Approaches to Constructing a Stratified Merged Knowledge Base

Anbu Yue[1], Weiru Liu[1], and Anthony Hunter[2]

[1] School of Electronics, Electrical Engineering and Computer Science,
Queen's University Belfast, Belfast BT7 1NN, UK
{a.yue, w.liu}@qub.ac.uk
[2] Department of Computer Science, University College London,
Gower Street, London WC1E 6BT, UK
a.hunter@cs.ucl.ac.uk

Abstract. Many merging operators have been proposed to merge either flat or stratified knowledge bases. The result of merging by such an operator is a flat base (or a set of models of the merged base) irrespective of whether the original ones are flat or stratified. The drawback of obtaining a flat merged base is that information about more preferred knowledge (formulae) versus less preferred knowledge is not explicitly represented, and this information can be very useful when deciding which formulae should be retained when there is a conflict. Therefore, it can be more desirable to return a stratified knowledge base as a merged result. A straightforward approach is to deploy the preference relation over possible worlds obtained after merging to reconstruct such a base. However, our study shows that such an approach can produce a poor result, that is, preference relations over possible worlds obtained after merging are not suitable for reconstructing a merged stratified base. Inspired by the Condorcet method in voting systems, we propose an alternative method to stratify a set of possible worlds given a set of stratified bases and take the stratification of possible worlds as the result of merging. Based on this, we provide a family of syntax-based methods and a family of model-based methods to construct a stratified merged knowledge base. In the syntax based methods, the formulae contained in the merged knowledge base are from the original individual knowledge bases. In contrast, in the model based methods, some additional formulae may be introduced into the merged knowledge base and no information in the original knowledge bases is lost. Since the merged result is a stratified knowledge base, the commonly agreed knowledge together with a preference relation over this knowledge can be extracted from the original knowledge bases.

1 Introduction

Preference (or priority) is very important in many fields of computer science, such as, in constraint satisfaction problems, in goal oriented decision making, and in system configurations. A preference relation can be used to represent an ordering over beliefs or goals. Preferences can be explicitly modelled in possibilistic logic (e.g. [5]) or using an ordinal conditional function (e.g., [9, 11]), or a preference relation. With explicit preference information, a flat knowledge base can be extended to a stratified knowledge base, that is, propositions are divided into different strata according to the preferences (or priorities) they are given [2, 3].

When multiple knowledge bases are available, one objective is to extract an overall view from them. This is known as knowledge base merging or belief merging. There are mainly two families of knowledge base merging operators: model-based ones which select some possible worlds (or interpretations) that are the *closest* to the original bases (e.g. [7, 8, 10]) and syntax-based ones which pick some formulae in the union of the original bases (e.g., [1, 4]). It is worth noting that most of these merging operators (e.g. [7, 8, 1, 4]) take flat knowledge bases as input, and only a few of them (e.g. [10]) allow the original knowledge bases to be stratified.

So far, all the merging operators, no matter whether model-based or syntax-based, return a flat knowledge base (or a set of models) as the merged result. We argue that, in practice, we may need a stratified knowledge base as the merged result which shows preference among formulae, irrespective of whether the original knowledge bases are flat or stratified. If a stratified knowledge base can be obtained as the result of merging, the preference over formulae can be very useful to resolve conflicts, that is, a more preferred formula should be retained in preference to a less preferred one if the two are in conflict.

Intuitively, it seems straightforward that a stratified merged base should be constructed easily from a model-based merging method using the preference relation over interpretations obtained after merging. However, our study of the merging methods in [8, 10] shows that such a method is not adequate.

In [8], the commensurability assumption is required, and so a number[1] assigned to an interpretation (possible world) can be compared with another number assigned to another interpretation. On the other hand, although the commensurability assumption is not explicitly required in [10], each interpretation is still assigned a vector of numbers, each of which is the *priority* (or the absolute position) of the interpretation in the stratification of interpretations in relation to an individual knowledge base, and the numbers assigned in different stratifications are assumed comparable in order to establish the pre-order relation over vectors of numbers.

We argue that the numbers about distances or priorities obtained from different stratifications should not be comparable if we do not have the commensurability assumption, especially when knowledge bases are designed independently.

Inspired by the Condorcet method in *voting systems*, we provide an alternative method to define a preference relation, called *relative preference relation*, over interpretations. A relative preference relation considers whether an interpretation is more preferred than another collectively from a set of pre-order relations (or stratifications) over interpretations obtained from individual knowledge bases, but is independent of the absolute *priority* ([10]) or the *distance* ([8]) of an interpretation in relation to each single knowledge base, since these *numbers* are not assumed comparable. Then, the stratification of interpretations is constructed from this relative preference relation.

[1] This number can be the sum of distances between an interpretation and all the original knowledge bases or a vector of numbers each of which is the distance between the interpretation and a knowledge base.

