

Revising Knowledge in Multi-Agent System Using Preferences

Inna Pivkina, Enrico Pontelli, and Tran Cao Son

Department of Computer Science
New Mexico State University
{ipivkina|epontelli|tson}@cs.nmsu.edu

Abstract. In this paper we extend the *Revision Programming* framework—a logic-based framework to express and maintain constraints on knowledge bases—with different forms of *preferences*. Preferences allow users to introduce a bias in the way agents update their knowledge to meet a given set of constraints. In particular, they provide a way to select one between alternative feasible revisions and they allow the generation of revisions in presence of conflicting constraints, by relaxing the set of satisfied constraints (*soft constraints*). A methodology for computing preferred revisions using answer set programming is presented.

1 Introduction

Multi-Agents Systems (MAS) require coordination mechanisms to facilitate dynamic collaboration of the intelligent components, with the goal of meeting local and/or global objectives. In the case of MAS, the coordination structure should provide communication protocols to link agents having inter-related objectives and it should facilitate mediation and integration of exchanged knowledge [6]. Centralized coordination architectures (e.g., mediator-based architectures) as well as fully distributed architectures (e.g., distributed knowledge networks) face the problem of non-monotonically updating agent’s theories to incorporate knowledge derived from different agents. The problem is compounded by the fact that incoming knowledge could be contradictory—either conflicting with the local knowledge or with other incoming items—incomplete, or unreliable. Recently a number of formalisms have been proposed [15, 3, 19, 7] to support dynamic updates of (propositional) logic programming theories; they provide convenient frameworks for describing knowledge base updates as well as constraints to ensure user-defined principles of consistency. These types of formalisms have been proved effective in the context of MAS (e.g., [11]).

One of such formalisms for knowledge base updates is *Revision Programming*. Revision programming is a formalism to describe and enforce constraints on belief sets, databases, and more generally, on arbitrary knowledge bases. The revision programming formalism was introduced in [14, 15]. In this framework, the *initial database* represents the initial state of a belief set or a knowledge base. A *revision program* is a collection of *revision rules* used to describe constraints on the content of the database. Revision rules could be quite complex and are usually in the form of conditions. For instance, a typical revision rule may express a condition that, if certain elements are

present in the database and some other elements are absent, then another given element must be absent from (or present in) the database. Revision rules offer a natural way of encoding policies for the integration of agent-generated knowledge (e.g., in a mediator-based architecture) or for the management of inter-agent exchanges.

In addition to being a declarative specification of a constraint on a knowledge base, a revision rule also has a computational interpretation—indicating a way to satisfy the constraint. Justified revisions semantics assigns to any knowledge base a (possibly empty) family of *revisions*. Each revision represents an updated version of the original knowledge base, that satisfies all the constraints provided by the revision program. Revisions are obtained by performing additions and deletions of elements from the original knowledge base, according to the content of the revision rules. Each revision might be chosen as an update of the original knowledge base w.r.t. the revision program.

The mechanisms used by revision programming to handle updates of a knowledge base or belief set may lead to indeterminate situations. The constraints imposed on the knowledge base are interpreted as *hard constraints*, that have to be met at all costs; nevertheless this is rather unnatural in domains where overlapping and conflicting consistency constraints may be present (e.g., legal reasoning [17], suppliers and broker agents in a supply chain [12])—leading to the generation of *no* acceptable revisions. Similarly, situations with under-specified constraints or incomplete knowledge may lead to revision programs that provide *multiple* alternative revisions for the same initial knowledge base. While such situations might be acceptable, there are many cases where a single revision is desired—e.g., agents desire to maintain a unique view of a knowledge base.

Preferences provide a natural way to address these issues; preferences allow the revision programmer to introduce a bias, and focus the generation of revisions towards more desirable directions. Preferences between revisions rules and/or preferences between the components of the revisions can be employed to select the way revisions are computed, ruling out undesirable alternatives and defeating conflicting constraints. The use of preference structures has been gaining relevance in the MAS community as key mechanism in negotiation models for MAS coordination architectures [8, 10].

In this work we propose extensions of revision programming that provide general mechanisms to express different classes of preferences—justified by the needs of knowledge integration in MAS. The basic underlying mechanism common to the extensions presented in this work is the idea of allowing classes of revision rules to be treated as *soft revision rules*. A revision might be allowed even if it does not satisfy all the soft revision rules but only selected subsets of them; user preferences express criteria to select the desired subsets of soft revision rules.

Our first approach (Section 3) is based on the use of *revision programs with preferences*, where dynamic partial orders are established between the revision rules. It provides a natural mechanism to select preferred ways of computing revisions, and to prune revisions that are not deemed interesting. This approach is analogous to the ordered logic program (a.k.a. prioritized logic program) approach explored in the context of logic programming (e.g., [5, 4]). In a labeled revision program, the revision program and the initial knowledge base are enriched by a *control program*, which expresses preferences on rules. The control program may include revision literals as well as conditions on the initial knowledge base. Given an initial knowledge base, the control program and

the revision program are translated into a revision program where regular justified revisions semantics is used. This approach provides preference capabilities similar to those supported by the MINERVA agent architecture [11].

The second approach (Section 4) generalizes revision programs through the introduction of *weights* (or *costs*) associated to the components of a revision program (revision rules and/or database atoms). The weights are aimed at providing general criteria for the selection of subsets of the soft revision rules to be considered in the computation of the revisions of the initial database. Different policies in assigning weights are considered, allowing the encoding of very powerful preference criteria (e.g., revisions that differ from the initial database in the least number of atoms). This level of preference management addresses many of the preference requirements described in the MAS literature (e.g., [10]).

For each of the proposed approaches to the management of preferences, we provide an effective implementation schema based on translation to answer set programming—specifically to the `smodels` [16] language. This leads to effective ways to compute *preferred* revisions for any initial database w.r.t. a revision program with preferences.

The main contribution of this work is the identification of forms of preferences that are specifically relevant to the revision programming paradigm and justified by the needs of knowledge maintenance and integration in MAS, and the investigation of the semantics and implementation issues deriving from their introduction.

