Superiority Based Revision of Defeasible Theories

Guido Governatori[•], Francesco Olivieri[†•],
Simone Scannapieco[†•] and Matteo Cristani[†]

[†]Department of Computer Science, University of Verona, Italy
[•]NICTA, Queensland Research Laboratory, Australia

Abstract. We propose a systematic investigation on how to modify a preference relation in a defeasible logic theory to change the conclusions of the theory itself. We argue that the approach we adopt is applicable to legal reasoning, where users, in general, cannot change facts and rules, but can propose their preferences about the relative strength of the rules.
We provide a comprehensive study of the possible combinatorial cases and we identify and analyse the cases where the revision process is successful.

1 Introduction

Typically skeptical non-monotonic formalisms are equipped with techniques to address conflicts, where a conflict is a combination of reasoning chains leading to a contradiction. The most common device to handle conflicts is a preference or superiority relation over the elements used by the formalism to reason. These elements can be formulae, axioms, rules or arguments, and the preference relation states that one of such elements is to be preferred to another one when both can be used.

In this research we concentrate on a specific rule-based non-monotonic formalism, Defeasible Logic, but the motivation behind the particular technical development applies in general to other rule-based formalisms. In a rule based formalism, typically knowledge is described in *facts* (describing immutable propositions/statements about a case), *rules* (describing relationships between a set of premises and a conclusion), and *preference relation* or *superiority relation* (describing the relative strength of rules). A revision operation transforms a theory by changing some of its elements, that is: facts, rules and superiority relation. Revision based on change of facts corresponds to an update operation [1], revision based on modification of rules has been investigated in [2], to the best of our knowledge, revision of non-monotonic theories based on modifications of the underlying superiority relation has been neglected so far. In this paper we concentrate on this issue, and we argue that, while little attention has been dedicated to this topic, it has natural correspondences to reasoning patterns in legal reasoning.

The paper is organised as follows: In Section 2 we motivate that reasoning over preferences on rules and on how to modify the preferences is a natural reasoning pattern in legal reasoning. Then in Section 3 we introduce Defeasible Logic, the formalism chosen for our investigation; in particular we introduce new auxiliary proof tags to describe derivations in Defeasible logic. The new proof tags do not modify the expressive power of the logic, but they identify patterns where instances of the superiority relation contribute to the derivation of a conclusion. Armed with this technical machinery,

M. Dean at al. (eds)
RuleML 2010, The 4th International Web Rule Symposium
LNCS 6403, pp. 104–118

we provide an exhaustive analysis of the cases and conditions under which revision operation modifying only the superiority relation are successful (Section 4). Section 5 concludes the paper with a short discussion of related and future wok.

2 Norms and Preferences in Legal Reasoning

It has been argued [3] that some aspects of legal reasoning can be captured by non-monotonic rule based formalisms. The main intuition is that norms can be represented by rules, facts to the evidence in cases, and the superiority relation is induced by legal principles determining how to solve conflicts between norms.

We take the stance that, typically in the legal reasoning domain, we do not have control over the rules (norms) and on how to modify them, but there is some control on how they can be used. A normal single citizen has no power to change the Law, and has no power on what norms are effective in the jurisdiction she is situated in. These powers instead are reserved to persons, entities and institutions specifically designated to do so, for example, the parliament, and, under some given constraints, also by judges (in Common Law juridical system, especially).

However, a citizen can argue that a norm instead of another norm applies in a specific case. This amounts to say that one norm is to be preferred to the other in the case.

Prima-facie conflicts appear in legal systems for a few main reasons, among which we can easily identify three major representatives: (1) norms from different sources, (2) norms emitted at different times, and (3) exceptions. These phenomena are well understood and principles to solve such issues existed for a long time in legal theory, and are still used, for instance, as an argument to drive constitutional judgement against a given norm or a given sentence. Here we list the three major legal principles, expressing preferences among rules to be applied [4].

Lex Superior When there is a conflict between two norms from different sources, the norms originating from the legislative source higher in the legislative source hierarchy takes precedence over the other norm.

Lex Posterior According to this principle a norm emitted after another norm takes precedence over the older norm.

Lex Specialis This principle states that when a norm is limited to a specific set of admissible circumstances, and under more general conditions another norm applies, the most specific norm prevails.

Besides the above principles a legislator can explicitly establish that one norm prevails over a conflicting norm.

The intuition behind the above principles (and eventually others) is that when there are two conflicting norms, and the two norms are applicable in a specific case, we can apply one of these principles to create an instance of a superiority relation that discriminates between the two conflicting norms. However, there is further complication. What about if several principles apply and these produce opposite preferences? This is when revision of preferences is relevant. The following example illustrates this situation.

Charlie is an immigrant living in Italy, who is interested in joining the Italian Army, based on Law 91 of 1992. However, his application is rejected, based upon a constitutional norm (Article 51 of the Italian Constitution). The two norms Law 91 and Article

51 are in conflict thus the Army's decision is based on the *lex superior* principle. Charlies appeals against the decision in court. The facts of the case are undisputed, and so are the norms to be applied and their interpretation. Thus the only chance for Bob, Charlie's lawyer, to overturn the decision is to argue that Law 91 overrides Article 51 of the constitution. Thus Bob, Charlie's advocate, counter-argues appealing to the *lex specialis* principle since Law 91 of 1992 explicitly covers the case of a foreigner who applies for joining the Army for the purpose of obtaining citizenship.

