Artificial Intelligence 175 (2011) 1815–1855

Contents lists available at ScienceDirect

Artificial Intelligence

www.elsevier.com/locate/artint

Solving conflicts in information merging by a flexible interpretation of atomic propositions

Steven Schockaert [a,*,1], Henri Prade [b]

[a] Ghent University, Department of Applied Mathematics and Computer Science, Krijgslaan 281, 9000 Gent, Belgium
[b] Toulouse University, Université Paul Sabatier, IRIT, CNRS, 118 Route de Narbonne, 31062 Toulouse Cedex 09, France

ARTICLE INFO

Article history:
Received 20 July 2010
Received in revised form 21 April 2011
Accepted 23 April 2011
Available online 28 April 2011

Keywords:
Information fusion
Similarity-based reasoning
Prioritized knowledge bases
Possibilistic logic
Penalty logic

ABSTRACT

Although many techniques for merging conflicting propositional knowledge bases have already been proposed, most existing work is based on the idea that inconsistency results from the presence of incorrect pieces of information, which should be identified and removed. In contrast, we take the view in this paper that conflicts are often caused by statements that are inaccurate rather than completely false, suggesting to restore consistency by interpreting certain statements in a flexible way, rather than ignoring them completely. In accordance with this view, we propose a novel approach to merging which exploits extra-logical background information about the semantic relatedness of atomic propositions. Several merging operators are presented, which are based on different formalizations of this background knowledge, ranging from purely qualitative approaches, related to possibilistic logic, to quantitative approaches with a probabilistic flavor. Both syntactic and semantic characterizations are provided for each merging operator, and the computational complexity is analyzed.

1. Introduction

In applications where information from different sources needs to be combined, conflicts are often the rule rather than the exception. The presence of conflicts requires special attention, as it casts doubt on the reliability of available information. Even worse, if information is encoded in classical logic, the combined pieces of information become trivial, as anything can be derived from contradiction. To accommodate the possibility of conflicts in a more useful way, a wide array of approaches has been proposed in the literature, ranging from purely syntactic approaches to semantic operators that manipulate sets of interpretations. In general, the problem of information merging has been studied both in logical and in numerical settings. An example of the latter case are situations where different probability or possibility distributions need to be fused [1,2]. This paper focuses exclusively on merging in a logical setting.

A common idea underlying many approaches to logical information merging is to get rid of the least reliable pieces of information. The exact mechanism being employed may, among others, be based on prior knowledge about the reliability of different pieces of information and of sources [3,4], on discriminating between pieces of information according to the number of supporting sources [5], or on the dialectical principles of argument and counter-argument [6]. The essential point of view underlying such approaches is that conflicts are caused by errors that are in some sense arbitrary: any piece of information has some chance of being wrong, and agreement between sources (or prior knowledge) is all that can help us to decide which pieces are more likely to be correct. A closely related idea is to weaken the information that is provided

* Corresponding author.
E-mail addresses: steven.schockaert@ugent.be (S. Schockaert), prade@irit.fr (H. Prade).
[1] Postdoctoral fellow of the Research Foundation – Flanders (FWO).

0004-3702/$ – see front matter
doi:10.1016/j.artint.2011.04.001

by each of the sources [7–10]. For instance, if some source claims $p \wedge q$, we may change this, among others to $p \vee q$. This may be motivated in several ways. If the sources express conflicting goals or preferences, for instance, we may consider that each competing source needs to concede to arrive at a feasible global strategy. Alternatively, when sources express beliefs, we may consider that weaker information is more reliable, and that by progressively weakening these beliefs we ultimately end up with beliefs that are correct, and thus consistent. Without extra-logical information, however, approaches based on this idea tend to be rather coarse, often depending on the assumption that all atoms in the language encode properties that are approximately of equal importance in the domain being modeled.

A quite different approach to dealing with conflicts is to relax the assumption that sources need to be combined conjunctively. Indeed, if individual sources are consistent, combining them disjunctively trivially restores consistency. Starting from this basic idea, more refined techniques have been developed, e.g. based on disjunctively combining conjunctions of maximal consistent subsets of knowledge bases [11,12]. Paraconsistent logics [13] offer yet another solution to the problem of conflict. Rather than trying to modify the knowledge bases, the logic itself is changed such that only non-trivial conclusions can be derived, even in the face of logical contradiction. The fact that both p and $\neg p$ may be entailed by a non-trivial theory may, in some paraconsistent logics (e.g. the logic of formal inconsistency [14]), be interpreted as evidence that the property modeled by p is controversial (or ill-defined, vague, etc.) in which case it is natural that different sources may have a different standpoint regarding p. A consequence of this methodology is that, contrarily to most other methods, there is no real loss of information when conflicts arise. Indeed, when p is asserted by some source and $\neg p$ by another source, we do not give up our belief in p but rather gain the insight that p is controversial.

In this paper, we are exploring a new direction, which is motivated by the fact that in real-world applications, random errors occur side-by-side with conflicts that are due to the use of properties that may be understood differently by different sources (e.g. vague properties such as 'tall'). The method being used to deal with conflicts is then not only determined by the nature of conflicts, but also — and perhaps even especially — by the nature of available background knowledge. Logics of formal inconsistency, for instance, require that some atoms are designated as uncontroversial, to ensure meaningful results. In general, most methods assume no, or very little prior knowledge, making them widely applicable, but at the same time limiting their ability to correctly identify the real cause of conflicts, and thus, ultimately, their usefulness. A similar consideration applies to the problem of belief revision, for which it is well known that extra-logical information about the epistemic state of an agent (e.g. in the form of an epistemic entrenchment ordering) is key to meaningful results [15]. For the task of merging the beliefs expressed by different sources, however, the use and importance of extra-logical information is less well-understood. Nonetheless, there are many situations where appropriate background knowledge is paramount in correctly dealing with conflicts.

Example 1. Consider a situation where predictions are available about tomorrow's weather from three well-reputed sources. Predictions from all sources are expressed in a propositional language over the set of atoms {*overcast*, *partiallyCloudy*, *openSky*}, subject to the integrity constraint that *overcast*, *partially cloudy*, and *open sky* are Jointly Exhaustive and Pairwise Disjoint atoms (a property called JEPD in the following):

$$K_1 = \{\mathit{partiallyCloudy} \vee \mathit{overcast}\} \tag{1}$$

$$K_2 = \{\mathit{openSky}\} \tag{2}$$

$$K_3 = \{\mathit{overcast}\} \tag{3}$$

If all three sources are considered equally reliable, classical merging strategies would either yield the trivial result $\mathit{overcast} \vee \mathit{partiallyCloudy} \vee \mathit{openSky}$, or conclude *overcast*, which is the only atom that is compatible with the majority of the sources. Since all three sources are well-reputed, however, the extreme conclusions *openSky* and *overcast* seem less plausible than the intermediate conclusion *partiallyCloudy*. However, without additional information, encoding this idea of being intermediate, there are no reasons to prefer *partiallyCloudy* over *overcast* or *openSky*.

The core idea in this example is that atoms such as *overcast* should be understood in a flexible way. The need for flexibility regarding the meaning of atoms may stem from different causes, including all of the following:

1. Sources are overconfident, and the assertions they make are too precise, given the knowledge they actually have. In the example above, it is clear that people prefer more informative weather reports (e.g. it will be sunny tomorrow), with the risk of being slightly wrong from time to time, over completely honest but less informative or uninformative reports (e.g. it may be sunny or cloudy).
2. Atoms refer to properties for which precise, generally accepted definitions are lacking. Typically these are properties that depend on threshold values in some continuous domain, or properties whose definition may slightly vary with the context. For instance, there may be situations that are described as an open sky by some people and partially cloudy by others.
3. Atoms refer to terms with a well-understood meaning, which are nonetheless used in a more liberal, or more restrictive way in certain contexts. It is not hard to imagine, for example, a person in a civil union answering affirmatively to the question "Are you married (yes/no)?", e.g. when filling in a web form. As another example, the term Asian is often

reserved to refer to people from South Asia in the UK, or the people from East Asia in the USA, thus restricting the actual meaning of the term.

4. Atoms refer to ambiguous terms, which may mean completely different things. The term "public school", for instance, is used with an entirely different meaning in the UK and USA; also, different people, places, or events are often described by the same name; etc.

The underlying idea common to the aforementioned causes of conflict is that there are some countermodels of a statement that are in some sense similar to models of the statement. Given that the statement is asserted by a reliable source, we may then consider that normally, the actual state of affairs is described by one of the models of that statement, although exceptionally, it might also be described by one of these similar countermodels. The aim of this paper is to formalize this intuition, leading to merging operators for propositional knowledge bases that effectively take available background information about the relatedness of atoms into account.

The paper is structured as follows. In the next section, we provide the required background on a number of well-known approaches to information merging in the propositional setting: distance-based merging, conflict-based merging and morpho-logical merging. Next, Section 3 introduces the basic ingredient of our approach, i.e. background information about the semantic relatedness of atoms. After outlining the main idea, it is shown how relatedness can be represented at the syntactic level and at the semantic level, and what the correspondences are between both representations. What results is a general framework, in which different types of relatedness can be expressed. To cope with this generality, four prototypical scenarios are presented in Section 4. By showing how the general framework can be instantiated to implement each of these scenarios, we bridge the gap between the general, but abstract framework and practical applications. In Section 5 we subsequently turn to the merging problem itself. We show how different merging operators naturally arise from different interpretations of the weighted knowledge bases that encode how different atoms are related. After defining the merging operators at the syntactic level, we provide semantic characterizations that reveal close links with existing merging operators. The computational complexity of the merging operators is studied in Section 6, after which we present our conclusions.

2. Background on propositional merging

In the following, we consider a propositional language built from a finite set of atoms A and the connectives $\vee, \wedge, \rightarrow, \equiv, \neg$ in the usual way. An interpretation I is defined as a subset of atoms a in A, where $I \models a$ for an atom a iff $a \in I$. An interpretation is said to be a model of a formula (resp. set of formulas) if it satisfies that formula (resp. every formula in the set) in the usual sense. We write $[\![\phi]\!]$ to denote the set of all models of a formula ϕ.

We will also need a few order-theoretic concepts. Given a relation $\leqslant$ in a universe U, we write $\min(U, \leqslant)$ for the set of elements from U that are minimal w.r.t. $\leqslant$, i.e.

$$\min(U, \leqslant) = \{u \in U \mid \neg\exists v \in U \,.\, v \leqslant u \wedge u \nleqslant v\}$$

Given a list of relations $\leqslant_1, \ldots, \leqslant_n$ in U, we write $par(\leqslant_1, \ldots, \leqslant_n)$ and $lex(\leqslant_1, \ldots, \leqslant_n)$ to denote their Pareto and lexicographic extensions respectively, i.e.

$$(u, v) \in par(\leqslant_1, \ldots, \leqslant_n) \quad \text{iff} \quad \forall i \in \{1, \ldots, n\} \,.\, u \leqslant_i v$$

and $(u, v) \in lex(\leqslant_1, \ldots, \leqslant_n)$ iff either $(u, v) \in par(\leqslant_1, \ldots, \leqslant_n)$ or

$$\exists k \in \{1, \ldots, n\} \,.\, \big(\forall i \in \{1, \ldots, k-1\} \,.\, u \leqslant_i v\big) \wedge (u <_k v)$$

where $u <_k v$ is used as a shorthand for $u \leqslant_k v \wedge v \nleqslant_k u$. Note that in the case of the lexicographic ordering, the numbering i of the relations $\leqslant_i$ encodes an ordering of their importance.

Let $K_1, \ldots, K_n$ be propositional knowledge bases that are individually consistent, and let C be a set of integrity constraints. The purpose of a merging process is to find one knowledge base $\Delta(K_1, \ldots, K_n)$ which is consistent with the integrity constraints (i.e. $[\![\Delta(K_1, \ldots, K_n)]\!] \subseteq [\![C]\!]$), and which integrates the information from the knowledge bases $K_1, \ldots, K_n$ to the best extent possible.

A variety of methods to obtain a suitable knowledge base $\Delta(K_1, \ldots, K_n)$ have already been studied. For instance, there is a long tradition in inconsistency management to restore consistency by identifying maximal consistent subsets [16]. In particular, we may define $\Delta(K_1, \ldots, K_n)$ to be the disjunction of (a subset of) these maximal consistent subsets (where a set of formulas is treated as a conjunction). Such approaches, however, mainly attempt to isolate the inconsistency, rather than reconciling any conflicting views that different sources may hold. Our approach will therefore be based on a different line of work, which tries to resolve inconsistencies using the idea that the models of $\Delta(K_1, \ldots, K_n)$ should be those interpretations that are in some sense close to the models of the knowledge bases $K_1, \ldots, K_n$. Next, we briefly recall three important classes of existing merging operators that are based on this idea, and will play a role in the remainder of this paper.

2.1. Distance-based merging

A common strategy is to define merging operators based on a pseudo-metric d on interpretations [8,5,17]. Although d is not required to satisfy the mathematical properties of a metric (e.g. triangle inequality or even symmetry), it is common to refer to d as a distance. Given two interpretations I and J, the well-known Hamming distance d_{Ham} is often used to this end:

$$d_{Ham}(I,J) = \big|(I \setminus J) \cup (J \setminus I)\big| = |I \setminus J| + |J \setminus I|$$

i.e. $d_{Ham}(I,J)$ is equal to the number of atoms on which interpretations I and J disagree. The Hamming distance is sometimes also called Dalal distance, as Dalal first proposed to use it in the context of belief revision [18]. A given distance d allows us to quantify the distance between an interpretation I and a propositional knowledge base K:

$$d(I,K) = \min_{J \in [\![K]\!]} d(I,J)$$

Given an appropriate aggregation operation f, this can be extended to lists of knowledge bases $\mathcal{K} = (K_1, \ldots, K_n)$:

$$d(I,\mathcal{K}) = f\big(d(I,K_1), \ldots, d(I,K_n)\big) \tag{4}$$

where f may, for instance, be the sum, a weighted sum, or (refinements of) the maximum. Now, a preorder $\leqslant_d$ between interpretations can be defined as

$$I \leqslant_d I' \quad \text{iff} \quad d(I,\mathcal{K}) \leqslant d\big(I',\mathcal{K}\big)$$

and the result of the merging process can semantically be defined as the knowledge base whose models are those interpretations that are minimal w.r.t. this preorder:

$$\big[\!\big[\Delta^{dist}(K_1, \ldots, K_n; C; \leqslant_d)\big]\!\big] = \min\big([\![C]\!], \leqslant_d\big) \tag{5}$$

Although this class of merging operators is quite general and has a strong intuitive appeal, in practical applications it is almost exclusively applied for $d = d_{Ham}$. An important limitation of d_{Ham} and related distances is that they are very sensitive to the particular translation of a given problem into propositional logic [19]. In addition, such distances do not allow us to express that two atoms have a meaning which is in some sense related.

2.2. Conflict-based merging

The motivating idea behind conflict-based merging [20] is that many merging operators derive from manipulating the conflict set $diff(I,J)$ between two interpretations I and J in the following sense:

$$diff(I,J) = (I \setminus J) \cup (J \setminus I) \tag{6}$$

The Hamming distance $d_{Ham}(I,J)$, for example, is equal to the cardinality of this conflict set. By comparing conflict sets in other ways than by their cardinality, other merging operators may be obtained. Similar to the distance-based merging scheme, from conflict sets between interpretations, we may define the conflict between an interpretation and a knowledge base [20]:

$$diff(I,K_i) = \min\big(\big\{diff(I,J) \bigm| J \in [\![K_i]\!]\big\}, \subseteq\big)$$

Note that since conflict sets may be incomparable with each other (w.r.t. set inclusion), $diff(I,K_i)$ may contain multiple conflict sets c_i. The conflict between an interpretation and a list of knowledge bases $\mathcal{K} = (K_1, \ldots, K_n)$ is then represented by a set of conflict vectors, i.e. vectors of conflict sets [20]:

$$diff(I,\mathcal{K}) = \big\{\langle c_1, \ldots, c_n\rangle \bigm| c_i \in diff(I,K_i)\big\}$$

The challenge now arises to define, from such conflict vectors, the models of the knowledge base that results from merging $K_1, \ldots, K_n$ under the integrity constraints C. First, we need a relation $\leqslant_{confl}$ comparing conflict vectors. Although many alternatives can be conceived, we will only consider the case where $\leqslant_{confl}$ is $par(\subseteq, \ldots, \subseteq)$ in this paper. Note that this choice essentially corresponds to a qualitative counterpart of the Hamming distance. Given a choice of $\leqslant_{confl}$, we may consider the relation $\leqslant^E_{confl}$ between interpretations, defined by $I \leqslant^E_{confl} J$ iff

$$\exists \mathbf{c} \in diff(I,\mathcal{K}) \,.\, \forall \mathbf{c}' \in diff(J,\mathcal{K}) \,.\, \mathbf{c} \leqslant_{confl} \mathbf{c}'$$

It is then proposed in [20] to use the following merging operator:

$$\big[\!\big[\Delta^{confl_1}(K_1, \ldots, K_n; C; \leqslant_{confl})\big]\!\big] = \min\big([\![C]\!], \leqslant^E_{confl}\big) \tag{7}$$

following the same intuition as the distance-based merging scheme: the models of the resulting knowledge base are those models of the integrity constraints that are closest to each of the given knowledge bases. The main difference is that the notion of *closest* is now seen in a more general perspective.

When applied to $\leqslant_{confl} = par(\subseteq, \ldots, \subseteq)$, however, this merging operator appears to be too tolerant, allowing more models than is intuitively required.

Example 2. Consider the situation of merging the single knowledge base K_1, whose only models are $J_1 = \{c, d, e\}$ and $J_2 = \{a, b, c\}$, with the integrity constraints C, whose models are $I_1 = \{b, c, d\}$, $I_2 = \{c, d\}$ and $I_3 = \{b, c\}$. The relevant conflict sets are given by

$$\begin{aligned} &diff(I_1, J_1) = \{b, e\}, && diff(I_1, J_2) = \{a, d\} \\ &diff(I_2, J_1) = \{e\}, && diff(I_2, J_2) = \{a, b, d\} \\ &diff(I_3, J_1) = \{b, d, e\}, && diff(I_3, J_2) = \{a\} \end{aligned}$$

Among all these conflict sets, only $\{e\}$ and $\{a\}$ are optimal w.r.t. $\subseteq$. This means that, of all models of C, I_1 is neither the closest to J_1 nor the closest to J_2. On the other hand I_2 is the closest to J_1 and I_3 is the closest to J_2. Therefore, it seems appropriate to take $\{I_2, I_3\}$ as models of the merged knowledge base. Using (7), on the other hand, leads to $\{I_1, I_2, I_3\}$. The reason is that I_1 is a minimal element of $[\![C]\!]$ since

$$\{e\} \nsubseteq \{a, d\} \quad \text{and} \quad \{a, b, d\} \nsubseteq \{a, d\}$$

which means $I_2 \nleqslant^E_{confl} I_1$, and

$$\{b, d, e\} \nsubseteq \{b, e\} \quad \text{and} \quad \{a\} \nsubseteq \{b, e\}$$

which means $I_3 \nleqslant^E_{confl} I_1$.

To address the issue illustrated in the previous example, we propose the following alternative:

$$I \in \left[\!\left[\Delta^{confl_2}(K_1, \ldots, K_n; C; diff, \leqslant_{confl})\right]\!\right] \quad \text{iff} \quad I \in [\![C]\!] \wedge \exists \mathbf{c} \in diff(I, \mathcal{K}) \,.\, \forall I' \in [\![C]\!] \,.\, \forall \mathbf{c}' \in diff\big(I', \mathcal{K}\big) \,.\, \mathbf{c}' \nless_{confl} \mathbf{c} \tag{8}$$

where $\mathbf{c}' <_{confl} \mathbf{c}$ iff $\mathbf{c}' \leqslant_{confl} \mathbf{c}$ and $\mathbf{c} \nleqslant_{confl} \mathbf{c}'$. Note that when we apply (8) to the scenario in Example 2, we indeed find the desired result $\{I_2, I_3\}$. The operator in (8) directly encodes the idea that models of the merged knowledge base should be those that are closest to the given knowledge bases, treating the conflict sets as qualitative (partially ordered) distances. Note that the operator is parametrized by the conflict-operator *diff*, which will also allow us to consider alternatives to (6).

In general, while providing extra flexibility, the conflict-based merging scheme has similar advantages and disadvantages as the distance-based approach. In particular, straightforward implementations do not allow to deal with semantic background information about the relatedness of atoms.

2.3. Morpho-logical merging

Another interesting view on propositional merging [10] is to weaken propositions using a logical counterpart of mathematical morphology [21]. In particular, the dilation $D_B(\phi)$ of a formula ϕ is defined by

$$\left[\!\left[D_B(\phi)\right]\!\right] = \left\{I \in 2^A \;\middle|\; B(I) \cap [\![\phi]\!] \neq \emptyset\right\} \tag{9}$$

where $B(I) \subseteq 2^A$ is a set of interpretations that are related to I in some way. Note that B, which is called the structuring element, can be regarded as a relation between interpretations. Again the Hamming distance is most commonly used, choosing

$$B(I) = \left\{I' \in 2^A \;\middle|\; d_{Ham}\big(I, I'\big) \leqslant t\right\} \tag{10}$$

for some $t \in \mathbb{N}$. This particular choice makes it straightforward to syntactically characterize $D_B(\phi)$ (see [10]). The knowledge bases $K_1, \ldots, K_n$ may be merged, subject to the integrity constraints C, using the operator Δ^{morph} defined by:

$$\left[\!\left[\Delta^{morph}(K_1, \ldots, K_n; C; B)\right]\!\right] = [\![C]\!] \cap \left[\!\left[D_B(K_1)\right]\!\right] \cap \cdots \cap \left[\!\left[D_B(K_n)\right]\!\right]$$

where the structuring element B is chosen such that consistency is effectively restored, i.e. $[\![\Delta^{morph}(K_1, \ldots, K_n; C; B)]\!] \neq \emptyset$. For instance, when using (10), we may choose the smallest value of t that restores consistency. Interestingly, in [22], the morpho-logical approach to merging has been generalized to first-order logic. In the propositional case, however, it is clear that morpho-logical merging is similar in spirit as distance-based and conflict-based merging.

3. Modeling heterogeneous vocabulary usage

Our approach to merging is based on interpreting statements that are provided by the sources in a flexible way. Conceptually this boils down to weakening each of the considered knowledge bases, i.e. increasing their set of models. Below we present the main ideas of our procedure. First, Section 3.1 elaborates on the relationship between weakening knowledge bases and interpreting atoms in a flexible way. Section 3.2 then discusses how this idea can be implemented at the syntactic level. Subsequently, Section 3.3 provides the semantic counterpart of the syntactic mechanism for weakening our understanding of statements.

3.1. Weakening knowledge bases

Let $K_1, \ldots, K_n$ be sets of statements (knowledge bases) that are asserted by distinct sources $s_1, \ldots, s_n$, where each source is assumed to use a propositional language over the *same* set of atoms A. We furthermore consider an additional knowledge base C containing the integrity constraints of the domain being modeled. Throughout this paper, we will tacitly assume that the set of integrity constraints C is consistent, and that each knowledge base K_i is individually consistent. Our goal then is to weaken the statements from the knowledge bases $K_1, \ldots, K_n$, resulting in knowledge bases $K'_1, \ldots, K'_n$ such that $K'_1 \cup \cdots \cup K'_n \cup C$ is consistent. The term *weakening* reflects the fact that all models of K_i are contained in the set of models of K'_i. While this general idea of weakening knowledge bases underlies many existing merging strategies, our approach is particular in its use of extra-logical information and the idea of similarity. Specifically, our point of departure is that some sources may have a slightly different understanding of the meaning of atom a. To formalize this intuition, let us write $a^{@s_i}$ to denote the understanding of atom a by source s_i. In other words, $a^{@s_i}$ is an artificial notation that we introduce to precisely capture what can be assumed to hold when source s_i claims a. Depending on the application context, the atom a itself could then correspond to the official meaning of a given term (assuming one exists), or to the particular way this term is to be understood w.r.t. the integrity constraints in C. Moreover, given a subset of atoms $X \subseteq A$, we write $X^{@s_i}$ for the set $\{x^{@s_i} \mid x \in X\}$. Given a knowledge base K_i, we write $K_i^{@s_i}$ to denote the knowledge base that results from substituting each occurrence of an atom $a \in A$ by the corresponding atom $a^{@s_i}$. Thus we tacitly admit that $(\neg a)^{@s_i} = \neg a^{@s_i}$, $(a \wedge b)^{@s_i} = a^{@s_i} \wedge b^{@s_i}$ and $(a \vee b)^{@s_i} = a^{@s_i} \vee b^{@s_i}$.

Example 3. Consider again the weather example (1)–(3). Consistency can be trivially restored by assuming that the three sources use a slightly different terminology:

$$K_1^{@s_1} = \{partiallyCloudy^{@s_1} \vee overcast^{@s_1}\}$$
$$K_2^{@s_2} = \{openSky^{@s_2}\}$$
$$K_3^{@s_3} = \{overcast^{@s_3}\}$$

As each of the knowledge bases $K_1^{@s_1}$, $K_2^{@s_2}$, $K_3^{@s_3}$ and C contain occurrences from disjoint sets of atoms, we clearly have that $K_1^{@s_1} \cup K_2^{@s_2} \cup K_3^{@s_3} \cup C$ is consistent.

To obtain more interesting conclusions than from the trivial solution in Example 3, we additionally encode how atoms of the form $a^{@s_i}$ relate to the atoms from A. This idea of using a disjoint vocabulary for each of the sources, and subsequently imposing constraints that allow useful results without introducing inconsistency can also be found in [23]. This latter approach, in our notation, boils down to adding a maximal set of equivalences of the form $a^{@s_i} \equiv a^{@s_j}$ which does not introduce inconsistency (a procedure called belief set merging), or a maximal set of equivalences of the form $a^{@s_i} \equiv a$ which does not introduce inconsistency (a procedure called belief set projection). The underlying idea is that the knowledge expressed by particular sources about particular variables is ignored. In this sense, the approach from [23] is also similar to the procedure proposed in [24], which is more explicitly based on this idea of ignoring variables.

In this paper, we propose a different solution, which may be understood as a refinement of the belief set projection approach from [23]. Our purpose in adding constraints, however, is not to ignore certain variables, but rather to encode in which way they may be understood. Although we may not know with certainty how atom a is understood by source s_i, in many contexts it seems reasonable to assume that information is available about which understandings are possible. In the case of Example 3, the understanding of *overcast* by source s_1 may be one of the following:

1. The intended meaning of *overcast* is adopted by source s_1, i.e. we have $overcast^{@s_1} \equiv overcast$, or in other words $overcast^{@s_1} \rightarrow overcast$ and $overcast \rightarrow overcast^{@s_1}$.
2. Source s_1 takes a more *liberal* interpretation of *overcast* which includes some of the cases that are normally described as *partiallyCloudy*, i.e. $overcast^{@s_1} \rightarrow overcast \vee partiallyCloudy$ and $overcast \rightarrow overcast^{@s_1}$.
3. Source s_1 takes a more *restrictive* interpretation of *overcast* which includes only some of the cases that are normally described as *overcast*, the remaining cases being described as *partiallyCloudy*. We then have $overcast^{@s_1} \rightarrow overcast$ and $overcast \rightarrow overcast^{@s_1} \vee partiallyCloudy^{@s_1}$.

Assume furthermore that the first situation is considered more plausible than the second situation, which is in turn considered to be more plausible than the last. Using possibilistic logic [25], for instance, we may describe our background assumptions on the possible state of affairs as follows:

$$M = \big\{\big(\big(oc^{@s_1} \equiv oc\big) \vee \big(\big(oc^{@s_1} \rightarrow oc \vee pc\big) \wedge \big(oc \rightarrow oc^{@s_1}\big)\big) \vee \big(\big(oc^{@s_1} \rightarrow oc\big) \wedge \big(oc \rightarrow oc^{@s_1} \vee pc^{@s_1}\big)\big), 1\big),$$
$$\big(\big(oc^{@s_1} \equiv oc\big) \vee \big(\big(oc^{@s_1} \rightarrow oc \vee pc\big) \wedge \big(oc \rightarrow oc^{@s_1}\big)\big), \lambda_1\big), \big(oc^{@s_1} \equiv oc, \lambda_0\big)\big\}$$

where we have abbreviated *overcast* and *partiallyCloudy* as oc and pc respectively, and $0 < \lambda_0 < \lambda_1 < 1$. The possibilistic knowledge base M encodes that we are certain that at least one of the three situations described above is correct. With less certainty, we believe that at least one of the first two situations is correct, and with even less certainty that the first situation is correct. It can be verified[2] that the possibilistic knowledge base M may be equivalently expressed as

$$M = \big\{\big(oc^{@s_1} \rightarrow oc \vee pc, 1\big), \big(oc \rightarrow oc^{@s_1} \vee pc^{@s_1}, 1\big), \big(oc \rightarrow oc^{@s_1}, \lambda_1\big), \big(oc^{@s_1} \rightarrow oc, \lambda_0\big)\big\} \tag{11}$$

Note how the knowledge base M encodes more or less plausible 'mistakes' that may explain conflicts between different sources. This idea of merging propositional knowledge bases by trying to identify the underlying mistakes that have been made by the sources has been advocated in [26].

