the set of positive conclusions, but none of these sets can contain c. Thus contraction, in this case, leaves the extension unchanged.

If $\nvdash c$ *then* $c \notin B^{\ominus}_{c}$. The intuition behind this postulate is that if a formula is not a theorem of the logic (e.g., a tautology) then, the removal of it is successful. Given the restriction of the language of Defeasible Logic there are formulas playing the same roles as tautologies. We can have two possible interpretations of this postulate in Defeasible Logic. In Defeasible Logic the notion of not being provable is captured by the negative proof tags. For this interpretation we have

If $c \in \#^{-}(T)$, then $c \in \#^{-}(T)^{\ominus}_{c}$

which follows immediately since $\Delta^{-}(T)$ and $\partial^{-}(T)$ are left unchanged. For the second interpretation we have that

If $c \in \partial^{-}(T)$ or $c \notin \partial^{+}(T)$ then $c \notin \partial^{+}(T)^{\ominus}_{c}$

which again follows immediately from the previous postulate.

If $c \in B$ *then* $B \subseteq (B^{\ominus}_c)^{\oplus}_c$. First of all we have to define the expansion of an extension, and here we consider the expansion as the extension obtained by replacing $\Delta^+(T)$ and $\partial^+(T)$ with $\Delta^+(T) \cup \{c\}$ and $\partial^+(T) \cup \{c\}$ respectively. The postulate does not hold in general; if we contract vacuously then the expansion can generate a structure that is not an extension. However, it holds with the restriction that $c \in \Delta^+(T)$.

If $c \in \Delta^+(T)$, then $E(T) \subset (E(T)^{\ominus}_c)^{\oplus}_c$.

This follows immediately since $\Delta^+(T) \subseteq \partial^+(T)$, and then

$$(\#^+(T)^{\ominus}_c)^{\oplus}_c = (\#^+(T) - \{c\}) \cup \{c\} = \#^+(T).$$

If $c \equiv d$, *then* $B^{\ominus c} = B^{\ominus d}$ The rewriting of this postulate is straightforward: we replace B by $E(T)$. The language of Defeasible Logic is restricted to literals, thus two literals are equivalent if they are the same literal, and two sets of literals are equivalent if they have the same elements, i.e., they are the same set.

$B^{\ominus}_a \cap B^{\ominus}_b \subseteq B^{\ominus}_{a \wedge b}$. In the context of Defeasible Logic a conjunction of literals is equivalent to the set of the literals. Accordingly we rephrase the postulate as[16]

$$E(T)^{\ominus}_a \cap E(T)^{\ominus}_b \subseteq E(T)^{\ominus}_{\{a,b\}}$$

To verify that the postulate holds we have to consider, for the negative proof literals

$$\#^-(T)^{\ominus}_a = \#^-(T)^{\ominus}_b = \#^-(T)^{\ominus}_{\{a,b\}} = \#(T)$$

and for the positive literals

$$\#^+(T)^{\ominus}_a = \#^+(T) - \{a\} \qquad \#^+(T)^{\ominus}_b = \#^+(T) - \{b\} \qquad \#^+(T)^{\ominus\{a,b\}} = \#^+(T) - \{a,b\}$$

but

$$(\#^+(T) - \{a\}) \cap (\#^+(T) - \{b\}) = \#^+(T) - \{a,b\}$$

If $b \notin B^{\ominus}_{a \wedge b}$ *then* $B^{\ominus}_{a \wedge b} \subseteq B^{\ominus}_b$. Similarly to the previous postulate we have

If $b \notin \partial^+(T)^{\ominus}_{\{a,b\}}$, then $E(T)^{\ominus}_{\{a,b\}} \subseteq E(T)^{\ominus}_b$.

It is immediate to check that this postulate holds: $\#^+(T)^{\ominus}_{\{a,b\}} = \#^+(T) - \{a,b\} \subseteq \#^+(T) - \{b\} = \#^+(T)^{\ominus}_b$.

[16] Again for the set theoretic operation we take the pointwise version of it.