On Strong and Default Negation in Logic Program Updates

Martin Slota
CENTRIA
New University of Lisbon

Martin Baláž
Faculty of Mathematics, Physics and Informatics
Comenius University

João Leite
CENTRIA
New University of Lisbon

Abstract
Existing semantics for answer-set program updates fall into two categories: either they consider only strong negation in heads of rules, or they primarily rely on default negation in heads of rules and optionally provide support for strong negation by means of a syntactic transformation.

In this paper we pinpoint the limitations of both these approaches and argue that both types of negation should be first-class citizens in the context of updates. We identify principles that plausibly constrain their interaction but are not simultaneously satisfied by any existing rule update semantics. Then we extend one of the most advanced semantics with direct support for strong negation and show that it satisfies the outlined principles as well as a variety of other desirable properties.

1 Introduction
The increasingly common use of rule-based knowledge representation languages in highly dynamic and information-rich contexts, such as the Semantic Web (Berners-Lee, Hendler, and Lassila 2001), requires standardised support for updates of knowledge represented by rules. Answer-set programming (Gelfond and Lifschitz 1988; 1991) forms the natural basis for investigation of rule updates, and various approaches to answer-set program updates have been explored throughout the last 15 years (Leite and Pereira 1998; Alferes et al. 1998; 2000; Eiter et al. 2002; Leite 2003; Sakama and Inoue 2003; Alferes et al. 2005; Banti et al. 2005; Zhang 2006; Šefránek 2006; Delgrande, Schaub, and Tompits 2007; Osorio and Cuevas 2007; Šefránek 2011; Krümpelmann 2012).

The most straightforward kind of conflict arising between an original rule and its update occurs when the original conclusion logically contradicts the newer one. Though the technical realisation and final result may differ significantly, depending on the particular rule update semantics, this kind of conflict is resolved by letting the newer rule prevail over the older one. Actually, under most semantics, this is also the only type of conflict that is subject to automatic resolution (Leite and Pereira 1998; Alferes et al. 2000; Eiter et al. 2002; Alferes et al. 2005; Banti et al. 2005; Delgrande, Schaub, and Tompits 2007; Osorio and Cuevas 2007).

From this perspective, allowing for both strong and default negation to appear in heads of rules is essential for an expressive and universal rule update framework (Leite 2003). While strong negation is the natural candidate here, used to express that an atom becomes explicitly false, default negation allows for more fine-grained control: the atom only ceases to be true, but its truth value may not be known after the update. The latter also makes it possible to move between any pair of epistemic states by means of updates, as illustrated in the following example:

Example 1.1 (Railway crossing (Leite 2003)). Suppose that we use the following logic program to choose an action at a railway crossing:

cross ← ¬train. wait ← train. listen ← ¬train, ¬¬train.

The intuitive meaning of these rules is as follows: one should cross if there is evidence that no train is approaching; wait if there is evidence that a train is approaching; listen if there is no such evidence.

Consider a situation where a train is approaching, represented by the fact (train.). After this train has passed by, we want to update our knowledge to an epistemic state where we lack evidence with regard to the approach of a train. If this was accomplished by updating with the fact (¬train.), we would cross the tracks at the subsequent state, risking being killed by another train that was approaching. Therefore, we need to express an update stating that all past evidence for an atom is to be removed, which can be accomplished by allowing default negation in heads of rules. In this scenario, the intended update can be expressed by the fact (¬¬train.).

With regard to the support of negation in rule heads, existing rule update semantics fall into two categories: those that only allow for strong negation, and those that primarily consider default negation. As illustrated above, the former are unsatisfactory as they render many belief states unreachable by updates. As for the latter, they optionally provide support for strong negation by means of a syntactic transformation.

Two such transformations are known from the literature, both of them based on the principle of coherence: if an atom \( p \) is true, its strong negation \( \neg p \) cannot be true simultaneously, so \( \neg \neg p \) must be true, and also vice versa, if \( \neg p \) is true, then so is \( \sim p \). The first transformation, introduced in (Alferes and Pereira 1996), encodes this principle directly
by adding, to both the original program and its update, the following two rules for every atom \( p \):

\[
\sim \neg p \leftarrow p, \quad \sim p \leftarrow \neg p.
\]

This way, every conflict between an atom \( p \) and its strong negation \( \neg p \) directly translates into two conflicts between the objective literals \( p, \neg p \) and their default negations. However, the added rules lead to undesired side effects that stand in direct opposition with basic principles underlying updates. Specifically, despite the fact that the empty program does not encode any change in the modelled world, the stable models assigned to a program may change after an update by the empty program.