Following this, we provide two families of methods, namely syntax-based and model-based, to construct a merged stratified knowledge base from the stratification of interpretations. The syntax-based methods assume that formulae in the stratified knowledge base are picked from the original bases. A disadvantage of these methods is that some implicit beliefs are lost. The model-based methods use models to reconstruct formulae in the merged base. These methods retain all the original knowledge and may also introduce additional formulae which do not appear in any original knowledge bases.

This paper is organized as follows. In Section 2, we introduce the preliminaries. In Section 3, we first explore a straightforward approach to constructing the stratification of interpretations after merging and to constructing a merged stratified knowledge base. We then provide an alternative approach to constructing the stratification of interpretations after merging by defining the concept of *relative preference relation*. In Section 4, we provide syntax-based methods for constructing a stratified knowledge base from a stratification of interpretations. Then, in Section 5, we provide model-based methods for constructing a stratified knowledge base. Finally, a brief comparison with related work and a short summary of the paper are given in Section 6.

2 Preliminary

2.1 Stratified knowledge base

We consider a propositional language $\mathcal{L}$ defined on a finite set $\mathcal{A}$ of propositional atoms, which are denoted by p, q, r etc. A proposition ϕ is constructed by propositional atoms with logic connections $\neg, \wedge, \vee, \rightarrow$ in the standard way. An interpretation ω (or possible world) is a function that maps $\mathcal{A}$ onto the set $\{0, 1\}$. The set of all possible interpretations on $\mathcal{A}$ is denoted as Ω. Function ω can be extended to any proposition in $\mathcal{L}$ in the usual way, $\omega : \mathcal{L} \rightarrow \{0, 1\}$. An interpretation ω is a model of (or satisfies) ϕ iff $\omega(\phi) = 1$, denoted as $\omega \models \phi$.

A *(flat) knowledge base* is a finite set of propositional formulae. A knowledge base K is consistent iff there is at least one interpretation that satisfies all propositions in K, and such interpretations are models of K. We use $Mod(K)$ to denote the set of models for K. $K \models \phi$ iff each model of K is a model of ϕ. For a set of models M, there exists a proposition ϕ_M s.t. $Mod(\phi_M) = M$. Theoretically, ϕ_M is non-deterministic since syntactically there can be more than one proposition that is satisfied by all the models in M.

A *stratified knowledge base* [3, 2] is a finite set K of propositional formulae with a total pre-order relation $\preceq$ on K (a pre-order relation is a reflective and transitive relation, and $\preceq$ on K is total iff for all ϕ and ψ in K, either $\phi \preceq \psi$ or $\psi \preceq \phi$ holds). Intuitively, if $\phi \preceq \psi$ then ϕ is regarded as more certain, more preferred or more important than ψ. From the pre-order relation $\preceq$ on K, K can be stratified as $K = (S_1, \ldots, S_n)$, where S_i contains all the minimal propositions of set $\bigcup_{j=i}^{n} S_j$ w.r.t. $\preceq$, i.e., $S_i = \{\phi \in K \setminus (\cup_{j=1}^{i-1} S_j) : \forall \varphi \in K \setminus (\cup_{j=1}^{i-1} S_i), \phi \leq \varphi\}$. Each S_i is called a stratum of K and is non-empty. In the rest of this paper, we denote $\bigcup K = \bigcup_{i=1}^{n} S_i$. It is clear that for all ϕ and ψ in K, $\phi \preceq \psi$ iff $\phi \in S_i$, $\psi \in S_j$, and $i \leq j$. For simplicity, when we mention a knowledge base it is actually a stratified knowledge base unless it is stated otherwise in the rest of this paper.

2.2 Model based semantics

In [3, 2], some model-based semantics are provided for stratified knowledge bases. In these methods, a pre-order relation on interpretations is induced from a knowledge base by an *ordering strategy*, and the minimal ones are regarded as the models of the knowledge base. Therefore, a non-classical consequence relation can be defined as $K \Vdash_X \phi$ iff $\omega(\phi) = 1$ for all ω such that ω is minimal w.r.t the pre-order relation $\preceq_X$ over Ω, where $\preceq_X$ is induced by K under the ordering strategy X. A strict relation $\prec_X$ is defined as $\omega \prec_X \omega'$ iff $\omega \preceq_X \omega'$ and $\omega' \not\preceq_X \omega$.

There are three widely used ordering strategies known as the *best out*, the *maxsat*, and the *leximin*. For a knowledge base $K = (S_1, \ldots, S_n)$, these ordering strategies are defined as follows.

