Online appendix for the paper

The Rationale behind the Concept of Goal

published in Theory and Practice of Logic Programming

Guido Governatori[1], Francesco Olivieri[2], Simone Scannapieco[2], Antonino Rotolo[3] and Matteo Cristani[2]

[1] *NICTA, Australia*
[2] *Department of Computer Science, Verona, Italy*
[3] *CIRSFID, University of Bologna, Bologna, Italy*

submitted 31 December 2014; revised 21 May 2015; accepted 14 December 2015

Appendix A Inferential mechanism example

This appendix is meant to offer more details for illustrating the inference mechanism proposed in this paper, and to consider Definition 13 more carefully. Therefore, first we report in Table A 1 the most interesting scenarios, where a rule r proves $+\partial_{\mathsf{SI}} q$ when attacked by an applicable rule s, which in turn is successfully counterattacked by an applicable rule t. Lastly, we end this appendix by reporting an example. The situation described there starts from a natural language description and then shows how it can be formalised with the logic we proposed.

For the sake of clarity, notation $\mathsf{B}, \Box$ (with $\Box \in \{\mathsf{O}, \mathsf{SI}\}$) represents belief rules which are Conv-applicable for mode $\Box$.

For instance, the sixth row of the table denotes situations like the following:

$$
\begin{aligned}
F &= \{a, b, \mathsf{O}c\}\\
R &= \{r : a \Rightarrow_{\mathsf{U}} q,\\
&\qquad s : b \Rightarrow_{\mathsf{O}} \neg q,\\
&\qquad t : c \Rightarrow q\}\\
> &= \{(t,s)\}.
\end{aligned}
$$

The outcome rule r for q is applicable for SI according to Definition 9. Since in our framework we have $\text{Conflict}(\mathsf{O}, \mathsf{SI})$, the rule s for $\neg q$ (which is applicable for O) does not satisfy condition (2.3.1) of Definition 13. As a result, s represents a valid attack to r. However, since we have $\text{Convert}(\mathsf{B}, \mathsf{O})$, rule t is Conv-applicable for O by Definition 7, with $t > s$ by construction. Thus, t satisfies condition (2.3.2.1) of Definition 13 and successfully counterattacks s. Consequently, r is able to conclude $+\partial_{\mathsf{SI}} q$.

Example 1
PeoplEyes is an eyeglasses manufacturer. Naturally, its final goal is to produce cool and

Mode of r	**Mode of** s	**Mode of** t	$+\partial_{\mathsf{SI}}q$ **because...**
U applicable for SI	U applicable for SI	U applicable for SI	$t > s$
U applicable for SI	U applicable for SI	O	Conflict(O, SI)
U applicable for SI	U applicable for SI	B, SI	$t > s$
U applicable for SI	U applicable for SI	B, O	Conflict(O, SI)
U applicable for SI	O	O	$t > s$
U applicable for SI	O	B, O	$t > s$
U applicable for SI	B, SI	U applicable for SI	$t > s$
U applicable for SI	B, SI	O	Conflict(O, SI)
U applicable for SI	B, SI	B, SI	$t > s$
U applicable for SI	B, SI	B, O	Conflict(O, SI)
U applicable for SI	B, O	B, O	$t > s$
B, SI	U applicable for SI	U applicable for SI	$t > s$
B, SI	U applicable for SI	O	Conflict(O, SI)
B, SI	U applicable for SI	B, SI	$t > s$
B, SI	U applicable for SI	B, O	Conflict(O, SI)
B, SI	O	O	$t > s$
B, SI	O	B, O	$t > s$
B, SI	B, SI	U applicable for SI	$t > s$
B, SI	B, SI	O	Conflict(O, SI)
B, SI	B, SI	B, SI	$t > s$
B, SI	B, SI	B, O	Conflict(O, SI)
B, SI	B, O	B, O	$t > s$

Table A 1. *Definition 13: Attacks and counterattacks for social intention*

perfectly assembled eyeglasses. The final steps of the production process are to shape the lenses to glasses, and mount them on the frames. To shape the lenses, *PeoplEyes* uses a very innovative and expensive laser machine, while for the final mounting phase two different machines can be used. Although both machines work well, the first and newer one is more precise and faster than the other one; *PeoplEyes* thus prefers to use the first machine as much as possible. Unfortunately, a new norm comes in force stating that no laser technology can be used, unless human staff wears laser-protective goggles.

If *PeoplEyes* has both human resources and raw material, and the three machines are fully working, but it has not yet bought any laser-protective goggles, all its goals would be achieved but it would fail to comply with the applicable regulations, since the norm for the no-usage of laser technology is violated and not compensated.

If *PeoplEyes* buys the laser-protective goggles, their entire production process also becomes norm compliant. If, at some time, the more precise mounting machine breaks, but the second one is still working, *PeoplEyes* can still reach some of its objectives since the usage of the second machine leads to a state of the world where the objective of mounting the glasses on the frames is accomplished. Again, if *PeoplEyes* has no protective laser goggles and both the mounting machines are out of order, *PeoplEyes*' production process is neither norm, nor outcome compliant.

The following theory is the formalisation into our logic of the above scenario.

$$F = \{lenses,\ frames,\ new\_safety\_regulation\}$$

$$
\begin{aligned}
R = \{ & r_1 : \Rightarrow_{\mathsf{U}} \mathit{eye\_Glasses} \\
& r_2 : \Rightarrow \mathit{laser} \\
& r_3 : \mathit{lenses}, \mathit{laser} \Rightarrow \mathit{glasses} \\
& r_4 : \Rightarrow \mathit{mounting\_machine1} \\
& r_5 : \Rightarrow \mathit{mounting\_machine2} \\
& r_6 : \mathit{mounting\_mach1} \Rightarrow \neg \mathit{mounting\_machine2} \\
& r_7 : \mathit{frames}, \mathit{glasses}, \mathit{mounting\_machine1} \Rightarrow \mathit{eye\_Glasses} \\
& r_8 : \mathit{frames}, \mathit{glasses}, \mathit{mounting\_machine2} \Rightarrow \mathit{eye\_Glasses} \\
& r_9 : \mathit{new\_safety\_regulation} \Rightarrow_{\mathsf{O}} \neg \mathit{laser} \otimes \mathit{goggles} \\
& r_{10} : \Rightarrow_{\mathsf{U}} \mathit{mounting\_machine1} \oplus \mathit{mounting\_machine2} \} \\
>^{sm} = \{ & r_6 > r_5 \}.
\end{aligned}
$$