This undesired behaviour is addressed in an alternative transformation from (Leite 2003) that encodes the coherence principle more carefully. Nevertheless, this transformation also leads to undesired consequences, as demonstrated in the following example:

**Example 1.2** (Faulty sensor). Suppose that we collect data from sensors and, for security reasons, multiple sensors are used to supply information about the critical fluent \( p \). In case of a malfunction of one of the sensors, we may end up with an inconsistent logic program consisting of the following two facts:

\[
p, \quad \sim p.
\]

At this point, no stable model of the program exists and action needs to be taken to find out what is wrong. If a problem is found in the sensor that supplied the first fact \( (p) \), after the sensor is repaired, this information needs to be reset by updating the program with the fact \( (\sim p) \). Following the universal pattern in rule updates, where recovery from conflicting states is always possible, we expect that this update is sufficient to assign a stable model to the updated program. However, the transformational semantics for strong negation defined in (Leite 2003) still does not provide any stable model – we remain without a valid epistemic state when one should in fact exist.

In this paper we address the issues with combining strong and default negation in the context of rule updates. Based on the above considerations, we formulate a generic desirable principle that is violated by the existing approaches. Then we show how two distinct definitions of one of the most well-behaved rule update semantics (Alferes et al. 2005; Banti et al. 2005) can be equivalently extended with support for strong negation. The resulting semantics not only satisfies the formulated principle, but also retains the formal and computational properties of the original semantics. More specifically, our main contributions are as follows:

- we show that the defined semantics enjoy the early recovery principle as well as a range of desirable properties for rule updates known from the literature.

This paper is organised as follows: In Sect. 2 we present the syntax and semantics of logic programs, generalise the well-supported semantics from the class of normal programs to extended ones and define the rule update semantics from (Alferes et al. 2005; Banti et al. 2005). Then, in Sect. 3, we formally establish the early recovery principle, define the new rule update semantics for strong negation and show that it satisfies the principle. In Sect. 4 we introduce other established rule update principles and show that the proposed semantics satisfies them. We discuss our findings and conclude in Sect. 5.1

## 2 Background

In this section we introduce the necessary technical background and generalise the well-supported semantics (Fages 1991) to the class of extended programs.

### 2.1 Logic Programs

In the following we present the syntax of non-disjunctive logic programs with both strong and default negation in heads and bodies of rules, along with the definition of stable models of such programs from (Leite 2003) that is equivalent to the original definitions based on reducts (Gelfond and Lifschitz 1988; 1991; Inoue and Sakama 1998). Furthermore, we define an alternative characterisation of the stable model semantics: the well-supported models of normal logic programs (Fages 1991).

We assume that a countable set of propositional atoms \( A \) is given and fixed. An *objective literal* is an atom \( p \in A \) or its strong negation \( \neg p \). We denote the set of all objective literals by \( \mathcal{L} \). A *default literal* is an objective literal preceded by \( \sim \) denoting default negation. A *literal* is either an objective or a default literal. We denote the set of all literals by \( \mathcal{L}^* \).

As a convention, double negation is absorbed, so that \( \neg \neg p \) denotes the atom \( p \) and \( \sim \sim l \) denotes the objective literal \( l \).

Given a set of literals \( S \), we introduce the following notation:

\[
S^+ = \{ l \in \mathcal{L} \mid l \in S \}, \quad S^- = \{ l \in \mathcal{L} \mid \sim l \in S \},
\]

\[
\sim S = \{ \sim L \mid L \in S \}.
\]

An *extended rule* is a pair \( \pi = (H_\pi, B_\pi) \) where \( H_\pi \) is a literal, referred to as the *head of \( \pi \), and \( B_\pi \) is a finite set of literals, referred to as the *body of \( \pi \). Usually we write \( \pi \) as \( (H_\pi \leftarrow B_\pi^+, \sim B_\pi^-) \). A *generalised rule* is an extended rule that contains no occurrence of \( \sim \), i.e., its head and body consist only of atoms and their default negations. A *normal rule* is a generalised rule that has an atom in the head. A *fact* is an extended rule whose body is empty and a *tautology* is any extended rule \( \pi \) such that \( H_\pi \in B_\pi \). An extended (generalised, normal) program is a set of extended (generalised, normal) rules.

An *interpretation* is a consistent subset of the set of objective literals, i.e., a subset of \( \mathcal{L} \) does not contain both \( p \) and \( \neg p \) for any atom \( p \). The satisfaction of an objective literal \( l \),

\[1\]An extended version of this paper with all the proofs is available as (Slota, Baláž, and Leite 2014).
default literal $\neg l$, set of literals $S$, extended rule $\pi$ and extended program $P$ in an interpretation $J$ is defined in the usual way: $J \models l$ iff $l \in J$; $J \models \neg l$ iff $l \notin J$; $J \models S$ iff $J \models l$ for all $l \in S$; $J \models \pi$ iff $J \models B_\pi$ implies $J \models H_\pi$; $J \models P$ iff $J \models \pi$ for all $\pi \in P$. Also, $J$ is a model of $P$ if $J \models P$, and $P$ is consistent if it has a model.