- **best out ordering [2]** Let $r_{BO}(\omega) = \min_i\{\omega \not\models S_i\}$. Define $min_i\emptyset = +\infty$. $\omega \preceq_{bo} \omega'$ iff $r_{BO}(\omega) \geq r_{BO}(\omega')$.
- **maxsat ordering [3]** Let $r_{MO}(\omega) = \min_i\{\omega \models S_i\}$. $\omega \preceq_{maxsat} \omega'$ iff $r_{MO}(\omega) \leq r_{MO}(\omega')$.
- **leximin ordering [2]**: let $K_i(\omega) = \{\phi \in S_i \mid \omega \models \phi_i\}$. Then the leximin ordering $\preceq_{leximin}$ on Ω is defined as: $\omega \preceq_{leximin} \omega'$ iff
  - $|K_i(\omega)| = |K_i(\omega')|$ for all i, or
  - there is an i s.t. $|K_i(\omega)| > |K_i(\omega')|$, and $|K_j(\omega)| = |K_j(\omega')|$ for all $j < i$;

  where $|K_i|$ denotes the cardinality of set K_i.

Example 1. Let $K = (\{p\}, \{q\})$ be a knowledge base.

Table 1. Ranks calculated by different ordering strategies

ω	r_{BO}	r_{MO}	$\bar{r}_{Leximin}$
00	1	$+\infty$	$\langle 00\rangle$
01	1	2	$\langle 01\rangle$
10	2	1	$\langle 10\rangle$
11	$+\infty$	1	$\langle 11\rangle$

The i-th digit in vector $\bar{r}_{Leximin}(\omega)$ represents $|K_i(\omega)|$. For example, let $\omega = \{01\}$ represent possible world $\neg pq$, then $\bar{r}_{Leximin}(\omega) = \langle 01\rangle$ means that $K_1(\omega) = 0$ and $K_2(\omega) = 1$. Then $\{11\}$ is the set of minimal models w.r.t the pre-order relations $\prec_{bo}$ and $\prec_{leximin}$, and set $\{11, 10\}$ contains all the minimal models w.r.t the pre-order relation $\prec_{maxsat}$. Therefore, $K \Vdash_{BO} p \wedge q$, $K \Vdash_{Maxsat} p$, and $K \Vdash_{Leximin} p \wedge q$.

From a pre-order relation $\preceq_X$ generated by the ordering strategy X from K, the interpretations in Ω can be stratified as $\Omega_{K,X} = (\Omega_1, \ldots, \Omega_m)$, where each Ω_i contains all the minimal interpretations from $\bigcup_{j=i}^{n} \Omega_j$ w.r.t. $\preceq_X$. Given $\Omega_{K_1} = (\Omega_1, \ldots, \Omega_n)$ and $\Omega_{K_2} = (\Omega'_1, \ldots, \Omega'_n)$, $\Omega_{K_1} = \Omega_{K_2}$ iff $\Omega_i = \Omega'_i$ for all i where $1 \leq i \leq n$. In the following, we may omit subscripts K, X from $\Omega_{K,X}$ when they are implicitly given.

As shown above, based on different ordering strategies, different conclusions are drawn from the same knowledge base K. Thus, selecting an ordering strategy for a given knowledge base is important. Also, it is possible that the same stratification on interpretations can be induced from different knowledge bases under different ordering strategies.

3 Approaches to Stratifying the Set of Interpretations

In a model-based merging method, a pre-order relation on interpretations is constructed and a set of models (the minimal ones) is obtained as the result of merging flat or stratified knowledge bases. The resulting knowledge base is a flat base. Intuitively, it seems reasonable to recover a stratified merged knowledge base from the pre-order relation over interpretations. Following this idea, we take merging operators in [10, 8] as examples and investigate if this approach is feasible. Our study below in Section 3.1 shows that a stratified merged knowledge base obtained this way can be counterintuitive. To overcome this problem, we propose an alternative method to stratify a set of interpretations using the concept of relative preference relation in Section 3.2.

3.1 A simple approach

In [10], a model-based merging method is proposed for merging stratified knowledge bases, however, the result is a flat knowledge base not a stratified one. The idea in the paper can be stated as follows. From each stratified knowledge base K with a chosen ordering strategy X, a stratification of interpretations is induced $\Omega_{K,X}$. In this way, an interpretation has a priority level w.r.t each K which is its priority level in $\Omega_{K,X}$. Then, each interpretation is associated with a vector of priority levels in relation to all the knowledge bases. Finally, a pre-order relation over interpretations is defined based on the lexicographical ordering over vectors of priorities and the interpretations which are minimal w.r.t this ordering relation are regarded as the models of the merged knowledge base. A straightforward approach to obtaining a stratified merged knowledge base is to construct strata directly from this pre-order relation over interpretations. Unfortunately, such a method is not as good as one may expect, as shown in the following example.

Example 2. Let $K_1 = (\{p\}, \{q\}, \{r\})$ and $K_2 = (\{r\}, \{q\}, \{p\})$ be two knowledge bases. Using the leximin ordering strategy, two pre-order relations on interpretations can be induced from them respectively, and a pre-order relation for the merged knowledge base can be calculated under the leximin aggregation function as shown in Table 2. In Table 2, values in vector $\bar{l}$ are obtained by concatenating the numbers in the second and the third columns in ascending order.