1.1 Related work

Since revision programming is strongly related to the logic programming formalisms [14, 13, 18], our work is related to several works on reasoning with preferences in logic programming. In this section, we discuss the differences between our approach and some of the research in this area. In logic programming, preferences have been an important source for “correct reasoning”. Intuitively, a logic program is developed to represent a problem, with the intention that its semantics (e.g., answer set or well-founded semantics) will yield correct answers to the specific problem instances. Adding preferences between rules is one way to eliminate counter-intuitive (or unwanted) results. Often, this also makes the program easier to understand and more elaboration tolerant. In the literature on logic programming with preferences, we can find at least two distinct ways to handle preferences. The first approach is to compile the preferences into the program (e.g., [9, 5]): given a program P with a set of preferences $pref$, a new program P_{pref} is defined whose answer set semantics is used as the preferred semantics of P with respect to $pref$. The second approach deals with preferences between rules by defining a new semantics for logic programs with preferences (e.g., [4]). The advantage of the first approach is that it does not require the introduction of a new semantics — thus, answer set solvers can be used to compute the preferred semantics. The second approach, on the other hand, provides a more direct treatment of preferences.¹

Section 3 of this paper follows the first approach. We define a notion of *revision program with preferences*, which is a labeled revision program with preferences between the rules. Given a revision program with preferences, we translate it into an ordinary revision program, and we define preferred justified revisions (w.r.t. the revision program

¹ Space limitations do not allow us to provide a comprehensive list of literature in this area.

with preferences) as the justified revisions w.r.t. the corresponding program. Our treatment of preferences is similar to that in [9, 5, 1]. In section 4, we introduce different types of preferences that can be dealt with more appropriately by following the second approach.

Our work in this paper is also strongly related to dynamic logic programming (DLP) [3]. DLP is introduced as a mean to update knowledge bases that might contain generalized logic programming rules. Roughly, a DLP is an ordered list of generalized logic programs, where each represents the properties of the knowledge base at a time moment. The semantics of a DLP – taking into consideration a sequence of programs up to a time point t – specifies which rules should be applied to derive the state of the knowledge base at t . It has been shown that DLP generalizes revision programming [3].

DLP has been extended to deal with preferences [2, 1]. A DLP with preferences, or a *prioritized DLP*, is a pair (P, R) of two DLPs; P is a labeled DLP whose language does not contain the binary predicate $<$ and R is a DLP whose language contains the binary predicate $<$ and whose set of constants includes all the rule labels from both programs. Intuitively, (P, Q) represents a knowledge at different time moments – the same way a DLP does – with the exception that there are preferences between rules in (P, Q) . An atom of the form $r_1 < r_2$ represents the fact that rule r_1 is preferred to rule r_2 . The semantics of prioritized DLP makes sure that the preference order between rules is reflected in the set of consequences derivable from the knowledge base. More precisely, for two conflicting rules r_1 and r_2 , if $r_1 < r_2$ is derived, then the consequence of the rule r_1 should be preferred over the consequence of r_2 . Prioritized DLP deals with preferences using the compilation approach. In fact, the approach coincides with that of preferred answer sets for extended logic programs [4] when the DLP consists of a single program. In this sense, the prioritized DLP approach is similar to the approach described in Section 3, in which we add to a revision program a preference relation between its rules and define the semantics of a revision program with preferences following the compilation approach. The main difference between our work and prioritized DLP lies in that we consider other types of preferences (e.g., maximal number of applicable rules, weighted rules, weighted atoms, or minimal size difference) and prioritized DLP does not. We plan to investigate the use of these types of preferences in DLP in the future.

Finally, DLP is also used as the main representation language for a multi-agent architecture in [11]. In this paper, we take the first step towards this direction by using revision programming with preferences to represent and reason about beliefs of multi-agents in a coordinated environment. A detailed comparison with MINERVA is planned in the near future.

2 Preliminaries: Revision Programming

In this section we present the formal definition of revision programs with justified revision semantics and some of their properties [15, 14, 13].

Elements of some finite universe U are called *atoms*. Subsets of U are called *databases*. Expressions of the form $\mathbf{in}(a)$ or $\mathbf{out}(a)$, where a is an atom, are called *revision literals*. For a revision literal $\mathbf{in}(a)$, its *dual* is the revision literal $\mathbf{out}(a)$. Similarly, the *dual* of $\mathbf{out}(a)$ is $\mathbf{in}(a)$. The dual of a revision literal α is denoted by α^D . A set of revision literals L is *coherent* if it does not contain a pair of dual literals. For any set of

atoms $B \subseteq U$, we denote $B^c = \{\mathbf{in}(a) : a \in B\} \cup \{\mathbf{out}(a) : a \notin B\}$. A *revision rule* is an expression of one of the following two types:

$$\mathbf{in}(a) \leftarrow \mathbf{in}(a_1), \dots, \mathbf{in}(a_m), \mathbf{out}(b_1), \dots, \mathbf{out}(b_n) \quad \text{or} \quad (1)$$

$$\mathbf{out}(a) \leftarrow \mathbf{in}(a_1), \dots, \mathbf{in}(a_m), \mathbf{out}(b_1), \dots, \mathbf{out}(b_n), \quad (2)$$

where a, a_i and b_i are atoms. A *revision program* is a collection of revision rules. Revision rules have a declarative interpretation as constraints on databases. For instance, rule (1) imposes the following condition: a is *in* the database, or at least one a_i , $1 \leq i \leq m$, is *not* in the database, or at least one b_j , $1 \leq j \leq n$, is *in* the database.

Revision rules also have a computational (imperative) interpretation that expresses a way to enforce a constraint. Assume that all data items a_i , $1 \leq i \leq m$, belong to the current database, say I , and none of the data items b_j , $1 \leq j \leq n$, belongs to I . Then, to enforce the constraint (1), the item a must be added to the database (removed from it, in the case of the constraint (2)), rather than removing (adding) some item a_i (b_j).

Given a revision rule r , by $\mathit{head}(r)$ and $\mathit{body}(r)$ we denote the literal on the left hand side and the set of literals on the right hand side of the \leftarrow , respectively.