**Definition 2.1** (Stable model). Let $P$ be an extended program. The set $\llbracket P \rrbracket_{SM}$ of stable models of $P$ consists of all interpretations $J$ such that

$$J^* = \text{least} (P \cup \text{def}(J))$$

where $\text{def}(J) = \{ \neg l \mid l \in L \setminus J \}$, $J^* = J \cup (L \setminus J)$ and least($\cdot$) denotes the least model of the argument program in which all literals are treated as propositional atoms.

A level mapping is a function that maps every atom to a natural number. Also, for any default literal $\neg p$, where $p \in A$, and finite set of atoms and their default negations $S$, $\ell(\neg p) = \ell(p)$, $\ell^l(S) = \min \{ \ell(L) \mid L \in S \}$ and $\ell^H(S) = \max \{ \ell(L) \mid L \in S \}$.

**Definition 2.2** (Well-supported model of a normal program). Let $P$ be a normal program and $\ell$ a level mapping. An interpretation $J \subseteq A$ is a well-supported model of $P$ w.r.t. $\ell$ if the following conditions are satisfied:

1. $J$ is a model of $P$;
2. For every objective literal $l \in J$ there exists a rule $\pi \in P$ such that

$$H_\pi = p \land J \models B_\pi \land \ell(H_\pi) > \ell^H(B_\pi).$$

The set $\llbracket P \rrbracket_{WS}$ of well-supported models of $P$ consists of all interpretations $J$ such that $J$ is a well-supported model of $P$ w.r.t. some level mapping.

We obtain a generalisation of Prop. 2.3 to the class of extended programs:

**Proposition 2.5.** Let $P$ be an extended program. Then, $\llbracket P \rrbracket_{WS} = \llbracket P \rrbracket_{SM}$.

### 2.3 Rule Updates

We turn our attention to rule updates, starting with one of the most advanced rule update semantics, the refined dynamic stable models for sequences of generalised programs (Alferes et al. 2005), as well as the equivalent definition of well-supported models (Banti et al. 2005). Then we define the transformations for adding support for strong negation to such semantics (Alferes and Pereira 1996; Leite 2003).

A rule update semantics provides a way to assign stable models to a pair or sequence of programs where each component represents an update of the preceding ones. Formally, a dynamic logic program (DLP) is a finite sequence of extended programs and by all($P$) we denote the multiset of all rules in the components of $P$. A rule update semantics $\mathbb{S}$ assigns a set of $\mathbb{S}$-models, denoted by $\llbracket P \rrbracket_\mathbb{S}$, to $P$.

We focus on semantics based on the causal rejection principle (Leite and Pereira 1998; Alferes et al. 2000; Eiter et al. 2002; Leite 2003; Alferes et al. 2005; Banti et al. 2005; Osorio and Cuevas 2007) which states that a rule is rejected if it is in a direct conflict with a more recent rule. The basic type of conflict between rules $\pi$ and $\sigma$ occurs when their heads contain complementary literals, i.e. when $H_\pi = \neg H_\sigma$. Based on such conflicts and on a stable model candidate, a set of rejected rules can be determined and it can be verified that the candidate is indeed stable w.r.t. the remaining rules.

We define the most mature of these semantics, providing two equivalent definitions: the refined dynamic stable models (Alferes et al. 2005), or RD-semantics, defined using a fixpoint equation, and the well-supported models (Banti et al. 2005), or WS-semantics, based on level mappings.

**Definition 2.6** (RD-semantics (Alferes et al. 2005)). Let $P = (P_i)_{i \leq n}$ be a DLP without strong negation. Given an interpretation $J$, the multisets of rejected rules $\mathbb{r}_J(P, J)$ and of default assumptions $\mathbb{d}_J(P, J)$ are defined as follows:

$$\mathbb{r}_J(P, J) = \{ \pi \in P_i \mid i < n \land \exists j \geq i \exists \sigma \in P_j : H_\pi = \neg H_\sigma \land J \models B_\sigma \},$$

$$\mathbb{d}_J(P, J) = \{ (\neg l) \mid l \in L \land \neg (\exists \pi \in \llbracket P \rrbracket : H_\pi = l \land J \models B_\pi) \}.$$
Definition 2.7 (WS-semantics (Banti et al. 2005)). Let \( P = \langle P_0 \rangle_{i<n} \) be a DLP without strong negation. Given an interpretation \( J \) and a level mapping \( \ell \), the multiset of rejected rules \( \text{rej}_\ell(P, J) \) is defined as follows:

\[
\text{rej}_\ell(P, J) = \{ \pi \in P_i | i < n \land \exists j > i \exists \sigma \in P_j : H_\sigma = \neg \neg \sigma \land J \models B_\sigma \land \ell(H_\sigma) > \ell(\neg \neg \sigma) \}.
\]