Based on the lexicographical ordering over vectors of priorities, a pre-order relation is defined on Ω and it is stratified as $\Omega = (\{111\}, \{110, 011\}, \{101\}, \{001, 100\}, \{010\}, \{000\})$. The minimal model is 111, so, the merged flat knowledge base is equivalent to $\{p \wedge q \wedge r\}$, which seems reasonable. Note, in this paper we denote each model by a bit vector consisting of truth values of atoms e.g. (p, q, r) in this example. So $\omega_1 = \{000\}$ means p, q and r are all false.

From this stratified Ω, we can infer that $p \wedge \neg q \wedge r$ is less preferable than $q \wedge (p \vee r) \wedge (\neg p \vee \neg r)$ which means that, if only two of p, q, r are true, then q must be true. In other words, q is more certain than both p and r. On the other hand, $\neg p \wedge q \wedge \neg r$ is less preferable than $\neg q \wedge (p \vee r) \wedge (\neg p \vee \neg r)$ which implies that when only one of p, q, r holds, it should not be q, that is, q is less certain than both p and r now. This contradicts with the previous inference. Therefore, taking the stratification of interpretations as a way to construct a merged knowledge base may imply counterintuitive results.

Table 2. Constructing the stratification of interpretations.

ω	$\Omega_{K_1,leximin}$	$\Omega_{K_2,leximin}$	$\bar{l}$
000	8	8	$\langle 88\rangle$
001	7	4	$\langle 47\rangle$
010	6	6	$\langle 66\rangle$
011	5	2	$\langle 25\rangle$
100	4	7	$\langle 47\rangle$
101	3	3	$\langle 33\rangle$
110	2	5	$\langle 25\rangle$
111	1	1	$\langle 11\rangle$

If we take a flat knowledge base as a special case of a stratified knowledge base with only one stratum, the merging methods in DA^2 family [8] can be viewed as special cases of merging stratified knowledge bases. Similarly, in these methods, a set of models for a flat knowledge base is given as the result of an operator.

Example 3. Let $K_1 = K_2 = K_3 = K_4 = K_5 = \{p\}$ and $K_6 = K_7 = K_8 = K_9 = \{\neg p\}$ be nine knowledge bases. Five of them say that p is true and four say that p is false. Merging by $\delta^{d_D, sum_1, sum_2}$, a specific operator in DA^2 [8], the model with p is true (having true value 1) is the only model for the merged knowledge base. In this merging operator, d_D is a distance measure between a formula and a possible world and it is defined as $d_D(\omega, \phi) = 0$ if $\omega \models \phi$, otherwise $d_D(\omega, \phi) = 1$. $\mathsf{sum}_1(\omega, K_i) = \Sigma_{\phi \in K_i} d_D(\omega, \phi)$ and $\mathsf{sum}_2(\omega, P) = \Sigma_{K_i \in P} \mathsf{sum}_1(\omega, K_i)$, which is the sum of distances between knowledge bases K_i in a knowledge profile P (a knowledge profile is a finite set of knowledge bases) and a possible world. In Table 3, values in the second to the tenth columns are the distances between an interpretation and a knowledge base (using sum_1), values in the eleventh column are from sum_2.

Table 3. Merging knowledge bases using an operator in DA^2

ω	K_1	K_2	K_3	K_4	K_5	K_6	K_7	K_8	K_9	sum_2
$\neg p$	1	1	1	1	1	0	0	0	0	5
p	0	0	0	0	0	1	1	1	1	4

Now if we revise $K_6, ..., K_9$ as $K'_6 = K'_7 = \{\neg p, \neg p \wedge q\}$ and $K'_8 = K'_9 = \{\neg p, \neg p \wedge \neg q\}$ respectively, then semantically, these four knowledge bases together state the same conclusion as K_6, ..., K_9 do. That is, both sets of knowledge bases say p is false. However, when we replace $K_6, ..., K_9$ with $K'_6, ..., K'_9$ and merge them with $K_1, ..., K_5$, we get a different merged knowledge base K', whose models are $\{00, 01\}$ as shown in Table 4. Obviously, $K' \models \neg p$, $K' \not\models q$ and $K' \not\models \neg q$.

The reason is that, in DA^2, the commensurability assumption is required. Under this assumption, although $K'_6, ..., K'_9$ collectively draw the same conclusion as $K_6, ..., K_9$, when they are merged with the other bases, the preferability of the statement p is implicitly decreased. That is why the merged result is changed.

To summarize, we believe that using the information on stratification over interpretations from a merging operator to construct a stratified knowledge base can not return a satisfactory result.