We assume *PeoplEye* has enough resources to start the process by setting *lenses* and *frames* as facts. Rule r_1 states that producing *eye_Glasses* is the main objective ($+\partial_{\mathsf{I}} \mathit{eye\_Glasses}$, we choose intention as the mental attitude to comply with/attain to); rules r_2, r_4 and r_5 describe that we can use, respectively, the laser and the two mounting machineries. Rule r_3 is to represent that, if we have lenses and a laser machinery available, then we can shape glasses; in the same way, rules r_7 and r_8 describe that whenever we have glasses and one of the mounting machinery is available, then we obtain the final product. Therefore, the positive extension for belief $+\partial$ contains *laser*, *glasses*, *mounting_machine1* and *eye_Glasses*. In that occasion, rule r_6 along with $>$ prevent the using of both machineries at the same time and thus $-\partial \mathit{mounting\_machine2}$ (we assumed, for illustrative purpose even if unrealistically, that a parallel execution is not possible). When a new safety regulation comes in force (r_9), the usage of the laser machinery is forbidden, unless protective goggles are worn ($+\partial_{\mathsf{O}} \neg \mathit{laser}$ and $+\partial_{\mathsf{O}} \neg \mathit{goggles}$). Finally, rule r_{10} is to describe the preference of using *mounting_machine1* instead of *mounting_machine2* (hence we have $+\partial_{\mathsf{I}} \mathit{mounting\_machine1}$ and $-\partial_{\mathsf{I}} \mathit{mounting\_machine2}$).

Since there exists no rule for goggles, the theory is outcome compliant (that is, it reaches some set of objectives), but not norm compliant (given that it fails to meet some obligation rules without compensating them). If we add *goggles* to the facts and we substitute r_2 with

$$r_2' : \mathsf{O} \mathit{goggles} \Rightarrow \mathit{laser}$$

then we are both norm and outcome compliant, as well as if we add

$$r_{11} : \mathit{mounting\_machine1\_broken} \Rightarrow \neg \mathit{mounting\_machine1}$$

to R and *mounting_machine1_broken* to F. Notice that, with respect to *laser*, we are intention compliant but *not* social intention compliant (given $\mathsf{O} \neg \mathit{lenses}$). This is a key characteristic of our logic: The system is informed that the process is compliant but some violations have occurred.

Appendix B Proofs of Propositions in Section 3

Proposition 1
Let D be a consistent, finite defeasible theory. For any literal l, it is not possible to have both

1. $D \vdash +\partial_X l$ and $D \vdash -\partial_X l$ with $X \in \text{MOD}$;
2. $D \vdash +\partial_X l$ and $D \vdash +\partial_X {\sim}l$ with $X \in \text{MOD} \setminus \{\mathsf{D}\}$.

Proof
1. (Coherence of the logic) The negative proof tags are the strong negation of the positive ones, and so are the conditions of a rule being discarded (Definition 10) for a rule being applicable (Definition 9). Hence, when the conditions for $+\partial_X$ hold, those for $-\partial_X$ do not.

2. (Consistency of the logic) We split the proof into two cases: (i) at least one of Xl and $X{\sim}l$ is in F, and (ii) neither of them is in F. For (i) the proposition immediately follows by the assumption of consistency. In fact, suppose that $Xl \in$ F. Then clause (1) of $+\partial_X$ holds for l. By consistency $X{\sim}l \notin$ F, thus clause (1) of Definition 13 does not hold for ${\sim}l$. Since $Xl \in$ F, also clause (2.1) is always falsified for ${\sim}l$, and the thesis is proved.

For (ii), let us assume that both $+\partial_X l$ and $+\partial_X {\sim}l$ hold in D. A straightforward assumption derived by Definitions 9 and 10 is that no rule can be at the same time applicable and discarded for X and l for any literal l and its complement. Thus, we have that there are applicable rules for X and l, as well as for X and ${\sim}l$. This means that clause (2.3.2) of Definition 13 holds for both l and ${\sim}l$. Therefore, for every applicable rule for l there is an applicable rule for ${\sim}l$ stronger than the rule for l. Symmetrically, for every applicable rule for ${\sim}l$ there is an applicable rule for l stronger than the rule for ${\sim}l$. Since the set of rules in D is finite by construction, this situation is possible only if there is a cycle in the transitive closure of the superiority relation, which is in contradiction with the hypothesis of D being consistent. □

Proposition 2
Let D be a consistent defeasible theory. For any literal l, the following statements hold:

1. if $D \vdash +\partial_X l$, then $D \vdash -\partial_X {\sim}l$ with $X \in \text{MOD} \setminus \{\mathsf{D}\}$;
2. if $D \vdash +\partial l$, then $D \vdash -\partial_{\mathsf{I}} {\sim}l$;
3. if $D \vdash +\partial l$ or $D \vdash +\partial_{\mathsf{O}} l$, then $D \vdash -\partial_{\mathsf{SI}} {\sim}l$;
4. if $D \vdash +\partial_{\mathsf{G}} l$, then $D \vdash +\partial_{\mathsf{D}} l$;
5. if $D \vdash -\partial_{\mathsf{D}} l$, then $D \vdash -\partial_{\mathsf{G}} l$.

Proof
For part 1., let D be a consistent defeasible theory, and $D \vdash +\partial_X l$. Literal ${\sim}l$ can be in only one of the following, mutually exclusive situations: (i) $D \vdash +\partial_X {\sim}l$; (ii) $D \vdash -\partial_X {\sim}l$; (iii) $D \not\vdash \pm\partial_X {\sim}l$. Part 2 of Proposition 1 allows us to exclude case (i), since $D \vdash +\partial_X l$ by hypothesis. Case (iii) denotes situations where there are loops in the theory involving literal ${\sim}l$,[1] but inevitably this would affect also the provability of Xl, i.e., we would not be able to give a proof for $+\partial_X l$ as well. This is in contradiction with the hypothesis. Consequently, situation (ii) must be the case.

[1] For example, situations like $X{\sim}l \Rightarrow_X {\sim}l$, where the proof conditions generate a loop without introducing a proof.

Parts 2. and 3. directly follow by Definitions 9 and 10, while Definitions 9 and 13 justify part 4., given that G is not involved in any conflict relation.

Part 5. Trivially, from part 4. □

Proposition 3
Let D be a consistent defeasible theory. For any literal l, the following statements *do not* hold:

6. if $D \vdash +\partial_{\mathsf{D}} l$, then $D \vdash +\partial_X l$ with $X \in \{\mathsf{G}, \mathsf{I}, \mathsf{SI}\}$;
7. if $D \vdash +\partial_{\mathsf{G}} l$, then $D \vdash +\partial_X l$ with $X \in \{\mathsf{I}, \mathsf{SI}\}$;
8. if $D \vdash +\partial_X l$, then $D \vdash +\partial_Y l$ with $X = \{\mathsf{I}, \mathsf{SI}\}$ and $Y = \{\mathsf{D}, \mathsf{G}\}$;
9. if $D \vdash -\partial_Y l$, then $D \vdash -\partial_X l$ with $Y \in \{\mathsf{D}, \mathsf{G}\}$ and $X \in \{\mathsf{I}, \mathsf{SI}\}$.