The two arguments do not discuss about facts and rules that hold in the case. They disagree about which rule prevails over the other, Article 51 of the Constitution or Law 91. In particular, Bob's argument can be see as an argument where the relative strength of the two rules is reversed compared to the argument of the Army's lawyer, and it is an argument to revise the previous decision.

The mechanism sketched above attains at the notion of strategic reasoning, where a discussant looks at the best argument to be used in a case to prove a given claim.

In the current literature about formalisms apt to model normative and legal reasoning, a simple and efficient non-monotonic formalism which has been discussed in the community is *defeasible logic*. This system is described in detail in the next section.

One of the strong aspects of defeasible logic is its characterisation in terms of argumentation semantics [5]. In other words, it is possible to relate it to general reasoning structure in non-monotonic reasoning, that is based on the notion of admissible reasoning chain. An admissible reasoning chain is an argument in favour of a thesis. For these reasons, much research effort has been spent upon defeasible logic, and once formulated in a complete way it encompasses other (skeptical) formalisms proposed for legal reasoning [5,6].

Most interestingly, in defeasible logic we can reach positive conclusions as well as negative conclusions, thus it gives understanding to both accept a conclusion as well as reject a conclusion. This is particularly advantageous when trying to address the issues determined by reasoning conflicts.

This paper provides a comprehensive study of the conditions under which it is possible to revise a defeasible theory by changing the superiority relation of the theory, that is changing the relative strength of conflicting rules.

3 Defeasible logics

A defeasible theory consists of five different kinds of knowledge: facts, strict rules, defeasible rules, defeaters, and a superiority relation [7]. Examples of facts and rules below are standard in the literature of the field.

Facts denote simple pieces of information that are considered always to be true. For example, a fact is that Sylvester is a cat: $cat(Sylvester)$. A *rule* r consists of its *antecedent* $A(r)$ which is a finite set of literals, an *arrow*, and its *consequent* (or *head*) $C(r)$, which is a single literal. A *strict rule* is a rule in which whenever the premises are indisputable (e.g. facts) then so is the conclusion, e.g.

$$cat(X) \rightarrow mammal(X),$$

which means “Every cat is a mammal”. A *defeasible rule* is a rule that can be defeated by contrary evidence: “Cats typically eat birds”, written formally:

$$cat(X) \Rightarrow eatBirds(X).$$

The underlying idea is that if we know that something is a cat, then we may conclude that it eats birds, unless there is other evidence that it may not. Defeasible rules with an empty antecedent are “almost” facts. *Defeaters* are rules that can not be used to draw any conclusions. Their only use is to prevent some conclusions, i.e. to defeat defeasible rules by producing evidence to the contrary. An example is “If a cat has just fed itself, then it might not eat birds”, formally

$$justFed(X) \rightsquigarrow \neg eatBirds(X).$$

The *superiority relation* among rules is used to define where one rule may override the conclusion of another one, e.g. given the defeasible rules

$$r : cat(X) \Rightarrow eatBirds(X)$$
$$r' : domesticCat(X) \Rightarrow \neg eatBirds(X)$$

which would contradict one another if Sylvester is both a cat and a domestic cat, they do not if we state that $r' > r$, leading Sylvester not to eat birds. Notice that in defeasible logic the superiority relation determines the relative strength of two conflicting rules.

Like in [7], we consider only a propositional version of this logic, and we do not take in account function symbols. Every expression with variables represents the finite set of its variable-free instances.

A *defeasible theory* D is a triple $(F,R,>)$, where F is a finite consistent set of literals called *facts*, R is a finite set of rules, and $>$ is an acyclic superiority relation on R. The set of all strict rules in R is denoted by R_s, and the set of strict and defeasible rules by R_{sd}. We name $R[q]$ the rule set in R with head q. A *conclusion* of D is a tagged literal and can have one of the following forms:

1. $+\Delta q$, which means that q is definitely provable in D, i.e. there is a definite proof for q, that is a proof using facts, and strict rules only;
2. $-\Delta q$, which means that q definitely not provable in D (i.e., a definite proof for q does not exist);
3. $+\partial q$, which means that q is defeasibly provable in D;
4. $-\partial q$, which means that q is defeasibly not provable in D.

A *proof* (or *derivation*) is a finite sequence $P = (P(1),\ldots,P(i))$ of tagged literals where for each n, $0 \leq n \leq i$ the following conditions (proof conditions) are satisfied.[1]

$+\Delta$: If $P(n+1) = +\Delta q$ then
  (1) $q \in F$ or
  (2) $\exists r \in R_s[q] \forall a \in A(r) : +\Delta a \in P(1..n)$

[1] $P(1..i)$ denotes the initial part of the sequence of length i, and $\sim p$ the complement of a literal p.