3.2. Syntactic encoding

Note that implications are used as the basic building blocks instead of equivalences in (11). This is due to the fact that the exact meaning of e.g. $oc^{@s_1}$ may not be expressible using the available terminology (i.e. the atoms in A). All we can express then, are necessary and sufficient conditions under which $oc^{@s_1}$ and $\neg oc^{@s_1}$ hold. This gives rise to four types of implications, which are used to encode the relationship between the atoms in $A^{@s_i}$ and those in A:

$$a^{@s_i} \rightarrow \bigvee\big\{w \mid w \in W_a^l\big\} \tag{12}$$
$$a \rightarrow \bigvee\big\{x^{@s_i} \mid x \in X_a^l\big\} \tag{13}$$
$$\neg a^{@s_i} \rightarrow \bigvee\big\{\neg y \mid y \in Y_a^l\big\} \tag{14}$$
$$\neg a \rightarrow \bigvee\big\{\neg z^{@s_i} \mid z \in Z_a^l\big\} \tag{15}$$

where $l \in \{0, \ldots, k\}$ and $W_a^l, X_a^l, Y_a^l, Z_a^l \subseteq A$ are sets of atoms such that

$$\{a\} = W_a^0 \subseteq W_a^1 \subseteq \cdots \subseteq W_a^k, \qquad \{a\} = X_a^0 \subseteq X_a^1 \subseteq \cdots \subseteq X_a^k$$
$$\{a\} = Y_a^0 \subseteq Y_a^1 \subseteq \cdots \subseteq Y_a^k, \qquad \{a\} = Z_a^0 \subseteq Z_a^1 \subseteq \cdots \subseteq Z_a^k$$

Note that an arbitrary number of disjuncts may appear in the right-hand side of the implications. The disjuncts that appear in the right-hand sides of the implications are denoted as W_a^l, X_a^l, Y_a^l and Z_a^l, where l is a tolerance parameter, i.e. rather than considering one implication of each type, we consider a sequence of implications which allow for an increasingly more liberal view. For $l = 0$, the implications (12)–(15) simply assert that $a^{@s_i} \equiv a$. For larger values of l, these implications correspond to increasingly weaker constraints on the exact logical relationship between a and $a^{@s_i}$. The underlying idea is that the larger the value of l, the more certain we are that the logical relations expressed by (12)–(15) are valid. By considering larger values of l, we effectively stretch the meaning of what is asserted by the source, until it becomes consistent with what is asserted by other sources. As the meaning of propositions cannot be stretched indefinitely, some fixed upper bound k is assumed. Note that in the approach presented in [23], we have either $a^{@s_i} \equiv a$, or $a^{@s_i} \equiv \top$ when the literal a should be ignored for the sake of consistency.

Note that (12) and (14) respectively correspond to necessary and sufficient conditions for having $a^{@s_i}$ true, expressed in the standard usage of the vocabulary. Similarly, (13) and (15) respectively correspond to necessary and sufficient conditions for having a true, assuming the vocabulary usage of source s_i. In practice, depending on the characteristics of the domain, several of the implications (12)–(15) may be trivial. To allow us to trivialize the implications in a convenient way, we treat $\top$ and $\bot$ as special atoms from A (rather than primitives in the language). We will tacitly assume that C contains at least the formula $\top \wedge \neg\bot$, enforcing that $\top$ is contained in any model of the integrity constraints, and $\bot$ is contained in no such model, and furthermore that $\top^{@s_i} = \top$ and $\bot^{@s_i} = \bot$ for every source s_i.

Example 4. In the case of (11), implications of the form (14) or (15) were not considered for $l > 0$. This is due to the fact that the vocabulary does not allow us to express sufficient conditions for a or $a^{@s_i}$, using respectively the vocabulary of source s_i and the standard vocabulary. This leads to:

2 By repeated application of the possibilistic resolution rule $(\neg p \vee q, \lambda), (p \vee r, \mu) \vdash (q \vee r, \min(\lambda, \mu))$.

$$
\begin{array}{llll}
W_{oc}^{0}=\{oc\}, & X_{oc}^{0}=\{oc\}, & Y_{oc}^{0}=\{oc\}, & Z_{oc}^{0}=\{oc\} \\
W_{oc}^{1}=\{oc,pc\}, & X_{oc}^{1}=\{oc\}, & Y_{oc}^{1}=\{oc,\bot\}, & Z_{oc}^{1}=\{oc,\bot\} \\
W_{oc}^{2}=\{oc,pc\}, & X_{oc}^{2}=\{oc,pc\}, & Y_{oc}^{2}=\{oc,\bot\}, & Z_{oc}^{2}=\{oc,\bot\}
\end{array}
$$

Note how the occurrence of $\bot$ in Y_{oc}^{l} and Z_{oc}^{l} for $l \geqslant 1$ effectively trivializes the corresponding implications. Furthermore, note that in this example, the upper bound k is 2.

The following example illustrates a situation where all four types of implications play a non-trivial role.

Example 5. The concept of a public school has a meaning in the UK which is different from the one in the USA. In particular, what is called a public school in the UK is called a private school in the USA, and what is called a public school in the USA is called a state school in the UK. Now assume that sources are supposed to use the UK terminology, but that occasionally, some source makes the mistake of using the term public school in its USA meaning. The possible understandings of the atom *pu* (public school) in terms of the remaining atoms *pr* (private school) and *st* (state school) can then be encoded in possibilistic logic as follows:

$$
\begin{aligned}
M=\big\{&\big(pu^{@s_i}\to pu\vee st,1\big),\big(pu\to pu^{@s_i}\vee pr^{@s_i},1\big),\big(\neg pu^{@s_i}\to\neg pu\vee\neg st,1\big),\big(\neg pu\to\neg pu^{@s_i}\vee\neg pr^{@s_i},1\big),\\
&\big(pu^{@s_i}\to pu,\lambda_0\big),\big(pu\to pu^{@s_i},\lambda_0\big)\big\}
\end{aligned}
$$

where $0<\lambda_0<1$. In this case, we get:

$$
\begin{array}{llll}
W_{pu}^{0}=\{pu\}, & X_{pu}^{0}=\{pu\}, & Y_{pu}^{0}=\{pu\}, & Z_{pu}^{0}=\{pu\} \\
W_{pu}^{1}=\{pu,st\}, & X_{pu}^{1}=\{pu,pr\}, & Y_{pu}^{1}=\{pu,st\}, & Z_{pu}^{1}=\{pu,pr\}
\end{array}
$$

Throughout the paper, certainty degrees which are attached to formulas will be interpreted in a variety of different ways (including necessities, priorities, and penalties). In each case, the knowledge base will syntactically be expressed in the same way, however. In particular, given the sets W_a^l, X_a^l, Y_a^l and Z_a^l, for each source s_i, we will consider the following weighted knowledge base:

$$
\begin{aligned}
M_{s_i}=&\Big\{\Big(a^{@s_i}\to\bigvee\big\{w\mid w\in W_a^l\big\},\lambda_{(l,s_i,a)}^{W}\Big)\Bigm| a\in A,\ l\in\{0,\dots,k\}\Big\}\\
&\cup\Big\{\Big(a\to\bigvee\big\{x^{@s_i}\mid x\in X_a^l\big\},\lambda_{(l,s_i,a)}^{X}\Big)\Bigm| a\in A,\ l\in\{0,\dots,k\}\Big\}\\
&\cup\Big\{\Big(\neg a^{@s_i}\to\bigvee\big\{\neg y\mid y\in Y_a^l\big\},\lambda_{(l,s_i,a)}^{Y}\Big)\Bigm| a\in A,\ l\in\{0,\dots,k\}\Big\}\\
&\cup\Big\{\Big(\neg a\to\bigvee\big\{\neg x^{@s_i}\mid z\in Z_a^l\big\},\lambda_{(l,s_i,a)}^{Z}\Big)\Bigm| a\in A,\ l\in\{0,\dots,k\}\Big\}
\end{aligned}
\tag{16}
$$

where the values $\lambda_{(l,s_i,a)}^{\times}$ (with $\times\in\{W,X,Y,Z\}$) are certainty values (or priorities) taken from a totally or partially ordered set $(\Lambda,\leqslant)$ such that $\lambda_{(l_1,s_i,a)}^{\times}\leqslant\lambda_{(l_2,s_i,a)}^{\times}$ when $l_1\leqslant l_2$, i.e. the more the constraints on the meaning of a certain atom a are relaxed, the more certain that they are correct. Note that the full generality of (16) is not always needed, and sometimes even a single sequence of certainty values λ_l would be enough. Among others this depends on how the certainty weights are interpreted (symbolic weights, penalties, priorities, etc.). In Section 5 we will discuss how different interpretations of these certainty values lead to merging operators with a different behavior.

In the following, we also consider the following alternative to the set M_{s_i}:

$$
\begin{aligned}
M'_{s_i}=\Big\{&\Big(\Big(a^{@s_i}\to\bigvee\big\{w\mid w\in W_a^l\big\}\Big)\wedge\Big(a\to\bigvee\big\{x^{@s_i}\mid x\in X_a^l\big\}\Big)\wedge\Big(\neg a^{@s_i}\to\bigvee\big\{\neg y\mid y\in Y_a^l\big\}\Big)\\
&\wedge\Big(\neg a\to\bigvee\big\{\neg x^{@s_i}\mid z\in Z_a^l\big\}\Big),\lambda_{(l,s_i,a)}\Big)\Bigm| a\in A,\ l\in\{0,\dots,k\}\Big\}
\end{aligned}
\tag{17}
$$

with the corresponding assumption that $l_1\leqslant l_2$ implies $\lambda_{(l_1,s_i,a)}\leqslant\lambda_{(l_2,s_i,a)}$. Note that in possibilistic logic, when $\lambda_{(l,s_i,a)}^{W}=\lambda_{(l,s_i,a)}^{X}=\lambda_{(l,s_i,a)}^{Y}=\lambda_{(l,s_i,a)}^{Z}=\lambda_{(l,s_i,a)}$, we have that M_{s_i} is equivalent to M'_{s_i}, while this is not the case in, for instance, penalty logic [27,28] due to the additive interpretation of the weights.

Clearly, the combined use of the four types of implications (12)–(15) results in a powerful mechanism which encompasses a wide variety of inconsistency scenarios. To bridge the gap between this general, but abstract mechanism and practical applications, we discuss in Section 4 four prototypical scenarios in which our framework could be applied. We then also introduce a graph representation, which allows to describe in a compact and intuitive way, how an atom may be understood flexibly, at a given level of tolerance.

3.3. Semantic encoding

At the semantic level, a flexible understanding of statements can be obtained by using suitable operators that manipulate sets of atoms (i.e. interpretations). Let R be a reflexive relation in A, and let $I \subseteq A$ be a set of atoms. Then we define the expansion $\langle I\rangle_R$ and contraction $[I]_R$ of I w.r.t. R as the following sets:

$$\langle I\rangle_R = \big\{a \in A \,\big|\, \exists b \in A\,.\,(a,b) \in R \wedge b \in I\big\} \tag{18}$$

$$[I]_R = co\langle coI\rangle_R = \big\{a \in A \,\big|\, \forall b \in A\,.\,(a,b) \in R \Rightarrow b \in I\big\} \tag{19}$$

where coX is understood as the set complement of X w.r.t. A, i.e. $coX = A \setminus X$. Since R was assumed to be reflexive, we clearly have $[I]_R \subseteq I \subseteq \langle I\rangle_R$. Also note that when R is an equivalence relation, expansion and contraction correspond to the notion of upper and lower approximation from rough set theory [29]. In that case, $\langle I\rangle_R$ is the union of the equivalence classes that overlap with I, whereas $[I]_R$ is the union of the equivalence classes that are included in I.

The intuition behind (18)–(19) is best seen when we interpret R as modeling some form of similarity. Then $\langle I\rangle_R$ contains those atoms that are similar to at least one atom from I. Hence $J \subseteq \langle I\rangle_R$ can be interpreted as asserting that J is approximately included in I, in the sense that every atom in J is similar to an atom in I. The set $[I]_R$, on the other hand, contains those atoms that are only similar to atoms from I. This means that $[I]_R \subseteq J$ can be interpreted as asserting that coJ is approximately included in coI, in the sense that every atom outside J is similar to an atom outside I. Let us now consider the following relation:

$$\sigma_{(R_1,R_2)}(I,J) \quad \text{iff} \quad [I]_{R_1} \subseteq J \subseteq \langle I\rangle_{R_2} \tag{20}$$

Then $\sigma_{(R_1,R_2)}(I,J)$ expresses some form of similarity between interpretations I and J. Indeed, since $J \subseteq \langle I\rangle_{R_2}$ we know that every atom interpreted as true by J is similar (w.r.t. R_2) to an atom interpreted as true by I, and since $[I]_{R_1} \subseteq J$ we know that every atom interpreted as false by J is similar (w.r.t. R_1) to an atom interpreted as false by I.

In the particular case where R_1 and R_2 are both equal to the identity relation (i.e. $(a,b) \in R_1$ iff $(a,b) \in R_2$ iff $a = b$), then also $\sigma_{(R_1,R_2)}$ degenerates to the identity relation (i.e. $\sigma_{(R_1,R_2)}(I,J)$ iff $I = J$). In general, when R_1 and R_2 capture some form of similarity between atoms, $\sigma_{(R_1,R_2)}$ captures a notion of similarity between interpretations. Note however that in general $\sigma_{(R_1,R_2)}$ is not symmetric, not even when R_1 and R_2 are symmetric and/or $R_1 = R_2$.

Example 6. Let pu and pr correspond to public and private schools, as before, and let un refer to a university. Then we may consider that public and private schools are similar to each other, but neither is similar to a university, i.e. we let $R = \{(pu,pu),(pr,pr),(un,un),(pu,pr),(pr,pu)\}$, $I = \{pu,un\}$ and $J = \{un\}$. Then we have

$$[I]_R = \{un\}, \qquad \langle I\rangle_R = \{pu,pr,un\}$$
$$[J]_R = \{un\}, \qquad \langle J\rangle_R = \{un\}$$

In particular we find that $[I]_R \subseteq J \subseteq \langle I\rangle_R$, but not $[J]_R \subseteq I \subseteq \langle J\rangle_R$. Thus $\sigma(R,R)(I,J)$ holds but not $\sigma(R,R)(J,I)$.

Thus the relatedness between interpretations I and J may be parametrized by four reflexive relations R_1, R_2, R_3, R_4 between atoms, by considering that I is similar to J when $\sigma_{(R_1,R_2)}(I,J)$ and $\sigma_{(R_3,R_4)}(J,I)$ both hold. From a practical point of view, however, it is not immediately clear why we need four different relations, and how they should be defined to model a particular scenario. This situation is reminiscent of the four types of implications (12)–(15) we have considered at the syntactic level. The intuitive correspondence between the aforementioned relations and the families W_a^l, X_a^l, Y_a^l and Z_a^l can be made explicit by interpreting the latter families as relations W^l, X^l, Y^l and Z^l, where $W^l = \{(a,b) \in A^2 \mid b \in W_a^l\}$, and analogously for the other relations. The exact correspondence is revealed by the following proposition.

Proposition 1. *Let W_a^l, X_a^l, Y_a^l, and Z_a^l for $a \in A$ and $l \in \{0,\dots,k\}$ be defined as before. Then we have the following characterizations*:

$$a \in \big(J \setminus \langle I\rangle_{W^l}\big) \quad \Leftrightarrow \quad I \cup J^{@s_i} \not\models \Big(a^{@s_i} \rightarrow \bigvee\big\{w \,\big|\, w \in W_a^l\big\}\Big) \tag{21}$$

$$a \in \big(I \setminus \langle J\rangle_{X^l}\big) \quad \Leftrightarrow \quad I \cup J^{@s_i} \not\models \Big(a \rightarrow \bigvee\big\{x^{@s_i} \,\big|\, x \in X_a^l\big\}\Big) \tag{22}$$

$$a \in \big([I]_{Y^l} \setminus J\big) \quad \Leftrightarrow \quad I \cup J^{@s_i} \not\models \Big(\neg a^{@s_i} \rightarrow \bigvee\big\{\neg y \,\big|\, y \in Y_a^l\big\}\Big) \tag{23}$$

$$a \in \big([J]_{Z^l} \setminus I\big) \quad \Leftrightarrow \quad I \cup J^{@s_i} \not\models \Big(\neg a \rightarrow \bigvee\big\{\neg z^{@s_i} \,\big|\, z \in Z_a^l\big\}\Big) \tag{24}$$

Intuitively, e.g. (21) teaches us that $I \cup J^{@s_i} \models (a^{@s_i} \rightarrow \bigvee\{w \mid w \in W_a^l\})$ means that whenever $a \in J$ for an atom a, we also have $a \in \langle I\rangle_{W^l}$. In other words, the fact that $I \cup J^{@s_i} \models (a^{@s_i} \rightarrow \bigvee\{w \mid w \in W_a^l\})$ for all $a \in A$ is equivalent to stating that $J \subseteq \langle I\rangle_{W^l}$. Thus, Proposition 1 essentially reveals that the relations R_1, R_2, R_3 and R_4 correspond to Y^l, W^l, Z^l and X^l

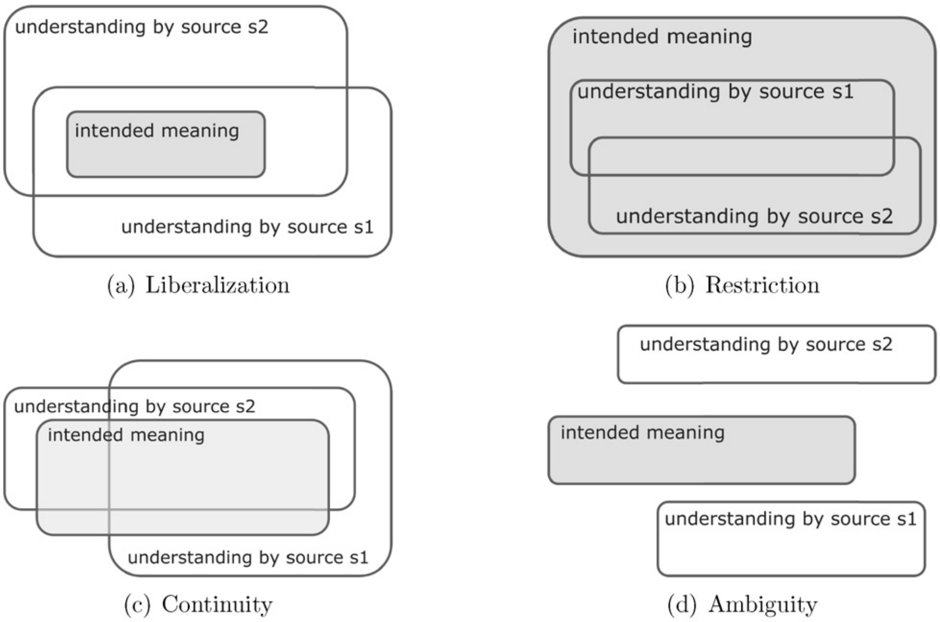

Fig. 1. Four prototypical types of weakening the meaning of an atom.

respectively. This observation also makes clear why we need, in general, the flexibility of parametrizing similarity between interpretations by four different relations between atoms. In practice, on the other hand, these four relations do not have to be specified explicitly most of the time, as we will explain in the following section.

4. Prototypical use cases

The use of four different families of nested sets, W_a^l, X_a^l, Y_a^l and Z_a^l, results in an expressive method to encode how knowledge bases may be weakened. This generality is needed, as the fact that an atom may be understood in a flexible way may mean different things in different contexts. For illustrative purposes, let us write $ext(a)$ to denote the set of situations in which property a holds (e.g. we may think of $ext(a)$ as some region of a conceptual space in the sense of Gärdenfors [30]). There seem to be four prototypical scenarios that are often encountered in practice:

liberalization In certain contexts, and depending on the view one takes, we may be more liberal regarding the exact meaning of a certain property. As a result, for borderline situations, a given property may be considered to be satisfied according to some sources, but not satisfied according to others. In this scenario, flexibility means that we need to admit that the sources may have a more liberal understanding of the meaning of a property: $ext(a) \subseteq \bigcap_i ext(a^{@s_i})$. This situation is illustrated in Fig. 1(a).

restriction The opposite situation also occurs: sources may hold a stricter view on the meaning of certain properties, for instance restricting the situations in which a property is considered to hold to the most typical situations. Being flexible then means that we may need to exclude certain borderline cases of the property: $ext(a) \supseteq \bigcup_i ext(a^{@s_i})$. This situation is illustrated in Fig. 1(b); note, however, that the understandings by different sources do not necessarily need to overlap.

continuity The meaning of a property may depend on some threshold value in a continuous domain, which is to some extent arbitrary. Statements provided by different sources may then be based on slightly different threshold values. This means that there are a number of situations in which the property would hold according to all sources, in addition to situations in which it is the particular definition adopted by a source that determines whether or not the property is considered to hold: $ext(a) \cap \bigcap_i ext(a^{@s_i}) \neq \emptyset$. Furthermore, due to the continuity there is potentially an infinite number of reasonable delineations of those situations in which the property is considered to hold. Due to the finite number of atoms that are considered in the language,[3] the exact meaning of a property, according to one source, is typically not expressible. Fig. 1(c) depicts this situation.

ambiguity The meaning of a property may be ambiguous. This means that some sources may interpret an atom in a completely different way than other sources. In such a case, it may happen that there is not a single situation in which the property is considered to hold according to two different sources. Intuitively, the meaning of an atom

[3] Note that even moving from a finite number of atoms to a countably infinite number of atoms would not change this situation.

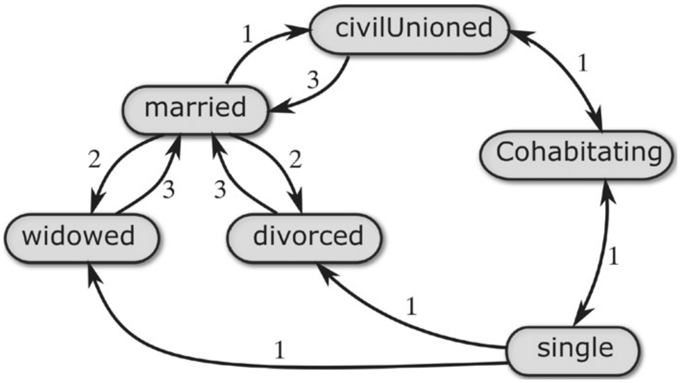

Fig. 2. Similarity graph for atoms related to marriage.

Table 1
Defining the sets W_a^l, X_a^l, Y_a^l and Z_a^l in terms of the sets of related atoms S_a^l.

	W_a^l	X_a^l	Y_a^l	Z_a^l
Liberalization	S_a^l	$\{a\}$	$\{a\}$	$\{a, \bot\}$
Restriction	$\{a\}$	$\{a, \top\}$	S_a^l	$\{a\}$
Continuity	S_a^l	S_a^l	$\{a, \bot\}$	$\{a, \bot\}$
Ambiguity	S_a^l	$\{b \mid a \in S_b^l\}$	S_a^l	$\{b \mid a \in S_b^l\}$

is then translated, rather than stretched: $ext(a) = ext(a^{@S_i})$ or $ext(a) \cap ext(a^{@S_i}) = \emptyset$. This situation is depicted in Fig. 1(d).

Note that from a formal point of view, in continuous domains, liberalization and restriction may be seen as special cases of continuity. However, we reserve the term continuity for scenarios in which it is not known a priori whether sources will assume a more liberal, more restrictive, or overlapping meaning. Moreover, when considering liberalization and restriction below, we will focus on discrete domains, and in particular assume that the exact meaning of $a^{@S_i}$ can be expressed as a disjunction, resp. conjunction of atoms from A.

In each of these four scenarios, it suffices to specify, for each given atom a and tolerance level l, a *single* set of atoms S_a^l. This set S_a^l may be understood as the set of atoms that are similar, or related to a at the given tolerance level, although the precise interpretation of this set will differ in each of the scenarios. In practical applications, the sets S_a^l may be conveniently and compactly described using a weighted, directed graph. As such graphs encode some notion of similarity, we will refer to them as similarity graphs. Figs. 2, 3 and 4 depict examples of such graphs. The underlying intuition is that whenever there is a path from node a to a node b, if some source claims that a holds, it is somewhat plausible that b is the case instead. The level of plausibility depends on the weights on the edges, which in turn correspond to the tolerance levels. Hence, like the tolerance levels, the weights may be given a number of different interpretations (qualitative as well as quantitative), as will become clear below. In many situations, these weights will be given an ordinal interpretation, in which case it suffices to rank the edges according to how likely the corresponding transition is.

Given a similarity graph, S_a^l can be defined as

$$S_a^l = \left\{ b \;\middle|\; b \in A,\ dist(a,b) \leqslant l \right\} \tag{25}$$

where $dist(a,b)$ is the sum of the weights on the shortest path between a and b in the similarity graph under consideration. Intuitively, the graph from Fig. 2 specifies, among others, that *married* can progressively be weakened to $married \vee civilUnioned$ and then $married \vee civilUnioned \vee widowed \vee divorced \vee cohabitating$. When the atoms in a similarity graph express jointly exhaustive and pairwise disjoint (JEPD) properties, i.e. when in every possible world exactly one of these atoms is true, symmetric and uniformly weighted similarity graphs essentially correspond to conceptual neighborhood graphs in the sense of Freksa [31,32].

It is important to note that similarity graphs are nothing more than a convenient vehicle to specify the sets S_a^l in some applications. Sometimes, it may not be appropriate to use such a graph, and it makes more sense to specify the sets S_a^l directly. For instance, the use of a similarity graph to specify the sets S_a^l implies some form of transitivity, i.e. $b \in S_a^{l_1}$ and $c \in S_b^{l_2}$ implies that $c \in S_a^{l_1+l_2}$.

Depending on the scenario, the exact nature of the sets S_a^l is different, resulting in different definitions of the sets W_a^l, X_a^l, Y_a^l and Z_a^l. Below we discuss in more detail how each given scenario dictates the definition of W_a^l, X_a^l, Y_a^l and Z_a^l. The results that are going to be established are summarized in Table 1.

4.1. Liberalization

The atoms that are connected to *married* in the graph from Fig. 2 represent situations that can be considered as special cases of being married, when assuming a liberal usage of this term. This means that non-standard understandings of *married* may only enlarge its intended meaning by designating extra atoms to be special cases of *married*. Such a reading of a similarity graph may be formalized as follows:

$$\bigvee_{\{a\}\subseteq S\subseteq S_a^l} \left(a^{@s_i} \equiv \bigvee S\right) \tag{26}$$

which directly translates the intuition that the understanding of a by source s_i can be expressed as the disjunction of a particular set S of atoms, i.e. $a^{@s_i} \equiv \bigvee S$. We do not know which atoms are in S, but we assume that all atoms in S are similar to a (at a given tolerance level l). This reading can be equivalently expressed using the following implications:

$$a^{@s_i} \to \bigvee S_a^l, \qquad a \to a^{@s_i} \tag{27}$$

Lemma 1. *The formula* (26) *is satisfied for all $a \in A$ iff the implications in* (27) *are satisfied for all $a \in A$.*

We may thus define:

$$W_a^l = S_a^l, \qquad X_a^l = Y_a^l = \{a\}, \qquad Z_a^l = \{a, \bot\}$$

where we used the fact that $a \to a^{@s_i}$ is equivalent to $\neg a^{@s_i} \to \neg a$. Alternatively, we may define $X_a^l = \{a, \top\}$ and $Y_a^l = \{a\}$, or $X_a^l = \{a\}$ and $Y_a^l = \{a, \bot\}$. Note in particular that negative literals of the form $\neg a^{@s_i}$ are always understood as $\neg a$. This corresponds to the observation that *married* in its standard understanding is truly atomic, i.e. there are no situations that are normally considered as *married* but may not be considered as such by certain sources. Conversely, when $\neg a$ holds, only trivial conclusions can be expressed using the atoms in $A^{@s_i}$, unless $W_a^l = \{a\}$, in which case $a^{@s_i} \equiv a$.