A set of atoms $B \subseteq U$ is a *model* of (or *satisfies*) a revision literal $\mathbf{in}(a)$ (resp., $\mathbf{out}(a)$), if $a \in B$ (resp., $a \notin B$). A set of atoms B is a *model* of (or *satisfies*) a revision rule r if either B is not a model of at least one revision literal from the body of r , or B is a model of $\mathit{head}(r)$. A set of atoms B is a *model* of a revision program P if B is a model of every rule in P . Let P be a revision program. The *necessary change* of P , $NC(P)$, is the least model of P , when treated as a Horn program built of independent propositional atoms of the form $\mathbf{in}(a)$ and $\mathbf{out}(b)$.

The collection of all revision literals describing the elements that do not change their status in the transition from a database I to a database R is called the *inertia set* for I and R , and is defined as follows:

$$I(I, R) = \{\mathbf{in}(a) : a \in I \cap R\} \cup \{\mathbf{out}(a) : a \notin I \cup R\}.$$

By the *reduct* of P with respect to a pair of databases (I, R) , denoted by $P_{I,R}$, we mean the revision program obtained from P by eliminating from the body of each rule in P all literals in $I(I, R)$. The necessary change of the program $P_{I,R}$ provides a justification for some insertions and deletions. These are exactly the changes that are *a posteriori* justified by P in the context of the initial database I and a putative revised database R .

Definition 1 ([15]). A database R is a P -justified revision of database I if the necessary change of $P_{I,R}$ is coherent and if

$$R = (I \setminus \{a \in U : \mathbf{out}(a) \in NC(P_{I,R})\}) \cup \{a \in U : \mathbf{in}(a) \in NC(P_{I,R})\}$$

Basic properties of justified revisions include the following [15]:

1. If a database R is a P -justified revision of I , then R is a model of P .
2. If a database B satisfies a revision program P then B is a unique P -justified revision of itself.
3. If R is a P -justified revision of I , then $R \div I$ is minimal in the family $\{B \div I : B \text{ is a model of } P\}$ —where $R \div I$ denotes the symmetric difference of R and I . In other words, justified revisions of a database differ minimally from the database.

Another important property of revision programs is that certain transformations (*shifts*) preserve justified revisions [13]. For each set $W \subseteq U$, a W -transformation is defined as follows. If α is a literal of the form $\mathbf{in}(a)$ or $\mathbf{out}(a)$, then

$$T_W(\alpha) = \begin{cases} \alpha^D, & \text{when } a \in W \\ \alpha, & \text{when } a \notin W. \end{cases}$$

Given a set L of literals, $T_W(L) = \{T_W(\alpha) : \alpha \in L\}$. Given a set A of atoms, $T_W(A) = \{a : \mathbf{in}(a) \in T_W(A^c)\}$. Given a revision program P , $T_W(P)$ is obtained from P by applying T_W to every literal in P . The Shifting theorem [13] states that for any databases I and J , database R is a P -justified revision of I if and only if $T_{I \div J}(R)$ is a $T_{I \div J}(P)$ -justified revision of J . The shifting theorem provides a practical way [13] to compute justified revisions using answer set programming engines (e.g., `smodels` [16]). It can be done by executing the following steps.

1. Given a revision program P and an initial database I , we can apply the transformation T_I to obtain the revision program $T_I(P)$ and the empty initial database.
2. $T_I(P)$ can be converted into a logic program with constraints by replacing revision rules of the type (1) by

$$a \leftarrow a_1, \dots, a_m, \text{not } b_1, \dots, \text{not } b_n \quad (3)$$

and replacing revision rules of the type (2) by constraints

$$\leftarrow a, a_1, \dots, a_m, \text{not } b_1, \dots, \text{not } b_n. \quad (4)$$

We denote the logic program with constraints obtained from a revision program Q via the above conversion by $lp(Q)$.

3. Given $lp(T_I(P))$ we can compute its answer sets.
4. Finally, the transformation T_I can be applied to the answer sets to obtain the P -justified revisions of I .

3 Revision programs with preferences

In this section, we introduce *revision programs with preferences*, that can be used to deal with preferences between rules of a revision program. We will begin with an example to motivate the introduction of preferences between revision rules. We then present the syntax and semantics and discuss some properties of revision programs with preferences.

3.1 Motivational example

Assume that we have a number of agents a_1, a_2, \dots, a_n . The environment is encoded through a set of parameters p_1, p_2, \dots, p_k . The agents perceive parameters of the environment, and provide perceived data (observations) to a controller. The observations are represented using atoms of the form: $observ(Par, Value, Agent)$, where Par is the name of the observed parameter, $Value$ is the value for the parameter, and $Agent$ is the name of the agent providing the observation.

The controller combines the data received from agents to update its view of the world, which includes exactly one value for each parameter. The views of the world are

described by atoms: $world(Par, Value, Agent)$, where $Value$ is the current value for the parameter Par , and $Agent$ is the name of the agent that provided the last accepted value for the parameter. The initial database contains a view of the world before the new observations arrive. A revision program, denoted by P , is used to update the view of the world, and is composed of rules of the type:

$$\mathbf{in}(observ(Par, Value, Agent)) \leftarrow$$

which describe all new observations; and rules of the following two types:

$$\mathbf{in}(world(Par, Value, Agent)) \leftarrow \mathbf{in}(observ(Par, Value, Agent)) \quad (\text{a})$$

$$\mathbf{out}(world(Par, Value, Agent)) \leftarrow \mathbf{in}(world(Par, Value1, Agent1)), \quad (\text{b})$$

(where $Agent \neq Agent1$ or $Value \neq Value1$).

Rules of type (a) allow to generate a new value for a parameter of a world view from a new observation. Rules of type (b) are used to enforce the fact that only one observation per parameter can be used to update the view.

It is easy to see that if the value of each parameter is perceived by *exactly one* agent and the initial world view of the controller is coherent, then each P -justified revision reflects the controller's world view that integrates its agent observations whenever they arrive. However, P does not allow any justified revisions when there are two agents which perceive different data for the same parameter at the same time. We illustrate this problem in the following scenario. Let us assume we have two agents a_1 and a_2 , both provide observations for the parameter named *temperature* denoting the temperature in the room. Initially, the controller knows that $world(temperature, 76, s_2)$. At a later time, it receives two new observations

$$\mathbf{in}(observ(temperature, 74, a_1)) \leftarrow$$

$$\mathbf{in}(observ(temperature, 72, a_2)) \leftarrow$$

There is no P -justified revision for this set of observations as the necessary change with respect to it is incoherent, it includes $\mathbf{in}(world(temperature, 74, a_1))$ (because of (a) and the first observation) and $\mathbf{out}(world(temperature, 74, a_1))$ (because of (a), (b), and the second observation).