Proof
Example 2 in the extended version offers counterexamples showing the reason why the above statements do not hold.

$$
\begin{aligned}
F &= \{saturday, John\_away, John\_sick\}\\
R &= \{r_2 : saturday \Rightarrow_{\mathsf{U}} visit\_John \odot visit\_parents \odot watch\_movie\\
&\qquad r_3 : John\_away \Rightarrow_{\mathsf{B}} \neg visit\_John\\
&\qquad r_4 : John\_sick \Rightarrow_{\mathsf{U}} \neg visit\_John \odot short\_visit\}\\
&\qquad r_7 : John\_away \Rightarrow_{\mathsf{B}} \neg short\_visit\}\\
> &= \{(r_2, r_4)\}.
\end{aligned}
$$

Given that $r_2 > r_4$, Alice has the desire to *visit_John*, and this is also her preferred outcome. Nonetheless, being *John_away* a fact, this is not her intention, while so are $\neg$*visit_John* and *visit_parents*. □

Appendix C Correctness and Completeness of DefeasibleExtension

In this appendix we give proofs of the lemmas used by Theorem 6 for the soundness and completeness of the algorithms proposed.

We recall that the algorithms in Section 4 are based on a series of transformations that reduce a given theory into an equivalent, simpler one. Here, equivalent means that the two theories have the same extension, and simpler means that the size of the target theory is smaller than that of the original one. Remember that the size of a theory is the number of instances of literals occurring in the theory plus the number of rules in the theory. Accordingly, each transformation either removes some rules or some literals from rules (specifically, rules or literals we know are no longer useful to produce new conclusions). There is an exception. At the beginning of the computation, the algorithm creates four rules (one for each type of goal-like attitude) for each outcome rule (and the outcome rule is then eliminated). The purpose of this operation is to simplify the transformation operations and the bookkeeping of which rules have been used and which rules are still able to produce new conclusions (and the type of conclusions). Alternatively, one could implement flags to achieve the same result, but in a more convoluted way. A consequence of this operation is that we no longer have outcome rules. This implies that we have (i) to adjust the proof theory, and (ii) to show that the adjusted proof theory and the theory with the various goal-like rules are equivalent to the original theory and original proof conditions.

The adjustment required to handle the replacement of each outcome rule with a set of rules of goal-like modes (where each new rule has the same body and consequent of the outcome rule it replaces) is to modify the definition of being applicable (Definition 9) and being discarded (Definition 10). Specifically, we have to replace

- $r \in R^{\mathsf{U}}$ in clause 3 of Definition 9 with $r \in R^{\mathsf{D}}$;
- $r \notin R^{\mathsf{U}}$ in clause 3 of Definition 10 with $r \notin R^{\mathsf{D}}$;
- $r \in R^{\mathsf{U}}$ in clause 4.1.1 of Definition 9 with $r \in R^{X}$; and
- $r \notin R^{\mathsf{U}}$ in clause 4.1.1 of Definition 10 with $r \notin R^{X}$.

Given a theory D with goal-like rules instead of outcome rules we will use $E_3(D)$ to refer to the extension of D computed using the proof theory obtained from the proof theory defined in Section 3 with the modified versions of the notions of applicable and discarded just given.

Lemma 7
Let $D = (F, R, >)$ be a defeasible theory. Let $D' = (F, R', >')$ be the defeasible theory obtained from D as follows:

$$
\begin{aligned}
R' &= R^{\mathsf{B}} \cup R^{\mathsf{O}} \cup \{r_X : A(r) \hookrightarrow_X C(r) | r : A(r) \hookrightarrow_{\mathsf{U}} C(r) \in R, X \in \{\mathsf{D}, \mathsf{G}, \mathsf{I}, \mathsf{SI}\}\} \\
>' &= \{(r,s) | (r,s) \in >,\ s, r \in R^{\mathsf{B}} \cup R^{\mathsf{O}}\} \cup \{(r_X, s_Y) | (r,s) \in >, r, s \in R^{\mathsf{U}}\} \cup \\
&\quad \{(r_X, s) | (r,s) \in >, r \in R^{\mathsf{U}}, s \in R^{\mathsf{B}} \cup R^{\mathsf{O}}\} \cup \{(r, s_X) | (r,s) \in >, r \in R^{\mathsf{B}} \cup R^{\mathsf{O}}, s \in R^{\mathsf{U}}\}
\end{aligned}
$$

Then, $E(D) = E_3(D')$.

Proof
The differences between D and D' are that each outcome-rule in D corresponds to four rules in D' each for a different mode and all with the same antecedent and consequent of the rule in D. Moreover, every time a rule r in D is stronger than a rule s in D, then any rule corresponding to r in D' is stronger than any rule corresponding to s in D'.

The differences in the proof theory for D and that for D' is in the definitions of applicable for X and discarded for X. It is immediate to verify that every time a rule r is applicable (at index n) for X, then r_X is applicable (at index n) for X (and the other way around). □

Given the functional nature of the transformations involved in the algorithms, we shall refer to the rules in the target theory with the same labels as the rules in the source theory. Thus, given a rule $r \in D$, we will refer to the rule corresponding to it in D' (if it exists) with the same label, namely r.

In the algorithms, belief rules may convert to another mode $\Diamond$ only through the support set $R^{\mathsf{B},\Diamond}$. Definition 7 requires $R^{\mathsf{B},\Diamond}$ to be initialised with a modal version of each belief rule with *non-empty* antecedent, such that every literal a in the antecedent is replaced by the corresponding modal literal $\Diamond a$.

In this manner, rules in $R^{\mathsf{B},\Diamond}$ satisfy clauses 1 and 2 of Definitions 7 and 8 by construction, while clauses 3 of both definitions are satisfied iff these new rules for $\Diamond$ are body-applicable (resp. body-discarded). Therefore, conditions for rules in $R^{\mathsf{B},\Diamond}$ to be applicable/discarded collapse into those of Definition 5 and 6, and accordingly these rules are applicable for mode $\Diamond$ only if they satisfy clauses clauses (2.1.1), (3.1), or (4.1.1) of Definitions 9 and 10, based on how $\Diamond$ is instantiated. That is to say, during the execution of the algorithms, we can empty the body of the rules in $R^{\mathsf{B},\Diamond}$ by iteratively proving all the modal literals in the antecedent to decide which rules are applicable at a given step.

Before proceeding with the demonstrations of the lemmas, we recall that in the formalisation of the logic in Section 3, we referred to modes with capital roman letters (X, Y, T) while the notation of the algorithms in Section 4 proposes the variant with $\Box$, $\blacksquare$ and $\Diamond$ since it was needed to fix a given modality for the iterations and pass the correct input for each call of a subroutine. Therefore, being that the hypotheses of the lemmas refer to the operations performed by the algorithms, while the proofs refer to the notation of Definitions 5–15, in the following the former ones use the symbol $\Box$ for a mode, the latter ones the capital roman letters notation.