Example 7. When a source claims that somebody is married, we may consider the possibility that he or she is actually in a civil union, whereas the standard understanding of marriage only permits the strict meaning (for $l = 1$):

$$W^1_{married} = \{married, civilUnion\}$$
$$X^1_{married} = \{married\}$$

When $\neg married$ is known to hold, no conclusions can be expressed using the atoms in $A^{@s_i}$ because source s_i may use the term *married* in a more liberal sense, such that $married^{@s_i}$ holds. On the other hand, when $\neg married^{@s_i}$ holds, we know that also $\neg married$ holds. This may be expressed as:

$$Y^1_{married} = \{married\}$$
$$Z^1_{married} = \{married, \bot\}$$

4.2. Restriction

The idea of restriction is dual to the idea of liberalization. Rather than admitting more borderline cases, here we need to consider the possibility of excluding borderline cases:

$$\bigvee_{\{a\}\subseteq S\subseteq S_a^l} \left(a^{@s_i} \equiv \bigwedge S\right) \tag{28}$$

which expresses the intuition that some source may only consider a to hold when some additional properties are satisfied. Although we do not know which set S of properties needs to be satisfied for $a^{@s_i}$ to hold, but we make the assumption that all these properties are similar to a. Although the sets S_a^l may, in principle, again be specified as a graph, such a graph would look less natural. Indeed, rather than specifying similar atoms, the set S_a^l here contains properties describing typicality. Again, implications of the form (12)–(15) may be used to describe the state of affairs:

$$\neg a^{@s_i} \to \bigvee \{\neg s \mid s \in S_a^l\}, \qquad \neg a \to \neg a^{@s_i} \tag{29}$$

Lemma 2. *The formula* (28) *is satisfied for all $a \in A$ iff the implications in* (29) *are satisfied for all $a \in A$.*

Fig. 3. Similarity graph for atoms related to the weather.

We may thus define:

$$W_a^l = \{a\}, \qquad X_a^l = \{a, \top\}, \qquad Y_a^l = S_a^l, \qquad Z_a^l = \{a\}$$

where we used the fact that $\neg a \to \neg a^{@s_i}$ is equivalent to $a^{@s_i} \to a$. Alternatively, we may define $W_a^l = \{a, \top\}$ and $Z_a^l = \{a\}$, or $W_a^l = \{a\}$ and $Z_a^l = \{a, \bot\}$. Note in particular that positive literals of the form $a^{@s_i}$ are always understood as a.

Example 8. The term polyhedron in mathematics is sometimes defined as the finite union of convex polyhedra, where a convex polyhedron is the intersection of a finite number of half-spaces in some particular dimension. Different definitions are used, however, where the term polyhedron is sometimes used only for convex polyhedra, for bounded polyhedra, or for three-dimensional polyhedra. Thus we have, for instance:

$$S^1_{polyhedron} = \{polyhedron, convex, bounded, threeDimensional\}$$

When a source claims $polyhedron^{@s_i}$ we may not know which particular definition was assumed, but in each case we may conclude *polyhedron*, as all possible definitions are more specific. If the source claims $\neg polyhedron^{@s_i}$, however, this may not be sufficient to conclude $\neg polyhedron$ as it may, for instance, be the case that the source is describing an unbounded polyhedron. Hence, all that may be concluded from $\neg polyhedron^{@s_i}$ is $\neg polyhedron \vee \neg convex \vee \neg bounded \vee \neg threeDimensional$.

Of course, it may be the case that both more liberal and more restrictive definitions of some property exist, depending on the considered source. As such hybrid situations can be treated entirely analogously as the scenarios of liberalization and restriction, we omit the details.

4.3. Continuity

The notion of similarity from Fig. 3 relates to the fact that the atoms involved are the result of a partially arbitrary discretization of a continuous domain. There is no well-defined, crisp boundary between situations that should be described as *overcast* and situations that should be described as *partiallyCloudy*. As a result, what is called *overcast* by one source may be called *partiallyCloudy* by another source. Due to their particular nature, such similarity graphs are symmetric. Formally we may describe the underlying intuition as:

$$a^{@s_i} \to \bigvee W_a^l, \qquad a \to \bigvee (W_a^l)^{@s_i}$$

While we may not be able to exactly express the meaning of $a^{@s_i}$ using the atoms in A, we know at least that when a source asserts $a^{@s_i}$, an atom similar to a will be the case. Conversely, when a holds, we may assume that $b^{@s_i}$ holds for some atom b similar to a. Note that the information described in a similarity graph, under this reading, is completely captured by implications of the form (12)–(13), hence we may take the implications (14)–(15) to be trivial:

$$W_a^l = X_a^l = S_a^l, \qquad Y_a^l = Z_a^l = \{a, \bot\}$$

In particular, this means that unless $a \to a^{@s_i}$ or $a^{@s_i} \to a$ is assumed, no conclusions can be drawn from $\neg a$ or $\neg a^{@s_i}$, using respectively atoms from $A^{@s_i}$ and atoms from A.

Example 9. The use of the atom *partiallyCloudy* may arise in conflict because the exact boundary between *overcast* situations and *partiallyCloudy* situations may be slightly different according to different sources, as does the exact boundary between *openSky* situations and *partiallyCloudy* situations. This situation is symmetric in that statements we believe to hold may assume a boundary which is more liberal or more restrictive than the boundary assumed by a given source, i.e.:

$$W^1_{partiallyCloudy} = X^1_{partiallyCloudy} = \{openSky, overcast, partiallyCloudy\}$$

Moreover, from $\neg partiallyCloudy^{@s_i}$ no conclusions can be formulated in terms of the atoms in A. The reason is that when $\neg partiallyCloudy^{@s_i}$ is claimed by a source, it may assume a more restrictive definition of the atom *partiallyCloudy*, and it may still be the case that *partiallyCloudy* holds in our standard understanding of this term. Thus we obtain:

$$Y^1_{partiallyCloudy} = Z^1_{partiallyCloudy} = \{partiallyCloudy, \bot\}$$

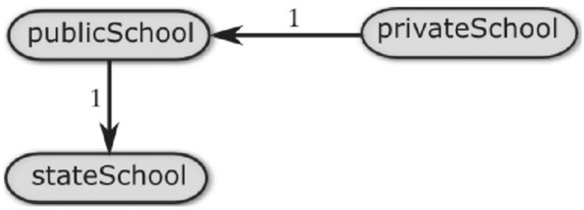

Fig. 4. Similarity graph for atoms related to school types.

4.4. Ambiguity

Fig. 4 illustrates a third possible reading of similarity graphs. Edges in this graph correspond to atom-pairs that are sometimes confused. Whereas the readings discussed in Sections 4.1 and 4.3 relate to the issue of vagueness, the similarity graph from Fig. 4 is centered around the ambiguity of the term *publicSchool*. While there is no straightforward possibility to be flexible about the meaning of a public school in either the UK or the USA meaning, confusion is caused by the fact that we do not know with absolute certainty that all sources conform to the requirement to adopt the UK interpretation. A similarity graph should then be understood as follows:

$$\bigvee_{b \in W_a^l} a^{@s_i} \equiv b, \qquad \bigvee_{a \in W_b^l} a \equiv b^{@s_i} \tag{30}$$

The formula on the left expresses that when a source asserts $a^{@s_i}$, this either means a or some atom with which a may have been confused. Conversely, the formula on the right expresses that the standard understanding of a corresponds to $a^{@s_i}$ or to an atom $b^{@s_i}$ such that sources may confuse b with a. It is not hard to see that (30) is equivalent to the following implications:

$$a^{@s_i} \to \bigvee \{ b \mid b \in W_a^l \}, \qquad a \to \bigvee \{ b^{@s_i} \mid a \in W_b^l \}$$
$$\neg a^{@s_i} \to \bigvee \{ \neg b \mid b \in W_a^l \}, \qquad \neg a \to \bigvee \{ \neg b^{@s_i} \mid a \in W_b^l \}$$

We obtain:

$$X_a^l = Z_a^l = \{ b \mid a \in S_b^l \}, \qquad Y_a^l = W_a^l = S_a^l$$

Example 10. While we can be confident that statements we believe to be true are based on a usage of the term *publicSchool* in its UK meaning, we need to consider the possibility that some source has assumed the USA meaning. Moreover, this confusion remains regardless of whether the atom *publicSchool* is used in a positive or in a negative literal:

$$W^1_{publicSchool} = Y^1_{publicSchool} = \{publicSchool, stateSchool\}$$
$$X^1_{publicSchool} = Z^1_{publicSchool} = \{publicSchool, privateSchool\}$$

5. Merging operators and interpretation of the weights

The methodology that was outlined above leaves us with a list of propositional knowledge bases $K_1^{@s_1}, \ldots, K_n^{@s_n}$, a set of integrity constraints C, and a list of weighted knowledge bases $M_{s_1}, \ldots, M_{s_n}$, defined by (16), that express flexible constraints on how atoms may be understood by the different sources. All these knowledge bases are built from different sets of atoms: $A^{@s_i}$ in the case of $K_i^{@s_i}$, A in the case of C, and $A^{@s_i} \cup A$ in the case of M_{s_i}. Given this point of departure, we are now interested in finding a single, consistent knowledge base K which encodes the combined beliefs of sources $s_1, \ldots, s_n$ to the best extent possible, such that all models of K satisfy the integrity constraints, i.e. $[\![K]\!] \subseteq [\![C]\!]$. We write $K = \Delta(K_1, \ldots, K_n; C; M_{s_1}, \ldots, M_{s_n})$ to denote the result of applying a specific merging operator Δ. When the sets $M_{s_1}, \ldots, M_{s_n}$ and C are clear from the context, we will also write this as $\Delta(K_1, \ldots, K_n)$. The exact behavior of the operator Δ primarily depends on the following two factors:

1. The weights $\lambda_{(l,s_i,a)}$ in the weighted knowledge bases M_{s_i} may be interpreted as necessities, priorities, or as penalties, among others. They may be totally ordered or partially ordered, and may either depend on all of l, s_i and a, only on l and a, only on l and s_i, or only on l. In each case, the nature of the knowledge bases M_{s_i} changes, and so should the result K.
2. There is a trade-off between having a more informative result and having a more cautious result. By appropriately tuning this trade-off, the result may be configured to match the needs of a particular application. Another solution is

to combine both informative and cautious results, by representing the result as a weighted knowledge base, giving the most cautious results the highest certainty or priority.

In this section, we analyze how exactly these two factors may influence the result K. First, we illustrate some of the possibilities in the following example.

Example 11. Consider again the weather example, but assume that the language contains 5 atoms: os, pc_1, pc_2, pc_3 and oc, where pc_1 intuitively corresponds to just a sky which is mostly open with the exception of a few clouds, pc_2 to about half-open and half-cloudy, and pc_3 to a sky which is mostly cloudy. Now consider the following three knowledge bases:

$$K_1 = \{os\}, \qquad K_2 = \{os\}, \qquad K_3 = \{oc\}$$

together with integrity constraints expressing that os, pc_1, pc_2, pc_3 and oc are JEPD properties. Then one may wonder what would intuitively be the most desirable result, assuming as before that all sources are well-reputed. We may consider that since the sources make claims which are completely opposite, we cannot draw any reliable conclusions and should therefore define the result as

$$\{os \vee pc_1 \vee pc_2 \vee pc_3 \vee oc\}$$

A different point of view is that the majority of the sources agrees on os and this should therefore be the result:

$$\{os\}$$

although we may also prefer the following, more cautious alternatives:

$$\{os \vee pc_1\}, \qquad \{os \vee pc_1 \vee pc_2\}$$

Another point of view is to consider that none of the sources would be completely wrong, and that the result should therefore be in (or close to) the middle of os and oc, e.g. one of the following alternatives:

$$\{pc_2\}, \qquad \{pc_1 \vee pc_2 \vee pc_3\}$$

Finally, because os is asserted by two out of three sources, we may also consider that the result should be between os and oc, but closer to os, e.g.:

$$\{pc_1\}, \qquad \{os \vee pc_1 \vee pc_2\}, \qquad \{os \vee pc_1 \vee pc_2 \vee pc_3\}$$

The type of flexibility illustrated by the previous example cannot be achieved using standard approaches, which are based on the Hamming distance. Indeed, the only conclusions that can reasonably be motivated in terms of the Hamming distance are $\{os \vee pc_1 \vee pc_2 \vee pc_3 \vee oc\}$, $\{os\}$, and $\{os \vee oc\}$. In this section, we study a number of different merging operators, based on the sets M_{s_i} and M'_{s_i} introduced in (16) and (17). Each of these merging operators naturally results from interpreting the weights in M_{s_i} or M'_{s_i} in a particular way. Moreover, in each case, we characterize the behavior of the merging operator at the semantic level, establishing close links with the existing frameworks for merging propositional knowledge bases that were recalled in Section 2. In some cases, our merging operators are a special case of distance-based or morphological merging operators, albeit w.r.t. a completely novel type of distance which is based on the idea of interpreting atoms in a flexible way. In other cases, merging operators can be described in natural extensions of distance-based or conflict-based merging operators. It is important to note here, however, that the aim of our paper is not to introduce a new family of merging operators per se, but rather to advocate a different way of measuring the similarity (or distance) between interpretations.

In the discussion that follows, we specifically consider the task of merging the *beliefs* held by different sources, rather than merging incompatible *goals* or *preferences* [33,34]. As such, the intuition behind each merging operator will be to find the interpretations that are most likely to be models of the real world, given the assertions of the different sources (rather than looking for a trade-off between conflicting objectives). By interpreting this notion of likelihood in different ways (ranging from purely qualitative to purely quantitative), and by assuming different levels of caution, it will become clear that a variety of behaviors can be obtained in a natural way, including those that are illustrated in Example 11.

5.1. Basic structure of the merging operators

Before going into the details of particular merging operators, we discuss the underlying idea which is common to each of the approaches that are introduced below. For the ease of presentation, we define the weighted knowledge bases P and P' as follows:

$$P = \bigcup_i M_{s_i} \cup \bigcup_i \{(\alpha, 1) \mid \alpha \in K_i^{@s_i}\} \cup \{(c, 1) \mid c \in C\} \tag{31}$$

$$P' = \bigcup_i M'_{s_i} \cup \bigcup_i \{(\alpha, 1) \mid \alpha \in K_i^{@s_i}\} \cup \{(c, 1) \mid c \in C\} \tag{32}$$

where we reinterpret the propositional knowledge bases $K_i^{@s_i}$ and C as weighted knowledge bases in which all weights are equal to 1, and M_{s_i} and M'_{s_i} are defined as in (16) and (17). Let us furthermore write P^* and P'^* for the sets of (unweighted) formulas that appear in P and in P' respectively, i.e. $P^* = \{\alpha \mid \exists\lambda\,.\,(\alpha, \lambda) \in P\}$ and similar for P'^*. Throughout this section, we implicitly assume that $\lambda_{(l,s_i,a)} < 1$ for all $l \in \{0, \ldots, k\}$, $i \in \{1, \ldots, n\}$ and $a \in A$. The idea is that the formulas in $K_i^{@s_i}$ and C are unconditionally true, hence they receive a maximal certainty level, while the formulas in M_{s_i} and M'_{s_i} are more or less plausible assumptions that may be violated, and thus receive a certainty level that is strictly smaller than 1. Moreover, unless otherwise stated, we assume that $\lambda_{(l,s_i,a)} = \lambda^W_{(l,s_i,a)} = \lambda^X_{(l,s_i,a)} = \lambda^Y_{(l,s_i,a)} = \lambda^Z_{(l,s_i,a)} = \lambda_l$ for every source s_i and atom a, which implies among others that the weights are totally ordered, and in particular that we may assume that the weights $\lambda_{(l,s_i,a)} = \lambda_l$ correspond to numbers in $[0, 1[$. As a notational convenience, let us furthermore introduce the following abbreviations:

$$\alpha^W_{(l,s_i,a)} = a^{@s_i} \to \bigvee\{w \mid w \in W^l_a\}$$
$$\alpha^X_{(l,s_i,a)} = a \to \bigvee\{x^{@s_i} \mid x \in X^l_a\}$$
$$\alpha^Y_{(l,s_i,a)} = \neg a^{@s_i} \to \bigvee\{\neg y \mid y \in Y^l_a\}$$
$$\alpha^Z_{(l,s_i,a)} = \neg a \to \bigvee\{\neg z^{@s_i} \mid z \in Z^l_a\}$$
$$\alpha_{(l,s_i,a)} = \alpha^W_{(l,s_i,a)} \wedge \alpha^X_{(l,s_i,a)} \wedge \alpha^Y_{(l,s_i,a)} \wedge \alpha^Z_{(l,s_i,a)}$$

i.e. $\alpha^\times_{(l,s_i,a)}$ is the formula which appears with weight $\lambda^\times_{(l,s_i,a)}$ in M_{s_i}, and $\alpha_{(l,s_i,a)}$ is the formula which appears with weight $\lambda_{(l,s_i,a)}$ in M'_{s_i}.

Each of the merging operators Δ that we will consider proceeds by selecting particular consistent subsets of P^* or P'^*, which contain at least all the formulas with weight 1, i.e. the formulas from C and $K_i^{@s_i}$. Let $\mathit{Pref}(P)$ and $\mathit{Pref}(P')$ be the subsets P^* and P'^* that are selected according to some criterion. A first idea might then be to define the corresponding merging operators Δ and Δ' as follows:

$$[\![\Delta(K_1, \ldots, K_n)]\!] = \bigvee \mathit{Pref}(P)$$
$$[\![\Delta'(K_1, \ldots, K_n)]\!] = \bigvee \mathit{Pref}(P')$$

where the sets of formulas in $\mathit{Pref}(P)$ and $\mathit{Pref}(P')$ are treated as conjunctions of formulas.

One possible drawback of this method is that the results $\Delta(K_1, \ldots, K_n)$ and $\Delta'(K_1, \ldots, K_n)$ still refer to atoms from $A^{@s_i}$. Recall that these atoms were introduced mainly for technical reasons, and their exact meaning is not known. Thus it seems reasonable to require that no occurrences of these atoms remain after merging (although they may provide insight to the user regarding the source of conflicts). To this end, we employ the notion of variable forgetting [35,24], defined for a set of propositional formulas Φ, an atom x and a set of atoms X as:

$$\mathit{forgetVar}(\Phi, \emptyset) = \emptyset$$
$$\mathit{forgetVar}(\Phi, \{x\}) = \Phi[x := \top] \vee \Phi[x := \bot]$$
$$\mathit{forgetVar}(\Phi, \{x\} \cup X) = \mathit{forgetVar}(\mathit{forgetVar}(\Phi, x), X)$$

where $\Phi[x := \phi]$ for an atom x and a formula ϕ denotes the set of formulas that is obtained from Φ by replacing every occurrence of x by ϕ; sets of formulas are interpreted as conjunctions. The propositional knowledge base $\mathit{forgetVar}(\Phi, X)$ can be seen as a projection of Φ which does not contain any occurrence of atoms from X and moreover, if ψ does not contain occurrences of atoms from X, we have $\Phi \models \psi$ iff $\mathit{forgetVar}(\Phi, X) \models \phi$. Thus, to merge the knowledge bases $K_1, \ldots, K_n$ we may take $\Delta(K_1, \ldots, K_n)$ or $\Delta'(K_1, \ldots, K_n)$ and forget the variables in $A^{@s_1} \cup \cdots \cup A^{@s_n}$:

$$\Delta_f(K_1, \ldots, K_n) = \mathit{forgetVar}(\Delta(K_1, \ldots, K_n), A^{@s_1} \cup \cdots \cup A^{@s_n})$$
$$\Delta'_f(K_1, \ldots, K_n) = \mathit{forgetVar}(\Delta'(K_1, \ldots, K_n), A^{@s_1} \cup \cdots \cup A^{@s_n})$$

While variable forgetting is computationally expensive in general, in certain cases efficient syntactic procedures can be obtained. In Appendix B, an example of such a procedure is provided.

In Sections 5.2–5.6 merging operators are introduced that result from interpreting the certainty weights in different ways. In Section 5.2, certainty weights are interpreted in a purely ordinal way, as possibilistic certainty weights. In Sections 5.3

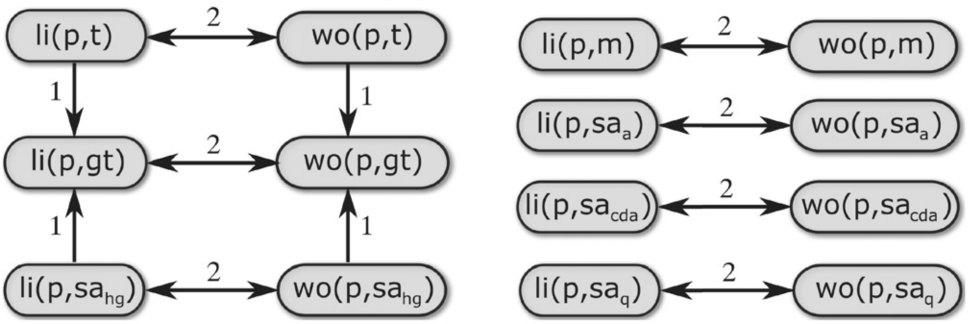

Fig. 5. Similarity graph for atoms related to Toulouse.

and 5.4 this approach is refined by using the concept of preferred subtheories, respectively defined w.r.t. subset-inclusion and w.r.t. cardinalities. Subsequently, in Section 5.5 a purely qualitative approach is presented, which uses symbolic certainty weights that are taken from an arbitrary partially ordered set. Finally, Section 5.6 discusses a quantitative approach with a probabilistic flavor based on penalty logic. Each of these five sections first introduces merging operators at the syntactic level, and subsequently provides a semantic characterization which clarifies their relationship to existing approaches for information merging. Moreover, the five sections can be read independently from each other.

5.2. Possibilistic certainty weights

5.2.1. Merging operators

In this section, we interpret the weighted knowledge bases M_{s_i} as possibilistic knowledge bases. In this case, the weights reflect lower bounds on a necessity measure. Following the standard treatment of inconsistencies in the possibilistic setting,[4] we obtain the following merging operators

$$\Delta^{poss}(K_1,\ldots,K_n) = forgetVar\left(P_{\underline{inc(P)}}, A^{@s_1}\cup\cdots\cup A^{@s_n}\right) \tag{33}$$

$$\Delta^{poss'}(K_1,\ldots,K_n) = forgetVar\left(P'_{\underline{inc(P')}}, A^{@s_1}\cup\cdots\cup A^{@s_n}\right) \tag{34}$$

where $P_{\underline{inc(P)}}$ is the set of formulas from P whose weight is strictly above the inconsistency level of P. In other words, we take $Pref(P) = \{P_{\underline{inc(P)}}\}$ and $Pref(P') = \{P'_{\underline{inc(P')}}\}$. As $P_{\underline{inc(P)}}$ is equivalent to $P'_{\underline{inc(P')}}$, we immediately find that $\Delta^{poss}(K_1,\ldots,K_n) \equiv \Delta^{poss'}(K_1,\ldots,K_n)$, hence we will only consider (33) henceforth.

Example 12. Consider two sources, where one source claims that Peter is married, lives in Toulouse, and works in Montauban (a city about 50 km from Toulouse). A second source also claims that Peter is married, and that he works in Saint-Alban. However, there are at least four places called Saint-Alban[5]: in the regions Ain, Côtes-d'Armor, Haute-Garonne (in France), and in Québec (in Canada). We assume that the source does not want to commit itself to one of the four places. We may then encode the two knowledge bases as follows:

$$K_1 = \left\{mar(p), li(p,t), wo(p,m)\right\}$$

$$K_2 = \left\{mar(p), wo(p,sa_a)\vee wo(p,sa_{cda})\vee wo(p,sa_{hg})\vee wo(p,sa_q)\right\}$$

where e.g. $li(p,t)$ means that Peter lives in Toulouse and $wo(p,m)$ means that Peter works in Montauban. The integrity constraints C specify the JEPD nature of the propositions related to marriage from Fig. 2, the fact that a person can only work in one place, and can only live in one place (e.g. $li(p,t)\to\neg li(p,m)$), and that Toulouse and Saint-Alban (Haute-Garonne) are both contained in the Urban Community of Greater Toulouse (e.g. $li(p,t)\to li(p,gt)$). The families W_a^l, X_a^l, Y_a^l and Z_a^l are defined according to the similarity graph from Fig. 2 for the atoms a related to marriage, and according to the similarity graph from Fig. 5 for the remaining atoms. In both cases, the similarity graphs are interpreted in terms of liberalization, and they are understood in a transitive way.

To obtain the result of merging K_1 and K_2 using Δ^{poss}, we may use Proposition 10:

$$K_1^{(1)} = \left\{mar(p)\vee civ(p), li(p,t)\vee li(p,gt), wo(p,m)\right\}$$

$$K_1^{(2)} = \left\{mar(p)\vee civ(p)\vee wid(p)\vee div(p)\vee coh(p), li(p,t)\vee wo(p,t)\vee li(p,gt), wo(p,m)\vee li(p,m)\right\}$$

[4] A brief introduction to possibilistic logic is provided in Appendix A.

[5] http://en.wikipedia.org/wiki/Saint-Alban, accessed May 31, 2010.

$$K_2^{(1)} = \{mar(p) \vee civ(p), wo(p, sa_a) \vee wo(p, sa_{cda}) \vee wo(p, sa_{hg}) \vee wo(p, gt) \vee wo(p, sa_q)\}$$
$$K_2^{(2)} = \{mar(p) \vee civ(p) \vee wid(p) \vee div(p) \vee coh(p), wo(p, sa_a) \vee li(p, sa_a) \vee wo(p, sa_{cda}) \vee li(p, sa_{cda}) \vee wo(p, sa_{hg}) \vee li(p, sa_{hg}) \vee wo(p, gt) \vee wo(p, sa_q) \vee li(p, sa_q)\}$$

where we write $K_i^{(j)}$ as an abbreviation of $K_i^{P_{\lambda_j}}$. Clearly $K_1^{(1)} \cup K_2^{(1)} \cup C$ is inconsistent, since Peter cannot, according to C, work in two disjoint places. On the other hand, $K_1^{(2)} \cup K_2^{(2)} \cup C$ is consistent, hence $\Delta^{poss}(K_1, K_2) = K_1^{(2)} \cup K_2^{(2)} \cup C$, where

$$\Delta^{poss}(K_1, K_2) \equiv C \cup \{mar(p) \vee civ(p) \vee wid(p) \vee div(p) \vee coh(p), (li(p, sa_{hg}) \wedge wo(p, m)) \vee (wo(p, t) \wedge li(p, m))\}$$

This solution reveals two possible hypotheses regarding the origin of the conflict between K_1 and K_2. The first hypothesis is that the claim of the first source that Peter lives in Toulouse, should be interpreted as Peter living in Greater Toulouse, and that the second source has confused the fact that Peter lives in Saint-Alban with the statement that Peter works in Saint-Alban; this leads to $li(p, sa_{hg}) \wedge wo(p, m)$. The second hypothesis is that the first source has swapped the places where Peter lives and works, i.e. in fact Peter works in Toulouse and lives in Montauban. In addition, the second source has confused Saint-Alban with Greater Toulouse. The second hypothesis leads to $wo(p, t) \wedge li(p, m)$.

Note that after merging, in the example, $ma(p)$ is also understood in a liberal sense, although this is not necessary for solving the inconsistency. This can be seen as a shortcoming of the possibilistic approach, which is overly cautious here. Intuitively, this issue can be addressed by first determining which atoms participate in the conflict, and only weakening these atoms. Several of the refined merging operators that are proposed below are based on this intuition.

5.2.2. Semantic characterization

Semantically, Δ^{poss} can be seen as a particular case of morpho-logical merging (cf. Section 2.3), using a structuring element that is defined in terms of the similarity relation σ defined in (20). In particular, we define the structuring element $B_{(W^l, X^l, Y^l, Z^l)}$ as

$$B_{(W^l, X^l, Y^l, Z^l)}(I) = \{J \in 2^A \mid \sigma_{(Y^l, W^l)}(I, J) \wedge \sigma_{(Z^l, X^l)}(J, I)\}$$

Then we have the following characterization.

Proposition 2. *Assume that $\lambda_{(l, s_i, a)} = \lambda_l$ for every source s_i, atom a and $l \in \{0, \ldots, k\}$, and assume $\lambda_k < 1$. Moreover, assume that P_{λ_k} is consistent, and let r be the smallest value from $\{0, \ldots, k\}$ such that P_{λ_r} is consistent, i.e. $inc(P) = \lambda_{r-1} < \lambda_k$. It holds that*

$$\Delta^{poss}(K_1, \ldots, K_n; C; M_{s_1}, \ldots, M_{s_n}) \equiv \Delta^{morph}(K_1, \ldots, K_n; C; B_{(W^r, X^r, Y^r, Z^r)})$$

Note however that this linkage can only be established once the structuring element of the morpho-logical approach is no longer defined by means of the Hamming distance, as usual, but in terms of the similarity relation underlying our approach.