The set \( [P]_{\text{ws}} \) of WS-models of \( P \) consists of all interpretations \( J \) such that for some level mapping \( \ell \), the following conditions are satisfied:

1. \( J \) is a model of \( \text{all}(P) \setminus \text{rej}_\ell(P, J) \);
2. For every \( l \in J \) there exists some rule \( \pi \in \text{all}(P) \setminus \text{rej}_\ell(P, J) \) such that

\[
H_\pi = l \land J \models B_\pi \land \ell(H_\pi) > \ell(\neg \neg \pi).
\]

Unlike most other rule update semantics, these semantics can properly deal with tautological and other irrelevant updates, as illustrated in the following example:

Example 2.8 (Irrelevant updates). Consider the DLP \( P = \langle P, U \rangle \) where programs \( P, U \) are as follows:

\[
P : \quad \text{day} \leftarrow \neg \text{night}. \quad \text{stars} \leftarrow \text{night}, \neg \text{cloudy}.
\]

\[
\text{night} \leftarrow \text{day}. \quad \neg \text{stars}.
\]

\[
U : \quad \text{stars} \leftarrow \text{stars}.
\]

Note that program \( P \) has the single stable model \( J_1 = \{ \text{day} \} \) and \( U \) contains a single tautological rule, i.e. it does not encode any change in the modelled domain. Thus, we expect that \( P \) also has the single stable model \( J_1 \). Nevertheless, many rule update semantics, such as those introduced in (Leite and Pereira 1998; Alferes et al. 2000; Eiter et al. 2002; Leite 2003; Sakama and Inoue 2003; Zhang 2006; Osorio and Cuevas 2007; Delgrande, Schaub, and Tompits 2007; Krümpelmann 2012), are sensitive to this or other tautological updates, introducing or eliminating models of the original program.

In this case, the unwanted model candidate is \( J_2 = \{ \text{night}, \text{stars} \} \) and it is neither an RD- nor a WS-model of \( P \), though the reasons for these are technically different under these two semantics. It is not difficult to verify that, given an arbitrary level mapping \( \ell \), the respective sets of rejected rules and the set of default assumptions are as follows:

\[
\text{rej}_\ell(P, J_2) = \{ \text{stars} \leftarrow \text{night}, \neg \text{cloudy}., (\neg \text{stars}. \} ,
\]

\[
\text{rej}_\ell(P, J_2) = \emptyset,
\]

\[
\text{def}(P, J_2) = \{ \text{\neg \neg \text{cloudy}.}, (\neg \text{stars}. \} .
\]

Note that \( \text{rej}_\ell(P, J_2) \) is empty because, independently of \( \ell \), no rule \( \pi \) in \( U \) satisfies the condition \( \ell(H_\pi) > \ell(\neg \neg \pi) \), so there is no rule that could reject another rule. Thus, the atom stars belongs to \( J_2 \) but does not belong to least \( \{ \text{\text{all}}(P) \setminus \text{rej}_\ell(P, J_2) \} \cup \text{def}(P, J_2) \), so \( J_2 \) is not an RD-model of \( P \).

Furthermore, no model of \( \text{all}(P) \setminus \text{rej}_\ell(P, J_2) \) contains stars, so \( J_2 \) cannot be a WS-model of \( P \).

Furthermore, the resilience of RD- and WS-semantics is not limited to empty and tautological updates, but extends to other irrelevant updates as well (Alferes et al. 2005; Banti et al. 2005). For example, consider the DLP \( P' = \langle P, U' \rangle \) where \( U' = \{ \text{stars} \leftarrow \text{venus}, \text{venus} \leftarrow \text{stars} \} \).

Though the updating program contains non-tautological rules, it does not provide a bottom-up justification of any model other than \( J_1 \) and, indeed, \( J_1 \) is the only RD- and WS-model of \( P' \).

We also note that the two presented semantics for DLPs without strong negation provide the same result regardless of the particular DLP to which they are applied.

Proposition 2.9 ((Banti et al. 2005)). Let \( P \) be a DLP without strong negation. Then, \( [P]_{\text{ws}} = [P]_{\text{so}} \).