Table 4. Merging knowledge bases with an operator in DA^2

ω	K_1	K_2	K_3	K_4	K_5	K'_6	K'_7	K'_8	K'_9	sum
00	1	1	1	1	1	1	1	0	0	7
01	1	1	1	1	1	0	0	1	1	7
10	0	0	0	0	0	2	2	2	2	8
11	0	0	0	0	0	2	2	2	2	8

3.2 Stratifying interpretations by relative preference relation

In the methods discussed above, an interpretation is associated with a number (or a vector of numbers) about its priority level(s) that determines which stratum (or strata) it is in and this number (or a vector of numbers) is the absolute position(s) (stratum/strata) it reflects. We argue that, the absolute position of an interpretation in a stratification is not so important, since one could not tell how the preferences among items of beliefs in other knowledge bases are determined. For instance, when one knowledge base regards that $Mod(p)$ are more preferable than $Mod(q)$ and $Mod(q)$ are more preferable than $Mod(r)$, then $Mod(r) \setminus (Mod(p) \cup Mod(q))$ are the third level of models to be preferred. If other knowledge bases do not consider q, then these models (which are for r) are underestimated if a merging process considers only the absolute position that a model occurs in each stratification of interpretations.

We believe that only the relative preferences between interpretations induced from a knowledge base by ordering strategy is meaningful in a merging process:

Definition 1 (Relative Preference Relation). *Let $\{\Omega_{K_1,X_1}, \dots, \Omega_{K_n,X_n}\}$ be a multi-set. We define a binary relative preference relation $R : \Omega \times \Omega$ as:*

$\mathsf{R}(\omega, \omega')$ *iff* $|\{\Omega_{K_i,X_i} s.t. \omega \prec_i \omega'\}| < |\{\Omega_{K_i,X_i} s.t. \omega' \prec_i \omega\}|$ *where $\prec_i$ is the strict partial order induced from Ω_{K_i,X_i}.*

$\mathsf{R}(\omega, \omega')$ means that more knowledge bases prefer ω than ω'. A relative preference relation is partial, anti-symmetric and irreflective, and it is not transitive, so it is not a total pre-order relation.

Definition 2 (Undominated Set). *Let R be a relative preference relation over $\Omega \times \Omega$ and let Q be a subset of Ω. Q is called an undominated set of Ω, if*

$$\forall \omega \in Q, \forall \omega' \in \Omega \setminus Q, \mathsf{R}(\omega', \omega) \text{ does not hold}$$

Q is a minimal undominated set *of Ω if for any undominated set P of Ω, $P \subset Q$ does not hold.*

We denote the set of minimal undominated sets of Ω w.r.t. R as U^R_Ω.

Definition 3. *Let R be a relative preference relation. A stratification of interpretations $\Omega = (\Omega_1, \dots, \Omega_n)$ can be obtained from R such that $\Omega_i = \cup Q$ where $Q \in U^R_{\Omega \setminus \cup_{j=1}^{i-1} \Omega_j}$ based on Definition 2.*

This way, the stratification of interpretations is independent of absolute priorities (or positions) of interpretations and the commensurability assumptions is not required.

Since from a stratification of interpretations, a total pre-order relation can be induced, the above definition also defines a total pre-order relation over interpretations.

Example 4. *Let $K_1 = (\{p\}, \{q\}, \{r\})$ and $K_2 = (\{r\}, \{q\}, \{p\})$ be two knowledge bases. If we apply ordering strategy leximin on both bases, we get two stratifications on Ω as*

$\Omega_{K_1,leximin} = (\{111\}, \{110\}, \{101\}, \{100\}, \{011\}, \{010\}, \{001\}, \{000\})$

$\Omega_{K_2,leximin} = (\{111\}, \{011\}, \{101\}, \{001\}, \{110\}, \{010\}, \{100\}, \{000\})$

Then a relative preference relation over Ω can be defined based on them. From this relative preference relation, we can get a final stratification on Ω as

$\Omega = (\{111\}, \{110, 011, 101\}, \{001, 010, 100\}, \{000\})$.

Obviously, p, q, r are symmetric and thus are equally preferred and this stratification is better than that obtained in Example 2.

4 Syntax-Based Approaches to Constructing Stratified Knowledge Bases

Based on the stratification of interpretations obtained in the above section, we explore approaches to stratifying a merged knowledge base. We discuss syntax-based methods in this section and investigate model-based methods in the next section.

In the syntax-based methods, we assume that we *pick* some (may not be all) propositions from the original knowledge bases and stratify them based on a stratification of interpretations.

Definition 4. *Let $\Omega = (\Omega_1, ..., \Omega_n)$ be a stratification of interpretations and S be a set of propositions. Let X be an ordering strategy. A stratified knowledge base $K_S^{X,\Omega} = (S_1, ..., S_m)$ is an X* **dominated construction** *from S w.r.t. Ω if $\bigcup_{i=1}^m S_i \subseteq S$ and $\Omega_{K_S^{X,\Omega},X} = \Omega$.*

Definition 5 (best out construction). *Let $\Omega = (\Omega_1, \ldots, \Omega_n)$ be a stratification of interpretations and S be a set of propositions. We define $K_S^{bo,\Omega} = (S_1, \ldots, S_{n-1})$ where $S_i = \{\phi \in S \mid \forall \omega \in \Omega_j, \omega \models \phi, \ \forall\, j \in [1, n-i]\} \setminus \bigcup_{j=1}^{i-1} S_j$ and $S_i \neq \emptyset$.*