Lemma 8
Let $D = (\mathrm{F}, R, >)$ be a defeasible theory such that $D \vdash +\partial_{\Box} l$ and $D' = (\mathrm{F}, R', >')$ be the theory obtained from D where

$$\begin{aligned} R' =& \{r : A(r) \setminus \{\Box l, \neg\Box{\sim}l\} \hookrightarrow C(r) |\ r \in R,\ A(r) \cap \widetilde{\Box l} = \emptyset\} \\ R'^{\mathsf{B},\Box} =& \{r : A(r) \setminus \{\Box l\} \hookrightarrow C(r) | r \in R^{\mathsf{B},\Box},\ A(r) \cap \widetilde{\Box l} = \emptyset\} \\ >' =& > \setminus \{(r,s),(s,r) \in > |\ A(r) \cap \widetilde{\Box l} \neq \emptyset\}. \end{aligned}$$

Then $D \equiv D'$.

Proof
The proof is by induction on the length of a derivation P. For the inductive base, we consider all possible derivations for a literal q in the theory.

$P(1) = +\partial_X q$*, with* $X \in \mathrm{MOD} \setminus \{\mathsf{D}\}$. This is possible in two cases: (1) $Xq \in \mathrm{F}$, or (2) $\widetilde{Yq} \cap \mathrm{F} = \emptyset$, for $Y = X$ or Conflict(Y,X), and $\exists r \in R^X[q,i]$ that is applicable in D for X at i and $P(1)$, and every rule $s \in R^Y[{\sim}q, j]$ is either (a) discarded for X at j and $P(1)$, or (b) defeated by a stronger rule $t \in R^T[q,k]$ applicable for T at k and $P(1)$ (T may conflict with Y).

Concerning (1), by construction of D', $Xq \in \mathrm{F}$ iff $Xq \in \mathrm{F}'$, thus if $+\partial_X q$ is provable in D then is provable in D', and vice versa.

Regarding (2), again by construction of D', $\widetilde{Yq} \cap \mathrm{F} = \emptyset$ iff $\widetilde{Yq} \cap \mathrm{F}' = \emptyset$. Moreover, r is applicable at $P(1)$ iff $i = 1$ (since lemma's operations do not modify the tail of the rules) and $A(r) = \emptyset$. Therefore, if $A(r) = \emptyset$ in D then $A(r) = \emptyset$ in D'. This means that if a rule is applicable in D at $P(1)$ then is applicable in D' at $P(1)$. In the other direction, if r is applicable in D' at $P(1)$, then either (i) $A(r) = \emptyset$ in D, or (ii) $A(r) = \{\Box l\}$, or $A(r) = \{\neg\Box{\sim}l\}$. For (i), r is straightforwardly applicable in D, as well as for (ii) since $D \vdash +\partial_{\Box} l$ by hypothesis.

When we consider possible attacks to rule r, namely $s \in R^Y[{\sim}q, j]$, we have to analyse cases (a) and (b) above.

(a) Since we reason about $P(1)$, it must be the case that no such rule s exists in R, and thus s cannot be in R' either. In the other direction, the difference between D and D' is that in R we have rules with $\widetilde{\Box l}$ in the antecedent, and such rules are not in R'. Since $D \vdash +\partial_{\Box} l$ by hypothesis, all rules in R for which there is no counterpart in R' are discarded in D.

(b) We modify the superiority relation by only withdrawing instances where one of the rules is discarded in D. But only when t is applicable then is active in the clauses of the proof conditions where the superiority relation is involved, i.e., (2.3.2) of Definition 13. We have just proved that if a rule is applicable in D then is applicable in D' as well, and if is discarded in D then is discarded in D'. If s is not discarded in D for Y at 1 and $P(1)$, then there exists an applicable rule t in D for q stronger than s. Therefore t is applicable in D' for T and $t >' s$ if $T = Y$, or Conflict(T,Y). Accordingly, $D' \vdash +\partial_X q$. The same reasoning applies in the other direction. Consequently, if we have a derivation of length 1 of $+\partial q$ in D', then we have a derivation of length 1 of $+\partial q$ in D as well.

Notice that in the inductive base by their own nature rules in $R^{\mathsf{B},\Diamond}$, even if can be modified or erased, cannot be used in a proof of length one.

$P(1) = +\partial_{\mathsf{D}} q$. The proof is essentially identical to the inductive base for $+\partial_X q$, with some slight modifications dictated by the different proof conditions for $+\partial_{\mathsf{D}}$: (1) $\mathsf{D}q \in \mathrm{F}$, or (2) $\neg\mathsf{D}q \notin \mathrm{F}$, and $\exists r \in R^{\mathsf{D}}[q,i]$ that is applicable for D at 1 and $P(1)$ and every rule $s \in R^{\mathsf{D}}[\sim q, j]$ is either (a) discarded for D at 1 and $P(1)$, or (b) s is not stronger than r.

$P(1) = -\partial_X q$ *with* $X \in \mathrm{MOD}$. Clearly conditions (1) and (2.1) of Definition 14 hold in D iff they do in D', given that $\mathrm{F} = \mathrm{F}'$. The analysis for clause (2.2) is the same of case (a) of $P(1) = +\partial_X q$, while for clause (2.3.1) the reader is referred to case (2), where in both cases r and s change their role. For condition (2.3.2) if $X = \mathsf{D}$, then $s > r$. Otherwise, either there is no $t \in R^T[q,k]$ in D (we recall that at $P(1)$, t cannot be discarded in D because that would imply a previous step in the proof), or $t \not> s$ and not Conflict(T,Y). Therefore $s \in R'$ by construction, and conditions on the superiority relation between s and t are preserved. Hence, $D' \vdash -\partial_X q$. For the other direction, we have to consider the case of a rule s in R but not in R'. As we have proved above, all rules discarded in D' are discarded in D, and all rules in R for which there is no corresponding rule in R' are discarded in D as well, and we can process this case with the same reasoning as above.

For the inductive step, the property equivalence between D and D' is assumed up to the n-th step of a generic proof for a given literal p.

$P(n+1) = +\partial_X q$*, with* $X \in \mathrm{MOD}$. Clauses (1) and (2.1) follow the same conditions treated in the inductive base for $+\partial_X q$. As regards clause (2.2), we distinguish if $X = \mathsf{B}$, or not. In the former case, if there exists a rule $r \in R[q,i]$ applicable for B in D, then clauses 1.–3. of Definition 5 are all satisfied. By inductive hypothesis, we conclude that the clauses are satisfied by r in D' as well no matter whether $\Box l \in A(r)$, or not.