In case of the stable model semantics for a single program, strong negation can be reduced away by treating all objective literals as atoms and adding, for each atom \( p \), the integrity constraint \( \langle \langle p, \neg p \rangle \rangle \) to the program (Gelfond and Lifschitz 1991). However, this transformation does not serve its purpose when adding support for strong negation to causal rejection semantics for DLPs because integrity constraints have empty heads, so according to these rule update semantics, they cannot be used to reject any other rule. For example, a DLP such as \( \{ \langle p, \neg p \rangle, \{ p \} \} \) would remain without a stable model even though the DLP \( \{ \langle p, \neg p \rangle, \} \{ p \} \) does have a stable model.

To capture the conflict between opposite objective literals \( l \) and \( \neg l \) in a way that is compatible with causal rejection semantics, a slightly modified syntactic transformation can be performed, translating such conflicts into conflicts between objective literals and their default negations. Two such transformations have been suggested in the literature (Alferes and Pereira 1996; Leite 2003), both based on the principle of coherence. For any extended program \( P \) and DLP \( P = \langle P_i \rangle_{i<n} \) they are defined as follows:

\[
P^\dagger = P \cup \{ \neg l \leftarrow l. | l \in \mathcal{L} \},
\]

\[
P^\dagger = \langle P^\dagger \rangle_{i<n} .
\]

\[
P^\dagger = P \cup \{ \neg H_\pi \leftarrow B_\pi. | \pi \in P \land H_\pi \in \mathcal{L} \},
\]

\[
P^\dagger = \langle P^\dagger \rangle_{i<n} .
\]

These transformations lead to four possibilities for defining the semantics of an arbitrary DLP \( P: [P^\dagger]_{\text{so}}, [P^\dagger]_{\text{so}}, [P^\dagger]_{\text{ws}} \) and \( [P^\dagger]_{\text{ws}} \). We discuss these in the following section.

3 Direct Support for Strong Negation in Rule Updates

The problem with existing semantics for strong negation in rule updates is that semantics based on the first transformation \( P^\dagger \) assign too many models to some DLPs, while semantics based on the second transformation \( P^\dagger \) sometimes do not assign any model to a DLP that should have one. The former is illustrated in the following example:

Example 3.1 (Undesired side effects of the first transformation). Consider the DLP \( P_1 = \langle P, U \rangle \) where \( P = \{ p, \neg p \} \) and \( U = \emptyset \). Since \( P \) has no stable model and \( U \) does not encode any change in the represented domain, it should follow that \( P_1 \) has no stable model either. However, \( [P_1]_{\text{so}} = \}

It is reasonable to expect that since formation) (Undesired side effects of the second trans-
is assigned when in fact a model should exist.

In other words, unless that program already contains some original rules.

Belong to the program ever, the vast majority of existing rule update semantics, including RD-semantics, we first need to look more closely at the set of facts that solves all conflicts in P, S addressed to the more recent program can be used to reject facts in the older one.

Thus, the problem with the first transformation is that an update by an empty program, which does not express that were added to the more recent program can be used to reject the facts in the older one.

Based on this example, in the following we formulate a generic early recovery principle that formally identifies conditions under which some stable model should be assigned to a DLP. For the sake of simplicity, we concentrate on DLPs of length 2 which are composed of facts. We discuss a generalisation of the principle to DLPs of arbitrary length and containing other rules than just facts in Sect. 5. After introducing the principle, we define a semantics for rule updates which directly supports both strong and default negation and satisfies the principle.

We begin by defining, for every objective literal l, the sets of literals l and ∼l as follows:

\[ I = \{ \sim l, \sim \sim l \} \quad \text{and} \quad \sim I = \{ l \} \ . \]

Intuitively, for every literal L, L denotes the set of literals that are in conflict with L. Furthermore, given two sets of facts P and U, we say that U solves all conflicts in P if for each pair of rules π, σ ∈ P such that Hπ ∈ Hσ there is a fact ρ ∈ U such that either Hπ ∈ Hρ or Hρ ∈ Hπ.

Considering a rule update semantics S, the new principle simply requires that when U solves all conflicts in P, S will assign some model to (P, U). Formally:

**Early recovery principle:** If P is a set of facts and U is a consistent set of facts that solves all conflicts in P, then \( [P,U]_S \neq \emptyset \).

We conjecture that rule update semantics should generally satisfy the above principle. In contrast with the usual behaviour of belief update operators, the nature of existing rule update semantics ensures that recovery from conflict is always possible, and this principle simply formalises and sharpens the sufficient conditions for such recovery.

Our next goal is to define a semantics for rule updates that not only satisfies the outlined principle, but also enjoys other established properties of rule updates that have been identified over the years. Similarly as for the original semantics for rule updates, we provide two equivalent definitions, one based on a fixed point equation and the other one on level mappings.