The above situation can be resolved by placing a preference between the values provided by the agents. For example, if we know that agent a_2 has a better temperature sensor than agent a_1 , then we should tell the controller that observations of a_2 are preferred to those of a_1 . This can be described by adding preferences of the form: $\mathbf{prefer}(r_2, r_1)$, where r_1 and r_2 are names of rules of type (a) containing a_1 and a_2 , respectively. With the above preference, the controller should be able to derive a justified revision which would contain $world(temperature, 72, a_2)$. If the agent a_2 has a broken temperature sensor and does not provide temperature observations, the value of *temperature* will be determined by a_1 and the world view will be updated correctly by P .

The above preference represents a fixed order of rule's application in creating revisions. Sometimes, preferences might be dynamic. As an example, we may prefer the controller to keep using temperature observations from the same agent if available. This can be described by preferences of the form:

$$\mathbf{prefer}(r1, r2) \leftarrow world(temperature, Value, a_1) \in I, \mathbf{in}(observ(temperature, NewValue, a_1));$$

$prefer(r2, r1) \leftarrow world(temperature, Value, a_2) \in I, \mathbf{in}(observ(temperature, NewValue, a_2));$

where $r1$ and $r2$ are names of rules of type (a) containing a_1 and a_2 respectively, and I is an initial database (a view of the world before the new observations arrive).

3.2 Syntax and Semantics

A *labeled revision program* is a pair (P, \mathcal{L}) where P is a revision program and \mathcal{L} is a function which assigns to each revision rule in P a unique name (label). The label of a rule $r \in P$ is denoted $\mathcal{L}(r)$. The rule with a label l is denoted $r(l)$. We will use $head(l)$, $body(l)$ to denote $head(r(l))$ and $body(r(l))$ respectively. The set of labels of all revision rules from P is denoted $\mathcal{L}(P)$. That is, $\mathcal{L}(P) = \{\mathcal{L}(r) : r \in P\}$. For simplicity, for each rule $\alpha_0 \leftarrow \alpha_1, \dots, \alpha_n$ of P , we will write:

$$l : \alpha_0 \leftarrow \alpha_1, \dots, \alpha_n$$

to indicate that l is the value assigned to the rule by the function \mathcal{L} .

A *preference* on rules in (P, \mathcal{L}) is an expression of the following form

$$prefer(l_1, l_2) \leftarrow initially(\alpha_1, \dots, \alpha_k), \alpha_{k+1}, \dots, \alpha_n, \quad (5)$$

where l_1, l_2 are labels of rules in P , $\alpha_1 \dots, \alpha_n$ are revision literals, $k \geq 0$, $n \geq k$.

Informally, the preference (5) mean that if revision literals $\alpha_1 \dots, \alpha_k$ are satisfied by the initial database and literals $\alpha_{k+1}, \dots, \alpha_n$ are satisfied by a revision, then we prefer to use rule $r(l_1)$ over rule $r(l_2)$. More precisely, if the body of rule $r(l_1)$ is satisfied then rule $r(l_2)$ is defeated and ignored. If $body(l_1)$ is not satisfied then rule $r(l_2)$ is used.

A *revision program with preferences* is a triple (P, \mathcal{L}, S) , where (P, \mathcal{L}) is a labeled revision program and S is a set of preferences on rules in (P, \mathcal{L}) . We refer to S as the control program since it plays an important role on what rules can be used in constructing the revisions.

A revision program with preferences (P, \mathcal{L}, S) can be translated into an ordinary revision program as follows. Let $U^{\mathcal{L}(P)}$ be the universe obtained from U by adding new atoms of the form $ok(l)$, $defeated(l)$, $prefer(l_1, l_2)$ for all $l, l_1, l_2 \in \mathcal{L}(P)$. Given an initial database I , we define a new revision program $P^{S, I}$ over $U^{\mathcal{L}(P)}$ as the revision program consisting of the following revision rules:

- for each $l \in \mathcal{L}(P)$, the revision program $P^{S, I}$ contains the two rules

$$head(l) \leftarrow body(l), \mathbf{in}(ok(l)) \quad (6)$$

$$\mathbf{in}(ok(l)) \leftarrow \mathbf{out}(defeated(l)) \quad (7)$$

- for each preference $prefer(l_1, l_2) \leftarrow initially(\alpha_1, \dots, \alpha_k), \alpha_{k+1}, \dots, \alpha_n$ in S such that $\alpha_1 \dots, \alpha_k$ are satisfied by I , $P^{S, I}$ contains the rules

$$\mathbf{in}(prefer(l_1, l_2)) \leftarrow \alpha_{k+1}, \dots, \alpha_n \quad (8)$$

$$\mathbf{in}(defeated(l_2)) \leftarrow body(l_1), \mathbf{in}(prefer(l_1, l_2)) \quad (9)$$

Following the compilation approach in dealing with preferences, we define the notion of (P, \mathcal{L}, S) -justified revisions of an initial database I as follows.

Definition 2. A database R is a (P, \mathcal{L}, S) -justified revision of I if there exists $R' \subseteq U^{\mathcal{L}(P)}$ such that R' is a $P^{S,I}$ -justified revision of I , and $R = R' \cap U$.

The next example illustrates the definition of justified revisions with respect to revision programs with preferences.