5.3. Inclusion-based preferred subtheories

5.3.1. Merging operators

As illustrated by Example 12, the standard possibilistic approach leads to results that may be deemed too cautious. In particular, the conflict about where Peter lives and works should, intuitively, not influence our belief that Peter is married, a belief which is shared by both sources. This behavior is due to the *drowning effect* in possibilistic logic, i.e. the fact that every statement whose certainty is not greater than the inconsistency level is ignored, regardless of whether it participates in any conflict. In Example 12, for instance, $mar(p)$ has a lower priority than $mar(p) \vee civ(p) \vee wid(p) \vee div(p) \vee coh(p)$; the former expression is below the inconsistency whereas the latter is above, which is why the disjunction is entailed by the result but $mar(p)$ is not. A refined merging strategy, in which this drowning effect no longer occurs, can be obtained by resorting to methods based on maximal consistent subsets. A standard approach was proposed by Brewka [36] within the context of default reasoning. The idea is to use priorities attached to formulas to designate particular maximal consistent subsets of formulas as preferred. In particular, let K be a prioritized knowledge base, where all priorities are taken from the set $\{\lambda_0, \ldots, \lambda_k\}$ with $\lambda_0 < \cdots < \lambda_k$. As for possibilistic knowledge bases, we write K_λ to denote the set of formulas in K whose priority is at least λ. A consistent set of formulas $F = F_k \cup \cdots \cup F_0$ is then called an (inclusion-based) preferred subtheory of K, with F_l a subset of the formulas that appear in K with priority λ_l ($l \in \{0, \ldots, k\}$), iff $F_k \cup \cdots \cup F_l$ is a maximal consistent subset of K_{λ_l} for all $l \in \{0, \ldots, k\}$, i.e. there is no consistent subset F' of K_{λ_l} such that $F' \supset F_k \cup \cdots \cup F_l$. This boils down to selecting as many formulas with priority λ_k as possible (without getting inconsistency), subsequently adding as many formulas with priority λ_{k-1} as possible, etc. Given a prioritized knowledge base K, we write $Pref_{\subseteq}(K)$ for the set of all inclusion-based preferred subtheories of K. A standard approach to deal with inconsistency in prioritized knowledge bases is to only consider formulas that are entailed by every preferred subtheory. Considering again the weighted knowledge base P defined in (31), this leads to the following merging operator:

$$\Delta^{prior_\subseteq}(K_1,\ldots,K_n) = forgetVar\Big(\bigvee Pref_\subseteq(P), A^{@s_1} \cup \cdots \cup A^{@s_n}\Big)$$

By considering P' instead of P, the merging operator $\Delta^{prior'_\subseteq}$ is obtained. Note that when we interpret P as a possibilistic knowledge base, whenever I is a model of a preferred subtheory of P we also have that $I \in [\![P_{\underline{inc(P)}}]\!]$. This means in particular that

$$\Delta^{prior_\subseteq}(K_1,\ldots,K_n) \models \Delta^{poss}(K_1,\ldots,K_n) \tag{35}$$

$$\Delta^{prior'_\subseteq}(K_1,\ldots,K_n) \models \Delta^{poss}(K_1,\ldots,K_n) \tag{36}$$

i.e. the use of preferred subtheories is indeed a refinement of the possibilistic treatment of inconsistency.

Example 13. Let us go back to the scenario of Example 12, and let us consider the weighted knowledge base P', which is not equivalent to considering P, in contrast to the possibilistic setting. In general, using P' may lead to more cautious results, whereas using P may lead to more informative results. Since P'_{λ_2} is consistent, we know that every preferred subtheory should in each case contain all formulas that appear in P' with weight at least λ_2. There are two different maximal consistent subsets of P'_{λ_1}, leading to the following two preferred subtheories

$$B_1 = P'^* \setminus \big\{\alpha_{(0,s_1,li(p,t))}, \alpha_{(0,s_2,wo(p,sa_{hg}))}, \alpha_{(1,s_2,wo(p,sa_{hg}))}\big\}$$
$$B_2 = P'^* \setminus \big\{\alpha_{(0,s_1,wo(p,m))}, \alpha_{(1,s_1,wo(p,m))}, \alpha_{(0,s_1,li(p,t))}, \alpha_{(1,s_1,li(p,t))}, \alpha_{(0,s_2,wo(p,sa_{hg}))}\big\}$$

We find

$$K_1^{B_1} = \big\{mar(p), li(p,t) \vee li(p,gt), wo(p,m)\big\}$$
$$K_1^{B_2} = \big\{mar(p), li(p,t) \vee li(p,gt) \vee wo(p,t), wo(p,m) \vee li(p,m)\big\}$$
$$K_2^{B_1} = \big\{mar(p), wo(p,sa_a) \vee wo(p,sa_{cda}) \vee wo(p,sa_{hg}) \vee wo(p,gt) \vee li(p,sa_{hg}) \vee wo(p,sa_q)\big\}$$
$$K_2^{B_2} = \big\{mar(p), wo(p,sa_a) \vee wo(p,sa_{cda}) \vee wo(p,sa_{hg}) \vee wo(p,gt) \vee wo(p,sa_q)\big\}$$

which, using Proposition 10 leads to

$$\begin{aligned}\Delta^{prior'_\subseteq}(K_1,K_2) &\equiv C \cup \big\{\big(K_1^{B_1} \wedge K_2^{B_1}\big) \vee \big(K_1^{B_2} \wedge K_2^{B_2}\big)\big\}\\ &\equiv C \cup \big\{mar(p), \big(li(p,sa_{hg}) \wedge wo(p,m)\big) \vee \big(wo(p,t) \wedge li(p,m)\big)\big\}\end{aligned}$$

Hence, as desired, the belief that Peter is married is kept, while obtaining the same conclusions about where Peter might live or work as in the possibilistic setting (i.e. Example 12).

5.3.2. Semantic characterization

Semantically, the merging operators $\Delta^{prior_\subseteq}$ and $\Delta^{prior'_\subseteq}$ are similar in spirit to the idea of conflict based merging. In the following, we reveal the exact relationship between both approaches. At a given tolerance level l we may define the following *partial conflict sets* for all interpretations I and J

$$diff^a_{W^l}(I,J) = J \setminus \langle I\rangle_{W^l} \tag{37}$$

$$diff^b_{X^l}(I,J) = I \setminus \langle J\rangle_{X^l} \tag{38}$$

$$diff^c_{Y^l}(I,J) = [I]_{Y^l} \setminus J \tag{39}$$

$$diff^d_{Z^l}(I,J) = [J]_{Z^l} \setminus I \tag{40}$$

Note in particular that for $l=0$, we recover the standard notion of conflict set, as defined in (6), i.e. $diff(I,J) = diff^a_{W^l}(I,J) \cup diff^b_{X^l}(I,J) \cup diff^c_{Y^l}(I,J) \cup diff^d_{Z^l}(I,J)$. An atom p will be contained in $diff^a_{W^l}(I,J)$ if it is in J and it is not similar to an atom in I (where similarity is defined w.r.t. the relation W^l). The intuition behind the other partial conflict sets is analogous. The higher the value of l, the more pairs of atoms are considered similar, and the fewer atoms remain in the partial conflict sets, i.e. for $l > 0$ we have

$$diff^a_{W^l}(I,J) \subseteq diff^a_{W^{l-1}}(I,J), \qquad diff^b_{X^l}(I,J) \subseteq diff^b_{X^{l-1}}(I,J)$$
$$diff^c_{Y^l}(I,J) \subseteq diff^c_{Y^{l-1}}(I,J), \qquad diff^d_{Z^l}(I,J) \subseteq diff^d_{Z^{l-1}}(I,J)$$

Intuitively, for larger values of l, the sets (37)–(40) only contain the most important conflicts between I and J. This idea of priorities among conflicts will allow us to refine the relation $\leqslant^E_{confl}$ that was defined in Section 2.2.

To gather all the information about the conflict at a given tolerance level l, two alternatives present themselves:

$$diff^1_{\mathcal{S}_l}(I,J) = \langle diff^a_{W^l}(I,J), diff^b_{X^l}(I,J), diff^c_{Y^l}(I,J), diff^d_{Z^l}(I,J)\rangle$$
$$diff^2_{\mathcal{S}_l}(I,J) = diff^a_{W^l}(I,J) \cup diff^b_{X^l}(I,J) \cup diff^c_{Y^l}(I,J) \cup diff^d_{Z^l}(I,J)$$

where we write $\mathcal{S}_l$ as a shorthand for (W^l, X^l, Y^l, Z^l). Note that $diff^1_{\mathcal{S}_l}$ is more informative, while $diff^2_{\mathcal{S}_l}$ stays closer to the standard approach of conflict-based merging. In particular, for $l = 0$, $diff^2_{\mathcal{S}_l}(I,J) = diff(I,J)$ with $diff(I,J)$ defined as in (6). In contrast to standard conflict-based merging, where the conflict between two interpretations is characterized using one conflict set, here we have a different conflict set for each tolerance level. All available information about the conflict between two interpretations is thus described as a vector of conflict sets:

$$diff^1_{\mathcal{S}}(I,J) = \langle diff^1_{\mathcal{S}_k}(I,J), \ldots, diff^1_{\mathcal{S}_0}(I,J)\rangle$$

and similar for $diff^2_{\mathcal{S}}$. We use $\mathcal{S}$ as a shorthand for $(\mathcal{S}_k, \ldots, \mathcal{S}_0)$. Analogously as for the standard conflict-based merging scheme, we may then define the conflict between an interpretation I and a knowledge base K. We consider the following two variants:

$$diff^1_{\mathcal{S}}(I,K) = \min\big(\{diff^1_{\mathcal{S}}(I,J) \mid J \in [\![K]\!]\}, lex(\subseteq^4, \ldots, \subseteq^4)\big)$$
$$diff^2_{\mathcal{S}}(I,K) = \min\big(\{diff^2_{\mathcal{S}}(I,J) \mid J \in [\![K]\!]\}, lex(\subseteq, \ldots, \subseteq)\big)$$

where $\subseteq^4 = par(\subseteq, \subseteq, \subseteq, \subseteq)$. Note the use of the lexicographic ordering here. Indeed, when comparing conflicts, it makes sense to first look at the most important conflicts, i.e. those that only disappear when considering high values for the tolerance level l.

Accordingly, we may define conflict vectors between an interpretation I and a list of knowledge bases $\mathcal{K} = (K_1, \ldots, K_n)$:

$$diff^1_{\mathcal{S}}(I,\mathcal{K}) = \big\{\langle c_1, \ldots, c_n\rangle \bigm| \forall i \,.\, c_i \in diff^1_{\mathcal{S}}(I, K_i)\big\}$$

and similar for $diff^2_{\mathcal{S}}$. The conflict vectors in $diff^1_{\mathcal{S}}(I,\mathcal{K})$ and $diff^2_{\mathcal{S}}(I,\mathcal{K})$ are now vectors of vectors of conflict representations between interpretations. To compare such conflict vectors, first let us define $\leqslant^1_l$ and $\leqslant^2_l$ as

$$\big\langle\langle \mathbf{c}^k_1, \ldots, \mathbf{c}^0_1\rangle, \ldots, \langle \mathbf{c}^k_n, \ldots, \mathbf{c}^0_n\rangle\big\rangle \leqslant^1_l \big\langle\langle \mathbf{c}'^k_1, \ldots, \mathbf{c}'^0_1\rangle, \ldots, \langle \mathbf{c}'^k_n, \ldots, \mathbf{c}'^0_n\rangle\big\rangle$$
$$\text{iff} \quad \big(\langle \mathbf{c}^l_1, \ldots, \mathbf{c}^l_n\rangle, \langle \mathbf{c}'^l_1, \ldots, \mathbf{c}'^l_n\rangle\big) \in par\big(\subseteq^4, \ldots, \subseteq^4\big)$$
$$\big\langle\langle c^k_1, \ldots, c^0_1\rangle, \ldots, \langle c^k_n, \ldots, c^0_n\rangle\big\rangle \leqslant^2_l \big\langle\langle c'^k_1, \ldots, c'^0_1\rangle, \ldots, \langle c'^k_n, \ldots, c'^0_n\rangle\big\rangle$$
$$\text{iff} \quad \big(\langle c^l_1, \ldots, c^l_n\rangle, \langle c'^l_1, \ldots, c'^l_n\rangle\big) \in par(\subseteq, \ldots, \subseteq)$$

where the $\mathbf{c}^l_i$ in the first inequality are 4-tuples of sets of atoms, while the c^l_i in the second inequality are simply sets of atoms. Thus, $\leqslant^1_l$ and $\leqslant^2_l$ compare conflict vectors, by considering the conflicts between interpretations and knowledge bases at a given tolerance level l. We now have the following characterization.

Proposition 3. *Let P and P' be defined as before and assume that $\lambda_{(l,s_i,a)} = \lambda_l$ for every source s_i and atom a, with $\lambda_k < 1$. It holds that*

$$\Delta^{prior_\subseteq}(K_1, \ldots, K_n) \equiv \Delta^{confl_2}\big(K_1, \ldots, K_n; C; diff^1_{\mathcal{S}}, lex\big(\leqslant^1_k, \ldots, \leqslant^1_0\big)\big) \tag{41}$$
$$\Delta^{prior'_\subseteq}(K_1, \ldots, K_n) \equiv \Delta^{confl_2}\big(K_1, \ldots, K_n; C; diff^2_{\mathcal{S}}, lex\big(\leqslant^2_k, \ldots, \leqslant^2_0\big)\big) \tag{42}$$

with Δ^{confl_2} defined as in (8)*.*

This means that semantically, the use of preferred subsets in $\Delta^{prior_\subseteq}$ and $\Delta^{prior'_\subseteq}$ is very close to the standard approach of conflict-based merging. The key difference is that the conflict between two interpretations is not represented as a single set, but as a vector of (vectors of) sets, discriminating between conflicts that are more important than others, in the sense that they are more difficult to explain in terms of flexible usage of atoms.

5.4. Cardinality-based preferred subtheories

5.4.1. Merging operators

A further refinement can be obtained by looking only at consistent subsets of maximal cardinality. A consistent set of formulas $F = F_k \cup \cdots \cup F_0$ is a cardinality-based preferred subtheory of K, with F_l a subset of the formulas that appear in K with priority λ_l ($l \in \{0, \ldots, k\}$), iff for every consistent set of formulas $F' = F'_k \cup \cdots \cup F'_0$, with F'_l a subset of the formulas that appear in K with priority λ_l, it holds that either $\forall l \,.\, |F_l| = |F'_l|$ or there exists some $l^* \in \{0, \ldots, k\}$ such that $|F_l| = |F'_l|$ for all $l > l^*$ and $|F_{l^*}| > |F'_{l^*}|$. We write $Pref_{card}(K)$ for the set of all cardinality-based preferred subtheories of K. The corresponding merging operators are given by

$$\Delta^{prior_{card}}(K_1, \ldots, K_n) = forgetVar\Big(\bigvee Pref_{card}(P), A^{@s_1} \cup \cdots \cup A^{@s_n}\Big) \tag{43}$$

Again we consider the variant using P' instead of P, which we write as $\Delta^{prior'_{card}}$. Clearly, $\Delta^{prior_{card}}$ is a refinement of $\Delta^{prior_{\subseteq}}$, i.e.

$$\Delta^{prior_{card}}(K_1, \ldots, K_n) \models \Delta^{prior_{\subseteq}}(K_1, \ldots, K_n)$$

The choice between Δ^{poss}, $\Delta^{prior_{\subseteq}}$ and $\Delta^{prior_{card}}$ depends, in practice, on how the weights can be interpreted. Evidently, the weaker the merging operator, the weaker the assumptions that need to be made on the meaning of the weights. In particular, in the possibilistic approach, weights may be interpreted in a purely qualitative way: formulas with a higher weight are assumed to be more plausible than formulas with a lower weight. The intuitive interpretation of the priorities, in the case of $\Delta^{prior_{\subseteq}}$ is in terms of order-of-magnitudes of probabilities [37,38], which can be formally studied using Spohn's ordinal conditional functions [39], or equivalently using quantitative possibility theory [40]. The fact that a formula ϕ has a higher priority than a formula ψ then intuitively means that the probability that ϕ is true is an order-of-magnitude higher than the probability that ψ is true. Under this interpretation, the preferred subtheories may be seen as those subsets of formulas that most likely correspond to the formulas that are correct. The restriction to consistent subsets of maximal cardinality, when using $\Delta^{prior_{card}}$, is motivated from the additional assumption that formulas with the same priority have approximately the same probability of being true, and an assumption of independence, i.e. the probability that one formula is true, is independent from the probability that other formulas with the same priority are true.

Example 14. Continuing on Example 13, we find that the inclusion-preferred subtheory B_2 is not cardinality preferred, since B_1 satisfies strictly more formulas from P' with weight λ_1. Thus we obtain

$$\Delta^{card'_{\subseteq}}(K_1, K_2) \equiv C \cup \big\{K_1^{B_1} \wedge K_2^{B_1}\big\} \equiv C \cup \big\{mar(p), li(p, sa_{hg}), wo(p, m)\big\}$$

which is more informative, and less cautious, than the result we found in Example 13.

5.4.2. Semantic characterization

Next, we establish a semantic counterpart to $\Delta^{prior_{card}}$ and $\Delta^{prior'_{card}}$, and show the connection between these operators and the distance based framework. First, note that two "distances" can naturally be defined between two interpretations, at each tolerance level l:

$$d_l^1(I, J) = \big|J \setminus \langle I\rangle_{W^l}\big| + \big|I \setminus \langle J\rangle_{X^l}\big| + \big|[I]_{Y^l} \setminus J\big| + \big|[J]_{Z^l} \setminus I\big|$$
$$d_l^2(I, J) = \big|\big(J \setminus \langle I\rangle_{W^l}\big) \cup \big(I \setminus \langle J\rangle_{X^l}\big) \cup \big([I]_{Y^l} \setminus J\big) \cup \big([J]_{Z^l} \setminus I\big)\big|$$

The notion of distance that is considered here is very weak, obeying neither the triangle inequality nor symmetry. Nonetheless, we will refer to d_l^1 and d_l^2 as distances, to preserve the terminology of the distance-based framework. Clearly, in the specific case where $l = 0$, we have

$$d_l^2(I, J) = \frac{1}{2} d_l^1(I, J) = d_{Ham}(I, J)$$

while for $l > 0$, we find

$$d_l^1(I, J) \leqslant d_{l-1}^1(I, J), \qquad d_l^2(I, J) \leqslant d_{l-1}^2(I, J)$$

That is, for larger values of l, the distances d_l^1 and d_l^2 only consider the most important conflicts. The distance between two interpretations, independent of a given tolerance level l, can be represented as a $(k+1)$-dimensional vector:

$$d^1(I, J) = \big\langle d_k^1(I, J), \ldots, d_0^1(I, J)\big\rangle \tag{44}$$

The distance between an interpretation and a knowledge base K is then given by

$$d^1(I, K) = \min\big(\big\{d^1(I, J) \bigm| J \in [\![K]\!]\big\}, lex(\leqslant, \ldots, \leqslant)\big)$$

In particular, note that to compare distance vectors, we first look at the distances for higher tolerance levels l, as these reveal the most important discrepancies between interpretations, using the distances at lower tolerance levels only as a further refinement. The distance between an interpretation and a list of knowledge bases $\mathcal{K} = (K_1, \ldots, K_n)$ is then given by

$$d^1(I, \mathcal{K}) = d^1(I, K_1) + \cdots + d^1(I, K_n)$$

The only difference with standard distance-based merging, when considering $f = \sum$ in (4), is that the additions in the right-hand side are vector additions instead of additions between real numbers. It turns out that $d^1(I, \mathcal{K})$ and $d^2(I, \mathcal{K})$ correspond to the notion of distance underlying $\Delta^{prior_{card}}$ and $\Delta^{prior'_{card}}$. First, note that the closeness of interpretations to $\mathcal{K}$ can be compared by

$$I \leqslant^1 I' \quad \text{iff} \quad \left(d^1(I,\mathcal{K}), d^1\left(I',\mathcal{K}\right)\right) \in lex(\leqslant,\ldots,\leqslant) \tag{45}$$

In entirely the same way, we arrive at the relation $\leqslant^2$ by considering d_l^2 instead of d_l^1. We then have the following proposition.

Proposition 4. *Let P and P' be defined as before and assume that $\lambda_{(l,s_i,a)} = \lambda_l$ for every source s_i and atom a, with $\lambda_k < 1$. It holds that*

$$\Delta^{prior_{card}}(K_1,\ldots,K_n) \equiv \Delta^{dist}(K_1,\ldots,K_n; C; \leqslant_1) \tag{46}$$

$$\Delta^{prior'_{card}}(K_1,\ldots,K_n) \equiv \Delta^{dist}(K_1,\ldots,K_n; C; \leqslant_2) \tag{47}$$

Note that for $k = 0$, we obtain a syntactic encoding of distance-based merging with the Hamming distance, for the special (but common) case where the sum is used as aggregation operator; existing syntactic encodings of Hamming-distance based merging can be found in [5] and [41].

5.5. Partially ordered weights

5.5.1. Merging operators

The results in Propositions 2, 3 and 4 are based on the assumption that $\lambda^{\times}_{(l,s_i,a)} = \lambda_{(l,s_i,a)} = \lambda_l$ for every $\times \in \{W, X, Y, Z\}$, source s_i and atom a. Due to the assumption that $l_1 \leqslant l_2$ implies $\lambda^{\times}_{(l_1,s_i,a)} \leqslant \lambda^{\times}_{(l_2,s_i,a)}$ and $\lambda_{(l_1,s_i,a)} \leqslant \lambda_{(l_2,s_i,a)}$, this means in particular that all priorities are totally ordered. In practice, however, this requirement is often too strong. Consider again the weather forecast scenario from Example 11. It may be the case that some source s_1 is known to be more reliable than another source s_2, in the sense that predictions from s_1 are likely to be closer to the truth than predictions from s_2. In such a situation, however, it is often not possible to exactly quantify the difference in reliability between s_1 and s_2. For example, we may be in a case where we can reasonably assume that $\lambda_{(1,s_1,os)} > \lambda_{(1,s_2,os)}$, without having sufficient information to decide whether $\lambda_{(1,s_1,os)} > \lambda_{(2,s_2,os)}$, $\lambda_{(1,s_1,os)} = \lambda_{(2,s_2,os)}$ or $\lambda_{(1,s_1,os)} < \lambda_{(2,s_2,os)}$. Moreover, even the assumption that, for a fixed source s, $\lambda_{(1,s,os)} = \lambda_{(1,s,oc)}$ may be considered too strict. It may, for instance, be the case that os is less likely to be accurate than oc (e.g. because the sources prefer an optimistic attitude when available evidence is inconclusive). In general, we will typically not be able to exactly quantify the "amount of stretching" that is needed to go from one atom to another, and as a consequence, insisting that $\lambda_{(l,s,a)} = \lambda_{(l,s,b)}$ for $a \neq b$ may be considered too imprudent.

We now show how the merging operators that have been proposed may be adapted to cope with partially ordered certainty weights. The possibilistic merging operators may still be used, provided that we extend the notion of λ-cut to partially ordered certainty weights. Several such extensions have been proposed in [42], based on the idea of selecting particular maximal consistent subsets. In particular if K_1 and K_2 are two subsets of a possibilistic knowledge base K with partially ordered weights, we consider the following relation:

$$K_1 \preccurlyeq K_2 \quad \text{iff} \quad \forall(\alpha_1,\lambda_1) \notin K_1 \,.\, \exists(\alpha_2,\lambda_2) \notin K_2 \,.\, \lambda_1 \leqslant \lambda_2$$

Let $Pref_{\preccurlyeq}(K)$ be the corresponding preferred subtheories, i.e.

$$Pref_{\preccurlyeq}(K) = \left\{B^* \,\middle|\, B \in \min\left(Cons(K), \preccurlyeq\right)\right\}$$

where we write $Cons(K)$ for the consistent subsets of K, and $B^* = \{\alpha \mid (\alpha,\lambda) \in K\}$. We may then consider the following merging operator:

$$\Delta^{poss}(K_1,\ldots,K_n) = forgetVar\left(\bigvee Pref_{\preccurlyeq}(P), A^{@s_1} \cup \cdots \cup A^{@s_n}\right)$$

and the variant $\Delta^{poss'}$ which is based on P'. It is easy to see that when $\lambda^{\times}_{(l,s_i,a)} = \lambda_{(l,s_i,a)} = \lambda_l$, we have indeed that both Δ^{poss} and $\Delta^{poss'}$ coincide with (33).

Example 15. When going back to the weather forecast of Example 11, we may notice that Δ^{poss}, $\Delta^{prior_{\subseteq}}$ and $\Delta^{prior_{card}}$ all lead to the same conclusion, viz. $\{pc_2\}$. While the use of inclusion- and cardinality-based preferred subtheories refines the possibilistic approach, by avoiding the drowning effect, these strategies are essentially based on the same intuition of priority. This intuition dictates that pc_2 is a more plausible conclusion than e.g. pc_1 as pc_1 is intuitively further away from oc than pc_2 is away from either os or oc. In other words, what is maximized is the minimal similarity between models of the resulting base and any given knowledge base K_1. In practice, however, we may be less confident in where exactly is the middle between os and oc. Assume that the weights are ordered as follows:

$$\lambda_{(l,s,a)} \leqslant \lambda_{(l',s',a')} \quad \text{iff} \quad \left(a = a' \wedge l \leqslant l'\right) \vee \left(l \leqslant l' - \theta\right) \tag{48}$$

with $\theta = 1$. Then $Pref_{\preccurlyeq}(P')$ contains three subsets:

$$B_{pc_1} = P'^* \setminus \{\alpha_{(0,s_1,os)}, \alpha_{(0,s_2,os)}, \alpha_{(0,s_3,oc)}, \alpha_{(1,s_3,oc)}, \alpha_{(2,s_3,oc)}\}$$
$$B_{pc_2} = P'^* \setminus \{\alpha_{(0,s_1,os)}, \alpha_{(1,s_1,os)}, \alpha_{(0,s_2,os)}, \alpha_{(1,s_2,os)}, \alpha_{(0,s_3,oc)}, \alpha_{(1,s_3,oc)}\}$$
$$B_{pc_3} = P'^* \setminus \{\alpha_{(0,s_1,os)}, \alpha_{(1,s_1,os)}, \alpha_{(2,s_1,os)}, \alpha_{(0,s_2,os)}, \alpha_{(1,s_2,os)}, \alpha_{(2,s_2,os)}, \alpha_{(0,s_3,oc)}\}$$

which leads to

$$\Delta^{poss}(K_1, K_2, K_3) \equiv C \cup \{pc_1 \vee pc_2 \vee pc_3\}$$

In the same way, choosing $\theta = 2$, we obtain the trivial conclusion

$$\Delta^{poss}(K_1, K_2, K_3) \equiv C \cup \{os \vee pc_1 \vee pc_2 \vee pc_3 \vee oc\}$$

Finally, note that by encoding information about the relative reliability of different sources, results may be obtained that are not centered around pc_2.

The idea of inclusion-based preferred subtheories of a prioritized knowledge base, in the case where priorities are partially ordered, was already proposed in [36]. The basic idea is to consider all possible linearizations κ. Specifically, let a linearization κ be any mapping from Λ to $[0, 1]$ satisfying $\lambda_1 \leqslant \lambda_2 \Rightarrow \kappa(\lambda_1) \leqslant \kappa(\lambda_2)$, and let us write $Lin(\Lambda, \leqslant)$ to denote the set of all such linearizations. Given a linearization κ, the linearized version $\kappa(K)$ of the prioritized knowledge base K is obtained by replacing all weights λ by their value in $[0, 1]$:

$$\kappa(K) = \{(\alpha, \kappa(\lambda)) \mid (\alpha, \lambda) \in K\}$$

The preferred subtheories of K are then simply the maximal consistent subsets that are preferred for $\kappa(K)$, for some linearization κ. Thus we arrive at the following merging operators:

$$\Delta^{prior_\subseteq}(K_1, \ldots, K_n) = forgetVar\left(\bigvee_\kappa \bigvee Pref_\subseteq(\kappa(P)), A^{@s_1} \cup \cdots \cup A^{@s_n}\right)$$
$$\Delta^{prior_{card}}(K_1, \ldots, K_n) = forgetVar\left(\bigvee_\kappa \bigvee Pref_{card}(\kappa(P)), A^{@s_1} \cup \cdots \cup A^{@s_n}\right)$$

as well as the variants $\Delta^{prior'_\subseteq}$ and $\Delta^{prior'_{card}}$ which are based on P'. It is trivial to see that when $\lambda^\times_{(l,s_i,a)} = \lambda_{(l,s_i,a)} = \lambda_l$, these definitions coincide with (35), (43), (35) and (43) respectively.

5.5.2. Semantic characterization

Thus, in the special case where all weights are totally ordered, the definitions of the operators we have presented here to cope with partially ordered weights coincide with those that were presented in the previous sections. At the other extreme, when all weights $\lambda_{(l_i,s_i,a_i)}$ and $\lambda_{(l_j,s_j,a_j)}$ are incomparable unless $s_i = s_j$ and $a_i = a_j$, it turns out that $\Delta^{prior_\subseteq}$, $\Delta^{prior_{card}}$ and Δ^{poss}, as well as $\Delta^{prior'_\subseteq}$, $\Delta^{prior'_{card}}$ and $\Delta^{poss'}$ coincide.