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

The proof conditions just given are meant to represent forward chaining of facts and strict rules ($+\Delta$), and that it is not possible to obtain a conclusion just by using forward chaining of facts and strict rules ($-\Delta$).

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

$-\partial$: If $P(n+1) = -\partial q$ then
(1) $-\Delta q \in P(1..n)$ and
(2) (2.1) $\forall r \in R_{sd}[q]\ \exists a \in A(r) : -\partial a \in P(1..n)$ or
(2.2) $+\Delta{\sim}q \in P(1..n)$ or
(2.3) $\exists s \in R[{\sim}q]$ such that
(2.3.1) $\forall a \in A(s) : +\partial a \in P(1..n)$ and
(2.3.2) $\forall t \in R_{sd}[q]$ either
$\exists a \in A(t) : -\partial a \in P(1..n)$ or $t \not> s$.

The main idea of the conditions for a defeasible proof ($+\partial$) is that there is an applicable rule, i.e., a rule whose all antecedents are already defeasibly provable and for every rule for the opposite conclusion either the rule is discarded, i.e., one of the antecedents is not defeasibly provable, or the rule is defeated by a stronger applicable rule for the conclusion we want to prove. The conditions for $-\partial$ show that any systematic attempt to defeasibly prove the conclusion fails.

In this paper, we do not make use of strict rules, nor defeaters[2], since every revision changes only priority among defeasible rules (the only rules that act in our framework), but we need to introduce eight new types of tagged literals. As it will be clear in the rest of the paper, they would be of significant utility in simplifying the categorisation process, and consequently, the revision calculus.

5. $+\Sigma q$, which means there is a reasoning chain supporting q;
6. $-\Sigma q$, which means there is not a reasoning chain supporting q;
7. $+\sigma q$, which means there exists a reasoning chain supporting q that is not defeated by any applicable reasoning chain attacking it;

[2] The restriction does not result in any loss of generality: (1) the superiority relation does not play any role in proving definite conclusions, and (2) for defeasible conclusions [7] proves that it is always possible to remove (a) strict rules from the superiority relation and (b) defeaters from the theory to obtain an equivalent theory without defeaters and where the strict rules are not involved in the superiority relation.

8. $-\sigma q$, which means that every reasoning chain supporting q is attacked by an applicable reasoning chain;
9. $+\varphi q$, which means there exists a reasoning chain that defeasibly proves q made of elements such that there does not exist any rule for the opposite conclusion;
10. $-\varphi q$, which means that for every reasoning chain supporting q there exists an element such that a rule for the opposite conclusion could fire;
11. $+\omega q$, which means there exists a reasoning chain supporting q that defeasibly proves every its antecedent;
12. $-\omega q$, which means that in every reasoning chain supporting q, at least one of its antecedents is not defeasibly provable.

The tagged literals can be formally defined by the following proof conditions as:

$+\Sigma$: If $P(n+1) = +\Sigma q$ then
(1) $q \in F$ or
(2) $\exists r \in R_{sd}[q] \forall a \in A(r) : +\Sigma a \in P(1..n)$

$-\Sigma$: If $P(n+1) = -\Sigma q$ then
(1) $q \notin F$ and
(2) $\forall r \in R_{sd}[q] \exists a \in A(r) : -\Sigma a \in P(1..n)$

$+\sigma$: If $P(n+1) = +\sigma q$ then
(1) $q \in F$ or
(2) (2.1) $\exists r \in R_{sd}[q] \forall a \in A(r) : +\sigma a \in P(1..n)$ and
(2.2) $\forall s \in R[\sim q] \exists a \in A(s)$ such that
$-\partial a \in P(1..n)$ or $s \not> r$.

$-\sigma$: If $P(n+1) = -\sigma q$ then
(1) $q \notin F$ and
(2) (2.1) $\forall r \in R_{sd}[q] \exists a \in A(r) : -\sigma a \in P(1..n)$ or
(2.2) $\exists s \in R[\sim q]$ such that
$\forall a \in A(s) : +\partial a \in P(1..n)$ and $s > r$.

Notice that the definitions given above for $\pm\sigma$ are weak forms of the notion of support proposed in [8,9] for the definition of an ambiguity propagating variant of defeasible logic, in the sense that these definitions are less selective than the ones of [8].

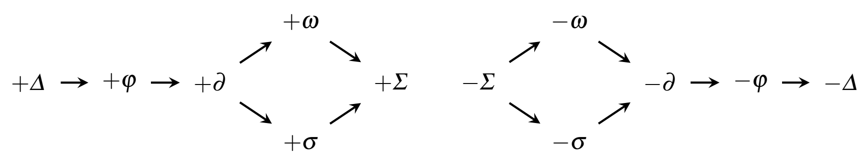
$+\varphi$: If $P(n+1) = +\varphi q$ then
(1) $q \in F$ or
(2) (2.1) $\exists r \in R_{sd}[q] \forall a \in A(r) : +\varphi a \in P(1..n)$ and
(2.2) $\forall s \in R[\sim q] \exists a \in A(s) : -\Sigma a \in P(1..n)$.