Otherwise, there exists a rule r applicable in D for X at $P(n+1)$ such that r is either in $R^X[q,i]$, or $R^{\mathsf{B},X}[q,1]$. By inductive hypothesis, we can conclude that: (i) if $r \in R^X[q,i]$ then r is body-applicable and the clauses of Definition 5 are satisfied by r in D' as well; (ii) if $r \in R^{\mathsf{B},X}[q,1]$ then r is Conv-applicable and the clauses of Definition 7 are satisfied by r in D' as well. As regards conditions (2.1.2) or (4.1.2), the provability/refutability of the elements in the chain prior to q is given by inductive hypothesis. The direction from rule applicability in D' to rule applicability in D follows the same reasoning and so is straightforward.

Condition (2.3.1) states that every rule $s \in R^{Y}[\sim q, j] \cup R^{\mathsf{B},Y}[\sim q, 1]$ is discarded in D for X at $P(n+1)$. This means that there exists an $a \in A(s)$ satisfying one of the clauses of Definition 6 if $s \in R^{\mathsf{B},Y}[\sim q, 1]$, or Definition 10 if $s \in R^{Y}[\sim q, j]$. Two possible situations arise. If $a \in \widetilde{\Box l}$, then $s \notin R'$; otherwise, by inductive hypothesis, either a satisfies Definition 6 or 8 in D' depending on $s \in R^{Y}[\sim q, j]$ or $s \in R^{\mathsf{B},Y}[\sim q, 1]$. Hence, s is discarded in D' as well. The same reasoning applies for the other direction. The difference between D and D' is that in R we have rules with elements of $\widetilde{\Box l}$ in the antecedent, and these rules are not in R'. Consequently, if s is discarded in D', then is discarded in D and all rules in R for which there is no corresponding rule in R' are discarded in D since $D \vdash +\partial_{\Box} l$ by hypothesis.

If $X \neq \mathsf{D}$, then condition (2.3.2) can be treated as case (b) of the corresponding inductive base except clause (2.3.2.1) where if $t > s$ then either: (i) $Y = T$, (ii) $s \in R^{\mathsf{B},T}[\sim q]$ and $t \in R^{T}[q]$ (Convert(Y, T)), or (iii) $s \in R^{Y}[\sim q]$ and $t \in R^{\mathsf{B},Y}[q]$ (Convert(T, Y)). Instead if $X = \mathsf{D}$, no modifications are needed.

$P(n+1) = -\partial_X q$, with $X \in$ MOD. The analysis is a combination of the inductive base for $-\partial_X q$ and inductive step for $+\partial_X q$ where we have already proved that a rule is applicable (discarded) in D iff is so in D' (or it is not contained in R'). Even condition (2.3.2.1) is just the strong negation of the reason in the above paragraph. □

Lemma 9
Let $D = (\mathrm{F}, R, >)$ be a defeasible theory such that $D \vdash -\partial_{\Box} l$ and $D' = (\mathrm{F}, R', >')$ be the theory obtained from D where

$$\begin{aligned} R' =& \{r : A(r) \setminus \{\neg \Box l\} \hookrightarrow C(r) |\ r \in R,\ \Box l \notin A(r)\} \\ R'^{\mathsf{B},\Box} =& \{r \in R^{\mathsf{B},\Box} |\ \Box l \notin A(r)\} \\ >' =& > \setminus \{(r,s),(s,r) \in > |\ \Box l \in A(r)\}. \end{aligned}$$

Then $D \equiv D'$.

Proof
We split the proof in two cases, depending on if $\Box \neq \mathsf{D}$, or $\Box = \mathsf{D}$.

As regards the former case, since Proposition 2 states that $+\partial_X m$ implies $-\partial_X {\sim} m$ then modifications on R', $R'^{\mathsf{B},\Box}$, and $>'$ represent a particular case of Lemma 8 where $m = {\sim} l$.

We now analyse the case when $\Box = \mathsf{D}$. The analysis is identical to the one shown for the inductive base of Lemma 8 but for what follows.

$P(1) = +\partial_X q$. Case (2)–(ii): $A(r) = \{\neg \Box l\}$ and since $D \vdash -\partial_{\Box} l$ by hypothesis, then if r is applicable in D' at $P(1)$ then is applicable in D at $P(1)$ as well.

Case (2)–(a): the difference between D and D' is that in R we have rules with $\Box l$ in the antecedent, and such rules are not in R'. Since $D \vdash -\partial_{\Box} l$ by hypothesis, all rules in R for which there is no counterpart in R' are discarded in D.

The same modification happens in the inductive step $P(n+1) = +\partial_X q$, where also the sentence 'If $a \in \widetilde{\Box l}$, then $s \notin R'$' becomes 'If $a = \Box l$, then $s \notin R'$'.

Finally, the inductive base and inductive step for the negative proof tags are identical to ones of the previous lemma. □

Hereafter we consider theories obtained by the transformations of Lemma 8. This means that all applicable rules are such because their antecedents are empty and every rule in R appears also in R' and vice versa, and there are no modifications in the antecedent of rules.

Lemma 10
Let $D = (\mathrm{F}, R, >)$ be a defeasible theory such that $D \vdash +\partial l$ and $D' = (\mathrm{F}, R', >)$ be the theory obtained from D where

$$R'^{\mathsf{O}} = \{A(r) \Rightarrow_{\mathsf{O}} C(r)!l |\ r \in R^{\mathsf{O}}[l,n]\} \tag{C1}$$

$$R'^{\mathsf{I}} = \{A(r) \Rightarrow_{\mathsf{I}} C(r)!l |\ r \in R^{\mathsf{I}}[l,n]\} \cup \{A(r) \Rightarrow_{\mathsf{I}} C(r) \ominus {\sim}l |\ r \in R^{\mathsf{I}}[{\sim}l,n]\} \tag{C2}$$

$$R'^{\mathsf{SI}} = \{A(r) \Rightarrow_{\mathsf{SI}} C(r) \ominus {\sim}l |\ r \in R^{\mathsf{SI}}[{\sim}l,n]\}. \tag{C3}$$

Moreover,

- if $D \vdash +\partial_{\mathsf{O}}{\sim}l$, then instead of (C1)
$$R'^{\mathsf{O}} = \{A(r) \Rightarrow_{\mathsf{O}} C(r)!l |\ r \in R^{\mathsf{O}}[l,n]\} \cup \{A(r) \Rightarrow_{\mathsf{O}} C(r) \ominus {\sim}l |\ r \in R^{\mathsf{O}}[{\sim}l,n]\}. \tag{C1}$$
- if $D \vdash -\partial_{\mathsf{O}}{\sim}l$, then instead of (C3)
$$R'^{\mathsf{SI}} = \{A(r) \Rightarrow_{\mathsf{SI}} C(r) \ominus {\sim}l |\ r \in R^{\mathsf{SI}}[{\sim}l,n]\} \cup \{A(r) \Rightarrow_{\mathsf{SI}} C(r)!l |\ r \in R^{\mathsf{SI}}[l,n]\}. \tag{C3}$$

Then $D \equiv D'$.