Proposition 5. *Let P and P' be defined as before and assume that $\lambda^{\times_i}_{(l_i,s_i,a_i)} \leqslant \lambda^{\times_j}_{(l_j,s_j,a_j)}$ iff $l_i \leqslant l_j$, $s_i = s_j$, $a_i = a_j$ and $\times_i = \times_j$. It holds that*

$$\Delta^{confl_2}(K_1, \ldots, K_n; C; diff^1_{\mathcal{S}}, par(\leqslant^1_k, \ldots, \leqslant^1_0)) \equiv \Delta^{poss}(K_1, \ldots, K_n) \tag{49}$$
$$\equiv \Delta^{prior_\subseteq}(K_1, \ldots, K_n) \tag{50}$$
$$\equiv \Delta^{prior_{card}}(K_1, \ldots, K_n) \tag{51}$$

Furthermore, when $\lambda_{(l_i,s_i,a_i)} \leqslant \lambda_{(l_j,s_j,a_j)}$ iff $l_i \leqslant l_j$, $s_i = s_j$ and $a_i = a_j$, it holds that

$$\Delta^{confl_2}(K_1, \ldots, K_n; C; diff^2_{\mathcal{S}}, par(\leqslant^2_k, \ldots, \leqslant^2_0)) \equiv \Delta^{poss'}(K_1, \ldots, K_n) \tag{52}$$
$$\equiv \Delta^{prior'_\subseteq}(K_1, \ldots, K_n) \tag{53}$$
$$\equiv \Delta^{prior'_{card}}(K_1, \ldots, K_n) \tag{54}$$

The fact that merging operators correspond to the Pareto extensions $par(\leqslant^1_k, \ldots, \leqslant^1_0)$ and $par(\leqslant^2_k, \ldots, \leqslant^2_0)$, rather than the lexicographic extensions as in Proposition 3 confirms our intuition that by being more cautious in defining the ordering on Λ, a more tolerant merging operator is obtained which provides more cautious results. In practice, useful merging operators may be somewhere in between, adopting a balance between cautiousness and informativity.

As a general remark, let us also mention that when the weights are maximally incomparable, as in Proposition 5, the different approaches recover the intersection of the maximal consistent subsets. In this way, these approaches implement the idea that was suggested after Example 12, to apply the possibilistic approach, but only looking at those formulas that are not involved in any conflict.

5.6. Penalties

5.6.1. Merging operators

The treatment of partially ordered weights in the previous section deals with situations where less information is available than is required by the merging operators from Sections 5.2–5.4. In this section, we show how penalty logic, conversely, allows to handle the case where precise information is available about the certainty of the statements in M_{s_i} or M'_{s_i}. To this end, we assume that the weights $\lambda^{\times}_{(l,s,a)}$ and $\lambda_{(l,s,a)}$ are non-negative integers, i.e. $0 \leqslant \lambda^{\times}_{(l,s,a)}, \lambda_{(l,s,a)} < +\infty$. The weighted knowledge bases M_{s_i} and M'_{s_i} then correspond to theories in penalty logic[6] [27,28]. In this case, rather than encoding an ordering, the weights have a numerical interpretation with a probabilistic flavor.

Let us define the following penalty logic bases:

$$Q = \bigcup_i (M_{s_i})_\lambda \cup \bigcup_i \{(\alpha, +\infty) \mid \alpha \in K_i^{@s_i}\} \cup \{(c, +\infty) \mid c \in C\} \tag{55}$$

$$Q' = \bigcup_i (M'_{s_i})_\lambda \cup \bigcup_i \{(\alpha, +\infty) \mid \alpha \in K_i^{@s_i}\} \cup \{(c, +\infty) \mid c \in C\} \tag{56}$$

and let us write Q^* and Q'^* for the corresponding classical knowledge bases that are obtained by ignoring the penalties. In [28], some relationships between penalty logic and Dempster–Shafer theory [43,44] are revealed, essentially suggesting to interpret a penalty logic formula (α, p) as an upper bound on the probability that α is violated: $P(\neg\alpha) \leqslant e^{-p}$. When the correctness of formulas appearing in a penalty logic base K is independent of the correctness of other formulas, the probability that a set of formulas (α_i, p_i) from K are violated in the true world is then upper bounded by $\prod_i e^{-p_i} = e^{-\sum_i p_i}$. This independence assumption is, however, quite strong, and presupposes that no formula is entailed by a subset of the remaining formulas. In particular, the independence assumption cannot be made in the case of Q and Q', since we already have $\alpha^{\times}_{(l-1,s,a)} \rightarrow \alpha^{\times}_{(l,s,a)}$ for any source s, atom a, $\times \in \{W, X, Y, Z\}$, and $l > 0$. However, in the case of Q', it seems natural to interpret $\lambda_{(l,s,a)}$ by

$$P(\neg\alpha_{(l,s,a)} \mid \neg\alpha_{(l-1,s,a)}) \leqslant e^{-\lambda_{(l,s,a)}}$$

for $l > 0$, and

$$P(\neg\alpha_{(0,s,a)}) \leqslant e^{-\lambda_{(0,s,a)}}$$

This only presupposes that the amount of tolerance required for a given atom when considering a given source, is independent of the amount of tolerance required for other atoms and other sources. A similar interpretation could be given to the penalties in Q', although this requires additional information about the probability that the four implications $\alpha^W_{(l,s,a)}$, $\alpha^X_{(l,s,a)}$, $\alpha^Y_{(l,s,a)}$ and $\alpha^Z_{(l,s,a)}$ are correct, and about the dependencies between these probabilities.

Given a penalty logic base K, we may consider the following relation between subsets B_1 and B_2 of formulas that appear in a weighted knowledge base K:

$$B_1 \leqslant_p B_2 \quad \text{iff} \quad pen_K(B_1) \leqslant pen_K(B_2)$$

where $pen_K(B_i) = \sum_{(\alpha,p)\in K, \alpha \notin B_i} p$ (see Appendix A for further details). We write $Pref_{pen}(K)$ to denote the subsets of formulas that are minimal w.r.t. $\leqslant_p$. Merging operators corresponding to Q and Q' can then be defined as follows:

$$\Delta^{pen}(K_1, \ldots, K_n) = forgetVar\Big(\bigvee Pref_{pen}(Q), A^{@s_1} \cup \cdots \cup A^{@s_n}\Big) \tag{57}$$

$$\Delta^{pen'}(K_1, \ldots, K_n) = forgetVar\Big(\bigvee Pref_{pen}(Q'), A^{@s_1} \cup \cdots \cup A^{@s_n}\Big) \tag{58}$$

Now assume, as in Sections 5.2–5.4 that $\lambda_{(l,s_i,a)}$ is independent of the source s_i and the atom a, and depends only on the tolerance level l, i.e. $\lambda_{(l,s_i,a)} = \lambda_l$. From an application point of view, a natural choice seems to assume that λ_l is equal to a constant. As the exact value of this constant does not affect the result of the merging operators (57) and (58), we may assume this constant to be 1. It is not hard to see that in this case, the elements of $Pref_{pen}(Q)$ and $Pref_{pen}(Q')$ simply are the consistent subsets with maximal cardinality of Q and Q'. Note that in this case, we have the following interpretation

[6] A brief introduction to penalty logic is provided in Appendix A.

$$P(\neg\alpha_{(l,s_i,a)}) = P(\neg\alpha_{(l,s_i,a)} \mid \neg\alpha_{(l-1,s_i,a)}) \cdot \cdots \cdot P(\neg\alpha_{(0,s_i,a)}) \leqslant \gamma^{l+1}$$

for some constant γ.

As a second important interpretation of the penalties λ_l, we consider the case where tolerance levels correspond to different order-of-magnitudes of probabilities. In this view, the probability $P(\neg\alpha_{(l,s_i,a)} \mid \neg\alpha_{(l-1,s_i,a)})$ should be decreasing in l. In particular, we assume that this conditional probability is upper bounded by ϵ^{r^l} for some $\epsilon \in]0,1[$ and a sufficiently large r. The corresponding penalty is of the form $\lambda_l = \gamma \cdot r^l$, where in fact $\gamma = -\log(\epsilon)$. Note that the exact value of $\gamma > 0$ does not influence the result, hence we may assume $\gamma = 1$. Recall that the motivation of using cardinality-based preferred subtheories was also in terms of order-of-magnitudes of probabilities. Accordingly, we have the following proposition.

Proposition 6. *Let Q and Q' be defined as before, and assume that $\lambda^{\times}_{(l,s,a)} = \lambda_{(l,s,a)} = \lambda_l = r^l$ for some $r \in \mathbb{N}$. If r is sufficiently large, it holds that*

$$\Delta^{pen}(K_1,\ldots,K_n) \equiv \Delta^{prior_{card}}(K_1,\ldots,K_n) \tag{59}$$

$$\Delta^{pen'}(K_1,\ldots,K_n) \equiv \Delta^{prior'_{card}}(K_1,\ldots,K_n) \tag{60}$$

In addition to constant weights ($\lambda_l = 1$) and exponential weights ($\lambda_l = r^l$), we may also consider weights that are proportional to $l+1$, i.e. choose $\lambda_l = l+1$. Note that $l+1$ is considered rather than l to ensure a non-zero weight for the case where $l = 0$.

Example 16. Consider again the weather scenario from Example 11. Recall that the approaches from Sections 5.2–5.4 all yield the same result pc_2, while the more cautious alternative $pc_1 \vee pc_2 \vee pc_3$ and the trivial conclusion $os \vee pc_1 \vee pc_2 \vee pc_3 \vee oc$ can be obtained using the partially ordered weights from Section 5.5. Example 11 suggests a number of other possibilities, which, as we will see, can be obtained by interpreting the penalties as weights.

There are five maximal consistent subsets of Q':

$$B_{os} = Q'^{*} \setminus \{\alpha_{(0,s_3,oc)}, \alpha_{(1,s_3,oc)}, \alpha_{(2,s_3,oc)}, \alpha_{(3,s_3,oc)}\}$$
$$B_{pc_1} = Q'^{*} \setminus \{\alpha_{(0,s_1,os)}, \alpha_{(0,s_2,os)}, \alpha_{(0,s_3,oc)}, \alpha_{(1,s_3,oc)}, \alpha_{(2,s_3,oc)}\}$$
$$B_{pc_2} = Q'^{*} \setminus \{\alpha_{(0,s_1,os)}, \alpha_{(1,s_1,os)}, \alpha_{(0,s_2,os)}, \alpha_{(1,s_2,os)}, \alpha_{(0,s_3,oc)}, \alpha_{(1,s_3,oc)}\}$$
$$B_{pc_3} = Q'^{*} \setminus \{\alpha_{(0,s_1,os)}, \alpha_{(1,s_1,os)}, \alpha_{(2,s_1,os)}, \alpha_{(0,s_2,os)}, \alpha_{(1,s_2,os)}, \alpha_{(2,s_2,os)}, \alpha_{(0,s_3,oc)}\}$$
$$B_{oc} = Q'^{*} \setminus \{\alpha_{(0,s_1,os)}, \alpha_{(1,s_1,os)}, \alpha_{(2,s_1,os)}, \alpha_{(3,s_1,os)}, \alpha_{(0,s_2,os)}, \alpha_{(1,s_2,os)}, \alpha_{(2,s_2,os)}, \alpha_{(3,s_2,os)}\}$$

When the penalties in M'_{s_i} are constant, the following penalties are obtained (for $\lambda_l = 1$):

$$pen_{Q'}(B_{os}) = 4, \qquad pen_{Q'}(B_{pc_1}) = 5, \qquad pen_{Q'}(B_{pc_2}) = 6, \qquad pen_{Q'}(B_{pc_3}) = 7, \qquad pen_{Q'}(B_{oc}) = 8$$

Hence only B_{os} is minimal which leads to

$$\Delta^{pen'}(K_1,K_2,K_3) \equiv C \cup \{os\}$$

When considering exponential penalties of the form $\lambda_l = r^l$, with a sufficiently large r, we obtain using (58)

$$\Delta^{pen'}(K_1,K_2,K_3) \equiv C \cup \{pc_2\}$$

Finally, when considering penalties of the form $\lambda_l = l+1$, we find

$$pen_{Q'}(B_{os}) = 10, \qquad pen_{Q'}(B_{pc_1}) = 8, \qquad pen_{Q'}(B_{pc_2}) = 9, \qquad pen_{Q'}(B_{pc_3}) = 13, \qquad pen_{Q'}(B_{oc}) = 20$$

leading to

$$\Delta^{pen'}(K_1,K_2,K_3) \equiv C \cup \{pc_1\}$$

Clearly, by using constant penalties of the form $\lambda_l = 1$, the result reflects the opinion of the majority. By choosing exponential penalties of the form $\lambda_l = r^l$, for a sufficiently large r, we obtain an opinion in the middle, as for the operators that were discussed in Sections 5.2–5.4. Finally, using linearly increasing penalties of the form $\lambda_l = r+1$, the result intuitively reflects the center-of-gravity of the opinions held by the sources. This latter behavior is related to least squares approximation, which can be made explicit as follows. For each atom a, let us write $l^*_{(a,s)}$ to denote the largest value of l for which $\alpha_{(l,s,a)}$ is violated in the real world, where $l^*_{(a,s)} = -1$ if even $\alpha_{(0,s,a)}$ is satisfied. Then $l^*_{(a,s)} + 1$ measures how abnormal the situation w.r.t. atom a and source s is. If we make the assumption that the abnormality $l^*_{(a,s)} + 1$ is, in fact, the discrete approximation of a continuous parameter which follows a normal distribution, and the degree of abnormality of each atom a and source

s is independent, the most likely situation is recovered by $\Delta^{pen'}$ when the penalties are such that $\sum_{i=0}^{l} \lambda_i = \gamma \cdot (l+1)^2$ for some constant γ. This is due to the well-known correspondence between least squares and maximum likelihood estimation of normally distributed variables. By taking $\lambda_l = l+1$, we find $\sum_{i=0}^{l} \lambda_i = \frac{1}{2} \cdot (l+1) \cdot (l+2)$, which may indeed serve as an approximation of $\gamma \cdot (l+1)^2$ for $\gamma = \frac{1}{2}$.

Depending on how the penalties are chosen, we either find a solution in the middle between the beliefs of the different sources, a solution expressing the majority opinion, or the centre-of-gravity of the beliefs of the sources. Formally, however, it is not hard to show that for every non-trivial choice of penalties (i.e. $0 < \lambda_{(l,s,a)} < +\infty$ for all l, s and a), Δ^{pen} and $\Delta^{pen'}$ are majority operators in the sense of [17]:

$$\forall K_1, K_2 \,.\, \exists n \in \mathbb{N} \,.\, \Delta^{pen}(\underbrace{K_1, \ldots, K_1}_{n}, K_2) \models K_1$$

and similar for $\Delta^{pen'}$. On the other hand, we do have the intuition that the more λ_l increases with l, the less the notion of majority plays a role. Interestingly, such *weak majority* operators are also studied in [45], considering a distance-based merging operator which is based on the square of the Hamming distance.

Finally, note that, due to their quantitative nature, a more cautious variant of Δ^{pen} and $\Delta^{pen'}$ can straightforwardly be defined, by being more tolerant in the definition of $Pref_{pen}(Q)$ and $Pref_{pen}(Q')$. For instance, we may define

$$B_1 \in Pref_{pen}(Q) \quad \text{iff} \quad pen_Q(B_1) \leqslant \gamma \cdot \min_B pen_Q(B)$$

for some constant $\gamma \geqslant 1$. Thus, in the scenario of Example 11, results such as $os \vee pc_1$, or $os \vee pc_1 \vee pc_2$ may be obtained.

5.6.2. Semantic characterization

Semantically, Δ^{pen} and $\Delta^{pen'}$ can be described in the distance-based framework, provided that the penalties do not depend on the underlying source,[7] i.e. for each $\times \in \{W, X, Y, Z\}$, tolerance level l, source s and atom a, $\lambda^{\times}_{(l,s,a)} = \lambda^{\times}_{(l,a)}$ for some $\lambda^{\times}_{(l,a)} \in [0, +\infty]$, and $\lambda_{(l,s,a)} = \lambda_{(l,a)}$ for some $\lambda_{(l,a)} \in [0, +\infty]$. Then, we may consider the following distances:

$$\begin{aligned} d^1_Q(I, J) &= \sum \bigl\{\lambda^W_{(l,a)} \bigm| l \in \{0, \ldots, k\},\ a \in J \setminus \langle I \rangle_{W^l}\bigr\} + \sum \bigl\{\lambda^X_{(l,a)} \bigm| l \in \{0, \ldots, k\},\ a \in I \setminus \langle J \rangle_{X^l}\bigr\} \\ &\quad + \sum \bigl\{\lambda^Y_{(l,a)} \bigm| l \in \{0, \ldots, k\},\ a \in [I]_{Y^l} \setminus J\bigr\} + \sum \bigl\{\lambda^Z_{(l,a)} \bigm| l \in \{0, \ldots, k\},\ a \in [J]_{Z^l} \setminus I\bigr\} \\ d^2_Q(I, J) &= \sum \bigl\{\lambda_{(l,a)} \bigm| l \in \{0, \ldots, k\}, a \in J \setminus \langle I \rangle_{W^l} \cup I \setminus \langle J \rangle_{X^l} \cup [I]_{Y^l} \setminus J \cup [J]_{Z^l} \setminus I\bigr\} \end{aligned}$$

Accordingly, we may define the distance between an interpretation and a knowledge base, or an interpretation and a list of knowledge bases, exactly as in the distance-based framework. This leads to the following proposition.

Proposition 7. *Let Q and Q' be defined as before and assume that $\lambda^{\times}_{(l,s_i,a)} = \lambda^{\times}_{(l,a)} < +\infty$ and $\lambda_{(l,s_i,a)} = \lambda_{l,a} < +\infty$ for every source s_i and atom a. It holds that*

$$\Delta^{pen}(K_1, \ldots, K_n) \equiv \Delta^{dist}(K_1, \ldots, K_n; C; \leqslant_{d^1_Q}) \tag{61}$$

$$\Delta^{pen'}(K_1, \ldots, K_n) \equiv \Delta^{dist}(K_1, \ldots, K_n; C; \leqslant_{d^2_Q}) \tag{62}$$

When comparing Proposition 7 with Proposition 4, it becomes obvious that Δ^{pen} and $\Delta^{prior_{card}}$ present two different solutions to deal with the fact that the distance between interpretations is most naturally represented as a vector in our setting, which was defined in (44). While Δ^{pen} uses a quantitative approach, which aggregates such vectors to scalar values, before proceeding in the standard distance-based framework (albeit with a non-standard distance), $\Delta^{prior_{card}}$ follows a more qualitative approach, in which a lexicographic extension of the distance-based framework is rather used.

6. Computational complexity

In this section, we investigate the computational complexity of our merging operators. First note that the proposed merging operators have syntactically been defined in existing logical formalisms. Hence existing algorithms for reasoning in e.g. possibilistic logic or penalty logic can readily be reused to implement the merging operators. Moreover, the membership results for these existing frameworks immediately carry over to the present setting.

Typically, propositional merging tasks are at the lower levels of the polynomial hierarchy. Recall that the complexity classes, Δ^P_k, Σ^P_k and Π^P_k, which constitute the polynomial hierarchy, are defined as follows ($i \in \mathbb{N}$) [46]:

[7] The general case can be treated by a straightforward generalization of the distance-based framework in which a difference distance is used for each source; we omit the details.

Table 2
Complexity of entailment checking for different merging operators.

Δ^{poss}	linear weights	Θ_2^P-complete
	partially ordered weights	Π_2^P-complete
$\Delta^{prior_\subseteq}$	linear weights	Π_2^P-complete
	partially ordered weights	Π_2^P-complete
$\Delta^{prior_{card}}$	linear weights	Δ_2^P-complete
	partially ordered weights	Π_2^P-complete
Δ^{pen}	exponentially bounded penalties	Δ_2^P-complete
	polynomially bounded penalties	Θ_2^P-complete

$$\Delta_0^P = \Sigma_0^P = \Pi_0^P = \mathrm{P}, \qquad \Delta_{i+1}^P = \mathrm{P}^{\Sigma_i^P}, \qquad \Sigma_{i+1}^P = \mathrm{NP}^{\Sigma_i^P}, \qquad \Pi_{i+1}^P = co\big(\Sigma_{i+1}^P\big)$$

where $\mathrm{NP}^{\Sigma_i^P}$ (resp. $\mathrm{P}^{\Sigma_i^P}$) is the class of problems that can be solved in polynomial time on a non-deterministic machine (resp. deterministic machine) with a Σ_i^P oracle, i.e. assuming a procedure that can solve Σ_i^P problems in constant time. All problems in the polynomial hierarchy are solvable with a polynomial amount of space and an exponential amount of time, e.g. using systematic search based on branch-and-bound. In particular, so far, no complexity class of the polynomial hierarchy is known to be strictly harder than P. Knowing which class of the polynomial hierarchy a given problem belongs to, however, is still important, as it serves as an indication of how hard we can expect the problem to be in practice. In the following, we also refer to the class Θ_i^P which contains the problems that are solvable in polynomial time using a logarithmic number of calls to a Σ_i^P oracle [47].

For each merging operator Δ, the main decision problem consists of checking whether $\Delta(K_1,\ldots,K_n; C; M_{s_1},\ldots,M_{s_n}) \models \phi$ for given knowledge bases K_i, integrity constraints C, weighted knowledge bases M_{s_i} and a propositional formula ϕ over the set of atoms A. Note that the computational complexity is not affected if the sets W_a^l, X_a^l, Y_a^l and Z_a^l are given as input instead of the weighted knowledge bases M_{s_i}. The complexity classes of the merging operators that were studied in Section 5 are summarized in Table 2. From a complexity point of view, it does not matter whether P or P' is considered (Q or Q' in the case of penalties). The following proposition summarizes the main results.

Proposition 8. *Let $K_1,\ldots,K_n$ be propositional knowledge bases over a set of atoms A, C a set of propositional integrity constraints over A, and let $M_{s_1},\ldots,M_{s_n}$ be defined as in* (16)*. For a propositional formula ϕ, the complexity of deciding whether*

$$\Delta(K_1,\ldots,K_n; C; M_{s_1},\ldots,M_{s_n}) \models \phi$$

holds is summarized for different merging operators Δ in Table 2*.*

Note that by exponentially bounded penalties, we mean that there exists an exponential function f of the problem size n, such that the value of all penalties is at most $f(n)$. In other words, the number of bits required to encode the penalties is polynomial in n. Polynomially bounded penalties are then upper bounded by $g(n)$ for a polynomial function g. In other words, the cases of exponentially and polynomially bounded queries respectively assume that penalties are encoded using binary and unary notation. Furthermore, note that we do not need to make such a distinction in the other cases, as possibilistic certainty weights, priorities and symbolic weights are all interpreted in an ordinal way.

In some sense, the complexity results are unsurprising, as it is well-known that entailment relations based on e.g. possibilistic logic or inclusion-based preferred subtheories are Θ_2^P and Π_2^P respectively. What the proposition shows is mainly that the restricted setting in which possibilistic logic, preferred subtheories, and penalty logic are used does not cause a decrease in complexity.

Overall, we can see that the possibilistic and penalty-logic based interpretations of the weights yield the most efficient procedures. Moreover, the complexity of Δ^{poss} depends on the number of different weights. The smaller this number, the fewer calls to the NP-oracle are required. In particular, in the case of $k+1$ different weights $\lambda_0,\ldots,\lambda_k$, entailment can be verified by solving $1 + \lceil \log_2(k+1) \rceil$ instances of the boolean satisfiability problem SAT, viz. $\lceil \log_2(k+1) \rceil$ instances for calculating the inconsistency level and 1 instance for entailment checking. In the same way, the number of required satisfiability checks for $\Delta^{prior_{card}}$ is given by $k+2$, when the weights are linearly ordered. On the other hand, $\Delta^{prior_\subseteq}$ (with linearly or partially ordered weights) and $\Delta^{prior_{card}}$ with partially ordered weights remain computationally hard, even in the case where $k=1$. In the case of Δ^{pen}, the number of required satisfiability checks depends on both k and the actual values of the penalties. In addition to these effects of restricting the value of k, in all cases, efficient procedures can be obtained by restricting the number of atoms that may be weakened, i.e. by taking $W_a^l = X_a^l = Y_a^l = Z_a^l = \{a\}$ for all l and for all but a few atoms a.

When weights are interpreted as priorities, a number of entailment relations may be considered that are not based on calculating the merged knowledge bases $\Delta^{prior_\subseteq}(K_1,\ldots,K_n)$ and $\Delta^{prior_{card}}(K_1,\ldots,K_n)$. When the weights are linearly ordered, well-known entailment relations $\approx_\subseteq^\exists$ and $\approx_{card}^\exists$ are defined as follows:

Table 3
Complexity of alternative entailment relations based on priorities.

Entailment relation	Complexity	Entailment relation	Complexity
$\approx_{\subseteq}^{\exists}$	Σ_2^P	$\approx_{card}^{\exists}$	Σ_2^P
$\approx_{\subseteq}^{\exists\exists}$	Σ_2^P	$\approx_{card}^{\exists\exists}$	Σ_2^P
$\approx_{\subseteq}^{\exists\forall}$	Σ_3^P	$\approx_{card}^{\exists\forall}$	Σ_2^P
$\approx_{\subseteq}^{\forall\exists}$	Π_3^P	$\approx_{card}^{\forall\exists}$	Π_3^P

$$(K_1,\ldots,K_n) \approx_{\subseteq}^{\exists} \phi \quad \text{iff} \quad \exists B \in \mathit{Pref}_{\subseteq}(P) \,.\, B \models \phi$$

$$(K_1,\ldots,K_n) \approx_{card}^{\exists} \phi \quad \text{iff} \quad \exists B \in \mathit{Pref}_{card}(P) \,.\, B \models \phi$$

Hence, rather than looking for conclusions that are common to all preferred subtheories, we only require that ϕ is a conclusion of one of the preferred subtheories. Of course, this means that it may happen that both ϕ and $\neg\phi$ can be derived. When weights are partially ordered, even more alternatives may be conceived:

$$(K_1,\ldots,K_n) \approx_{\subseteq}^{\exists\exists} \phi \quad \text{iff} \quad \exists\kappa\exists B \in \mathit{Pref}_{\subseteq}\big(\kappa(P)\big) \,.\, B \models \phi$$

$$(K_1,\ldots,K_n) \approx_{\subseteq}^{\exists\forall} \phi \quad \text{iff} \quad \exists\kappa\forall B \in \mathit{Pref}_{\subseteq}\big(\kappa(P)\big) \,.\, B \models \phi$$

$$(K_1,\ldots,K_n) \approx_{\subseteq}^{\forall\exists} \phi \quad \text{iff} \quad \forall\kappa\exists B \in \mathit{Pref}_{\subseteq}\big(\kappa(P)\big) \,.\, B \models \phi$$

and similar for the alternatives based on Pref_{card}. Note that $\approx_{\subseteq}^{\exists\forall}$ could, for instance, be used to verify whether there exists an assessment of the relative reliability of the sources, such that a formula ϕ could be concluded after merging the knowledge bases. Similarly, $\approx_{\subseteq}^{\forall\exists}$ could be used to verify whether ϕ is a plausible consequence, independent of which sources are considered most reliable.

Interestingly, in some cases, the complexity goes up a level in the polynomial hierarchy when these alternative entailment relations are used.

Proposition 9. *Let $K_1,\ldots,K_n$ be propositional knowledge bases over a set of atoms A, C a set of propositional integrity constraints over A, and let $M_{s_1},\ldots,M_{s_n}$ be defined as before. For a propositional formula ϕ, the complexity of deciding whether*

$$(K_1,\ldots,K_n; C; M_{s_1},\ldots,M_{s_n}) \approx \phi$$

holds is summarized for different entailment relations $\approx$ in Table 3.