Example 1. Let P be the program containing the rules

$$\begin{aligned} r_1 &: \mathbf{in}(\mathit{world}(\mathit{temperature}, 76, a_1)) \leftarrow \mathbf{in}(\mathit{observ}(\mathit{temperature}, 76, a_1)). \\ r_2 &: \mathbf{in}(\mathit{world}(\mathit{temperature}, 77, a_2)) \leftarrow \mathbf{in}(\mathit{observ}(\mathit{temperature}, 77, a_2)). \end{aligned}$$

and the set S of preferences consists of a single preference $\mathit{prefer}(r_1, r_2)$. Let $I_1 = \{\mathit{observ}(\mathit{temperature}, 76, a_1), \mathit{observ}(\mathit{temperature}, 77, a_2)\}$ be the initial database. The revision program P^{S,I_1} is the following:

$$\begin{aligned} \mathbf{in}(\mathit{world}(\mathit{temperature}, 76, a_1)) &\leftarrow \mathbf{in}(\mathit{observ}(\mathit{temperature}, 76, a_1)), \mathbf{in}(\mathit{ok}(r_1)) \\ \mathbf{in}(\mathit{world}(\mathit{temperature}, 77, a_2)) &\leftarrow \mathbf{in}(\mathit{observ}(\mathit{temperature}, 77, a_2)), \mathbf{in}(\mathit{ok}(r_2)) \\ &\quad \mathbf{in}(\mathit{ok}(r_1)) \leftarrow \mathbf{out}(\mathit{defeated}(r_1)) \\ &\quad \mathbf{in}(\mathit{ok}(r_2)) \leftarrow \mathbf{out}(\mathit{defeated}(r_2)) \\ \mathbf{in}(\mathit{prefer}(r_1, r_2)) &\leftarrow \\ \mathbf{in}(\mathit{defeated}(r_2)) &\leftarrow \mathbf{in}(\mathit{observ}(\mathit{temperature}, 76, a_1)), \\ &\quad \mathbf{in}(\mathit{prefer}(r_1, r_2)) \end{aligned}$$

Since I_1 has only one P^{S,I_1} -justified revision,

$$R_1 = \left\{ \begin{array}{l} \mathit{observ}(\mathit{temperature}, 76, a_1), \mathit{observ}(\mathit{temperature}, 77, a_2), \\ \mathit{world}(\mathit{temperature}, 76, a_1), \mathit{prefer}(r_1, r_2), \mathit{ok}(r_1), \mathit{defeated}(r_2) \end{array} \right\},$$

then I_1 has only one (P, \mathcal{L}, S) -justified revision, $\{\mathit{world}(\mathit{temperature}, 76, a_1)\}^2$.

Now, consider the case where the initial database is

$$I_2 = \{\mathit{observ}(\mathit{temperature}, 77, a_2)\}.$$

The revision program $P^{S,I_2} = P^{S,I_1}$. Since I_2 has only one P^{S,I_2} -justified revision,

$$R_2 = \left\{ \begin{array}{l} \mathit{world}(\mathit{temperature}, 77, a_2), \mathit{observ}(\mathit{temperature}, 77, a_2), \\ \mathit{prefer}(r_1, r_2), \mathit{ok}(r_1), \mathit{ok}(r_2) \end{array} \right\},$$

we can conclude that I_2 has only one (P, \mathcal{L}, S) -justified revision,

$$\{\mathit{world}(\mathit{temperature}, 77, a_2)\}^2.$$

Notice the difference in the two cases: in the first case, rule r_2 is defeated and cannot be used in generating the justified revision. In the second case both rules can be used. \square

3.3 Properties

Justified revision semantics for revision programs with preferences extends justified revision semantics for ordinary revision programs. More precisely:

Theorem 1. A database R is a $(P, \mathcal{L}, \emptyset)$ -justified revision of I if and only if R is a P -justified revision of I .

We will now investigate other properties of revision programs with preferences. Because of the presence of preferences, it is expected that not every (P, \mathcal{L}, S) -justified revision of I is a model of P . This can be seen in the next example.

² We omit the observations from the revised database.

Example 2. Let P be the program

$$r_1 : \mathbf{in}(a) \leftarrow \mathbf{out}(b) \quad r_2 : \mathbf{in}(b) \leftarrow \mathbf{out}(a)$$

and the set S consists of two preferences: $prefer(r_1, r_2)$ and $prefer(r_2, r_1)$. Then, \emptyset is (P, \mathcal{L}, S) -justified revision of \emptyset (both rules are defeated) but not a model of P . \square

The next theorem shows that for each (P, \mathcal{L}, S) -preferred justified revision R of I , the subset of rules in P that are satisfied by R , is uniquely determined. To formulate the theorem, we need some more notation. Let J be a subset of $U^{\mathcal{L}(P)}$. By $P|_J$ we denote the program consisting of the rules r in P such that

- $ok(\mathcal{L}(r)) \in J$, or
- $ok(\mathcal{L}(r)) \notin J$ and $body(r) \setminus J^c \neq \emptyset$.

Theorem 2. *For every $P^{S,I}$ -justified revision R of I , the corresponding (P, \mathcal{L}, S) -justified revision $R \cap U$ of I is a model of program $P|_R$.*

In the rest of this subsection, we discuss some properties that guarantee that each (P, \mathcal{L}, S) -justified revision of I is a model of the program P . We concentrate on conditions on the set of preferences S . Obviously, Example 2 suggests that S should not contain a cycle between rules. The next example shows that if preferences are placed on a pair of rules such that the body of one of them is satisfied when the other rule is fired, may result in revisions that are not models of the program.

Example 3. Let P be the program

$$r_1 : \mathbf{in}(a) \leftarrow \mathbf{in}(b) \quad r_2 : \mathbf{in}(d) \leftarrow \mathbf{out}(a)$$

and the set of preferences S consists of $prefer(r_2, r_1)$. Then, $\{b, d\}$ is (P, \mathcal{L}, S) -justified revision of $\{b\}$ but is not a model of P . \square

We now define precisely the conditions that guarantee that preferred justified revisions are models of the revision programs as well. First, we define when two rules are disjoint, i.e., when two rules cannot be used at the same time in creating revisions.

Definition 3. *Let (P, \mathcal{L}, S) be a revision program with preferences. Two rules r, r' of P are disjoint if one of the following conditions is satisfied:*

1. $(head(r))^D \in body(r')$ and $(head(r'))^D \in body(r)$; or
2. $body(r) \cup body(r')$ is incoherent.

We say that a set of preferences is *selecting* if it contains only preferences between disjoint rules.