Proof
The demonstration follows the inductive base and inductive step of Lemma 8 where we consider the particular case $\Box = \mathsf{B}$. Since here operations to obtain D' modify only the consequent of rules, verifying conditions when a given rule is applicable/discarded reduces to clauses (2.1.2) and (4.1.2) of Definitions 9–10, while conditions for a rule being body-applicable/discarded are trivially treated. Moreover, the analysis is narrowed to modalities O, I, and SI since rules for the other modalities are not affected by the operations of the lemma. Finally, notice that the operations of the lemma do not erase rules from R to R' but it may be the case that, given a rule r, if removal or truncation operate on an element c_k in $C(r)$, then $r \in R[l]$ while $r \notin R'[l]$ for a given literal l (removal of l or truncation at c_k).

$P(1) = +\partial_X q$, with $X \in \{\mathsf{O}, \mathsf{I}, \mathsf{SI}\}$. We start by considering condition (2.2) of Definition 13 where a rule $r \in R^X[q,i]$ is applicable in D at $i = 1$ and $P(1)$. In both cases when $q = l$ or $q \neq l$, q is the first element of $C(r)$ since either we truncate chains at l, or we remove ${\sim}l$ from them. Therefore, r is applicable in D' as well. In the other direction, if r is applicable in D' at 1 and $P(1)$, then $r \in R$ has either q as the first element, or only ${\sim}l$ precedes q. In the first case r is trivially applicable, while in the second case the applicability of r follows from the hypothesis that $D \vdash +\partial l$ and $D \vdash +\partial_{\mathsf{O}}{\sim}l$ if $r \in R^{\mathsf{O}}$, or $D \vdash +\partial l$ and $D \vdash -\partial_{\mathsf{O}}{\sim}l$ if $r \in R^{\mathsf{SI}}$.

Concerning condition (2.3.1) of Definition 13 there is no such rule s in R, hence s cannot be in R' (we recall that at $P(1)$, s cannot be discarded in D because that would imply a previous step in the proof). Regarding the other direction, we have to consider the situation where there is a rule $s \in R^Y[{\sim}q, j]$ which is not in $R'^Y[{\sim}q]$. This is the case when the truncation has operated on $s \in R^Y[\sim q, j]$ since l preceded $\sim q$ in $C(s)$, making s discarded in D as well (either when (i) $Y = \mathsf{O}$ or $Y = \mathsf{I}$, or (ii) $D \vdash -\partial_{\mathsf{O}} \sim l$ and $Y = \mathsf{SI}$).

For (2.3.2) the reasoning is the same of the equivalent case in Lemma 8 with the additional condition that rule t may be applicable in D' at $P(1)$ but q appears at index 2 in $C(t)$ in D.

$P(n+1) = +\partial_X q$, with $X \in \{\mathsf{O}, \mathsf{I}, \mathsf{SI}\}$. Again, let us suppose $r \in R[q,i]$ to be applicable in D for X at i and $P(n+1)$. By hypothesis and clauses (2.1.2) or (4.1.2) of Definition 9, we conclude that $c_k \neq l$ and $q \neq \sim l$ (Conflict(B, I) and Conflict$(\mathsf{B}, \mathsf{SI})$). Thus, r is applicable in D' by inductive hypothesis. The other direction sees $r \in R'[q,i]$ applicable in D' and either $\sim l$ preceded q in $C(r)$ in D, or not. Since in the first case, the corresponding operation of the lemma is the removal of $\sim l$ from $C(r)$, while in the latter case no operations on the consequent are done, the applicability of r in D at $P(n+1)$ is straightforward.

For condition (2.3.1), the only difference between the inductive base is when there is a rule s in $R^Y[\sim q, j]$ but $s \notin R'^Y[\sim q, k]$. This means that l precedes $\sim q$ in $C(s)$ in D, and thus by hypothesis s is discarded in D. Notice that if $q = l$, then $R'^Y[\sim l, k] = \emptyset$ for any k by the removal operation of the lemma, and thus condition (2.3.1) is vacuously true.

$P(1) = -\partial_X q$ and $P(n+1) = -\partial_X q$, with $X \in$ MOD. They trivially follow from the inductive base and inductive step. ☐

Lemma 11
Let $D = (\mathrm{F}, R, >)$ be a defeasible theory such that $D \vdash -\partial l$ and $D' = (\mathrm{F}, R', >)$ be the theory obtained from D where

$$R'^{\mathsf{I}} = \{A(r) \Rightarrow_{\mathsf{I}} C(r)!\sim l \mid r \in R^{\mathsf{I}}[\sim l, n]\}.$$

Moreover,

- if $D \vdash +\partial_{\mathsf{O}} l$, then
$$R'^{\mathsf{O}} = \{A(r) \Rightarrow_{\mathsf{O}} C(r) \ominus l \mid r \in R^{\mathsf{O}}[l, n]\};$$
- if $D \vdash -\partial_{\mathsf{O}} l$, then
$$R'^{\mathsf{SI}} = \{A(r) \Rightarrow_{\mathsf{SI}} C(r)!\sim l \mid r \in R^{\mathsf{SI}}[\sim l, n]\}.$$

Then $D \equiv D'$.

Proof
The demonstration is a mere variant of that of Lemma 10 since: (i) Proposition 2 states that $+\partial_X m$ implies $-\partial_X \sim m$ (mode D is not involved), and (ii) operations of the lemma are a subset of those of Lemma 10 where we switch l with $\sim l$, and the other way around. ☐

Lemma 12
Let $D = (\mathrm{F}, R, >)$ be a defeasible theory such that $D \vdash +\partial_{\mathsf{O}} l$ and $D' = (\mathrm{F}, R', >)$ be the theory obtained from D where

$$R'^{\mathsf{O}} = \{A(r) \Rightarrow_{\mathsf{O}} C(r)!\sim l \ominus \sim l \mid r \in R^{\mathsf{O}}[\sim l, n]\} \quad \text{(C1)}$$

$$R'^{\mathsf{SI}} = \{A(r) \Rightarrow_{\mathsf{SI}} C(r) \ominus \sim l \mid r \in R^{\mathsf{SI}}[\sim l, n]\}. \quad \text{(C2)}$$

Moreover,

- if $D \vdash -\partial l$, then instead of (C1)

$$R'^{\mathsf{O}} = \{A(r) \Rightarrow_{\mathsf{O}} C(r)!{\sim}l \ominus {\sim}l | \ r \in R^{\mathsf{O}}[{\sim}l,n]\} \cup \{A(r) \Rightarrow_{\mathsf{O}} C(r) \ominus l | \ r \in R^{\mathsf{O}}[l,n]\}; \tag{C1}$$

- if $D \vdash -\partial{\sim}l$, then instead of (C2)

$$R'^{\mathsf{SI}} = \{A(r) \Rightarrow_{\mathsf{SI}} C(r) \ominus {\sim}l | \ r \in R^{\mathsf{SI}}[{\sim}l,n]\} \cup \{A(r) \Rightarrow_{\mathsf{SI}} C(r)!l | \ r \in R^{\mathsf{SI}}[l,n]\}. \tag{C2}$$

Then $D \equiv D'$.