Proposition 1. *Let Ω be a stratification of interpretations and S be a set of propositions. If there exists a stratified knowledge base K s.t. $\Omega_{K,bo} = \Omega$ and $\bigcup K \subseteq S$, then $K_S^{bo,\Omega}$ defined in Definition 5 is a best out dominated construction from S w.r.t. Ω, that is $\bigcup_{i=1}^{n-1} S_i \subseteq S$ and $\Omega_{K_S^{bo,\Omega},bo} = \Omega$.*

Example 5. *Let $\Omega = (\{011\}, \{111\}, \{101\}, \{000, 010, 100, 110, 001\})$ and let the set of propositions be $S = \{p \vee q, r, q \vee \neg r, \neg p \vee \neg r, \neg p \wedge \neg q\}$, then we can get a stratified knowledge base based on S as $K_S^{bo,\Omega} = (\{p \vee q, r\}, \{q \vee \neg r\}, \{\neg p \vee \neg r\})$ which satisfies $\Omega_{K_S^{bo,\Omega},bo} = \Omega$. This implies that there is a stratified knowledge base K such that $\bigcup K \subseteq S$ and $\Omega_{K,bo} = \Omega$.*

However, if we have $S' = \{p \vee q, r, \neg q \vee \neg r, \neg p \vee \neg r, \neg p \wedge \neg q, \neg p \vee q\}$, then we have $K_{S'}^{bo,\Omega} = (\{p \vee q, r\}, \{\neg p \vee q\}, \{\neg p \vee \neg r\})$. In this case, $\Omega_{K_{S'}^{bo,\Omega},bo} = (\{011\}, \{111\}, \{001, 101\}, \{000, 010, 100, 110\}) \neq \Omega$, which means that $\nexists K$ such that $\Omega_{K,bo} = \Omega$ and $\bigcup K \subseteq S'$.

Definition 6 (maxsat construction). *Let $\Omega = (\Omega_1, \ldots, \Omega_n)$ be a stratification of interpretations and S be a set of propositions. We define $K_S^{maxsat,\Omega} = (S_1, \ldots, S_n)$ where $S_i = \{\phi \in S \mid \forall \omega \in \Omega_i, \omega \models \phi\} \setminus \bigcup_{j=1}^{i-1} S_j$ and $S_i \neq \emptyset$.*

Proposition 2. *Let Ω be a stratification of interpretations and S be a set of propositions. If there exists a stratified knowledge base K s.t. $\Omega_{K,maxsat} = \Omega$ and $\bigcup K \subseteq S$, then $K_S^{maxsat,\Omega}$ is a maxsat-dominated construction from S w.r.t. Ω, that is $\bigcup_{i=1}^{n-1} S_i \subseteq S$ and $\Omega_{K_S^{maxsat,\Omega},maxsat} = \Omega$.*

Definition 7 (leximin construction). *Let $\Omega = (\Omega_1, \ldots, \Omega_n)$ be a stratification of interpretations and S be a set of propositions. We define $K_S^{leximin,\Omega} = (S_1, \ldots, S_n)$ where $S_i = \{\phi \in S \mid \forall \omega \in \Omega_i, \omega \models \phi$, where $\forall j > i, \forall \omega \in \Omega_j, \omega \not\models \phi\}$ and $S_i \neq \emptyset$.*

Proposition 3. *Let Ω be a stratification of interpretations and S be a set of propositions. If there exists a stratified knowledge base $K = (S_1, \ldots, S_n)$ s.t. each S_i is a singleton set, $\Omega_{K,leximin} = \Omega$ and $\bigcup K = S$, then $K_S^{leximin,\Omega}$ is a leximin dominated construction from Ω, i.e. $\Omega_{K_S^{leximin,\Omega},leximin} = \Omega$.*

In this proposition, it is required that $\bigcup K = S$, because the leximin ordering strategy is more syntax sensitive than best out and maxsat.

Example 6. Let $\Omega = (\{011\}, \{101, 111\}, \{000, 010, 100, 110\}, \{001\})$ and let $S = \{(p \vee q) \wedge r, q \vee \neg r, \neg p \vee \neg r\}$. Then we have a stratified knowledge base $K_S^{leximin,\Omega} = (\{(p \vee q) \wedge r\}, \{q \vee \neg r\}, \{\neg p \vee \neg r\})$ which satisfies $\Omega_{K_S^{leximin,\Omega},leximin} = \Omega$. This implies that $\exists K$ such that $\Omega_{K,leximin} = \Omega$ and $\bigcup K = S$.