Definition 4. *Let (P, \mathcal{L}, S) be a revision program with preferences. S is said to be a set of selecting preferences if for every rule*

$$prefer(r, r') \leftarrow l_1, \dots, l_k$$

in S , rules r and r' are disjoint.

Finally, we say that a set of preferences is cycle-free if the transitive closure of the preference relation *prefer* does not contain a cycle.

Definition 5. Let (P, \mathcal{L}, S) be a revision program with preferences and $<_S = \{(r_1, r_2) \mid \text{prefer}(r_1, r_2) \text{ occurs in } S\}$. S is said to be cycle-free if for every rule r of P , (r, r) does not belong to the transitive closure $<_S^*$ of $<_S$.

The next theorem shows that the conditions on the set of preferences S guarantee that preferred justified revisions are models of the original revision program.

Theorem 3. Let (P, \mathcal{L}, S) be a revision program with preferences where S is a set of selecting preferences and is cycle-free. For every (P, \mathcal{L}, S) -justified revision R of I , R is a model of P .

The next theorem discusses the shifting property of revision programs with preferences. We extend the definition of *W-transformation* to a set of preferences on rules. Given a preference on rules p of the form (5), its *W-transformation* is the preference

$$T_W(p) = \text{prefer}(l_1, l_2) \leftarrow \text{initially}(T_W(\alpha_1), \dots, T_W(\alpha_k)), T_W(\alpha_{k+1}), \dots, T_W(\alpha_n).$$

Given a set of preferences S , its *W-transformation* is $T_W(S) = \{T_W(p) : p \in S\}$.

Theorem 4. Let (P, \mathcal{L}, S) be a revision program with preferences. For every two databases I_1 and I_2 , a database R_1 is a (P, \mathcal{L}, S) -justified revision of I_1 if and only if $T_{I_1 \div I_2}(R_1)$ is a $(T_{I_1 \div I_2}(P), \mathcal{L}, T_{I_1 \div I_2}(S))$ -justified revision of I_2 .

4 Soft revision rules with weights

Preferences between rules (Section 3) can be useful in at least two ways. They can be used to recover from incoherency when agents provide inconsistent data, as in the example from Section 3.1. They can also be used to eliminate some revisions. The next example shows that in some situations, this type of preferences is rather weak.

Example 4. Consider again the example from Section 3.1, with two agents a_1 and a_2 whose observations are used to determine the value of the parameter *temperature*. Let us assume now that a_1 and a_2 are of the same quality, i.e., *temperature* can be updated by one of the observations yielded by a_1 and a_2 . This means that there is no preference between the rule of type (a) (for a_1) and the rule of type (a) (for a_2) and vice versa. Yet, as we can see, allowing both rules to be used in computing the revisions will not allow the controller to update its world view when the observations are inconsistent. \square

The problem in the above example could be resolved by grouping the rules of the type (a) into a set and allowing only one rule from this set to be used in creating revisions if the presence of all the rules does not allow justified revisions.

Inspired by the research in constraint programming, we propose to address the situation when there are no justified revisions by dividing a revision program P in two parts, HR and SR , i.e., $P = HR \cup SR$. Rules from HR and SR are called *hard rules* and *soft rules*, respectively. The intuition is that rules in HR must be satisfied by each revision, while revisions may satisfy only a subset of SR if it is impossible to satisfy all of them. The subset of soft rules that is satisfied, say S , should be optimal with respect to some comparison criteria. In this section, we investigate several criteria—each one is discussed in a separate subsection.

4.1 Maximal number of rules

Let $P = HR \cup SR$. Our goal is to find revisions that satisfy all rules from HR and the most number of rules from SR . Example 4 motivates the search for this type of revisions. In the next definition, we make this precise.

Definition 6. R is a (HR, SR) -preferred justified revision of I if R is a $(HR \cup S)$ -justified revision of I for some $S \subseteq SR$, and for all $S' \subseteq SR$ such that S' has more rules than S , there are no $(HR \cup S')$ -justified revisions of I .

Preferred justified revision can be computed, under the maximal number of rules criteria, by extending the translation of revision programs to answer set programming, to handle the distinction between hard and soft rules. The objective is to determine $(HR \cup S)$ -justified revisions of an initial database I , where S is a subset of SR of maximal size such that $(HR \cup S)$ -justified revisions exist.

The idea is to make use of two language extensions proposed by the `smodels` system: choice rules and `maximize` statements. Intuitively, each soft rule can be either accepted or rejected in the program used to determine revisions. Let us assume that the rules in $T_I(SR)$ have been uniquely numbered. For each initial database I , we translate $P = HR \cup SR$ into an `smodels` program $lp(T_I(HR)) \cup lp'(T_I(SR))$ where $lp'(T_I(SR))$ is defined as follows. If the rule number i in $T_I(SR)$ is

$$\mathbf{in}(a) \leftarrow \mathbf{in}(p_1), \dots, \mathbf{in}(p_m), \mathbf{out}(s_1), \dots, \mathbf{out}(s_n)$$

then the following rules are added to $lp'(T_I(SR))$

$$\begin{aligned} \{rule_i\} &: - p_1, \dots, p_m, not\ s_1, \dots, not\ s_n. \\ a &: - rule_i \end{aligned}$$

where $rule_i$ is a distinct new atom. Similarly, if

$$\mathbf{out}(a) \leftarrow \mathbf{in}(p_1), \dots, \mathbf{in}(p_m), \mathbf{out}(s_1), \dots, \mathbf{out}(s_n)$$

is the rule number i in $T_I(SR)$, then the following rules are added to $lp'(T_I(SR))$

$$\begin{aligned} \{rule_i\} &: - p_1, \dots, p_m, not\ s_1, \dots, not\ s_n. \\ &: - rule_i, a. \end{aligned}$$

where $rule_i$ is a distinct new atom. Finally, we need to enforce the fact that we desire to maximize the number of SR rules that are satisfied. This corresponds to maximizing the number of $rule_i$ that are true in the computed answer sets. This can be directly expressed by the following statement:

$$\text{maximize}\{rule_1, \dots, rule_k\}. \quad (10)$$

where k is the number of rules in SR . The way how `smodels` system processes `maximize` statement is as follows. It first searches a single model and prints it. After that, `smodels` prints only "better" models. The last model that `smodels` prints will correspond to a (HR, SR) -preferred justified revision of I . Notice that this is the only occurrence of `maximize` in the translation which is a requirement for the correct handling of this construct in `smodels`.