7. Related work

Although many approaches to merging propositional knowledge bases have already been studied, we are not aware of any proposals that model the fact that different atoms may be closely related in meaning, or indeed, that we may be uncertain about how exactly we should understand a given assertion. On the other hand, the fact that uncertainty in meaning (or if we prefer 'vagueness') is pervading social interaction has been recognized early in AI [48]. This phenomenon has also been extensively studied by (cognitive) linguists, especially in the context of dialogues, which can be considered as the simplest form of social interaction. The reasons for adopting vague language in conversations may be many. For instance, vague language may help the listener to determine how much processing effort should be devoted to a given utterance, focusing him or her to the most relevant information; it may indicate a lack of certainty about the exact state of affairs; or it may be used to serve social functions, such as softening implicit complaints and criticism [49]. Whether or not vagueness in conversations lead to misconceptions depends on the participants' ability to relate what is expressed to their common ground. Thus, successful communication depends on the establishment of such a common ground, a process called grounding [50]. This common ground allows for a common language on which the points of view of the speaker and listener in a dialogue can be aligned. A key issue in communication is that this alignment between speaker and listener may be faulty, in which case a repair mechanism is needed [51]. In some cases, the misalignment is deliberate. This aspect of communication is stressed in [52], where a bipolar view on assertability is put forward, distinguishing between situations in which a statement is definitively assertable, situations in which it is merely acceptable to assert it, and situations in which it cannot be asserted without condemnation. As an example where misalignment is deliberate, [48] considers the example of a hotel manager, who claims that a room of his hotel is 'quite large', despite knowing that the client may not agree with this. In general, misalignments can be discovered because of inconsistencies, but, due to the interactive nature of dialogues also by means of explicit clarification requests and reformulation. Clearly, the issue of merging multiple-source information can be related to this view. However, as the interactive component is missing, the main vehicle to establish plausible alignments is by detecting and interpreting inconsistencies. In this sense, the integrity constraints C could be seen as explicit common ground between the sources, which may or may not be properly aligned with them. The weighted knowledge bases M_{s_i} then encode strategies for repairing situations of detected misalignment. This point of view is also reminiscent of approaches to causal diagnosis [53], which attempt to find the most likely disorders, given a set of observed symptoms. Here, conflicts take the role of symptoms and the possible misalignments take the role of disorders.

Another line of related work is the use of similarity relations between interpretations to define approximate entailment relations. This idea was first proposed in [54] where "p approximately entails q" is understood as "every model of p is similar to a model of q". This idea can be contrasted with non-monotonic reasoning, where "if p, generally q" iff the most plausible models of p are also models of q [55], i.e. in the former case the set of models of q is expanded, whereas in the latter case the set of models of p is restricted. A detailed comparison between similarity-based and non-monotonic entailment was made in [56,57], e.g. pointing out that, while similarity-based entailment is monotonic, it fails to satisfy the so-called cut property: from $p \models\!\approx r$ and $p \wedge r \models\!\approx q$ it does not follow that $p \models\!\approx q$. The use of a similarity relation between interpretations was also briefly discussed from a belief revision point of view in [58].

In certain application domains, ideas have been examined that are to some extent similar to the motivation of our approach, although only within more restrictive settings. To repair inconsistent description logic ontologies, for instance, [59] looks for overgeneralized concept inclusions that create conflicts, and accordingly replaces concepts by more liberal or more restrictive variants. For instance, the conflict that arises from 'children only like ice cream' and 'children only like chocolate' is resolved by replacing these assertions by 'children only like sweeties', thus taking advantage of known concept relations to resolve the conflict in a meaningful way. In this approach, we can clearly recognize the view that inconsistencies are often caused by sources that are not sufficiently cautious when asserting information, although what they claim may not be far from the truth. Another related application is presented in [60], where the problem of merging networks of qualitative temporal and spatial relations is considered. The idea is to deal with conflicts by finding spatial or temporal scenarios that are similar to scenarios that are compatible with what is claimed by each of the sources. For instance, if one source claims that spatial regions a and b are disjoint while another asserts that in fact a is a part of b, we could conclude that a overlaps with b. The notion of similarity here operates at the level of the spatial or temporal relations, and is directly related to the conceptual neighborhood diagrams of Freksa [31]. Somewhat related, [61] discusses an application in which temporal relations are extracted from web documents, and conflicts are solved by reinterpreting the temporal relationships as fuzzy temporal relations that only hold to some degree. In the presence of a conflict, 'a happened during b' may then be interpreted as, e.g., 'a happened during b to degree 0.6'. Such fuzzy temporal relations can be modeled in a generalization of Allen's interval algebra [62], which was proposed in [63]. Intuitively, 'a happened during b to degree 0.6' means that the degree of similarity between the actual state of affairs and a temporal configuration in which a happened during b is 0.6. A similar approach in the spatial domain, based on a generalization of the region connection calculus [64], was proposed in [65].

The idea that flexibility of terms may be interpreted in different ways (cf. the four scenarios from Section 4) is somewhat reminiscent of the study of linguistic hedges such as 'very' in the framework of fuzzy set theory [66,67]. In particular, a linguistic phrase such as 'more or less old' can be understood in at least two different ways. There is an inclusive reading, in which 'more or less old' is a liberalization of 'old', i.e. everybody who is considered old, is also considered more or less old. However, there also is an exclusive reading in which the meaning of 'more or less old' does not encompass the meaning of 'old', i.e. people who are very old are not considered to be 'more or less old'. In other words, when modeling linguistic hedges, there is a choice between expanding (or contracting, depending on the type of modifier) the meaning of a term and shifting it [68].

Finally, note that preliminary versions of the approach we have presented in this paper can be found in [69] and [70]. In particular, [69] has introduced the idea of using a similarity graph and of defining merging operators, at the semantic level, in terms of the operators $\langle . \rangle_R$ and $[.]_R$ defined in (18)–(19). In [70], merging operators were defined at the syntactic level. The syntactic procedures that were proposed there essentially correspond to the application of Proposition 10 in the specific case where the weights are interpreted in a possibilistic setting, or in the setting of inclusion-based preferred subtheories.

8. Conclusions

In this paper, we have proposed a novel approach to merging, which differs from existing approaches in its use of extra-logical background information about the semantic relatedness of atoms. The central idea is that in many applications, atoms correspond to natural language terms that may be understood in a slightly different way by different sources. By exploiting available knowledge about which of these terms have a similar meaning (or may otherwise be confused), consistency can be restored in a more informed way. The requirement for this extra-logical information may be seen as a disadvantage, in the sense that we cannot expect such information to be available in every application. However, in such cases, the merging operators we have proposed degenerate to existing approaches such as morphological, conflict-based or distance-based merging. In general, our operators refine (special cases of) these existing approaches using whatever information is available about the relatedness of terms.

Rather than defining one specific merging operator, we have proposed a general framework which is based on (i) assuming a disjoint vocabulary for each source to trivially restore consistency, and (ii) introducing an additional weighted knowledge base to encode flexible constraints on how the vocabularies of different sources relate to each other. Subsequently, we have analyzed how different interpretations of these weights naturally lead to merging operators with a different behavior. The interpretations we have considered range from purely qualitative approaches, in which weights are taken from an abstract, partially-ordered scale, to purely quantitative approaches with a probabilistic flavor. This diversity allows us to appropriately handle cases where more, or less information is available about the reliability of sources, the strength of known semantic relationships, etc.

Each of the proposed merging operators is based on the intuition of finding the interpretations that are most plausible, given what is asserted by the sources. In this sense, they are tailored towards merging beliefs, rather than towards merging preferences or goals. While the exploitation of semantic relatedness between terms is clearly of interest for the latter problem as well, the scenario of merging preferences additionally requires to consider principles from social choice theory to ensure that the result is fair.

Acknowledgements

We would like to thank the anonymous reviewers, whose detailed comments have greatly improved the readability of this paper.

Appendix A. Tools for managing inconsistency

A.1. Possibilistic logic

A possibility distribution in a universe X is an $X - [0, 1]$ mapping, used as a convenient way to encode a complete ordering that models plausibility or preference. Given a possibility distribution π on the set of all possible interpretations 2^A, the possibility $\Pi(\phi)$ and necessity $N(\phi)$ of a formula ϕ are defined as

$$\Pi(\phi) = \max_{I \in [\![\phi]\!]} \pi(I)$$
$$N(\phi) = 1 - \Pi(\neg\phi) = \min_{I \in [\![\neg\phi]\!]} 1 - \pi(I)$$

A possibilistic logic [25] formula (ϕ, λ) is a pair made of a classical logic formula ϕ, and a weight $\lambda \in]0, 1]$ expressing its certainty or priority. The formula (ϕ, λ) is semantically interpreted as $N(\phi) \geqslant \lambda$, where N is a necessity measure. Necessity measures N are characterized by the decomposability property $N(\phi \wedge \psi) = \min(N(\phi), N(\psi))$. A possibilistic logic base, i.e. a set of possibilistic logic formulas can always be put in clausal form thanks to this property. The basic inference rule is the following resolution rule, here written in the propositional case:

$$(\neg\phi \vee \psi, \lambda); (\phi \vee \gamma, \mu) \vdash \big(\psi \vee \gamma, \min(\lambda, \mu)\big)$$

Given a possibilistic logic base K and a certainty level λ, two particular classical knowledge bases can be defined:

$$K_\lambda = \big\{\phi \bigm| (\phi, \mu) \in K \text{ and } \mu \geqslant \lambda\big\}$$
$$K_{\underline{\lambda}} = \big\{\phi \bigm| (\phi, \mu) \in K \text{ and } \mu > \lambda\big\}$$

A possibilistic logic base $K = \{(\phi_i, \lambda_i) \mid i = 1, n\}$ is semantically equivalent to a possibility distribution π_K which restricts the set of interpretations that are more or less compatible with K. The possibility distribution π_K is the min-based conjunctive combination of the representations $\pi_{(\phi_i,\lambda_i)}$ of each formula in K. An interpretation I is all the more possible as it does not violate any formula ϕ_i with a high certainty level λ_i:

$$\pi_{(\phi_i,\lambda_i)}(I) = \begin{cases} 1 & \text{if } I \models \phi_i \\ 1 - \lambda_i & \text{if } I \models \neg\phi_i \end{cases}$$
$$\pi_K(I) = \min_i \pi_{(\phi_i,\lambda_i)}(I)$$

An important feature of possibilistic logic is its ability to cope with inconsistency. The inconsistency level of K is defined as

$$inc(K) = \max\{\lambda \mid K_\lambda \text{ inconsistent}\} \tag{A.1}$$

with $\max\emptyset = 0$. We have $inc(K) = \lambda$ iff $\max_I \pi_K(I) = 1 - \lambda$. Formulas in $K_{\underline{inc(K)}}$ are safe from inconsistency.

A.2. Penalty logic

Penalty logic [27,28] is similar in spirit to possibilistic logic, but with an additive interpretation of the weights. A penalty logic formula (ϕ, p) consists of a classical formula ϕ and a penalty $p \in [0, +\infty]$. In particular, weights are not restricted to the unit interval anymore, but can be any non-negative real number or $+\infty$. The intuition behind a penalty logic formula (ϕ, p) is that p is the cost that has to be paid for having ϕ violated. Given a penalty logic base K, we can naturally associate penalties to classical interpretations I:

$$pen_K(I) = \sum_{(\phi,p) \in K,\, I \not\models \phi} p$$

Let us write K^* for the set of classical formulas that appear in K, i.e. $K^* = \{\phi \mid \exists p \,.\, (\phi, p) \in K\}$. The penalty of a set of formulas $B \subseteq K^*$ can then be defined as

$$pen_K(B) = \sum_{(\alpha,p)\in K,\, \alpha\notin B} p$$

Like possibilistic logic, penalty logic can naturally deal with inconsistency. In particular, a consistent subset $B \subseteq K^*$ is called preferred if its penalty $pen_K(B)$ is minimal among all consistent subsets of K^*. An inconsistency-tolerant entailment relation $\approx$ can then be defined as $K \approx \phi$ iff $B \models \phi$ for all preferred subsets of K^*. Note that in this setting, multiplying all penalties by a strictly positive constant does not affect the treatment of inconsistency. Without any practical restriction, we may thus assume that all penalties are taken from $\mathbb{N} \cup \{+\infty\}$ which has clear computational advantages.

Appendix B. Variable forgetting

In this appendix, we introduce an efficient syntactic procedure that can be used, under some restrictions, to implement the step of variable forgetting. In particular, for a source s_i, an atom a, $\times \in \{W, X, Y, Z\}$, $B \in Pref(P)$ and $B' \in Pref(P')$, let $w(a, s_i, \times; B)$ and $w(a, s_i; B')$ be defined as follows:

$$w(a, s_i, \times; B) = \begin{cases} \min\{l \in \{0, \ldots, k\} \mid \alpha^{\times}_{(l,s_i,a)} \in B\} & \text{if } \alpha^{\times}_{(k,s_i,a)} \in B \\ k+1 & \text{otherwise} \end{cases}$$
$$w\big(a, s_i; B'\big) = \begin{cases} \min\{l \in \{0, \ldots, k\} \mid \alpha_{(l,s_i,a)} \in B'\} & \text{if } \alpha_{(k,s_i,a)} \in B' \\ k+1 & \text{otherwise} \end{cases}$$

Now, let us assume that all knowledge bases are in conjunctive-normal form. For any given $B \in Pref(P)$, we then define the B-weakened version K_i^B of knowledge base K_i as the propositional knowledge base which is obtained from K_i by replacing all occurrences of positive literals a by $\bigvee W_a^{w(a,s_i,W;B)}$ and all occurrences of negative literals $\neg a$ by $\bigvee\{\neg y \mid y \in Y_a^{w(a,s_i,Y;B)}\}$, where we define $W_a^{k+1} = \{\top\}$ and $Y_a^{k+1} = \{\bot\}$. Similarly, for $B' \in Pref(P')$, we define the B'-weakened version $K_i^{B'}$ as the propositional knowledge base which is obtained from K_i by replacing all occurrences of positive literals a by $\bigvee W_a^{w(a,s_i;B')}$ and all occurrences of negative literals $\neg a$ by $\bigvee\{\neg y \mid y \in Y_a^{w(a,s_i;B')}\}$. It is easy to see that the operations of B-weakening and B'-weakening indeed weaken a knowledge base K_i, in the sense that its set of models is increased. Intuitively, it is thus clear that by sufficiently weakening all knowledge bases in this way, consistency can be restored. In some cases, the merging operators $\Delta_f(K_1, \ldots, K_n)$ and $\Delta'_f(K_1, \ldots, K_n)$, in fact, implement this idea, as made explicit in the following proposition.

Proposition 10. *Assume that for each $a \in A$ and $l \in \{0, \ldots, k\}$, it holds that $X_a^l = \{a, \top\}$ and $Z_a^l = \{a, \bot\}$. Furthermore, assume that each knowledge base K_i is equal to its set of prime implicates. It holds that*

$$\Delta_f(K_1, \ldots, K_n) \equiv \bigvee_{B\in Pref(P)} K_1^B \cup \cdots \cup K_n^B \cup C \tag{B.1}$$

$$\Delta'_f(K_1, \ldots, K_n) \equiv \bigvee_{B'\in Pref(P')} K_1^{B'} \cup \cdots \cup K_n^{B'} \cup C \tag{B.2}$$

Recall that a clause γ is an implicate of propositional knowledge base Φ iff $\Phi \models \gamma$, and a prime implicate if furthermore for every other implicate γ', if $\gamma' \models \gamma$ then also $\gamma \models \gamma'$. Moreover, we additionally require that no literals are repeated in prime implicates (e.g. if $a \vee b$ is a prime implicate of Φ then $a \vee a \vee b$ is not considered as prime implicate). Every propositional knowledge base can then be represented by its set of prime implicates. Also note that the condition that $X_a^l = \{a, \top\}$ and $Z_a^l = \{a, \bot\}$ is, in fact, satisfied in two of the four scenarios considered in Section 4: liberalization and restriction. Indeed, although Table 1 e.g. suggest rather $X_a^l = \{a\}$ in the case of liberalization, we may equivalently take $X_a^l = \{a, \top\}$ since $Y_a^l = \{a\}$ and $a \to a^{@s_i}$ is equivalent to $\neg a^{@s_i} \to \neg a$.

Interestingly, the technique we have proposed in [70] to weaken propositional knowledge bases corresponds to a special case of Proposition 10, thus revealing that the approach we take in this paper is compatible with the approach from [70], although it is more general.

Note that the condition from Proposition 10 that each knowledge base K_i should be equal to its set of prime implicates is not redundant, as illustrated by the following example.

Example 17. Let $C = \emptyset$, $K_1 = \{a \vee b, \neg b\}$, $K_2 = \{\neg a\}$, where $W_a^1 = Y_a^1 = \{a, c\}$ and $W_b^1 = Y_b^1 = \{b, c\}$ and the sets X_a^1, Z_a^1, X_b^1, Z_b^1, W_c^1, X_c^1, Y_c^1 and Z_c^1 are trivial. Clearly, K_1 is not equal to its set of prime implicates, which is given by $\{a, \neg b\}$. Let $pref(P') = \{B'\}$, where

$$B' = K_1^{@s_1} \cup K_2^{@s_2} \cup \{\alpha_{(1,s_1,a)}, \alpha_{(1,s_1,b)}, \alpha_{(1,s_1,c)}, \alpha_{(0,s_1,c)}, \alpha_{(1,s_2,a)}, \alpha_{(1,s_2,b)}, \alpha_{(1,s_2,c)}, \alpha_{(0,s_2,a)}, \alpha_{(0,s_2,b)}, \alpha_{(0,s_2,c)}\}$$

Then we have

$$K_1^{B'} = \{a \vee b \vee c, \neg b \vee \neg c\}$$
$$K_2^{B'} = \{\neg a\}$$

The right-hand side of (B.2) is then equivalent to $\neg a \wedge ((b \wedge \neg c) \vee (\neg b \wedge c))$. To find the left-hand side of (B.2), which corresponds to forgetting the variables from $A^{@s_1} \cup A^{@s_2}$ in B, we may write K_1 as $\{a\}$ and apply Proposition 10 to obtain

$$\Delta'_f(K_1, K_2) \equiv (a \vee c) \wedge \neg a \equiv c \wedge \neg a$$

Clearly we have $\neg a \wedge ((b \wedge \neg c) \vee (\neg b \wedge c)) \not\equiv c \wedge \neg a$.

Note that due to the restriction to sets of prime implicates, the result of the merging operation does not depend on the syntactic structure of the knowledge bases. Finally note that when the families W_a^l and Y_a^l are trivial, instead of X_a^l and Z_a^l, a counterpart to Proposition 10 can be obtained, which is based on knowledge bases in disjunctive-normal form. Because W_a^l and Y_a^l are not simultaneously trivial in any of the four scenarios discussed in Section 4, we omit the details.

Appendix C. Proofs

C.1. Proof of Proposition 1

1. We first show (21). Assume that $a \in (J \setminus \langle I \rangle_{W^l})$; we find $a \notin \langle I \rangle_{W^l}$ and hence for every $w \in I$, $(a, w) \notin W^l$. Therefore, if $w \in W_a^l$ we have $w \notin I$, which means $I \not\models \bigvee\{w \mid w \in W_a^l\}$ and also $(I \cup J^{@s_i}) \not\models \bigvee\{w \mid w \in W_a^l\}$. From $a \in J$ on the other hand we find $J^{@s_i} \models a^{@s_i}$ and $I \cup J^{@s_i} \models a^{@s_i}$. Together this means that $I \cup J^{@s_i} \not\models (a^{@s_i} \rightarrow \bigvee\{w \mid w \in W_a^l\})$.
Conversely, assume $I \cup J^{@s_i} \not\models (a^{@s_i} \rightarrow \bigvee\{w \mid w \in W_a^l\})$. This implies that $I \cup J^{@s_i} \models a^{@s_i}$ and $I \cup J^{@s_i} \not\models \bigvee\{w \mid w \in W_a^l\}$, or equivalently $J^{@s_i} \models a^{@s_i}$ and $I \not\models \bigvee\{w \mid w \in W_a^l\}$. From $J^{@s_i} \models a^{@s_i}$ we immediately derive $a \in J$. From $I \not\models \bigvee\{w \mid w \in W_a^l\}$ we find that there is no $w \in I$ such that $(a, w) \in W^l$, and thus $a \notin \langle I \rangle_{W^l}$. We conclude that $a \in (J \setminus \langle I \rangle_{W^l})$.
2. In entirely the same way, we can show (22).
3. To show (23), first note that $([I]_{Y^l} \setminus J) = co\langle coI \rangle_{Y^l} \cap co\,J = co\,J \setminus \langle coI \rangle_{Y^l}$. If $a \in (co\,J \setminus \langle coI \rangle_{Y^l})$, we find $a \notin \langle coI \rangle_{Y^l}$ and hence for every $y \in coI$, $(a, y) \notin Y^l$. Therefore, if $y \in Y_a^l$ we have $y \in I$, which means $I \not\models \bigvee\{\neg y \mid y \in Y_a^l\}$ and also $(I \cup J^{@s_i}) \not\models \bigvee\{\neg y \mid y \in Y_a^l\}$. From $a \in co\,J$ on the other hand we find $J^{@s_i} \models \neg a^{@s_i}$ and $I \cup J^{@s_i} \models \neg a^{@s_i}$. Together this means that $I \cup J^{@s_i} \not\models (\neg a^{@s_i} \rightarrow \bigvee\{\neg y \mid y \in Y_a^l\})$.
Conversely, assume $I \cup J^{@s_i} \not\models (\neg a^{@s_i} \rightarrow \bigvee\{\neg y \mid y \in Y_a^l\})$. This implies that $I \cup J^{@s_i} \models \neg a^{@s_i}$ and $I \cup J^{@s_i} \not\models \bigvee\{\neg y \mid y \in Y_a^l\}$, or equivalently $J^{@s_i} \models \neg a^{@s_i}$ and $I \not\models \bigvee\{\neg y \mid y \in Y_a^l\}$. From $J^{@s_i} \models \neg a^{@s_i}$ we immediately derive $a \in co\,J$. From $I \not\models \bigvee\{\neg y \mid y \in Y_a^l\}$ we find that there is no $y \in coI$ such that $(a, y) \in Y^l$, and thus $a \notin \langle coI \rangle_{Y^l}$. We conclude that $a \in (co\,J \setminus \langle coI \rangle_{Y^l})$.
4. Finally, (24) is shown in the same way as (23).

C.2. Proof of Lemma 1

($\Rightarrow$) Assume that (26) holds for all $a \in A$. First assume that $a^{@s_i}$ is true. Then we clearly have that $\bigvee S$ is true for some S satisfying $\{a\} \subseteq S \subseteq S_a^l$, which means that $\bigvee S_a^l$ is the case. We conclude $a^{@s_i} \rightarrow \bigvee S_a^l$ using the deduction theorem. Next assume that a is true. This means that $\bigvee S$ will be true for every S satisfying $\{a\} \subseteq S \subseteq S_a^l$ and consequently that $a^{@s_i}$ will be true. Hence we have shown $a \rightarrow a^{@s_i}$.

($\Leftarrow$) Assume that the implications in (27) hold for all $a \in A$. First assume that $a^{@s_i}$ holds. Then we know from $a^{@s_i} \rightarrow \bigvee S_a^l$ that some $b \in S_a^l$ is satisfied, and hence also (26). Conversely, if $a^{@s_i}$ is false it is sufficient to show that $\bigvee S$ is false for some S satisfying $\{a\} \subseteq S \subseteq S_a^l$. In particular for $S = \{a\}$ we immediately have $\bigvee S$ false from $\neg a^{@s_i} \rightarrow \neg a$.

C.3. Proof of Lemma 2

($\Rightarrow$) Assume that (28) holds for all $a \in A$. First assume that $a^{@s_i}$ is false. Then we clearly have that $\bigwedge S$ is false for some S satisfying $\{a\} \subseteq S \subseteq S_a^l$, which means that for some $s \in S_a^l$, $\neg s$ is the case, or in other words $\bigvee\{\neg s \mid s \in S_a^l\}$. We conclude $\neg a^{@s_i} \rightarrow \bigvee\{\neg s \mid s \in S_a^l\}$ using the deduction theorem. Next assume that a is false. This means that $\bigwedge S$ will be false for every S satisfying $\{a\} \subseteq S \subseteq S_a^l$ and consequently that $a^{@s_i}$ will be false. Hence we obtain $\neg a \rightarrow \neg a^{@s_i}$.

($\Leftarrow$) Assume that the implications in (29) hold for all $a \in A$. First assume that $a^{@s_i}$ holds. Then we know from $\neg a \rightarrow \neg a^{@s_i}$, which is equivalent to $a^{@s_i} \rightarrow a$, that a is the case. This means that for $S = \{a\}$, $\bigwedge S$ holds, and thus $a^{@s_i} \equiv \bigwedge S$. Conversely, if $a^{@s_i}$ is false, we know from (29) that there is an $s \in S_a^l$ such that s is false. This means that $\bigwedge S$ will be false for some S satisfying $\{a\} \subseteq S \subseteq S_a^l$, and thus $a^{@s_i} \equiv \bigwedge S$.

C.4. Proof of Proposition 2

We show that

$$I \in \left[\!\left[\Delta^{poss}(K_1, \ldots, K_n; C; M_{s_1}, \ldots, M_{s_n}) \right]\!\right] \tag{C.1}$$

is equivalent to

$$I \in \left[\!\left[\Delta^{morph}(K_1, \ldots, K_n; C; B_{(W^r, X^r, Y^r, Z^r)}) \right]\!\right] \tag{C.2}$$

Clearly, by definition of variable forgetting, (C.1) is equivalent to

$$\exists J_1, \ldots, J_n \in 2^A \,.\, I \cup J_1^{@s_1} \cup \cdots \cup J_n^{@s_n} \in P_{\underline{inc(P)}}$$

or, by definition of r

$$\exists J_1, \ldots, J_n \in 2^A \,.\, I \cup J_1^{@s_1} \cup \cdots \cup J_n^{@s_n} \in P_{\lambda_r}$$

Applying the definition of P, this corresponds to

$$\exists J_1, \ldots, J_n \in 2^A \,.\, I \in [\![C]\!] \wedge \forall i \,.\, I \cup J_i^{@s_i} \in \left[\!\left[(M_{s_i})_{\lambda_r}\right]\!\right] \wedge J_i^{@s_i} \in \left[\!\left[K_i^{@s_i}\right]\!\right]$$

which is equivalent to

$$I \in [\![C]\!] \wedge \exists J_1, \ldots, J_n \in 2^A \,.\, \forall i \,.\, I \cup J_i^{@s_i} \in \left[\!\left[(M_{s_i})_{\lambda_r}\right]\!\right] \wedge J_i \in [\![K_i]\!]$$

From Proposition 1, we know that this is equivalent to

$$I \in [\![C]\!] \wedge \exists J_1, \ldots, J_n \in 2^A \,.\, \forall i \,.\, J_i \subseteq \langle I \rangle_{W^l} \wedge I \subseteq \langle J_i \rangle_{X^r} \wedge [I]_{Y^r} \subseteq J_i \wedge [J_i]_{Z^r} \subseteq I \wedge J_i \in [\![K_i]\!]$$

or in other words

$$I \in [\![C]\!] \wedge \exists J_1, \ldots, J_n \in 2^A \,.\, \forall i.\, \sigma_{(Y^r, W^r)}(I, J_i) \wedge \sigma_{(Z^r, X^r)}(J_i, I) \wedge J_i \in [\![K_i]\!]$$

By definition of $B_{(W^r, X^r, Y^r, Z^r)}$ we find

$$I \in [\![C]\!] \wedge \exists J_1, \ldots, J_n \in 2^A \,.\, \forall i \,.\, J_i \in B_{(W^r, X^r, Y^r, Z^r)}(I) \wedge J_i \in [\![K_i]\!]$$

or, equivalently

$$I \in [\![C]\!] \wedge \forall i \,.\, B_{(W^r, X^r, Y^r, Z^r)}(I) \cap [\![K_i]\!] \neq \emptyset$$

which corresponds to (C.2).