$-\varphi$: If $P(n+1) = -\varphi q$ then
(1) $q \notin F$ and
(2) (2.1) $\forall r \in R_{sd}[q] \exists a \in A(r) : -\varphi a \in P(1..n)$ or
(2.2) $\exists s \in R[\sim q] \forall a \in A(s) : +\Sigma a \in P(1..n)$.

$+\omega$: If $P(n+1) = +\omega q$ then
(1) $q \in F$ or
(2) $\exists r \in R_{sd}[q] \forall a \in A(r) : +\partial a \in P(1..n)$.

$-\omega$: If $P(n+1) = -\omega q$ then
(1) $q \notin F$ and
(2) $\forall r \in R_{sd}[q] \exists a \in A(r) : -\partial a \in P(1..n)$.

By the above definitions, it is straightforward to derive the implication chains reported below in Figure 1(a) -(b) .



(a) Positive implication chain **(b)** negative implication chain

Fig. 1: Implication chains.

One could think that $+\sigma$ implies $+\omega$ (and symmetrically, $-\omega$ implies $-\sigma$). It is not so. To better explain this fact, and the meaning of the proof conditions, we present an illustrative example.

Example 1

$$\begin{array}{llllll}
\Rightarrow_{r_1} & a & \Rightarrow_{r_2} & c & \Rightarrow_{r_3} & d \\
\vee & & & & \wedge & \\
\Rightarrow_{r_4} & \neg a & & & \Rightarrow_{r_5} & \neg d \Rightarrow_{r_6} p \\
\Rightarrow_{r_7} & b & \Rightarrow_{r_8} & \neg c & & \\
\Rightarrow_{r_9} & \neg b & & & & \\
\Rightarrow_{r_{10}} & e & \Rightarrow_{r_{11}} & f & &
\end{array}$$

with $r_1 > r_4$, and $r_5 > r_3$. In this theory, we can obtain the following conclusions:

	a	b	c	d	e	f	p
$+$	$+\partial$	$+\sigma$	$+\partial$	$+\omega$	$+\phi$	$+\phi$	$+\partial$
$-$	$-\phi$	$-\partial$	$-\partial$	$-\sigma$			$-\phi$

	$\neg a$	$\neg b$	$\neg c$	$\neg d$	$\neg e$	$\neg f$	$\neg p$
$+$	$+\omega$	$+\sigma$	$+\sigma$	$+\partial$			
$-$	$-\partial$	$-\partial$	$-\omega$	$-\phi$	$-\Sigma$	$-\Sigma$	$-\Sigma$

From the definitions above and the example, we can take some theoretical results about the proof tags that will be used during the revision process described in Section 4.

Proposition 1 *Given a consistent defeasible theory D, if we have $+\varphi p$ for a literal p, then $-\Sigma{\sim}p$.*

Proof. Let us suppose D is a consistent defeasible theory, and $+\varphi p$ holds for a literal p. Now, if we assume that $+\Sigma{\sim}p$, we say that there exists a reasoning chain supporting ${\sim}p$ which fails somewhere, leading also to $-\varphi p$ to hold, against the hypothesis.[3] A contradiction.

The opposite does not hold (literal p in Example 1). The next proposition states formally the following idea: if we can defeasibly prove a literal p, and we know also that there exists a chain leading to ${\sim}p$ with all the antecedents defeasibly proved, then such a chain has to be defeated by a priority rule at the last proof step (by the rule proving p).

Proposition 2 *Given a consistent defeasible theory D, if $+\partial p \wedge +\omega{\sim}p$ holds for a literal p, then $-\sigma{\sim}p$.*

Proof. By definition of $+\partial$, we have that condition below

(2.3) $\forall s \in R[{\sim}q]$ either
  (2.3.1) $\exists a \in A(s) : -\partial a \in P(1..n)$ or
  (2.3.2) $\exists t \in R_{sd}[q]$ such that
    $\forall a \in A(t) : +\partial a \in P(1..n)$ and $t > s$.

holds for p. In fact condition (2.3.2) has to be true since we know condition (2.3.1) is not, because

$$\left.\begin{array}{r}+\partial p \implies \exists r \in R[p].\forall a \in A(r) : +\partial a \\ +\omega{\sim}p \implies \exists s \in R[{\sim}p].\forall a \in A(s) : +\partial a\end{array}\right\} \implies$$

$$\exists t \in R[p].\forall a \in A(t) : +\partial a \text{ and } t > s.$$

This is the definition of $-\sigma{\sim}p$. Since all the premises of ${\sim}p$ are defeasibly proved by hypothesis, and we have proved that the chain is defeated, then it has to loose on the last proof step.

4 Preference defeasible revision

Here we analyse the processes of revision in a defeasible theory, when no changes to the rules and facts are allowed. Henceforth, when no confusion arises, every time we speak about a (revision) transformation we refer to a (revision) transformation acting only on the superiority relation.

In the legal domain, when two lawyers dispute a case, there are four situations in which each of them can be if she revises the superiority relation employed by the other one.