To directly accommodate strong negation in the RD-semantics, we first need to look more closely at the set of rejected rules rej2(P, J), particularly at the fact that it allows conflicting rules within the same component of P to reject one another. This behaviour, along with the constrained set of defaults def(P, J), is used to prevent tautological and other irrelevant cyclic updates from affecting the semantics. However, in the presence of strong negation, rejecting conflicting rules within the same program has undesired side effects. For example, the early recovery principle requires that some model be assigned to the DLP \( \{ p, \neg p \} \cup \{ \neg p \} \) from Example 3.2, but if the rules in the initial program reject each other, then the only possible stable model to assign is ∅. However, such a stable model would violate the causal rejection principle since it does not satisfy the initial rule \( \neg p \) and there is no rule in the updating program that overrides it.

To overcome the limitations of this approach to the prevention of tautological updates, we disentangle rule rejection per se from ensuring that rejection is done without
cyclic justifications. We introduce the set of rejected rules \( \text{rej}_i(P, S) \) which directly supports strong negation and does not allow for rejection within the same program. Prevention of cyclic rejections is done separately by using a customised immediate consequence operator \( T_{P, J} \). Given a stable model candidate \( J \), instead of verifying that \( J^* \) is the least fixed point of the usual consequence operator, as done in the RD-semantics using \( \text{least}(-) \), we verify that \( J^* \) is the least fixed point of \( T_{P, J} \).

**Definition 3.3 (Extended RD-semantics).** Let \( P = \{P_i\}_{i<n} \) be a DLP. Given an interpretation \( J \) and a set of literals \( S \), the multiset of rejected rules \( \text{rej}_i(P, S) \), and the consequence operator \( T_{P, J} \) are defined as follows:

\[
\text{rej}_i(P, S) = \{ \pi \in P_i | i < n \land \exists j > i \exists \sigma \in P_j : H_\sigma \in \bar{T}_{\pi} \land B_\sigma \subseteq S \},
\]

\[
\text{rem}(P, S) = \text{all}(P) \setminus \text{rej}_i(P, S),
\]

\[
T_{P, J}(S) = \{ H_\pi | \pi \in (\text{rem}(P, J^*) \cup \text{def}(J)) \land B_\pi \subseteq S \land \neg \exists \sigma \in \text{rem}(P, S) : H_\sigma \in \bar{T}_{\pi} \land B_\sigma \subseteq J^* \}.
\]

Furthermore, \( T_{P, J}^0(S) = S \) and for every \( k \geq 0 \),

\[
T_{P, J}^{k+1}(S) = T_{P, J}(T_{P, J}^k(S)).
\]

The set \( [P]_{\text{RD}}^+ \) of extended RD-models of \( P \) consists of all interpretations \( J \) such that

\[
J^* = \bigcup_{k \geq 0} T_{P, J}^k(\emptyset).
\]

Adding support for strong negation to the WS-semantics is done by modifying the set of rejected rules \( \text{rej}_i(P, J) \) to account for the new type of conflict. Additionally, in order to ensure that rejection of a literal \( L \) cannot be based on the assumption that some conflicting literal \( L' \in \bar{T} \) is true, a rejecting rule \( \sigma \) must satisfy the stronger condition \( \ell^i(\bar{H}) > \ell^i(B_\sigma) \). Finally, to prevent defeated rules from affecting the resulting models, we require that all supporting rules belong to \( \text{rem}(P, J^*) \).

**Definition 3.4 (Extended WS-semantics).** Let \( P = \{P_i\}_{i<n} \) be a DLP. Given an interpretation \( J \) and a level mapping \( \ell \), the multiset of rejected rules \( \text{rej}_{i}^\ell(P, J) \) is defined by:

\[
\text{rej}_{i}^\ell(P, J) = \{ \pi \in P_i | i < n \land \exists j > i \exists \sigma \in P_j : H_\sigma \in \bar{H}_{\pi} \land J \models B_\sigma \land \ell^i(\bar{H}_{\pi}) > \ell^i(B_\sigma) \}.
\]

The set \( [P]_{\text{WS}}^+ \) of extended WS-models of \( P \) consists of all interpretations \( J \) such that for some level mapping \( \ell \), the following conditions are satisfied:

1. \( J \) is a model of \( \text{all}(P) \setminus \text{rej}_{i}^\ell(P, J) \);
2. For every \( \pi \in J \) there exists some rule \( \pi \in \text{rem}(P, J^*) \) such that

\[
H_\pi = \ell(J) = B_\pi \land \ell(H_\pi) > \ell(B_\pi).
\]

The following theorem establishes that the two defined semantics are equivalent:

**Theorem 3.5.** Let \( P \) be a DLP. Then, \( [P]_{\text{WS}}^+ = [P]_{\text{RD}}^+ \).

Also, on DLPs without strong negation they coincide with the original semantics.