C.5. Proof of Proposition 3

As an example, we show (41); the proof of (42) is entirely analogous.
Note that $I \in [\![\Delta^{prior_\subseteq}(K_1, \ldots, K_n)]\!]$ is equivalent to

$$\exists J_1, \ldots, J_n \in 2^A \,.\, I \cup \bigcup_i J_i^{@s_i} \models \bigvee Pref_\subseteq(P)$$

which means that for some $B \in Pref_\subseteq(P)$

$$\exists J_1, \ldots, J_n \in 2^A \,.\, I \cup \bigcup_i J_i^{@s_i} \models B$$

Let us write B_l for the formulas from B that appear in M_{s_i} for some i with weight λ_l. The fact that B is a preferred subtheory means that $I \models C$ (since the integrity constraints C are assumed to be consistent), $J_i \models K_i$ (since each knowledge base is assumed to be individually consistent), and furthermore, that there can be no $r \in \{0, \ldots, k\}$ and $I', J_1', \ldots, J_n' \in 2^A$ such that $I' \models C$, $J_i' \models K_i$ and such that $B_r \subset B_r'$ while $B_l' = B_l$ for all $l > r$, where we write B_l' for the set of formulas which appear in P with weight λ_l and that are satisfied by $I' \cup \bigcup_i J_i'^{@s_i}$. Using Proposition 1, $B_l' \subseteq B_l$ is equivalent to asserting that for all $i \in \{1, \ldots, n\}$ the following four inclusions hold:

$$J_i \setminus \langle I \rangle_{W^l} \subseteq J_i' \setminus \langle I' \rangle_{W^l}, \qquad I \setminus \langle J_i \rangle_{X^l} \subseteq I' \setminus \langle J_i' \rangle_{X^l}$$
$$[I]_{Y^l} \setminus J_i \subseteq [I']_{Y^l} \setminus J_i', \qquad [J_i]_{Z^l} \setminus I \subseteq [J_i']_{Z^l} \setminus I'$$

In other words, for all i

$$diff^a_{W_l}(I, J_i) \subseteq diff^a_{W_l}\big(I', J'_i\big), \qquad diff^b_{X_l}(I, J_i) \subseteq diff^b_{X_l}\big(I', J'_i\big)$$
$$diff^c_{Y_l}(I, J_i) \subseteq diff^c_{Y_l}\big(I', J'_i\big), \qquad diff^d_{Z_l}(I, J_i) \subseteq diff^d_{Z_l}\big(I', J'_i\big)$$

or, equivalently

$$\forall i \in \{1, \ldots, n\} \,.\, diff^1_{\mathcal{S}_l}(I, J_i) \subseteq^4 diff^1_{\mathcal{S}_l}\big(I', J'_i\big)$$

which means

$$\big\langle diff^1_{\mathcal{S}}(I, J_1), \ldots, diff^1_{\mathcal{S}}(I, J_n)\big\rangle \leqslant^1_l \big\langle diff^1_{\mathcal{S}}\big(I', J'_1\big), \ldots, diff^1_{\mathcal{S}}\big(I', J'_n\big)\big\rangle$$

Thus, we obtain

$$\exists \mathbf{c} \in diff^1_{\mathcal{S}}(I, \mathcal{K}) \,.\, \neg\exists r \,.\, \exists I' \in 2^A \,.\, \exists \mathbf{c}' \in diff^1_{\mathcal{S}}\big(I', \mathcal{K}\big) \,.\, \big(\forall l > r \,.\, \mathbf{c} =^1_l \mathbf{c}'\big) \wedge \mathbf{c}' <^1_r \mathbf{c}$$

where we write $\mathbf{c} =^1_l \mathbf{c}'$ for $\mathbf{c} \leqslant^1_l \mathbf{c}' \wedge \mathbf{c}' \leqslant^1_l \mathbf{c}$, and we write $\mathbf{c}' <^1_r \mathbf{c}$ for $\mathbf{c}' \leqslant^1_r \mathbf{c} \wedge \mathbf{c} \nleqslant^1_r \mathbf{c}'$. But this is nothing else than saying

$$\exists \mathbf{c} \in diff^1_{\mathcal{S}}(I, \mathcal{K}) \,.\, \forall I' \in 2^A \,.\, \forall \mathbf{c}' \in diff^1_{\mathcal{S}}\big(I', \mathcal{K}\big) \,.\, \big(\mathbf{c}, \mathbf{c}'\big) \in lex\big(\leqslant^1_k, \ldots, \leqslant^1_0\big) \vee \big(\mathbf{c}', \mathbf{c}\big) \notin lex\big(\leqslant^1_k, \ldots, \leqslant^1_0\big)$$

and thus, since I moreover satisfies the integrity constraints C, we have

$$I \in \big[\!\big[\Delta^{confl_2}\big(K_1, \ldots, K_n; C; diff^1_{\mathcal{S}}, lex\big(\leqslant^1_k, \ldots, \leqslant^1_0\big)\big)\big]\!\big]$$

C.6. Proof of Proposition 4

As an example, we show (46); the proof of (47) is entirely analogous.
Note that $I \in [\![\Delta^{prior_{card}}(K_1, \ldots, K_n)]\!]$ is equivalent to

$$\exists J_1, \ldots, J_n \in 2^A \,.\, I \cup \bigcup_i J_i^{@s_i} \models \bigvee Pref_{card}(P)$$

which means that for some $B \in Pref_{card}(P)$

$$\exists J_1, \ldots, J_n \in 2^A \,.\, I \cup \bigcup_i J_i^{@s_i} \models B$$

Let us write B_l for the formulas from B that appear in M_{s_i} for some i with weight λ_l. The fact that B is a cardinality-based preferred subtheory means that $I \models C$ (since the integrity constraints C are assumed to be consistent), $J_i \models K_i$ (since each knowledge base is assumed to be individually consistent), and furthermore, that there can be no $r \in \{0, \ldots, k\}$ and $I', J'_1, \ldots, J'_n \in 2^A$ such that $I' \models C$, $J'_i \models K_i$ and such that $|B_r| < |B'_r|$ while $|B'_l| = |B_l|$ for all $l > r$, where we write B'_l for the set of formulas which appear in P with weight λ_l and that are satisfied by $I' \cup \bigcup_i J'^{@s_i}_i$. Using Proposition 1, we find that:

$$|B_l| = \sum_i \big|J_i \setminus \langle I\rangle_{W^l}\big| + \big|I \setminus \langle J_i\rangle_{X^l}\big| + \big|[I]_{Y^l} \setminus J_i\big| + \big|[J_i]_{Z^l} \setminus I\big|$$
$$\big|B'_l\big| = \sum_i \big|J'_i \setminus \langle I'\rangle_{W^l}\big| + \big|I' \setminus \langle J'_i\rangle_{X^l}\big| + \big|[I']_{Y^l} \setminus J'_i\big| + \big|[J'_i]_{Z^l} \setminus I'\big|$$

In other words,

$$|B_l| = d^1_l(I, \mathcal{K}), \qquad |B'_l| = d^1_l\big(I', \mathcal{K}\big)$$

From which we immediately find that I being a model of a cardinality-based preferred subtheory is equivalent to the fact that there cannot be an interpretation I' such that $I' \leqslant^1 I$ while $I \nleqslant^1 I'$, or in other words, $I \in [\![\Delta^{dist}(K_1, \ldots, K_n; C; \leqslant_1)]\!]$.

C.7. Proof of Proposition 5

As an example, we show (52)–(54). The proof of (49)–(51) is analogous.
First, note that $I \in [\![\Delta^{confl_2}(K_1, \ldots, K_n; C; diff^2_{\mathcal{S}}, par(\leqslant^2_k, \ldots, \leqslant^2_0))]\!]$ means that $I \in [\![C]\!]$ and that there is a $\langle c_1, \ldots, c_n\rangle \in diff^2_{\mathcal{S}}(I, \mathcal{K})$ such that for every $I' \in [\![C]\!]$, we have either for every $i \in \{0, \ldots, n\}$ and $l \in \{0, \ldots, k\}$

$$c_i \subseteq diff^2_{\mathcal{S}_l}\big(I', K_i\big)$$

or for at least one $i \in \{0, \dots, n\}$ and $l \in \{0, \dots, k\}$

$$diff^2_{\mathcal{S}_l}(I', K_i) \not\subseteq c_i$$

Furthermore, by construction, each c_i corresponds to a model J_i of K_i. Thus, we have that there exist models $J_1, \dots, J_n$ of $K_1, \dots, K_n$ such that for all models $J'_1, \dots, J'_n$ of $K_1, \dots, K_n$, it holds that either for every $i \in \{0, \dots, n\}$ and $l \in \{0, \dots, k\}$

$$\begin{aligned} &diff^a_{W^l}(I, J_i) \cup diff^b_{X^l}(I, J_i) \cup diff^c_{Y^l}(I, J_i) \cup diff^d_{Z^l}(I, J_i) \\ &\quad \subseteq diff^a_{W^l}(I', J'_i) \cup diff^b_{X^l}(I', J'_i) \cup diff^c_{Y^l}(I', J'_i) \cup diff^d_{Z^l}(I', J'_i) \end{aligned}$$

or for some $i \in \{0, \dots, n\}$ and $l \in \{0, \dots, k\}$

$$\begin{aligned} &diff^a_{W^l}(I', J'_i) \cup diff^b_{X^l}(I', J'_i) \cup diff^c_{Y^l}(I', J'_i) \cup diff^d_{Z^l}(I', J'_i) \\ &\quad \not\subseteq diff^a_{W^l}(I, J_i) \cup diff^b_{X^l}(I, J_i) \cup diff^c_{Y^l}(I, J_i) \cup diff^d_{Z^l}(I, J_i) \end{aligned}$$

Using Proposition 1, we find that this is equivalent to stating that the set $B(I, J_1, \dots, J_n)$ of formulas from $M_{s_1} \cup \dots \cup M_{s_n}$ that are satisfied by $I \cup J_1^{@s_1} \cup \dots \cup J_n^{@s_n}$ is not properly included in the set $B(I', J'_1, \dots, J'_n)$ of formulas from $M_{s_1} \cup \dots \cup M_{s_n}$ that are satisfied by $I' \cup {J'_1}^{@s_1} \cup \dots \cup {J'_n}^{@s_n}$.

We now show (52), (53) and (54).

1. The fact that $B(I, J_1, \dots, J_n) \not\subset B(I', J'_1, \dots, J'_n)$ means that

$$B(I, J_1, \dots, J_n) \supseteq B(I', J'_1, \dots, J'_n) \quad \text{or} \quad B(I, J_1, \dots, J_n) \not\subseteq B(I', J'_1, \dots, J'_n)$$

Now due to the structure of the possibilistic knowledge bases M_{s_i} and the choice of the ordering on Λ, for $(\alpha, \lambda_{(l,s,a)})$ and $(\alpha', \lambda_{(l',s',a')})$ in $M_{s_1} \cup \dots \cup M_{s_n}$, it holds that

$$\lambda_{(l,s,a)} \leqslant \lambda_{(l',s',a')} \quad \text{implies} \quad \alpha \to \alpha' \tag{C.3}$$

In particular, this means that whenever a formula α is satisfied which appears with weight $\lambda_{(l,s,a)}$ in $M_{s_1} \cup \dots \cup M_{s_n}$, all formulas with a higher weight are also satisfied. Therefore, $B(I, J_1, \dots, J_n) \supseteq B(I', J'_1, \dots, J'_n)$ is the same as $B(I, J_1, \dots, J_n) \preccurlyeq B(I', J'_1, \dots, J'_n)$, and $B(I, J_1, \dots, J_n) \not\subseteq B(I', J'_1, \dots, J'_n)$ is the same as $B(I', J'_1, \dots, J'_n) \not\preccurlyeq B(I, J_1, \dots, J_n)$. Since $I \in [\![\Delta^{poss}(K_1, \dots, K_n)]\!]$ is equivalent to saying that $I \in [\![C]\!]$ and there exist models $J_i \in [\![K_i]\!]$ such that for all $I' \in [\![C]\!]$ and all models $J'_i \in [\![K_i]\!]$, $B(I, J_1, \dots, J_n) \preccurlyeq B(I', J'_1, \dots, J'_n)$ or $B(I', J'_1, \dots, J'_n) \not\preccurlyeq B(I, J_1, \dots, J_n)$, the stated follows.
2. It is straightforward to construct a linearization κ such that the set $B(I, J_1, \dots, J_n)$ is a preferred subtheory of $\kappa(P')$. Indeed, it suffices to rank the weights such that $\kappa(\lambda_{(l,s,a)}) > \kappa(\lambda_{(l',s',a')})$ whenever the formula with weight $\lambda_{(l,s,a)}$ appears in $B(I, J_1, \dots, J_n)$ and the formula with weight $\lambda_{(l',s',a')}$ does not. The fact that such a linearization can always be obtained follows straightforwardly from (C.3). Conversely, it is furthermore clear that for every linearization κ, and every model $I \cup J_1^{@s_1} \cup \dots \cup J_n^{@s_n}$ of a preferred subtheory of $\kappa(P')$, it holds that the set $B(I, J_1, \dots, J_n)$ is a maximal consistent subset of formulas from $M_{s_1} \cup \dots \cup M_{s_n}$. In other words, there is a one-on-one correspondence between the models of $Pref_{\subseteq}(P')$ and those of $Pref_{\preccurlyeq}(P')$.
3. By construction we already have $Pref_{card}(P') \subseteq Pref_{\subseteq}(P')$. To show (54), it therefore suffices to show $Pref_{\subseteq}(P') \subseteq Pref_{card}(P')$. Now let $B \in Pref_{\subseteq}(P')$ and let B_l be the formulas from B that appear in $\kappa(P')$ with weight l, and let $\overline{B}_l$ the formulas outside B that appear in $\kappa(P')$ with weight l. Let the different weights appearing in $\kappa(P')$ be $l_1 < \dots < l_r$. Without lack of generality, we may assume that $l_1 > 0$, since only the relative ordering of these weights is important. We now construct a new linearization κ' as follows:

$$\kappa'(\lambda_{(l,s,a)}) = \begin{cases} l_i & \text{if the unique formula from } P' \text{ with weight } \lambda_{(l,s,a)} \text{ is in } B \\ \frac{l_i + l_{i-1}}{2} & \text{otherwise} \end{cases}$$

where $l_i = \kappa(\lambda_{(l,s,a)})$ and $l_0 = 0 < l_1$. By construction, B is still a preferred subtheory of $\kappa'(P')$. Moreover, for each priority level l in $\kappa'(P')$ it holds that either all formulas belong to B or none of these formulas belong to B. As a consequence, B is also a cardinality-based preferred subtheory of $\kappa'(P')$.

C.8. Proof of Proposition 6

As an example, we show (59), the proof of (60) being analogous. If $I \in [\![\Delta^{pen}(K_1, \dots, K_n)]\!]$, we have $I \in [\![C]\!]$ and for some models $J_i \in [\![K_i]\!]$ we have $B(I, J_1, \dots, J_n) \in Pref_{pen}(Q)$, with $B(I, J_1, \dots, J_n)$ the formulas from Q that are satisfied by $I \cup J_1^{@s_1} \cup \dots \cup J_n^{@s_n}$. By construction, all formulas that appear in Q with weight $+\infty$ are contained in $B(I, J_1, \dots, J_n)$ (which is the case exactly when $I \in [\![C]\!]$ and $J_i \in [\![K_i]\!]$ for all i). Let us write $B_l(I, J_1, \dots, J_n)$ for the set of formulas in $B(I, J_1, \dots, J_n)$ that have weight λ_l in Q. Now suppose that there was an interpretation $I' \in [\![C]\!]$ such that for some models $J'_i \in [\![K_i]\!]$, $|B_k(I', J'_1, \dots, J'_n)| > |B_k(I, J_1, \dots, J_n)|$. Then we have

$$\sum\{p \mid (\alpha,p) \in Q,\ \alpha \notin B_k(I, J_1, \ldots, J_n)\} = \gamma_k r^k + \gamma_{k-1} r^{k-1} + \cdots + \gamma_0$$
$$\sum\{p \mid (\alpha,p) \in Q, \alpha \notin B_k(I', J'_1, \ldots, J'_n)\} = \gamma'_k r^k + \gamma'_{k-1} r^{k-1} + \cdots + \gamma'_0$$

where

$$\gamma_l = |B_l(I, J_1, \ldots, J_n)|, \qquad \gamma'_l = |B_l(I', J'_1, \ldots, J'_n)|$$

From $|B_k(I', J'_1, \ldots, J'_n)| > |B_k(I, J_1, \ldots, J_n)|$ we thus know that $\gamma'_k > \gamma_k$, which implies that $\sum\{p \mid (\alpha,p) \in Q,\ \alpha \notin B_k(I', J'_1, \ldots, J'_n)\} > \sum\{p \mid (\alpha,p) \in Q,\ \alpha \notin B_k(I, J_1, \ldots, J_n)\}$ as soon as r is sufficiently large. Indeed, if $\gamma'_{k-1} r^{k-1} + \cdots + \gamma'_0 \geqslant \gamma_{k-1} r^{k-1} + \cdots + \gamma_0$, this is trivial. On the other hand, if $\gamma'_{k-1} r^{k-1} + \cdots + \gamma'_0 < \gamma_{k-1} r^{k-1} + \cdots + \gamma_0$, it suffices to choose

$$r > \sqrt[k]{\frac{(\gamma_{k-1} r^{k-1} + \cdots + \gamma_0) - (\gamma'_{k-1} r^{k-1} + \cdots + \gamma'_0)}{\gamma'_k - \gamma_k}}$$

We may proceed in entirely the same fashion when $|B_l(I', J'_1, \ldots, J'_n)| = |B_l(I, J_1, \ldots, J_n)|$ for all $l > l_0$ and $|B_{l_0}(I', J'_1, \ldots, J'_n)| > |B_{l_0}(I, J_1, \ldots, J_n)|$. Thus, provided that r is sufficiently large, $B(I, J_1, \ldots, J_n)$ is minimal w.r.t. $\leqslant_p$ iff $B(I, J_1, \ldots, J_n) \in \Delta^{prior_{card}}$, from which the stated immediately follows.

C.9. Proof of Proposition 7

As an example, we show (61); (62) is shown entirely analogously.

The fact that $I \in [\![\Delta^{pen}(K_1, \ldots, K_n)]\!]$ means that $I \in [\![C]\!]$ and there exist models $J_i \in [\![K_i]\!]$ such that the subset $B(I, J_1, \ldots, J_n)$ of formulas from Q that are satisfied by $I \cup J_1^{@s_1} \cup \cdots \cup J_n^{@s_n}$ is preferred, i.e. for any other $I' \in [\![C]\!]$ and $J'_i \in [\![K_i]\!]$, it holds that $B(I, J_1, \ldots, J_n) \leqslant_p B(I', J'_1, \ldots, J'_n)$. From Proposition 1, we know that $(\alpha^W_{(l,s_i,a)}, p) \notin B(I, J_1, \ldots, J_n)$ is equivalent to $a \in J_i \setminus \langle I \rangle_{W^l}$, and similar for formulas of the form $(\alpha^X_{(l,s_i,a)}, p)$, $(\alpha^Y_{(l,s_i,a)}, p)$ and $(\alpha^Z_{(l,s_i,a)}, p)$, and for $B(I', J'_1, \ldots, J'_n)$. Thus, we immediately have

$$\sum\{p \mid (\alpha,p) \in Q,\ \alpha \notin B(I, J_1, \ldots, J_n)\} = d^1_Q(I, J_1) + \cdots + d^n_Q(I, J_1)$$

and consequently, that $I \leqslant_{d^1_Q} I'$ iff there exist models $J_i \in [\![K_i]\!]$ such that for all models $I' \in [\![C]\!]$ and $J'_i \in [\![K_i]\!]$ it holds that $B(I, J_1, \ldots, J_n) \leqslant_p B(I', J'_1, \ldots, J'_n)$. In other words, we have that $I \in [\![\Delta^{dist}(K_1, \ldots, K_n; C; \leqslant_{d^1_Q})]\!]$ iff $I \in [\![\Delta^{pen}(K_1, \ldots, K_n)]\!]$.

C.10. Proof of Proposition 8

1. Let us first consider the case of Δ^{poss} with linear weights. Membership in Θ^P_2 follows straightforwardly from the fact that deciding whether $K_{\underline{inc(K)}} \models \phi$ for a possibilistic knowledge base K and propositional formula ϕ is in Θ^P_2 (see [71]). We show Θ^P_2-hardness by reduction from PARITY(SAT) [47,72]. Given n instances $S_1, \ldots, S_n$ of the boolean satisfiability problem SAT, the problem PARITY(SAT) consists of deciding whether the number of satisfiable instances among $S_1, \ldots, S_n$ is odd. This problem is Θ^P_2-hard, even under the assumption that the instances are such that whenever S_i is unsatisfiable, $S_{i+1}, \ldots, S_n$ are unsatisfiable as well. We will now show that this problem, under the latter assumption, can be reduced to the problem of deciding $\Delta^{poss}(K_1, \ldots, K_n; C; M_{s_1}, \ldots, M_{s_n}) \models \phi$ in polynomial time.
   Let us choose $K_i = S_i \cup \{a_i\}$ for all $i \in \{1, \ldots, n\}$, where $a_1, \ldots, a_n$ are atoms which do not occur in any of the instances $S_1, \ldots, S_n$. Furthermore, we choose $C = \emptyset$, and the weighted knowledge bases M_{s_i} such that
   $$\bigwedge_{\times \in \{W,X,Y,Z\}} \alpha^{\times}_{(l,s_i,a)} \equiv \begin{cases} \top & \text{if } l > k - i + 1 \\ (a^{@s_i} \equiv a) & \text{otherwise} \end{cases}$$
   The weights are chosen such that $\lambda^{\times}_{(l,s_i,a)} = \lambda_l$ for some $\lambda_l \in [0,1]$. Clearly, P_{λ_k} is satisfiable iff S_1 is satisfiable, $P_{\lambda_{k-1}}$ is satisfiable iff $S_1 \cup S_2$ is satisfiable, which, due to the assumption on the instances S_i, is equivalent to the condition that S_2 is satisfiable. In general, we find that $P_{\lambda_{k-i}}$ is satisfiable iff S_{i+1} is satisfiable. By taking $\phi = (a_1 \wedge \neg a_2) \vee (a_3 \wedge \neg a_4) \vee \cdots$ we have that $\Delta^{poss}(K_1, \ldots, K_n; C; M_{s_1}, \ldots, M_{s_n}) \models \phi$ iff the number of λs in $\{\lambda_1, \ldots, \lambda_k\}$ for which P_λ is satisfiable is odd, which is in turn equivalent to the fact that an odd number among $S_1, \ldots, S_k$ are satisfiable.
2. Next, we consider Δ^{poss} with partially ordered weights. To prove membership in Π^P_2, we provide a Σ^P_2 procedure for checking
   $$\Delta^{poss}(K_1, \ldots, K_n; C; M_{s_1}, \ldots, M_{s_n}) \not\models \phi$$
   - Guess a subset K of P^*;
   - Verify that K is consistent using one call to the NP oracle;

- Verify that $K \not\models \phi$ using one call to the NP oracle;
- Verify that K is minimal w.r.t. $\preccurlyeq$. First, construct in polynomial time a minimal subset K' of K such that $K' \preccurlyeq K$ and $K \preccurlyeq K'$, by removing from K all formulas (α, λ) such that for some $(\alpha', \lambda') \in (P^* \setminus K)$, it holds that $\lambda \leqslant \lambda'$. It is clear that removing any formula from K' will then lead to a subset which is strictly less preferred than K or K'. To verify that K is minimal, it now suffices to consider all weights λ such that K' contains no formula with weight λ, and, in each case, check whether adding all formulas (α, λ') with $\lambda' \geqslant \lambda$ to K' leads to an inconsistent subset. Clearly, the number of required satisfiability checks is at most polynomial in the size of P^*.

To show that entailment checking for Δ^{poss} with partially ordered weights is Π_2^P-hard, we provide a reduction from quantified boolean formula problems of the form $\forall \mathbf{a} \,.\, \exists \mathbf{b} \,.\, F(\mathbf{a}, \mathbf{b})$.

Define $K_1 = \{a_1, \ldots, a_m, c\}$, $K_2 = \{\neg a_1, \ldots, \neg a_m\}$, $C = \{F(\mathbf{a}, \mathbf{b}) \equiv c\}$, $k = 1$, $W^1_{a_i} = X^1_{a_i} = \{a_i, \top\}$ and $Y^1_{a_i} = Z^1_{a_i} = \{a_i, \bot\}$ and $W^1_c = X^1_c = \{c, \top\}$ and $Y^1_c = Z^1_c = \{c, \bot\}$. Let P be defined in terms of K_1, K_2, C, M_{s_1} and M_{s_2} as before, and let the ordering on Λ be defined as in Proposition 5. It is clear that for any $\preccurlyeq$-preferred subset B, and for every i, we either have $B \models a_i$ or $B \models \neg a_i$. Indeed, if neither $B \models a_i$ nor $B \models \neg a_i$ we could construct a new consistent subset $B' = B \cup \{\alpha^W_{(0,s_1,a_i)}\}$, for which it holds that $B' \models a_i$ and $B' \preccurlyeq B$ but not $B \preccurlyeq B'$. Furthermore, it is clear that for any subset $A_0 \subseteq \{a_1, \ldots, a_m\}$ there exists a consistent subset B of P^* such that $B \models a_i$ if $a_i \in A_0$ and $B \models \neg a_i$ otherwise. We furthermore have $B \models c$ for such a minimal subset B iff $\alpha^W_{(0,s_1,c)} \in B$ in which case we also have $B \models F(a_1, \ldots, a_m, b_1, \ldots, b_r)$. Due to the minimality of B w.r.t. $\preccurlyeq$, $\alpha^W_{(0,s_1,c)} \in B$ will hold as soon as there exists a truth valuation for the atoms $b_1, \ldots, b_r$ which makes $F(a_1, \ldots, a_m, b_1, \ldots, b_r)$ true, given that the atoms in A_0 are true and those in $\{a_1, \ldots, a_m\} \setminus A_0$ are not. Thus, we can conclude that $\Delta^{poss}(K_1, K_2; C; M_{s_1}, M_{s_2}) \models F(a_1, \ldots, a_m, b_1, \ldots, b_r)$ iff $B \models F(a_1, \ldots, a_m, b_1, \ldots, b_r)$ for every $\preccurlyeq$-preferred B, which is equivalent to $\forall a_1, \ldots, a_m \,.\, \exists b_1, \ldots, b_r \,.\, F(a_1, \ldots, a_m, b_1, \ldots, b_r)$.

Note that using Proposition 5, it follows from this result that also the following decision problems are Π_2^P-hard when the weights are partially ordered:

$$\Delta^{prior_\subseteq}(K_1, \ldots, K_n; C; M_{s_1}, \ldots, M_{s_n}) \models \phi \tag{C.4}$$

$$\Delta^{prior_{card}}(K_1, \ldots, K_n; C; M_{s_1}, \ldots, M_{s_n}) \models \phi \tag{C.5}$$

3. For the operator $\Delta^{prior_\subseteq}$ with linear weights, membership in Π_2^P follows from the Π_2^P-completeness of the problem UNI-INCL considered in [73]. The Π_2^P-hardness is shown in entirely the same way as the Π_2^P-hardness of entailment checking for Δ^{poss} with partially ordered weights.
4. Entailment checking for $\Delta^{prior_\subseteq}$ with partially ordered weights is Π_2^P-hard, which follows by restricting to the special case where all weights are totally ordered. Membership in Π_2^P is proven by showing that the complement problem is in Σ_2^P. Indeed, to decide whether

$$\Delta^{prior_\subseteq}(K_1, \ldots, K_n; C; M_{s_1}, \ldots, M_{s_n}) \not\models \phi$$

we may use the following Σ_2^P procedure (inspired by the Π_2^P-hardness proof of PBR revision [74]):
- Guess a linearization κ;
- Guess a subset K of $\kappa(P_0)$;
- Verify that K is consistent by calling the NP oracle;
- Verify that $K \not\models \phi$ by calling the NP oracle;
- Verify that K is preferred w.r.t. inclusion. For this step, let $\lambda_1 < \cdots < \lambda_r \in [0, 1]$ be the weights that appear in $\kappa(P)$:
  - For each formula ϕ in $P_{\lambda_r} \setminus K$, check that $(K \cap P_{\lambda_r}) \cup \{\phi\}$ is inconsistent;
  - For each formula ϕ in $P_{\lambda_{r-1}} \setminus (K \cup P_{\lambda_r})$, check that $(K \cap P_{\lambda_{r-1}}) \cup \{\phi\}$ is inconsistent;
  - …
  - For each formula ϕ in $P_{\lambda_1} \setminus (K \cup P_{\lambda_2})$, check that $K \cup \{\phi\}$ is inconsistent;

  Clearly, the number of calls to the NP oracle is at most linear in the number of formulas in P.
5. Membership in Δ_2^P of entailment checking for $\Delta^{prior_{card}}$ with linearly ordered weights follows from the Δ_2^P-completeness of the UNI-LEX problem considered in [73]. We show hardness by reduction from the problem ALM, following an analogous approach as the proof of Δ_2^P-hardness in [73]. Given a satisfiable set of clauses Ψ over the variables $a_1, \ldots, a_n$, models of Ψ can be ordered as follows: $I \leqslant I'$ iff $I = I'$ or there is an i such that $I \cap \{a_1, \ldots, a_i\} = I' \cap \{a_1, \ldots, a_i\}$, $a_{i+1} \notin I$ and $a_{i+1} \in I'$. ALM is then the problem of deciding whether the model I of Ψ that is maximal w.r.t. this ordering makes a_n true.

   In particular, let $C = \Psi$, $K_1 = \{a_1, \ldots, a_n\}$, and $k = n - 1$. We furthermore define $W^l_{a_i} = X^l_{a_i} = Y^l_{a_i} = Z^l_{a_i} = \{a_i\}$ for $i \leqslant n - l$ and $W^l_{a_i} = X^l_{a_i} = \{a_i, \top\}$ and $Y^l_{a_i} = Z^l_{a_i} = \{a_i, \bot\}$ otherwise. Then there is only one preferred subset of P according to which it will first be tried to satisfy a_1, then a_2, etc. Thus we find that $\Delta^{prior_{card}}(K_1; C; M_{s_1}) \models a_n$ iff the maximal model of Ψ in the sense described above is such that a_n is true.
6. For partially ordered weights, the fact that entailment checking for $\Delta^{prior_{card}}$ is Π_2^P-hard was already established in (C.5). Membership in Π_2^P follows analogously as for $\Delta^{prior_\subseteq}$, using the fact that the MAX-GSAT-ARRAY problem considered in [73] is NP-complete.