[3] All proof conditions given in this paper obey the principle of strong negation, thus for any literal p and any proof tag # it is not possible to have both $+\#p$ and $-\#p$. [9]

(a) The revision process supports the argument of *reasonable doubt*. Someone proves that the rules imply a given conclusion. If the preference is revised then we can derive that this is not the case, showing thus that the conclusion was not beyond reasonable doubt.
(b) The revision process beats the argument of *beyond reasonable doubt*. Analogously to situation (a), someone proves that the rules do not imply a given conclusion. If the preference is revised then we can derive that this is indeed the case.
(c) The revision process supports the argument of *proof of innocence/guilt*. Someone proves that the rules imply a given conclusion. If the preference is revised then we can derive that the opposite holds.
(d) The revision process cannot support a given thesis.

Revising a defeasible theory by changing only the priority among its rules means studying how an hypothetic revision operator works in the three cases reported below:

(1) how to obtain $-\partial p$, starting from $+\partial p$;
(2) how to obtain $+\partial{\sim}p$, starting from $+\partial p$;
(3) how to obtain $+\partial p$, starting from $-\partial p$.

We name these three revisions *canonical*. We provide an exhaustive analysis, based on the definitions above, in the next subsections.
The situation (a) is represented by the canonical case (1). The situation (b) is represented by the canonical case (3). Situation (c) is represented by the canonical case (2). The situation (d) arises when the condition $+\varphi p$ holds.

In this case, if one of the parties argues in favour of a thesis in a defeasible way, then the counter-part cannot exhibit a proof of the opposite, independently of the changes in the superiority relation.

In the cases (1) and (2) analysed below, we know that $-\partial\neg p$ holds, since D is a consistent theory in which $+\partial p$ holds. Furthermore, Proposition 3 allows us not to consider a tree with branches tagged by $\pm\varphi$.

Notice that some revisions do not indeed modify the knowledge in the system. For instance, revising a theory from $+\partial p$ to $-\partial{\sim}p$ is useless.

The above reasoning proves that we have canonical revisions, revisions that are equivalent to canonical ones and useless revisions, thus we have the following theorem.

Theorem 1 *The revision of the preference relation in a defeasible theory is either canonical or useless.*

Proposition 3 states that if there is no way to defeat a chain supporting a literal p, there is no revision transformation which leads to defeasibly derive ${\sim}p$.

Proposition 3 *Given a consistent defeasible theory D, if for a literal p holds $+\varphi p$, then there does not exist a transformation to obtain $+\partial{\sim}p$.*

Proof. Given any theory, to obtain a defeasible proof of a literal q, there must exist at least a reasoning chain for q, i.e. $+\Sigma q$. This is in contradiction with Proposition 1 which states that if $+\varphi{\sim}q$ holds, also $-\Sigma{\sim}q$ does.

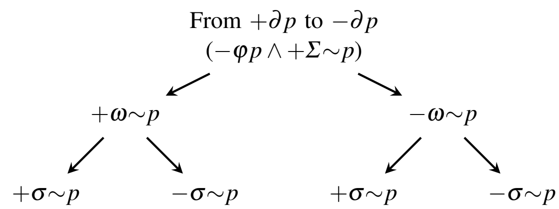


Fig. 2: From $+\partial p$ to $-\partial p$: revision cases.

For every consistent theory, $+\partial p \implies -\partial{\sim}p$, [7], Proposition 3 states also that with the same premises it is impossible to revise the theory in order to obtain $-\partial p$.

We are now ready to go onto the systematic analysis of the combinations arising from the above defined model. We list the cases by tagging each macroscopic case by the name *Canonical case* and the combinations depending upon the analytical schema introduced above by the name *Instance*.

4.1 Canonical case: from $+\partial p$ to $-\partial p$

Instance $-\Sigma{\sim}p \wedge +\partial p$**:** this first case is not reported in Figure 2 since the premises are not true ($-\Sigma{\sim}p$ holds). This means there is no supporting chains for ${\sim}p$, so we can not operate on them. Holding $-\varphi p$, this means there exists at least one of its premises that could be defeated by a rule leading to the opposite conclusion. Thus, in order to obtain $-\partial p$, we have to revise the theory putting at least one of such rules be able to fire (to defeat, or at least to have the same power of a rule which actually proves one of the antecedents in the chain supporting p).

Instance $+\omega{\sim}p \wedge +\sigma{\sim}p$**:** as stated in Proposition 2 this branch represents an impossible case for any consistent defeasible theory.

Instance $+\omega{\sim}p \wedge -\sigma{\sim}p$**:** by the straightforward implication of Proposition 2, the chain supporting ${\sim}p$ fails on the last proof step defeated by priorities for rules which defeasibly prove p. Thus, we have only to erase these priorities.

Instance $-\omega{\sim}p \wedge +\sigma{\sim}p$**:** since there exists a chain P_{np} (whilst P_p denotes the proof for p) supporting ${\sim}p$ which is never defeated ($-\omega{\sim}p$ condition tells us only that such a chain fails before the last proof step), a revision process does not have to operate on a chain supporting p. We have to strengthen P_{np} changing so many priorities to let a rule in P_{np}, which leads to an opposite conclusion of a rule in P_p, have at least the same strength of such a rule in P_p. In this process, we do not remove any priority rule among elements in P_p, but only add priority rules to let a rule in P_{np} win.