**Theorem 3.6.** Let \( P \) be a DLP without strong negation. Then, \( [P]_{\text{WS}}^+ = [P]_{\text{RD}}^+ \).

Furthermore, unlike the transformational semantics for strong negation, the new semantics satisfy the early recovery principle.

**Theorem 3.7.** The extended RD-semantics and extended WS-semantics satisfy the early recovery principle.

4 Properties
In this section we take a closer look at the formal and computational properties of the proposed rule update semantics.

The various approaches to rule updates (Leite and Pereira 1998; Alferes et al. 2000; Eiter et al. 2002; Leite 2003; Sakama and Inoue 2003; Alferes et al. 2005; Banti et al. 2005; Zhang 2006; Šefránek 2006; Osorio and Cuevas 2007; Delgrande, Schaub, and Tompits 2007; Šefránek 2011; Krümpelmann 2012) share a number of basic characteristics. For example, all of them generalise stable models, i.e., the models they assign to a sequence \( P \) (of length 1) are exactly the stable models of \( P \). Similarly, they adhere to the principle of primacy of new information (Dalal 1988), so models assigned to \( P_i \), satisfy the latest program \( P_{i-1} \). However, they also differ significantly in their technical realisation and classes of supported inputs, and desirable properties such as immunity to tautologies are violated by many of them.

Table 1 lists many of the generic properties proposed for rule updates that have been identified and formalised throughout the years (Leite and Pereira 1998; Eiter et al. 2002; Leite 2003; Alferes et al. 2005). The rule update semantics we defined in the previous section enjoys all of them.

**Theorem 4.1.** The extended RD-semantics and extended WS-semantics satisfy all properties listed in Table 1.

Our semantics also retains the same computational complexity as the stable models.

**Theorem 4.2.** Let \( P \) be a DLP. The problem of deciding whether some \( J \in [P]_{\text{WS}}^+ \) exists is \( \text{NP} \)-complete. Given a literal \( L \), the problem of deciding whether for all \( J \in [P]_{\text{WS}}^+ \) it holds that \( J \models L \) is \( \text{coNP} \)-complete.

5 Concluding Remarks
In this paper we have identified shortcomings in the existing semantics for rule updates that fully support both strong and default negation, and proposed a generic early recovery principle that captures them formally. Subsequently, we provided two equivalent definitions of a new semantics for rule updates.

We have shown that the newly introduced rule update semantics constitutes a strict improvement upon the state of the art in rule updates as it enjoys the following combination of characteristics, unmatched by any previously existing semantics:

- It allows for both strong and default negation in heads of rules, making it possible to move between any pair of epistemic states by means of updates;
such that for every solved \( \ell \) of \( P \) there is the single model \( \{ l \in \mathcal{L} : \exists < n : (l) \in P_i \land (\forall j > i : \{ \neg l_i, \neg l_j \} \cap P_j = \emptyset) \}. \) We formalise these ideas in what follows.

We safely conclude that the rule update semantics should assign all potential conflicts within it certainly get resolved, we can render it too strong? Certain caution is appropriate here, another set of facts. Can it be generalised further without in Sect. 3, only covers a single update of a set of facts by due to the latter two rules in the updating program. This another set of facts. Can it be generalised further without it is updated by \( P \) has no stable model. The single conflict in \( P \) is solved after it is updated by \( U \), but then another conflict is introduced because the DLP in question has more than two components due to the latter two rules in the updating program. This second conflict can be solved after another update by \( V \). Consequently, we expect that some stable model be assigned to the DLP \( \langle P, U, V \rangle \).

The original early recovery principle does not impose this because the DLP in question has more than two components and the rules within it are not only facts. However, the DLP is acyclic, as shown by any level mapping \( \ell \) with \( \ell(p) = 3 \), \( \ell(q) = 0 \), \( \ell(r) = 2 \) and \( \ell(s) = 1 \), so the generalised early recovery principle does apply. Furthermore, we also find the single extended RD-model of \( \langle P, U, V \rangle \) is \( \{ \neg p, q, \neg r, s \} \), i.e. the semantics respects the stronger principle in this case.

Moreover, as established in the following theorem, it is no coincidence that the extended RD-semantics respects the stronger principle in the above example – the principle is generally satisfied by the semantics introduced in this paper.