Proof
Again, the proof is a variant of that of Lemma 10 that differs only when truncation and removal operate on a consequent at the same time.

A CTD is relevant whenever its elements are proved as obligations. Consequently, if D proves $\mathsf{O}l$, then $\mathsf{O}{\sim}l$ cannot hold. If this is the case, then $\mathsf{O}{\sim}l$ cannot be violated and elements following ${\sim}l$ in obligation rules cannot be triggered. Nonetheless, the inductive base and inductive step do not significantly differ from those of Lemma 10. In fact, even operation (1) involving truncation and removal of ${\sim}l$ does not affect the equivalence of conditions for being applicable/discarded between D and D'. □

Proofs for Lemmas 13–17 are not reported. As stated for Lemma 12, they are variants of that for Lemma 10 where the modifications concern the set of rules on which we operate. The underlying motivation is that truncation and removal operations affect when a rule is applicable/discarded as shown before where we have proved that, given a rule s and a literal ${\sim}q$, it may be the case that ${\sim}q \notin C(s)$ in R' while the opposite holds in R. Such modifications reflect only the nature of the operations of truncation and removal while they do not depend on the mode of the rule involved.

Lemma 13
Let $D = (\mathrm{F}, R, >)$ be a defeasible theory such that $D \vdash -\partial_{\mathsf{O}} l$ and $D' = (\mathrm{F}, R', >)$ be the theory obtained from D where

$$R'^{\mathsf{O}} = \{A(r) \Rightarrow_{\mathsf{O}} C(r)!l \ominus l | \ r \in R^{\mathsf{O}}[l,n]\}.$$

Moreover,

- if $D \vdash -\partial l$, then

$$R'^{\mathsf{SI}} = \{A(r) \Rightarrow_{\mathsf{SI}} C(r)!{\sim}l | \ r \in R^{\mathsf{SI}}[{\sim}l,n]\}.$$

Then $D \equiv D'$.

Lemma 14
Let $D = (\mathrm{F}, R, >)$ be a defeasible theory such that $D \vdash +\partial_{\mathsf{D}} l$, $D \vdash +\partial_{\mathsf{D}} {\sim}l$, and $D' = (\mathrm{F}, R', >)$ be the theory obtained from D where

$$R'^{\mathsf{G}} = \{A(r) \Rightarrow_{\mathsf{G}} C(r)!l \ominus l | \ r \in R^{\mathsf{G}}[l,n]\} \cup \{A(r) \Rightarrow_{\mathsf{G}} C(r)!{\sim}l \ominus {\sim}l | \ r \in R^{\mathsf{G}}[{\sim}l,n]\}.$$

Then $D \equiv D'$.

Lemma 15
Let $D=(\mathrm{F},R,>)$ be a defeasible theory such that $D\vdash -\partial_{\mathsf{D}} l$ and $D'=(\mathrm{F},R',>)$ be the theory obtained from D where

$$\begin{aligned}R'^{\mathsf{D}} =&\{A(r)\Rightarrow_{\mathsf{D}} C(r)\ominus l|\ r\in R^{\mathsf{D}}[l,n]\}\\ R'^{\mathsf{G}} =&\{A(r)\Rightarrow_{\mathsf{G}} C(r)\ominus l|\ r\in R^{\mathsf{G}}[l,n]\}.\end{aligned}$$

Then $D\equiv D'$.

Lemma 16
Let $D=(\mathrm{F},R,>)$ be a defeasible theory such that $D\vdash +\partial_X l$, with $X\in\{\mathsf{G},\mathsf{I},\mathsf{SI}\}$, and $D'=(\mathrm{F},R',>)$ be the theory obtained from D where

$$\begin{aligned}R'^{X} =&\{A(r)\Rightarrow_X C(r)!l|\ r\in R^X[l,n]\}\cup\\ &\{A(r)\Rightarrow_X C(r)\ominus{\sim}l|\ r\in R^X[{\sim}l,n]\}.\end{aligned}$$

Then $D\equiv D'$.

Lemma 17
Let $D=(\mathrm{F},R,>)$ be a defeasible theory such that $D\vdash -\partial_X l$, with $X\in\{\mathsf{G},\mathsf{I},\mathsf{SI}\}$, and $D'=(\mathrm{F},R',>)$ be the theory obtained from D where

$$R'^{X} =\{A(r)\Rightarrow_X C(r)\ominus l|\ r\in R^X[l,n]\}.$$

Then $D\equiv D'$.

Lemma 18
Let $D=(\mathrm{F},R,>)$ be a defeasible theory and $l\in\mathrm{Lit}$ such that (i) $Xl\notin\mathrm{F}$, (ii) $\neg Xl\notin\mathrm{F}$ and $Y{\sim}l\notin\mathrm{F}$ with $Y=X$ or Conflict(Y,X), (iii) $\exists r\in R^X[l,1]\cup R^{\mathsf{B},X}[l,1]$, (iv) $A(r)=\emptyset$, and (v) $R^X[{\sim}l]\cup R^{\mathsf{B},X}[{\sim}l]\cup R^Y[{\sim}l]\setminus R_{infd}\subseteq r_{inf}$, with $X\in\mathrm{MOD}\setminus\{\mathsf{D}\}$. Then $D\vdash +\partial_X l$.

Proof
To prove Xl, Definition 13 must be taken into consideration: since hypothesis (i) falsifies clause (1), then clause (2) must be the case. Let r be a rule that meets the conditions of the lemma. Hypotheses (iii) and (iv) state that r is applicable for X. In particular, if $r=s^{\diamond}\in R^{\mathsf{B},X}$ then s is Conv-applicable. Finally, for clause (2.3) we have that all rules for ${\sim}l$ are inferiorly defeated by an appropriate rule with empty antecedent for l, but a rule with empty body is applicable. Consequently, all clauses for proving $+\partial_X$ are satisfied. Thus, $D\vdash +\partial_X l$. □

Lemma 19
Let $D=(\mathrm{F},R,>)$ be a defeasible theory and $l\in\mathrm{Lit}$ such that (i) $\mathsf{D}l\notin\mathrm{F}$, (ii) $\neg\mathsf{D}l\notin\mathrm{F}$, (iii) $\exists r\in R^{\mathsf{D}}[l,1]\cup R^{\mathsf{B},\mathsf{D}}[l,1]$, (iv) $A(r)=\emptyset$, and (v) $r_{sup}=\emptyset$. Then $D\vdash +\partial_{\mathsf{D}} l$.

Proof
The demonstration is analogous to that for Lemma 18 since all lemma's hypotheses meet clause (2) of Definition 11. □

Lemma 20
Let $D=(\mathrm{F},R,>)$ be a defeasible theory and $l\in\mathrm{Lit}$ such that $l,Xl\notin\mathrm{F}$ and $R^X[l]\cup R^{\mathsf{B},X}[l]=\emptyset$, with $X\in\mathrm{MOD}$. Then $D\vdash -\partial_X l$.