Instance $-\omega{\sim}p \wedge -\sigma{\sim}p$**:** the reasoning chain P_{np} supporting ${\sim}p$ is defeated, but not necessarily by a chain proving p (P_p). The case is analogous of the above, but: probably we have to act not only on P_{np}, but also on P_p; we do not have only to introduce priority rules, but also to erase (invert) them. This case represents the most general situation, where less information is given: a revision is possible, but we do not know *a priori* where to change the theory.

4.2 Canonical case: from $+\partial p$ to $+\partial{\sim}p$

We follow the cases depicted in the search tree in Figure 2, in order to explain how a revision operator should work. We change the root label when revising from $+\partial p$ to $+\partial{\sim}p$, taking in account the same premises $(-\varphi p \wedge +\Sigma{\sim}p)$. Once more, our revision tree does not take in account tags $\pm\varphi$ for the same reasons explained in Section 4.

Instance $+\omega{\sim}p \wedge +\sigma{\sim}p$: as stated in Proposition 2 this branch represents an impossible case for any consistent defeasible theory.

Instance $+\omega{\sim}p \wedge -\sigma{\sim}p$: Proposition 2 states that the chain supporting ${\sim}p$ fails on the last proof step. This, combined with $-\sigma{\sim}p$, implies this last step is defeated by a priority for the rule which defeasibly proves p. In fact, there would exist more than one chain that fails on the last step, and also more than one chain which proves p. We propose two different approaches. We name P the set of chains proving defeasibly p, $P_{ls} \subseteq P$ the chains that prove defeasibly p for which there is a priority rule that applies at the last proof step (against a chain that proves ${\sim}p$), and N the set of chains for which the premises hold:

1. We choose a chain in N. We invert the priority rule for every chain in P_{ls} that wins at the last proof step. We introduce a new priority for making it win against any remaining chain in P.
2. In this approach we have two neatly distinguished cases:
   (a) $||P_{ls}|| > ||N||$: for every chain in N we invert the priority rules on the last proof step. For every remaining chain in P, we add a priority rule between the defeasible rule used in the last proof step of a chain in N and the rule used in the last proof step of a chain in P (possibly different for each chain in N) such that the chain in P looses.
   (b) $||N|| > ||P_{ls}||$: firstly we choose a number $||P_{ls}||$ of chains in N and invert the priority rule on the step that makes them loose. If at the end of this step there are still chains in P that defeasibly prove p, we go on with the method used for the case (2)(a), only looking at the subset of chains in N on which we operated at the first step.

The two approaches rely on different underlying ideas. In the first case we want a unique winning chain. This makes the revision procedure faster than the second method, we do not have to choose every time a different chain where to act. Moreover, it guarantees to make at most many changes as the second one (in general, it revises the theory with the minimum number of changes).

The strength of the second method relies on the concept of *team defeaters*: we give power not only to a single element, but to a team of rules. Thus, in the first method if the only winning chain would be defeated, the entire revision process must be repeated, whilst in the second method if one of the winning rule would be beaten, we have to repair only for it, but not for all the other chains that continue to win.

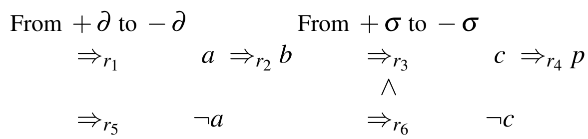
Let us consider the following simple example:

$$\begin{array}{ll} \Rightarrow_{r_1} p & \Rightarrow_{r_2} p \\ \vee & \vee \\ \Rightarrow_{r_3} \neg p & \Rightarrow_{r_4} \neg p \end{array}$$

The first approach would give in output: $\{r_1 > r_3, r_4 > r_1, r_4 > r_2\}$ (if the second chain for $\neg p$ would be chosen to win), erasing one priority rule and introducing two, whilst the second approach would lead to have the following priority rule set: $\{r_3 > r_1, r_4 > r_2\}$, erasing two priority rules, and introducing two. It is easy to see that if r_4 would be defeated by a rule r_s, in the first case we have to entirely revise the theory, for example, let r_3 win among r_1 and r_2, while in the second case we have only to introduce $r_3 > r_2$.

Instance $-\omega{\sim}p \wedge +\sigma{\sim}p$: there exists at least a chain supporting ${\sim}p$, which is not defeated. To revise the theory, we have to choose one of them and, starting from ${\sim}p$ go back in the chain to the ambiguity point (where holds $P(i) = +\partial p_i \wedge P(i+1) = -\partial p_{i+1}$), strengthen the chain adding a priority rule where a rule leading to an antecedent in the chain for ${\sim}p$ and a rule for the opposite have the same strength.

Instance $-\omega{\sim}p \wedge -\sigma{\sim}p$: every chain supporting ${\sim}p$ is defeated at least one time. A first approach one could be tempted to use is to go back in the chain searching for the point where $P(i) = +\sigma p_i \wedge P(i+1) = -\sigma p_{i+1}$. Note that this is not enough to guarantee the chain to win. Let us consider the following example.