**Theorem 5.2.** The extended RD-semantics and extended WS-semantics satisfy the generalised early recovery principle.

<table>
<thead>
<tr>
<th>Table 1: Desirable properties of rule update semantics</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Generalisation of stable models</strong></td>
</tr>
<tr>
<td><strong>Primacy of new information</strong></td>
</tr>
<tr>
<td><strong>Fact update</strong></td>
</tr>
<tr>
<td><strong>Support</strong></td>
</tr>
<tr>
<td><strong>Idempotence</strong></td>
</tr>
<tr>
<td><strong>Absorption</strong></td>
</tr>
<tr>
<td><strong>Augmentation</strong></td>
</tr>
<tr>
<td><strong>Non-interference</strong></td>
</tr>
<tr>
<td><strong>Immunity to tautologies</strong></td>
</tr>
<tr>
<td><strong>Causal rejection principle</strong></td>
</tr>
</tbody>
</table>

- It satisfies the *early recovery principle* which guarantees the existence of a model whenever all conflicts in the original program are satisfied;
- It enjoys all rule update principles and desirable properties reported in Table 1;
- It does not increase the computational complexity of the stable model semantics upon which it is based.

However, the early recovery principle, as it is formulated in Sect. 3, only covers a single update of a set of facts by another set of facts. Can it be generalised further without rendering it too strong? Certain caution is appropriate here, since in general the absence of a stable model can be caused by odd cycles or simply by the fundamental differences between different approaches to rule update, and the purpose of this principle is not to choose which approach to take.

Nevertheless, one generalisation that should cause no harm is the generalisation to iterated updates, i.e. to sequences of sets of facts. Another generalisation that appears very reasonable is the generalisation to acyclic DLPs, i.e. DLPs such that \( \text{all}(P) \) is an acyclic program. An acyclic program has at most one stable model, and if we guarantee that all potential conflicts within it certainly get resolved, we can safely conclude that the rule update semantics should assign some model to it. We formalise these ideas in what follows.

We say that a program \( P \) is acyclic (Apt and Bezem 1991) if for some level mapping \( \ell \), such that for every \( l \in \mathcal{L}, \ell(l) = \ell(\neg l) \), and every rule \( \pi \in P \) it holds that \( \ell(H_\pi) > \ell(S_\pi) \). Given a DLP \( P = \langle P_i \rangle_{i<n} \), we say that \( \text{all conflicts in } P \) are solved if for every \( i < n \) and each pair of rules \( \pi, \sigma \in P_i \) such that \( H_\pi \in H_\sigma \), there is some \( j > i \) and a fact \( \rho \in P_j \) such that either \( H_\rho \in H_\pi \) or \( H_\rho \in H_\sigma \).

**Generalised early recovery principle:** If \( \text{all}(P) \) is acyclic and all conflicts in \( P \) are solved, then \( \llbracket P \rrbracket_s \neq \emptyset \).

Note that this generalisation of the early recovery principle applies to a much broader class of DLPs than the original one. We illustrate this in the following example:

**Example 5.1 (Recovery in a stratified program).** Consider the following programs programs \( P, U \) and \( V \):

\[
P : \quad p \leftarrow q, \neg r. \quad \neg p \leftarrow s. \quad q. \quad s \leftarrow q.
\]

\[
U : \quad \neg p. \quad r \leftarrow q. \quad \neg r \leftarrow q, s.
\]

\[
V : \quad \neg r.
\]

Looking more closely at program \( P \), we see that atoms \( q \) and \( s \) are derived by the latter two rules inside it while atom \( r \) is false by default since there is no rule that could be used to derive its truth. Consequently, the bodies of the first two rules are both satisfied and as their heads are conflicting, \( P \) has no stable model. The single conflict in \( P \) is solved after it is updated by \( U \), but then another conflict is introduced due to the latter two rules in the updating program. This second conflict can be solved after another update by \( V \). Consequently, we expect that some stable model be assigned to the DLP \( \langle P, U, V \rangle \).
Both the original and the generalised early recovery principle can guide the future addition of full support for both kinds of negations in other approaches to rule updates, such as those proposed in (Sakama and Inoue 2003; Zhang 2006; Delgrande, Schaub, and Tompits 2007; Krümpelmann 2012), making it possible to reach any belief state by updating the current program. Furthermore, adding support for strong negation is also interesting in the context of recent results on program revision and updates that are performed on the semantic level, ensuring syntax-independence of the respective methods (Delgrande et al. 2013; Slota and Leite 2014; 2012a; 2010), in the context of finding suitable condensing operators (Slota and Leite 2013), and unifying with updates in classical logic (Slota and Leite 2012b).

**Acknowledgments**

João Leite was partially supported by Fundação para a Ciência e a Tecnologia under project “AS-CCO/121823/2010). Martin Slota was partially supported by Fundação para a Ciência e a Tecnologia under project “ERRO – Efficient Reasoning with Rules and Ontologies” (PTDC/EIA-CCO/110921/2009). The collaboration between the co-authors resulted from the Slovak–Portuguese bilateral project “ReDIK – Reasoning with Dynamic Inconsistent Knowledge”, supported by APVV agency under SK-PT-0028-10 and by Fundação para a Ciência e a Tecnologia (FCT/2487/3/6/2011/S).

**References**