For the above three methods, it is assumed that we know what propositions should appear in the merged stratified knowledge base. This assumption comes from the intuition that when merging stratified knowledge bases, only those propositions that appear in some knowledge bases would be considered. This is consistent with the ideas in syntax-based merging operators. However when merging knowledge bases, some implicit knowledge can be drawn and such knowledge does not necessarily appear in any of the individual knowledge bases.

Example 7. Let $K_1 = (\{p \wedge q\})$ and $K_2 = (\{\neg p \wedge q\})$. From this, we can get two stratifications on Ω using ordering strategy leximin as

$$\Omega_{K_1,leximin} = (\{11\}, \{00, 01, 10\})$$
$$\Omega_{K_2,leximin} = (\{01\}, \{00, 10, 11\})$$

Based on these two stratifications, it is possible to define a relative preference relation R, and then Ω can be stratified as

$$\Omega_K = (\{01, 11\}, \{00, 10\})$$

Through this stratification, we can infer that q should be true and p be unknown (or undefined) in the merged base, if we take the models in the first stratum as the models of merging. But q as a proposition does not appear in K_1 or K_2. If we attempt to reconstruct a stratified merged knowledge base from the set $S = K_1 \cup K_2 = \{p \wedge q, \neg p \wedge q\}$ directly with either best out, or maxsat, or leximin, we can only get $K_S^{\Omega_K, X} = (\emptyset)$.

So, if we restrict S to be as $S \subseteq \bigcup_i (\cup K_i)$ then implicit knowledge will be lost. One way to overcome this is to allow S to be a bigger set, such as (a trivial one) S could be $\bigcup Cn(\bigcup K)$, where Cn is the classical deductive closure operator.

5 Model Based Approaches to Constructing Stratified Knowledge Bases

An alternative to the syntax-based family of methods is to construct propositions for the merged knowledge base directly from the stratification of interpretations $\Omega = (\Omega_1, \ldots, \Omega_n)$, rather than *picking* propositions from the original knowledge bases. In this section, we investigate how such an approach can be established.

Definition 8. *Let $\Omega = (\Omega_1, \ldots, \Omega_n)$ be stratification of interpretations. Define $K^{bo,\Omega} = (S_1, \ldots, S_{n-1})$, where $S_i = \phi_{(\bigcup_{j=1}^{n-i} \Omega_j)}, i = 1, \ldots, n-1$.*

Proposition 4. *Let Ω be a stratification of interpretations. Then $\Omega_{K^{bo,\Omega},bo} = \Omega$.*

Example 8. *Let $\Omega = (\{11\}, \{10\}, \{00, 01\})$. Then $K^{bo,\Omega} = (\{p\}, \{p \wedge q\})$, and $\Omega_{K^{bo,\Omega},bo} = \Omega$.*

Definition 9. *Let $\Omega = (\Omega_1, \ldots, \Omega_n)$ be a stratification of interpretations. Define $K^{maxsat,\Omega} = (S_1, \ldots, S_{n-1})$, where $S_i = \phi_{(\bigcup_{j=1}^{i} \Omega_j)}$, $i = 1, \ldots, n-1$.*

Proposition 5. *Let Ω be a stratification of interpretations. Then $\Omega_{K^{maxsat,\Omega},maxsat} = \Omega$.*

Example 9. *Let $\Omega = (\{11\}, \{10\}, \{00, 01\})$. Then $K^{maxsat,\Omega} = (\{p \wedge q\}, \{p\})$, and $\Omega_{K^{maxsat,\Omega},maxsat} = \Omega$.*

Definition 10. *Let $\Omega = (\Omega_1, \ldots, \Omega_n)$ be a stratification of interpretations. Define $\Omega' = (\Omega'_0, \ldots, \Omega'_{2^l-1})$ as:*

1. *$\Omega'_i = \emptyset$, $i \in [0, 2^l - n - 1]$*
2. *$\Omega'_{2^l-n+i-1} = \Omega_i$, $i \in [1, n]$*

where l is the smallest number s.t. $2^l \geq n$.

Let $S_i = \phi_{\bigcup_{\pi(j,i)=0} \Omega'_j} (1 \leq i \leq l)$, where $\pi(j, i) = 0$ if $(j \mod 2^{l-i+1}) < 2^{l-i}$, otherwise $\pi(j, i) = 1$.

Then we define $K^{Leximin,\Omega} = (S_1, \ldots, S_l)$.

In this definition, $\pi(j,i)$ is in fact the value of i^{th} (from the left hand) digit of j when j is represented as a binary value with l-bits. For example, if we set $l = 3$ and we have $j = 3$, then j can be represented as a binary value 011, so $\pi(j,2) = 1$, since the second digit of 011 is 1. We also have $(3 \mod 2^{3-2+1}) = 3$ and $3 \geq 2^{3-2}$, so $\pi(3,2) = 1$ too.