4.2 Maximal subset of rules

A variation of definition from Section 4.1 can be obtained when instead of satisfying maximal number of soft rules it is desired to satisfy a maximal subset of soft rules. In other words, given $P = HR \cup SR$, the goal is to find revisions that satisfy all rules from HR and a maximal subset (with respect to set inclusion) of rules from SR . The precise definition follows.

Definition 7. R is a (HR, SR) -preferred $^{\subseteq}$ justified revision of I if R is a $(HR \cup S)$ -justified revision of I for some $S \subseteq SR$, and for all S' if $S \subset S' \subseteq SR$, then there are no $(HR \cup S')$ -justified revisions of I .

The procedure described in Section 4.1 allows to compute only one of (HR, SR) -preferred $^{\subseteq}$ justified revisions which has maximal number of soft rules satisfied.

4.3 Weights

An alternative to the maximal subset of soft rules is to assign weights to the revisions and then select those with the maximal (or minimal) weight. In this section, we consider two different ways of assigning weights to revisions. First, we assign weight to rule. Next, we assign weight to atoms. In both cases, the goal is to find a subset S of SR such that the program $HR \cup S$ has revisions whose weight is maximal.

Weighted rules. Each rule r in SR is assigned a weight (a number), $w(r)$. Intuitively, $w(r)$ represents the importance of r , i.e., the more the weight of a rule the more important it is to satisfy it.

Example 5. Let us reconsider the example from Section 3.1. Rules of the type (a) are treated as soft rules, while the rules of type (b) are treated as hard rules. We can make use of rule weights to select desired revisions. For example, if an observed parameter value falls outside expected value range for the parameter, it may suggest that an agent that provided the observation has a faulty sensor. Thus, we may prefer observations that are closer to the expected range. This can be expressed by associating to each rule r of the type (a) the weight

$$w(r) = \min\{0, MaxEV - Value\} + \min\{0, Value - MinEV\},$$

where $MaxEV$ and $MinEV$ are maximum and minimum expected values for Par . \square

Let us define the rule-weighted justified revision of a program with weights for rules.

Definition 8. R is called a rule-weighted (HR, SR) -justified revision of I if the following two conditions are satisfied:

1. there exists a set of rules $S \subseteq SR$ s.t. R is a $(HR \cup S)$ -justified revision of I , and
2. for any set of rules $S' \subseteq SR$, if R' is a $(HR \cup S')$ -justified revision of I , then the sum of weights of rules in S' is less or equal than the sum of weights of rules in S .

Let us generalize the implementation in the previous sections to consider weighted rules. The underlying principle is similar, with the difference that the selection of soft rules to include is driven by the goal of maximizing the total weight of the soft rules that are satisfied by the justified revision. The only change we need to introduce w.r.t.

the implementation is in the `maximize` statement. Let us assume that $w(i)$ denotes the weight associated to the i th SR rule. Then, instead of the rule (10) the following `maximize` statement is generated:

$$\text{maximize}[rule_1 = w(1), rule_2 = w(2), \dots, rule_k = w(k)].$$

Weighted atoms. Instead of assigning weights to rules, we can also assign weights to atoms in the universe U . Each atom a in the universe U is assigned a weight $w(a)$ which represented the degree we would like to keep it unchanged, i.e., the more the weight of an atom the less we want to change its status in a database. The next example presents a situation where this type of preferences is desirable.

Example 6. Let us return to the example from Section 3.1 with the same partition of rules in hard and soft rules as in Example 5. Let us assume that the choice of which observation to use to update the view of the world is based on the principle that stronger values for the parameters are preferable, as they denote a stronger signal. This can be encoded by associating weights of the form

$$w(\text{world}(\text{Param}, \text{Value}, \text{Sensor})) = -\text{Value}$$

and minimizing the total weight of the revision. \square

Let us define preferred justified revision for programs with weight atoms.

Definition 9. R is called an *atom-weighted* (HR, SR) -justified revision of I if the following two conditions are satisfied:

1. there exists a set of rules $S \subseteq SR$ s.t. R is a $(HR \cup S)$ -justified revision of I , and
2. for any set of rules $S' \subseteq SR$, if Q is a $(HR \cup S')$ -justified revision of I , then the sum of weights of atoms in $I \div Q$ is greater than or equal to the sum of weights of atoms in $I \div R$.

Atom-weighted revisions can be computed using `smodels`. In this case the selection of the SR rules to be included is indirectly driven by the goal of minimizing the total weight of the atoms in $I \div R$, if I is the initial database and R is the justified revision. We make use of the following observation: given a revision program P and an initial database I , if $T_I(R)$ is a $T_I(P)$ -justified revision of \emptyset , then $T_I(R) = I \div R$. Thanks to this observation, the computation of the weight of the difference between the original database and a P -justified revision can be computed by determining the *total* weight of the true atoms obtained from the answer set generated for the $T_I(P)$ program.

The program used to compute preferred revisions is encoded similarly to what is done for the case of maximal number of rules or weighted rules, i.e., each soft rule is encoded using `smodels`'s choice rules. The only difference is that instead of making use of a `maximize` statement, we make use of a `minimize` statement of the form:

$$\text{minimize}[a_1 = w(a_1), a_2 = w(a_2), \dots, a_n = w(a_n)] \quad (11)$$

where a_1, \dots, a_n are all the atoms in U .

4.4 Minimal size difference

In this subsection, we consider justified revisions that have minimal size difference with initial database. The next example shows that this is desirable in different situations.

Example 7. Assume that a department needs to form a committee to work on some problem. Each of the department faculty members has his or her own conditions on the committee members which need to be satisfied. The head of the department provided an initial proposal for members of the committee. The task is to form a committee which will satisfy all conditions imposed by the faculty members and will differ the least from the initial proposal — the size of the symmetric difference between the initial proposal and its revision is minimal.