7. Finally, let us consider the problem of checking entailment for Δ^{pen}. Let $m = \sum_{(\alpha,p)\in Q,\, p<+\infty} p$. By construction, the only formulas in Q with weight $+\infty$ are those in the knowledge bases $K_i^{@s_i}$ and C, which together are consistent by assumption. This means that the optimal consistent subsets of Q^* will have a penalty which is at most m. Checking whether there exists a consistent subset B of Q^* such that $pen(B) \leqslant p$ for a given p is clearly in NP. Using a number of calls to an NP-oracle which is logarithmic in m, we can thus find the smallest value p^* such that a consistent subset B of Q^* exists with $pen(B) = p^*$. To decide $\Delta^{pen}(K_1,\ldots,K_n; C; M_{s_1},\ldots,M_{s_n}) \models \phi$ it then suffices to verify whether $B \models \phi$ for all consistent subsets of Q^* for which $pen(B) = p^*$. This can be done in *co*NP; indeed, the complement can be shown by guessing an interpretation I, verifying that $pen(I) = p^*$ and that $I \not\models \phi$. This means that overall, we have a procedure which takes polynomial time on a deterministic machine, and makes a number of calls to an NP oracle which is logarithmic in m. Hence, if the penalties are bounded by an exponential function of the problem size, then $\log(m)$ is polynomial in the problem size, yielding membership in Δ_2^P, and if the penalties are bounded by a polynomial function, then $\log(m)$ is logarithmic in the problem size, yielding membership in Θ_2^P.
In the case of exponentially bounded penalties, hardness follows from Proposition 6 and the fact that entailment checking for $\Delta^{prior_{card}}$ with linear weights was shown to be Δ_2^P-complete. The Θ_2^P-hardness is the case of polynomially bounded penalties follows from the Θ_2^P-completeness of distance-based merging with the Hamming distance and the sum as aggregation operator [9]. Indeed, by restricting to the case $k = 0$, we know from Proposition 7 that Δ^{pen} degenerates to standard Hamming-distance based merging.

C.11. Proof of Proposition 9

1. Membership of $\approx_\subseteq^\exists$ and $\approx_{card}^\exists$ in Σ_2^P follows from the Σ_2^P-completeness of the EXI-INCL and EXI-LEX problems studied in [73]. To show hardness, we give a reduction from quantified boolean formula problems of the form $\exists\mathbf{a}\,.\,\forall\mathbf{b}\,.\,F(\mathbf{a},\mathbf{b})$. In particular let $K_1 = \{a_1,\ldots,a_m\}$, $K_2 = \{\neg a_1,\ldots,\neg a_m\}$, $C = \emptyset$, $\lambda^\times_{(l,s,a)} = \lambda_l \in]0,1[$, $k = 1$, and $W^1_{a_i} = X^1_{a_i} = \{a_i, \top\}$ and $Y^1_{a_i} = Z^1_{a_i} = \{a_i, \bot\}$. Then it is easy to see that there is a preferred subtheory B such that $B \models F(\mathbf{a},\mathbf{b})$ iff the QBF $\exists\mathbf{a}\,.\,\forall\mathbf{b}\,.\,F(\mathbf{a},\mathbf{b})$ is satisfied. Moreover, all preferred subtheories will satisfy exactly half of the formulas that appear in M_{s_1} and M_{s_2} with weight λ_0, and all formulas that appear with weight λ_1 (as the latter are all trivial). As a consequence the hardness results holds both for $\approx_\subseteq^\exists$ and $\approx_{card}^\exists$.
2. Σ_2^P membership of $\approx_\subseteq^{\exists\exists}$ and $\approx_{card}^{\exists\exists}$ is shown in the same way as for $\approx_\subseteq^\exists$ and $\approx_{card}^\exists$. Hardness follows immediately by restriction to the linearly ordered case.
3. Membership of $\approx_\subseteq^{\exists\forall}$ in Σ_3^P is straightforward: guess a linearization κ and verify whether the entailment holds for all $B \in Pref_{card}(\kappa(P))$ using a Σ_2^P oracle. Hardness is shown by reduction of quantified boolean formula problems of the form $\exists\mathbf{a}\,.\,\forall\mathbf{b}\,.\,\exists\mathbf{c}\,.\,H(\mathbf{a},\mathbf{b},\mathbf{c})$. Consider knowledge bases

$$K_1 = \{a_1,\ldots,a_m,b_1,\ldots,b_r,z\}$$
$$K_2 = \{\neg a_1,\ldots,\neg a_m,\neg b_1,\ldots,\neg b_r\}$$
$$C = \big\{z \equiv H(\mathbf{a},\mathbf{b},\mathbf{c})\big\}$$

and $W^1_x = X^1_x = \{x,\top\}$ and $Y^1_x = Z^1_x = \{x,\bot\}$ for all atoms x. Furthermore, let us take $\lambda^\times_{(1,s,x)} = \lambda^*$ for all atoms x, $\lambda^\times_{(0,s_1,a_i)} = \lambda_i^+$, $\lambda^\times_{(0,s_2,a_i)} = \lambda_i^-$, $\lambda^\times_{(0,s_j,b_i)} = \mu$ and $\lambda^\times_{(0,s_1,z)} = \mu$, where the ordering on the weights is such that $\mu < \lambda_i^-, \lambda_i^+ < \lambda^*$ for all i, and the weights λ_i^-, λ_i^+, λ_j^-, λ_j^+ are all incomparable for $i \neq j$. Then it is clear that any choice for the truth values of the atoms a_i is enforced by some specific linearization, and thus that $(K_1,\ldots,K_n) \approx H(\mathbf{a},\mathbf{b},\mathbf{c})$ iff there is a choice of truth values for the atoms a_i such that $H(\mathbf{a},\mathbf{b},\mathbf{c})$ is consistent with every choice of truth values for the atoms b_i, or indeed iff $\exists\mathbf{a}\,.\,\forall\mathbf{b}\,.\,\exists\mathbf{c}\,.\,H(\mathbf{a},\mathbf{b},\mathbf{c})$ is satisfiable.
4. Membership of $\approx_{card}^{\exists\forall}$ follows from the following procedure: guess a linearization κ and verify that the entailment holds for all $B \in Pref_{card}(\kappa(P))$. The complexity of this last step is Δ_2^P (see Proposition 8), hence it can be performed in polynomial time using an NP-oracle. Hardness is shown by taking

$$K_1 = \{a_1,\ldots,a_m\}, \qquad K_2 = \{\neg a_1,\ldots,\neg a_m\}, \qquad C = \emptyset$$

and $W^1_{a_i} = X^1_{a_i} = \{a,\top\}$ and $Y^1_{a_i} = Z^1_{a_i} = \{a,\bot\}$. Furthermore, let $\lambda^\times_{(1,s,a_i)} = \lambda^*$ and $\lambda^\times_{(0,s,a_i)} = \lambda_i$ such that $\lambda_i < \lambda^*$ and λ_i is incomparable to λ_j for $i \neq j$. Then every choice of truth values for the atoms a_i is enforced by some specific linearization, hence we clearly have that $\exists\mathbf{a}\,.\,\forall\mathbf{b}\,.\,F(\mathbf{a},\mathbf{b})$ iff there is a linearization κ such that $B \models F(\mathbf{a},\mathbf{b})$ for every $B \in Pref_{card}(\kappa(P))$.
5. Membership in Π_3^P for $\approx_\subseteq^{\forall\exists}$ is easily seen, by showing that the complement problem is in Σ_3^P. Indeed, it suffices to guess a linearization κ and then use a Σ_2^P oracle to verify wether for every preferred subtheory B of $\kappa(P)$, it is the case that $B \not\models \phi$. This is not the case if $\exists B \in Pref_\subseteq(\kappa(P))\,.\,B \models \phi$, and verifying this latter expression was already shown to be in Σ_2^P (in the first item). For the same reason, we have that $\approx_{card}^{\forall\exists}$ is in Π_3^P.
To show Π_3^P-hardness we simulate QBFs of the form $\forall\mathbf{a}\,.\,\exists\mathbf{b}\,.\,\forall\mathbf{c}\,.\,H(\mathbf{a},\mathbf{b},\mathbf{c})$. The knowledge bases here are chosen as

$$K_1 = \{a_1, \dots, a_m, b_1, \dots, b_r\}$$
$$K_2 = \{\neg a_1, \dots, \neg a_m, \neg b_1, \dots, \neg b_r\}$$
$$C = \emptyset$$

where the weights and the families W_x^l, X_x^l, Y_x^l and Z_x^l are defined as in the hardness proof of $\approx_{\subseteq}^{\exists\forall}$. Again we have that every choice of the truth values of the atoms a_i is enforced by some linearization κ. Given such a linearization κ, each preferred subtheory corresponds to a choice of the truth values of the atoms b_i. Hence, we have $(K_1, \dots, K_n) \approx_{\subseteq}^{\forall\exists} H(\mathbf{a}, \mathbf{b}, \mathbf{c})$ iff $\forall \mathbf{a} . \exists \mathbf{b} . \forall \mathbf{c} . H(\mathbf{a}, \mathbf{b}, \mathbf{c})$ holds. Since every preferred subtheory satisfies the same number of formulas at each priority level, Π_3^P-hardness for $\approx_{card}^{\forall\exists}$ follows as well.

C.12. Proof of Proposition 10

As an example, we show (B.1), the proof of (B.2) being entirely analogous. Since clearly $forgetVar(\Phi_1 \vee \Phi_2, X) \equiv forgetVar(\Phi_1, X) \vee forgetVar(\Phi_2, X)$, it suffices to show that for every $B \in Pref(P)$, we have

$$forgetVar\big(B, A^{@s_1} \cup \dots \cup A^{@s_n}\big) \equiv K_1^B \cup \dots \cup K_n^B \cup C$$

Let us write B_i for the formulas from M_{s_i} that are contained in B. We then find

$$\begin{aligned} forgetVar\big(B, a^{@s_i}\big) \equiv{}& K_1^{@s_1} \wedge \dots \wedge K_{i-1}^{@s_{i-1}} \wedge K_{i+1}^{@s_{i+1}} \wedge \dots \wedge K_n^{@s_n} \wedge B_1 \wedge \dots \wedge B_{i-1} \wedge B_{i+1} \wedge \dots \wedge B_n \\ & \wedge C \wedge forgetVar\big(K_i^{@s_i} \wedge B_i, a^{@s_i}\big) \end{aligned}$$

since only $K_i^{@s_i}$ and B_i contain occurrences of $a^{@s_i}$. Without lack of generality, we may assume that $K_i^{@s_i}$ is of the form $\{\alpha_i \vee a^{@s_i} \mid 1 \leqslant i \leqslant s\} \cup \{\beta_i \vee \neg a^{@s_i} \mid 1 \leqslant i \leqslant t\} \cup \Phi$, where none of the formulas α_i, β_i or the formulas in Φ contain occurrences of a. Moreover, note that the only non-trivial formulas in B_i that contain occurrences of $a^{@s_i}$ are $a^{@s_i} \to \bigvee W_a^r$ and $\neg a^{@s_i} \to \bigvee\{\neg y \mid y \in Y_a^{r'}\}$ for $r \geqslant w(a, s_i, W; B)$ and $r' \geqslant w(a, s_i, Y; B)$; let Ψ be the set of formulas from B_i which do not refer to $a^{@s_i}$. We find

$$\begin{aligned} & forgetVar\big(K_i^{@s_i} \wedge B_i, a^{@s_i}\big) \\ &\equiv \big(K_i^{@s_i}\big[a^{@s_i} := \top\big] \wedge B_i\big[a^{@s_i} := \top\big]\big) \vee \big(K_i^{@s_i}\big[a^{@s_i} := \bot\big] \wedge B_i\big[a^{@s_i} := \bot\big]\big) \\ &\equiv \left(\bigwedge_{i=1}^{t} \beta_i \wedge \bigwedge \Phi \wedge \bigwedge_{r=w(a,s_i,W;B)}^{k} \bigvee W_a^r \wedge \bigwedge \Psi\right) \vee \left(\bigwedge_{i=1}^{s} \alpha_i \wedge \bigwedge \Phi \wedge \bigwedge_{r=w(a,s_i,Y;B)}^{k} \bigvee\big\{\neg y \bigm| y \in Y_a^r\big\} \wedge \bigwedge \Psi\right) \\ &\equiv \left(\bigwedge_{i=1}^{t} \beta_i \wedge \bigwedge \Phi \wedge \bigvee W_a^{w(a,s_i,W;B)} \wedge \bigwedge \Psi\right) \vee \left(\bigwedge_{i=1}^{s} \alpha_i \wedge \bigwedge \Phi \wedge \bigvee\big\{\neg y \bigm| y \in Y_a^{w(a,s_i,Y;B)}\big\} \wedge \bigwedge \Psi\right) \\ &\equiv \bigwedge \Phi \wedge \bigwedge \Psi \wedge \left(\left(\bigwedge_{i=1}^{t} \beta_i \wedge \bigvee W_a^{w(a,s_i,W;B)}\right) \vee \left(\bigwedge_{i=1}^{s} \alpha_i \wedge \bigvee\big\{\neg y \bigm| y \in Y_a^{w(a,s_i,Y;B)}\big\}\right)\right) \\ &\equiv \bigwedge \Phi \wedge \bigwedge \Psi \wedge \bigwedge_{i=1}^{t} \bigwedge_{j=1}^{s} (\beta_i \vee \alpha_j) \wedge \bigwedge_{i=1}^{t} \Big(\beta_i \vee \bigvee\big\{\neg y \bigm| y \in Y_a^{w(a,s_i,Y;B)}\big\}\Big) \\ &\quad \wedge \bigwedge_{j=1}^{s} \Big(\bigvee W_a^{w(a,s_i,W;B)} \vee \alpha_i\Big) \wedge \Big(\bigvee W_a^{w(a,s_i,W;B)} \vee \bigvee\big\{\neg y \bigm| y \in Y_a^{w(a,s_i,Y;B)}\big\}\Big) \end{aligned}$$

Note that $\bigvee W_a^{w(a,s_i,W;B)} \vee \bigvee\{\neg y \mid y \in Y_a^{w(a,s_i,Y;B)}\}$ is trivially satisfied since $a \in W_a^{w(a,s_i,W;B)}$ and $a \in Y_a^{w(a,s_i,Y;B)}$. Furthermore, for each i and j, we have that Φ contains a formula which is equivalent to $\beta_i \vee \alpha_j$, from the assumption that K_i is equal to its set of prime implicates. This leads to

$$\bigwedge \Phi \wedge \bigwedge \Psi \wedge \bigwedge_{i=1}^{t} \Big(\beta_i \vee \bigvee\big\{\neg y \bigm| y \in Y_a^{w(a,s_i,Y;B)}\big\}\Big) \wedge \bigwedge_{j=1}^{s} \Big(\bigvee W_a^{w(a,s_i,W;B)} \vee \alpha_j\Big)$$

which corresponds exactly to replacing every occurrence of a positive literal $a^{@s_i}$ by $\bigvee W_a^{w(a,s_i,W;B)}$ and every occurrence of a negative literal $\neg a^{@s_i}$ by $\bigvee\{\neg y \mid y \in Y_a^{w(a,s_i,Y;B)}\}$. This process can be repeated for all other atoms from $A_1^{@s_1} \cup \dots \cup A_n^{@s_n}$, leading to the stated equivalence.

References

[1] M. Abidi, R. Gonzalez, Data Fusion in Robotics and Machine Intelligence, Academic Press, 1992.
[2] I. Matzkevich, B. Abramson, The topological fusion of Bayes nets, in: Proceedings of the Eighth Conference on Uncertainty in Artificial Intelligence, 1992, pp. 191–198.
[3] L. Cholvy, A logical approach to multi-sources reasoning, in: International Conference on Knowledge Representation and Reasoning Under Uncertainty, Logic at Work, 1994, pp. 183–196.
[4] S. Benferhat, D. Dubois, H. Prade, How to infer from inconsistent beliefs without revising, in: IJCAI'95: Proceedings of the 14th International Joint Conference on Artificial Intelligence, 1995, pp. 1449–1455.
[5] J. Lin, Integration of weighted knowledge bases, Artificial Intelligence 83 (2) (1996) 363–378.
[6] L. Amgoud, S. Kaci, An argumentation framework for merging conflicting knowledge bases, International Journal of Approximate Reasoning 45 (2) (2007) 321–340.
[7] R. Booth, Social contraction and belief negotiation, Information Fusion 7 (1) (2006) 19–34.
[8] P. Revesz, On the semantics of theory change: arbitration between old and new information, in: Proceedings of the 12th ACM SIGACT-SIGMOD-SIGART Symposium on Principles of Database Systems, 1993, pp. 71–82.
[9] S. Konieczny, J. Lang, P. Marquis, DA2 merging operators, Artificial Intelligence 157 (1–2) (2004) 49–79.
[10] I. Bloch, J. Lang, Towards mathematical morpho-logics, in: Technologies for Constructing Intelligent Systems, Physica-Verlag GmbH, 2002, p. 380.
[11] D. Dubois, H. Prade, Possibility theory and data fusion in poorly informed environments, Control Engineering Practice 2 (1994) 811–823.
[12] D. Dubois, H. Fargier, H. Prade, Multiple-sources informations fusion – a practical inconsistency-tolerant approach, in: Proc. 8th International Conference on Information Processing and Management of Uncertainty in Knowledge-Based Systems (IPMU'00), Madrid, July 13–17, 2000, pp. 1047–1054.
[13] G. Priest, Paraconsistent logic, in: Handbook of Philosophical Logic, vol. 6, Kluwer Academic Publishers, 2002, pp. 287–393.
[14] W. Carnielli, M. Coniglio, J. Marcos, Logics of formal inconsistency, in: Handbook of Philosophical Logic, vol. 14, Springer, 2005, pp. 1–93.
[15] P. Gärdenfors, Knowledge in Flux, MIT Press, 1988.
[16] N. Rescher, R. Manor, On inference from inconsistent premisses, Theory and Decision 1 (1970) 179–217.
[17] S. Konieczny, R. Pino Pérez, Merging information under constraints: a logical framework, Journal of Logic and Computation 12 (5) (2002) 773–808.
[18] M. Dalal, Investigations into a theory of knowledge base revision: preliminary report, in: Proceedings of the 7th National Conference on Artificial Intelligence, 1988, pp. 475–479.
[19] C. Lafage, J. Lang, Propositional distances and compact preference representation, European Journal of Operational Research 160 (3) (2005) 741–761.
[20] P. Everaere, S. Konieczny, P. Marquis, Conflict-based merging operators, in: Eleventh International Conference on Principles of Knowledge Representation and Reasoning (KR'08), 2008, pp. 348–357.
[21] J. Serra, Image Analysis and Mathematical Morphology, Academic Press, 1982.
[22] N. Gorogiannis, A. Hunter, Merging first-order knowledge using dilation operators, in: Proceedings of the 5th International Conference on Foundations of information and knowledge systems, 2008, pp. 132–150.
[23] J. Delgrande, T. Schaub, A consistency-based framework for merging knowledge bases, Journal of Applied Logic 5 (3) (2007) 459–477.
[24] J. Lang, P. Marquis, Reasoning under inconsistency: A forgetting-based approach, Artificial Intelligence 174 (12–13) (2010) 799–823.
[25] D. Dubois, J. Lang, H. Prade, Possibilistic logic, in: D.N.D. Gabbay, J. Robinson, C. Hogger (Eds.), Handbook of Logic in Artificial Intelligence and Logic Programming, vol. 3, Oxford University Press, 1994, pp. 439–513.
[26] P. Liberatore, Merging locally correct knowledge bases: A preliminary report, Tech. rep., Computing Research Repository (CoRR), arXiv:cs.AI/0212053, 2002.
[27] G. Pinkas, Propositional non-monotonic reasoning and inconsistency in symmetric neural networks, in: Proceedings of the 12th International Joint Conference on Artificial Intelligence, vol. 1, 1991, pp. 525–530.
[28] F. Dupin de Saint-Cyr, J. Lang, T. Schiex, Penalty logic and its link with Dempster–Shafer theory, in: Proceedings of the 10th International Conference on Uncertainty in Artificial Intelligence, 1994, pp. 204–211.
[29] Z. Pawlak, Rough sets, International Journal of Parallel Programming 11 (5) (1982) 341–356.
[30] P. Gärdenfors, Conceptual Spaces: The Geometry of Thought, MIT Press, 2000.
[31] C. Freksa, Conceptual neighborhood and its role in temporal and spatial reasoning, in: M. Singh, L. Travé-Massuyès (Eds.), Decision Support Systems and Qualitative Reasoning, North-Holland, Amsterdam, 1991, pp. 181–187.
[32] C. Freksa, Temporal reasoning based on semi-intervals, Artificial Intelligence 54 (1–2) (1992) 199–227.
[33] D. Dubois, H. Prade, R. Yager, Merging fuzzy information, in: J. Bezdek, D. Dubois, H. Prade (Eds.), Fuzzy Sets in Approximate Reasoning and Information Systems, in: The Handbooks of Fuzzy Sets Series, Kluwer, Boston, Mass., 1999, pp. 335–401.
[34] D. Gabbay, O. Rodrigues, G. Pigozzi (Eds.), Journal of Logic and Computation, special issue on Connections between Belief Revision, Belief Merging and Social Choice 19 (3) (2009).
[35] F. Lin, R. Reiter, Forget it, in: Working Notes of AAAI Fall Symposium on Relevance, 1994, pp. 154–159.
[36] G. Brewka, Preferred subtheories: An extended logical framework for default reasoning, in: Proc. of the 11th Int. Joint Conf. on, Artificial Intelligence, 1989, pp. 1043–1048.
[37] T. Eiter, G. Gottlob, The complexity of logic-based abduction, Journal of the ACM 42 (1) (1995) 3–42.
[38] M. Goldszmidt, J. Pearl, Qualitative probabilities for default reasoning, belief revision, and causal modeling, Artificial Intelligence 84 (1–2) (1996) 57–112.
[39] W. Spohn, Ordinal conditional functions: A dynamic theory of epistemic states, Causation in Decision, Belief Change and Statistics 2 (1988) 105–134.
[40] D. Dubois, H. Prade, Epistemic entrenchment and possibilistic logic, Artificial Intelligence 50 (2) (1991) 223–239.
[41] S. Benferhat, D. Dubois, S. Kaci, H. Prade, Possibilistic merging and distance-based fusion of propositional information, Annals of Mathematics and Artificial Intelligence 34 (1–3) (2002) 217–252.
[42] S. Benferhat, S. Lagrue, O. Papini, Reasoning with partially ordered information in a possibilistic logic framework, Fuzzy Sets and Systems 144 (1) (2004) 25–41.
[43] A. Dempster, Upper and lower probabilities induced by a multivalued mapping, The Annals of Mathematical Statistics 38 (2) (1967) 325–339.
[44] G. Shafer, A Mathematical Theory of Evidence, Princeton University Press, Princeton, NJ, 1976.
[45] S. Konieczny, R. Pino-Pérez, On the frontier between arbitration and majority, in: Eighth International Conference on Principles of Knowledge Representation and Reasoning, 2002, pp. 109–120.
[46] C. Papadimitriou, Computational Complexity, John Wiley and Sons Ltd., 2003.
[47] T. Eiter, G. Gottlob, The complexity class Θ_2^p: Recent results and applications in AI and modal logic, in: Proceedings of the 11th International Symposium on Fundamentals of Computation Theory, in: LNCS, vol. 1279, Springer, 1997, pp. 1–18.
[48] W. Wahlster, Implementing fuzziness in dialogue systems, in: B. Rieger (Ed.), Empirical Semantics, Brockmeyer, Bochum, 1980.
[49] A.H. Jucker, S.W. Smith, T. Lüdge, Interactive aspects of vagueness in conversation, Journal of Pragmatics 35 (12) (2003) 1737–1769.
[50] D. Traum, A computational theory of grounding in natural language conversation, PhD thesis, University of Rochester, 1994.

[51] M. Pickering, S. Garrod, Toward a mechanistic psychology of dialogue, Behavioral and Brain Sciences 27 (2) (2004) 169–190.
[52] J. Lawry, Imprecise bipolar belief measures based on partial knowledge from agent dialogues, in: Scalable Uncertainty Management, in: Lecture Notes in Computer Science, vol. 6379, Springer, 2010, pp. 205–218.
[53] J. Reggia, D. Nau, P. Wang, H. Peng, A formal model of diagnostic inference, Information Sciences 37 (1985) 227–285.
[54] E. Ruspini, On the semantics of fuzzy logic, International Journal of Approximate Reasoning 5 (1991) 45–88.
[55] S. Kraus, D. Lehmann, M. Magidor, Nonmonotonic reasoning, preferential models and cumulative logics, Artificial Intelligence 44 (1–2) (1990) 167–207.
[56] D. Dubois, H. Prade, F. Esteva, P. Garcia, L. Godo, A logical approach to interpolation based on similarity relations, International Journal of Approximate Reasoning 17 (1) (1997) 1–36.
[57] D. Dubois, H. Prade, Similarity versus preference in fuzzy set-based logics, in: E. Orlowska (Ed.), Modelling Incomplete Information: Rough Set Analysis, Physica-Verlag, Heidelberg, 1998, pp. 441–461.
[58] R. Rodriguez, P. Garcia, L. Godo, Relating similarity-based models, counterfactuals and theory change, in: Proc. EUFIT' 95, 1995, pp. 230–234.
[59] E. Ovchinnikova, T. Wandmacher, K.-U. Kühnberger, Solving terminological inconsistency problems in ontology design, Interoperability in Business Information Systems 2 (1) (2007) 65–80.
[60] J.-F. Condotta, S. Kaci, N. Schwind, A framework for merging qualitative constraint networks, in: Proc. of the 21st International FLAIRS Conf., 2008, pp. 586–591.
[61] S. Schockaert, M. De Cock, E.E. Kerre, Reasoning about fuzzy temporal information from the web: towards retrieval of historical events, Soft Computing 14 (8) (2010) 869–886.
[62] J. Allen, Maintaining knowledge about temporal intervals, Communications of the ACM 26 (11) (1983) 832–843.
[63] S. Schockaert, M. De Cock, Temporal reasoning about fuzzy intervals, Artificial Intelligence 172 (2008) 1158–1193.
[64] D. Randell, Z. Cui, A. Cohn, A spatial logic based on regions and connection, in: Proceedings of the 3rd International Conference on Knowledge Representation and Reasoning, 1992, pp. 165–176.
[65] S. Schockaert, M. De Cock, E.E. Kerre, Spatial reasoning in a fuzzy region connection calculus, Artificial Intelligence 173 (258–298).
[66] G. Lakoff, Hedges: A study in meaning criteria and the logic of fuzzy concepts, Journal of Philosophical Logic 2 (4) (1973) 458–508.
[67] L. Zadeh, A fuzzy-set-theoretic interpretation of linguistic hedges, Cybernetics and Systems 2 (3) (1972) 4–34.
[68] P. MacVicar-Whelan, Fuzzy sets, the concept of height and the hedge very, IEEE Transactions on Systems, Man and Cybernetics 8 (1978) 507–511.
[69] S. Schockaert, H. Prade, Merging conflicting propositional knowledge by similarity, in: Proceedings of the 2009 21st IEEE International Conference on Tools with Artificial Intelligence, 2009, pp. 224–228.
[70] S. Schockaert, H. Prade, An inconsistency-tolerant approach to information merging based on proposition relaxation, in: Proceedings of the Twenty-Fourth AAAI Conference on Artificial Intelligence, 2010.
[71] J. Lang, Possibilistic logic: complexity and algorithms, in: J. Kohlas, S. Moral, D. Gabbay, P. Smets (Eds.), Algorithms for Uncertainty and Defeasible Reasoning, in: Handbook of Defeasible Reasoning and Uncertainty Management Systems, vol. 5, Kluwer Academic Publishers, 2001, pp. 179–220.
[72] K.W. Wagner, More complicated questions about maxima and minima, and some closures of NP, in: International Colloquium on Automata, Languages and Programming on Automata, Languages and Programming, Springer-Verlag, New York, 1986, pp. 434–443.
[73] C. Cayrol, M.-C. Lagasquie-Schiex, T. Schiex, Nonmonotonic reasoning: from complexity to algorithms, Annals of Mathematics and Artificial Intelligence 22 (3–4) (1998) 207–236.
[74] B. Nebel, Syntax-based approaches to belief revision, in: P. Gärdenfors (Ed.), Belief Revision, Cambridge University Press, 1992, pp. 52–88.