Proof
Conditions (1) and (2.2) of Definitions 12 and 14 are vacuously satisfied with the same comment for $R^{\mathsf{B},X}$ in Lemma 18. □

Lemma 21
Let $D = (\mathrm{F}, R, >)$ be a defeasible theory and $l \in \mathrm{Lit}$ such that (i) $X{\sim}l \notin \mathrm{F}$, (ii) $\neg X{\sim}l \notin \mathrm{F}$ and $Yl \notin \mathrm{F}$ with $Y = X$ or Conflict(Y,X), (iii) $\exists r \in R^X[l,1] \cup R^{\mathsf{B},X}[l,1]$, (iv) $A(r) = \emptyset$, and (v) $r_{sup} = \emptyset$, with $X \in \mathrm{MOD}$. Then $D \vdash -\partial_X{\sim}l$.

Proof
Let r be a rule in a theory D for which the conditions of the lemma hold. It is easy to verify that clauses (1) and (2.3) of Definitions 12 and 14 are satisfied for ${\sim}l$. □

Theorem 4
Given a finite defeasible theory D with size S, Algorithms 2 PROVED and 3 REFUTED terminate and their computational complexity is $O(S)$.

Proof
Every time Algorithms 2 PROVED or 3 REFUTED are invoked, they both modify a subset of the set of rules R, which is finite by hypothesis. Consequently, we have their termination. Moreover, since $|R| \in O(S)$ and each rule can be accessed in constant time, we obtain that their computational complexity is $O(S)$. □

Theorem 5
Given a finite defeasible theory D with size S, Algorithm 1 DEFEASIBLEEXTENSION terminates and its computational complexity is $O(S)$.

Proof
The most important part to analyse concerning termination of Algorithm 1 DEFEASIBLE-EXTENSION is the **repeat/until** cycle at lines 12–37. Once an instance of the cycle has been performed, we are in one of the following, mutually exclusive situations:

1. No modification of the extension has occurred. In this case, line 37 ensures the termination of the algorithm;
2. The theory has been modified with respect to a literal in HB. Notice that the algorithm takes care of removing the literal from HB once the suitable operations have been performed (specifically, at line 3 of Algorithm 2 PROVED and 3 REFUTED). Since this set is finite, the process described above eventually empties HB and, at the next iteration of the cycle, the extension of the theory cannot be modified. In this case, the algorithm ends its execution as well.

Moreover, Lemma 4 proved the termination of its internal sub-routines.

In order to analyse complexity of the algorithm, it is of the utmost importance to correctly comprehend Definition 19. Remember that the size of a theory is the number of *all occurrences* of each literal in every rule plus the number of the rules. The first term is usually (much) bigger than the latter. Let us examine a theory with x literals and whose size is S, and consider the scenario when an algorithm A, looping over all x literals of the theory, invokes an inner procedure P which selectively deletes a literal given as input from all the rules of the theory (no matter to what end). A rough computational complexity would be

$O(S^2)$, given that, when one of the $x \in O(S)$ literal is selected, P removes all its occurrences from every rule, again $O(S)$.

However, a more fined-grained analysis shows that the complexity of A is lower. The mistake being to consider the complexity of P separately from the complexity of the external loop, while instead they are strictly dependent. Indeed, the overall number of operations made by the sum of all loop iterations cannot outrun the number of occurrences of the literals, $O(S)$, because the operations in the inner procedure directly decrease, iteration after iteration, the number of the remaining repetitions of the outmost loop, and the other way around. Therefore, the overall complexity is not bound by $O(S) \cdot O(S) = O(S^2)$, but by $O(S) + O(S) = O(S)$.

We can now contextualise the above reasoning to Algorithm 1 DEFEASIBLEEXTENSION, where D is the theory with size S. The initialisation steps (lines 1–5 and 10–11) add an $O(S)$ factor to the overall complexity. The main cycle at lines 12–37 is iterated over HB, whose cardinality is in $O(S)$. The analysis of the preceding paragraph implies that invoking Algorithm 2 PROVED at lines 7 and 29 as well as invoking Algorithm 3 REFUTED at lines 8, 15, 26 and 27 represent an additive factor $O(S)$ to the complexity of **repeat/until** loop and **for** cycle at lines 6–9 as well. Finally, all operations on the set of rules and the superiority relation require constant time, given the implementation of data structures proposed. Therefore, we can state that the complexity of the algorithm is $O(S)$. □

Theorem 6
Algorithm 1 DEFEASIBLEEXTENSION is sound and complete.

Proof
As already argued at the beginning of the section, the aim of Algorithm 1 DEFEASIBLEEXTENSION is to compute the defeasible extension of a given defeasible theory D through successive transformations on the set of facts and rules, and on the superiority relation: at each step, they compute a simpler theory while retaining the same extension. Again, we remark that the word 'simpler' is used to denote a theory with fewer elements in it. Since we have already proved the termination of the algorithm, it eventually comes to a fixed-point theory where no more operations can be made.

In order to demonstrate the soundness of Algorithm 1 DEFEASIBLEEXTENSION, we show in the list below that all the operations performed by the algorithm are justified by Proposition 2 and described in Lemmas 7–21, where we prove the soundness of each operation involved.

1. Algorithm 1 DEFEASIBLEEXTENSION:
   - Lines 2–3 and 5: Lemma 7;
   - Line 7: item 2. below;
   - Line 8: item 3. below;
   - Line 15: Lemma 20 and item 3. below;
   - Line 24: Lemma 19 and item 2. below;
   - Lines 26–27: Lemma 21 and item 3. below;
   - Line 29: Lemma 18 and item 2. below;
2. Algorithm 2 PROVED:
   - Line 4: Lemma 21 and item 3. below;

- Line 5: Part 2. of Proposition 2 and item 3. below;
- Line 6: Part 3. of Proposition 2 and item 3. below;
- Lines 7–9: Lemma 8;
- CASE B at lines 11–14: Lemma 10;
- CASE O at lines 15–18: Lemma 12;
- CASE D at lines 19–23: Lemma 14;
- OTHERWISE at lines 24–26: Lemma 16;

3. Algorithm 3 REFUTED:
   - Lines 4–6: Lemma 9;
   - CASE B at lines 8–11: Lemma 11;
   - CASE O at lines 12–14: Lemma 13;
   - CASE D at lines 15–16: Lemma 15;
   - OTHERWISE at lines 17–18: Lemma 17;

The result of these lemmas is that whether a literal is defeasibly proved or not in the initial theory, so it will be in the final theory. This proves the soundness of the algorithm.

Moreover, since (i) all lemmas show the equivalence of the two theories, and (ii) the equivalence relation is a bijection, this also demonstrates the completeness of Algorithm 1 DEFEASIBLEEXTENSION. □