In this problem we have a set of agents (faculty members) each of which provides its set of requirements (revision rules). The goal is to satisfy all agent's requirements in such a way that the least number of changes is made to the initial database (proposal).

Assume that faculty members are Ann, Bob, Chris, David, Emily and Frank. Conditions that they impose on the committee are the following:

$$\begin{aligned}
 \text{Ann} : & \quad \mathbf{in}(\text{Bob}) \leftarrow \mathbf{out}(\text{Chris}) \\
 & \quad \mathbf{in}(\text{Chris}) \leftarrow \mathbf{out}(\text{Bob}) \\
 \text{Bob} : & \quad \mathbf{out}(\text{David}) \leftarrow \mathbf{in}(\text{Bob}) \\
 \text{Chris} : & \quad \mathbf{out}(\text{Ann}) \leftarrow \mathbf{out}(\text{David}) \\
 \text{David} : & \quad \mathbf{in}(\text{David}) \leftarrow \mathbf{in}(\text{Chris}), \mathbf{out}(\text{Ann})
 \end{aligned}$$

The initial proposal is $I = \{\text{Ann}, \text{David}\}$. Then, there is one minimal size difference P -justified revisions of I , which is $R_1 = \{\text{Ann}, \text{David}, \text{Bob}\}$. The size of the difference $R_1 \div I$ is 1.

Ordinary P -justified revisions of I also include $R_2 = \{\text{Bob}\}$ with size of the difference $R_2 \div I$ equal to 3. \square

The next definition captures what is a minimal size different justified revision.

Definition 10. R is called a minimal size difference P -justified revision of I if the following two conditions are satisfied:

1. R is a P -justified revision of I , and
2. for any P -justified revision R' , the number of atoms in $R \div I$ is less than or equal to the number of atoms in $R' \div I$.

Minimal size difference justified revision can be computed in almost the same way as for atom-weighted justified revisions. The intuition is that instead of minimizing the total weight of $I \div R$ (where I is the initial database and R is a P -justified revision), we would like to minimize the size of $I \div R$. This can be accomplished by replacing the minimize statement (11) with the following minimize statement:

$$\text{minimize}\{a_1, a_2, \dots, a_n\}$$

where a_1, \dots, a_n are all the atoms in U .

5 Conclusions

The notion of preference has found pervasive applications in the context of knowledge representation and commonsense reasoning in MAS. Indeed a large number of approaches have been proposed to improve the knowledge representation capabilities of logic programming by introducing different forms of preferences. In this paper, we

presented a novel extension of the revision programming framework which provides the foundations for expressing very general types of preferences. Preferences provide the ability to “defeat” the use of certain revision rules in the computation of the revisions; this allows us to either reduce the number of revisions generated (eventually leading to a single revision), or to generate revisions even in presence of conflicting revision rules.

We proposed different preference schemes, starting from a relatively dynamic partial order between revision rules (*revision programs with preferences*), and then moving to a more general notion of weights, associated to revision rules and/or database atoms. Soft revision rules can be dynamically included or excluded from the generation of revisions depending on optimization criteria based on the weights of the revision (e.g., minimization of the total weight associated to the revision). We provided motivating examples for the different preference schemes, along with a precise description of how preferred revisions can be computed using the `smodels` answer set inference engine.

References

1. J.J. Alferes, P. Dell’Acqua, and L.M. Pereira. A compilation of updates plus preferences. In *Logics in Artificial Intelligence, European Conference*, pages 62–73. Springer, 2002.
2. J.J. Alferes and L.M. Pereira. Updates plus preferences. In *Logics in Artificial Intelligence, European Workshop (JELIA)*, pages 345–360. Springer, 2000.
3. J.J. Alferes et al. Dynamic Updates of Non-monotonic Knowledge Bases. *JLP*, 45, 2000.
4. G. Brewka and T. Eiter. Preferred answer sets for extended logic programs. *Artificial Intelligence*, 109(1–2):297–356, 1999.
5. E. Delgrande, T. Schaub, and H. Tompits. A framework for compiling preferences in logic programs. *Theory and Practice of Logic Programming*, 3(2):129–187, March 2003.
6. H.E. Durfee. *Coordination of Distributed Problem Solvers*. Kluwer Academic Press, 1988.
7. T. Eiter, M. Fink, G. Sabbatini, and H. Tompits. Using Methods of Declarative Logic Programming for Intelligent Information Agents. *TPLP*, 2(6), 2002.
8. P. Faratin and B. Van de Walle. Agent Preference Relations: Strict, Indifferent, and Incomparable. In *AAMAS*. ACM, 2002.
9. M. Gelfond and T.C. Son. Prioritized default theory. In *Selected Papers from the Workshop on Logic Programming and Knowledge Representation 1997*, 164–223. LNAI 1471, 1998.
10. P. La Mura and Y. Shoham. Conditional, Hierarchical, Multi-agent Preferences. In *TARK VII*, 1998.
11. J.A. Leite, J.J. Alferes, and L.M. Pereira. MINERVA: a Dynamic Logic Programming Agent Architecture. In *Intelligent Agents VIII*. Springer Verlag, 2002.
12. J. Liu and Y. Ye. *E-Commerce Agents*. Lecture Notes in AI, Springer Verlag, 2001.
13. W. Marek, I. Pivkina, and M. Truszczyński. Revision programming = logic programming + integrity constraints. In *Computer Science Logic*, Springer Verlag, 1999.
14. W. Marek and M. Truszczyński. Revision programming, database updates and integrity constraints. In *ICDT*, pages 368–382. Springer Verlag, 1995.
15. W. Marek and M. Truszczyński. Revision programming. *Theoretical Computer Science*, 190(2):241–277, 1998.
16. I. Niemelä and P. Simons. Efficient implementation of the well-founded and stable model semantics. In *LPNMR*, pages 289–303. MIT Press, 1996.
17. H. Prakken. *Logical Tools for Modeling Legal Arguments*. Kluwer Publishers, 1997.
18. T. Przymusiński and H. Turner. Update by means of Inference rules. In *LPNMR*, 1995.
19. C. Sakama and K. Inoue. Updating Extended Logic Programs through Abduction. In *LPNMR*. Springer Verlag, 1999.