As it can be easily seen, letting r_3 win over r_6 is not sufficient. We have also to introduce a priority rule between r_1 and r_5. Thus, we have to act exactly as in the previous case, with the solely difference that every time a rule in the chain supporting $\neg p$ is defeated, the priority rule has to be inverted.

4.3 Canonical case: from $-\partial p$ to $+\partial p$

We start this case, saying that $-\partial{\sim}p$ has to hold since, if it is not so, the case is analogous of the previous revision from $+\partial q$ to $+\partial{\sim}q$. Moreover, we do not take in consideration the case when $-\Sigma p$ holds, as if there are no chains leading to p, there will be no revision to obtain $+\partial p$. The cases are the ones reported in Figure 3.

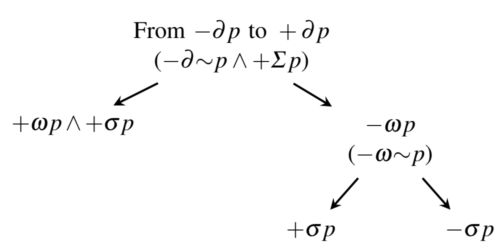


Fig. 3: From $-\partial p$ to $+\partial p$: revision cases.

Note that $+\omega p$ and $-\sigma p$ can not hold at the same time: as all the premises for p are proven, the chain has to fail on the last step, i.e. it has to be defeated by a firing rule for $\sim p$. This would defeasibly prove $\sim p$, but this can not happen since we have stated that $-\partial\sim p$ holds. Furthermore, $-\omega p$ implies that also $-\omega\sim p$ holds, since if it is not so, we have either $+\omega p$, or $+\partial\sim p$, both of them against the hypothesis.

Instance $+\omega p \wedge +\sigma p$: since there would exist more than one chain such that $+\omega p \wedge +\sigma p$ holds, we have to choose one of them, and introduce as many priority rules as the number of chains where $+\omega\sim p$ holds.

Instance $-\omega p \wedge +\sigma p$: this case is analogous to the revision case: From $+\partial p$ to $+\partial\sim p$: $-\omega\sim p \wedge +\sigma\sim p$.

Instance $-\omega p \wedge -\sigma p$: this case is analogous to the case: From $+\partial p$ to $+\partial\sim p$: $-\omega\sim p \wedge -\sigma\sim p$.

We have to remark that conditions $\pm\sigma\sim p$ do not give information on the revision process, since they do not tell if the changes will apply on chains for $\sim p$, or not. Referring to the example proposed below, we can see that, holding $+\sigma\sim p$, there exists a revision which involves the chain for $\sim p$ (introducing $r_1 > r_3$, and $r_2 > r_4$), and the other one that does not (introducing $r_5 > r_6$).

$$\begin{array}{l}\Rightarrow_{r_1} a \Rightarrow_{r_2} p\\ \Rightarrow_{r_3} \neg a \Rightarrow_{r_4} \neg p\\ \\ \Rightarrow_{r_5} b \Rightarrow_{r_6} p\\ \Rightarrow_{r_6} \neg b\end{array}$$

An analogous situation can be proposed for $-\sigma\sim p$.

$$\begin{array}{l}\Rightarrow_{r_1} a \Rightarrow_{r_2} p\\ \Rightarrow_{r_3} \neg a \Rightarrow_{r_4} b \Rightarrow_{r_5} \neg p\\ \qquad\quad \wedge\\ \qquad\quad \Rightarrow_{r_6} \neg b\\ \Rightarrow_{r_7} c \Rightarrow_{r_8} p\\ \Rightarrow_{r_9} \neg c\end{array}$$

In here, there exist two revisions: one introducing $r_1 > r_3$ and $r_2 > r_4$, and the other one which introduces $r_7 > r_9$.

Note that in all the canonical cases, the revision mechanism guarantees that no new cycle can be introduced. We can formulate the above result, that is a straightforward consequence of the case analysis presented here.

Theorem 2 *Revising superiority relation generates a superiority relation.*

5 Conclusions and further work

A large number of real-life cases in legal reasoning, information security, digital forensic, and even engineering or medical diagnosis, exhibit the two circumstances: (a) different persons have different preferences, and (b) decision making depends upon the order the rules are applied. When defeasible rules are in conflict, and then potentially generate inconsistencies, decision making may require preferences. In the same way, belief revision in presence of inconsistent information requires preference revision.

Notice that in non-monotonic reasoning, revision is not necessarily triggered by inconsistencies. [2] investigates revision for defeasible logic and relationships with AGM postulates. While the ultimate aim is similar to that of the present paper – i.e., transforming a theory to make a previously provable (resp. non provable), non provable (resp. provable) – the approach is different, and more akin to standard belief revision. More precisely, revision is achieved by introducing new exceptional rules. Furthermore they discuss how to adapt the AGM postulates for non-monotonic reasoning.

In this work we are not interested in examining conformance with the AGM postulates. [10] show that, typically, belief revision methodologies are not suitable to changes in theories intended for legal reasoning, and similarly they show that it is possible to revise theories fully satisfying the AGM postulates, but then the outcome is totally meaningless from a legal point of view. Anyway, to investigate the relationships between AGM and the approach presented here one has to adjust the AGM postulates to be meaningful (e.g., what is the meaning of expansion or contraction, when the operation is defined on instances of the preference relation).

Preference revision is just one of the aspects of legal interpretation. [11,12] propose a defeasible logic framework to model extensive and restrictive legal interpretation. This is achieved by using revision mechanisms on constitutive rules, where the mechanism is defined to change the strength of existing constitutive rules. It is an interesting question whether extensive and restrictive interpretation can be modelled as preference revision operators. An important aspect of legal interpretation is finding the legal rules to be applied in a case, in this work we assumed that the relevant rules have already been discovered, and in case of conflicts, preference revision can be used to solve them.

Closely related to our work are [13,14]. They propose extensions of an argumentation framework and defeasible logic, where the superiority relation is dynamically derived from arguments and rules in given theories. The main difference with these works is that we investigate general conditions under which it is possible to modify the superiority relation to change the conclusions of a theory, while they provide specific mechanisms but no guarantees that a change will happen. [13] is motivated, as us, by legal reasoning, and they use rules to encode the legal principles we mentioned in the introduction. We leave the investigation to the relationships with these works as future research.

Apart from the applications sketched above we shall investigate two limits to the revision operator:

- Revision of preference should not involve minimal defeasible rules. This constraint captures the idea that a rule that wins against all other rules is a basic juridical principle;
- Under given circumstances the revision process should not, for at least a subset of "protected" pairs violate the original preferential order. For instance we should not revise those preferences that are unquestioned because derived by commonly accepted principles or explicitly expressed by the legislator, as discussed in the introduction.

We unashamedly avoided, in this phase, any computational analysis of the introduced operator, but clearly a deeper investigation will include also the definition of that aspects.

Acknowledgements

A previous version of the paper has been presented at the 13th International Workshop on Non-monotonic Reasoning (NMR 2010) [15]. We thank the anonymous referees for NMR 2010 and RuleML 2010 for their valuable comments and feedback.

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

References

1. Katsuno, H., Mendelzon, A.O.: Propositional knowledge base revision and minimal change. Artificial Intelligence **52** (1991) 263–294
2. Billington, D., Antoniou, G., Governatori, G., Maher, M.J.: Revising nonmonotonic belief sets: The case of defeasible logic. In: KI-99: Advances in Artificial Intelligence, Springer-Verlag (1999) 101–112
3. Gordon, T.F., Governatori, G., Rotolo, A.: Rules and norms: Requirements for rule interchange languages in the legal domain. In Governatori, G., Hall, J., Paschke, A., eds.: Rule Representation, Interchange and Reasoning on the Web (RuleML 2009). Number 5858 in LNCS, Springer (2009) 282–296
4. Sartor, G.: Legal Reasoning. Springer, Dordrecht (2005)
5. Governatori, G., Maher, M.J., Billington, D., Antoniou, G.: Argumentation semantics for defeasible logics. Journal of Logic and Computation **14** (2004) 675–702
6. Antoniou, G., Maher, M.J., Billington, D.: Defeasible logic versus logic programming without negation as failure. J. Log. Program. **42** (2000) 47–57
7. Antoniou, G., Billington, D., Governatori, G., Maher, M.J.: Representation results for defeasible logic. ACM Transactions on Computational Logic **2** (2001) 255–287
8. Antoniou, G., Billington, D., Governatori, G., Maher, M.J.: A flexible framework for defeasible logics. In: AAAI. (2000) 401–405
9. Antoniou, G., Billington, D., Governatori, G., Maher, M.J., Rock, A.: A family of defeasible reasoning logics and its implementation. In: ECAI. (2000) 459–463
10. Governatori, G., Rotolo, A.: Changing legal systems: Legal abrogations and annulments in defeasible logic. Logic Journal of the IGPL **18** (2010) 157–194
11. Boella, G., Governatori, G., Rotolo, A., van der Torre, L.: *Lex minus dixit quam voluit*, *lex magis dixit quam voluit*: A formal study on legal compliance and interpretation. In: AI approaches to the complexity of legal systems, Berlin, Springer (2010)
12. Boella, G., Governatori, G., Rotolo, A., van der Torre, L.: A logical understanding of legal interpretation. In: Proceedings of KR 2010. (2010)
13. Prakken, H., Sartor, G.: Argument-based extended logic programming with defeasible priorities. Journal of Applied Non-Classical Logics **7** (1997)
14. Antoniou, G.: Defeasible logic with dynamic priorities. Int. J. Intell. Syst. **19** (2004) 463–472
15. Governatori, G., Olivieri, F., Scannapieco, S., Cristani, M.: Superiority based revision of defeasible theories. In Meyer, T., Ternovska, E., eds.: 13 International Workshop on Non-Monotonic Reasoning (NMR 2010). CEUR Workshop Proceedings (2010)