Proposition 6. *Let Ω be a stratification of interpretations. Then $\Omega_{K^{leximin,\Omega},leximin} = \Omega$.*

Example 10. *Let $\Omega = (\{11\}, \{10\}, \{00, 01\})$. Then $K^{leximin,\Omega} = (\{p \wedge q\}, \{p \wedge \neg q\})$, and $\Omega_{K^{leximin,\Omega},leximin} = \Omega$.*

When the interpretations are stratified into relatively a large number of strata, the $leximin$ dominated construction method can drastically reduce the number of strata of the merged knowledge base compared to both best out and maxsat.

Example 11. *Let $\Omega = (\{111\}, \{110\}, \{101\}, \{100\}, \{011\}, \{010\}, \{001\}, \{000\})$. Then*

$$K^{leximin,\Omega} = (\{p\}, \{q\}, \{r\})$$

However, the other two strategies both return a knowledge base with a lot more propositions. That is

$$K^{bo,\Omega} = (\{p \vee q \vee r\}, \{p \vee q\}, \{(p \vee q) \wedge (p \vee r)\}, \{p\}, \{p \wedge (q \vee r)\}, \{p \wedge q\}, \{p \wedge q \wedge r\})$$

and

$K^{maxsat,\Omega} = (\{p \wedge q \wedge r\}, \{p \wedge q\}, \{p \wedge (q \vee r)\}, \{p\}, \{(p \vee q) \wedge (p \vee r)\}, \{p \vee q\}, \{p \vee q \vee r\})$

6 Conclusion

In knowledge base merging, most existing methods merge either flat or stratified knowledge bases and produce a flat base (or a set of models) as the result (e.g., [6–8, 10]). We argue that ideally a stratified merged base would be better since it has additional information about which formulae are more preferred than others. Motivated by this, we investigated how such a stratified merged base can be constructed.

We first looked at the possibility of recovering a stratified base based on the stratification of interpretations obtained after applying a merging operator. However, the results show that such a straightforward method can produce counterintuitive results. The main reason is that almost all the merging methods, especially the model-based ones (which return a set of models as the merged results), require an assumption of commensurability, so the absolute position of each interpretation in each stratification of interpretations is important. We argue that only the relative position of an interpretation w.r.t other interpretations is important if we do not require this commensurability. Based on this, we proposed a method to define a binary relative preference relation between interpretations and then this relation is used to stratify interpretations given a set of individual stratifications of interpretations induced from the original knowledge bases.

Following this, we proposed a family of syntax-based and model-based approaches to stratifying a merged knowledge base. Properties of these stratification approaches are also studied.

The idea of constructing a relative preference relation is inspired by Condorcet methods in *voting systems* or the *social choice theory*. The winner of votes by the Schulze method, an instance of Condorcet methods, is exactly the same as the most preferred interpretations in our approach to generate the stratification of interpretations using the relative preference relation, when we treat a candidate as an interpretation and a boll as a stratification of interpretations.

For future work, we will further investigate appropriate approaches to stratifying merged knowledge bases and to discuss additional logical properties of our methods. In contrast with merging flat knowledge bases, we believe that merging stratified knowledge bases should put more emphasize on considering preferences of propositions and thus should satisfy a different set of constraints or postulates to those for merging flat bases.

Acknowledgement: This work is funded by the EPSRC projects with reference numbers: EP/D070864/1 and EP/D074282/1

References

1. Baral, C., Kraus, S., Minker, J. Combining multiple knowledge bases. *IEEE Transactions on Knowledge Data Engineering* 3:208-220, 1991
2. Benferhat, S., Cayrol, C., Dubois, D., Lang, J., Prade, H. Inconsistency management and prioritized syntax-based entailment. *Proc. Of IJCAI'93*:640-647, 1993
3. Brewka, G. A rank based description language for qualitative preferences. *The 16th European Conference on Artificial Intelligence (ECAI'04)*:303-307, 2004
4. Delgrande, J.P., Schaub, T. Consistency-based approaches to merging knowledge bases: preliminary report. *The 10th International Workshop on Non-Monotonic Reasoning (NMR2004)*:126-133, 2004
5. Dubois, D., Lang, J., Prade, H. Possibilistic logic. *Handbook of Logic in Artificial Intelligence and Logic Programming*, Nomonotonic Reasoning and Uncertain Reasoning, Vol 3: 439-513. In Gabbay, D., Hogger, C.J., Robinson, J.A., eds. Oxford University Press, 1994.
6. Everaere, P., Konieczny, S., Marquis, P. On merging strategy-proofness. *The 9th International Conference on Principles of Knowledge Representation and Reasoning (KR04)*:357-368, 2004
7. Konieczny, S., Pérez, R.P. On the logic of merging. *Proc. Of KR'98*:488-498, 1998
8. Konieczny, S., Lang, J., Marquis, P. DA^2 merging operators. *Artificial Intelligence* 157(1-2):49-79, 2004
9. Meyer, T. On the semantics of combination operations. *Journal of Applied Non-Classical Logics* 11 (1-2):59-84, 2001
10. Qi, G., Liu, W., Bell, D.A. Merging stratified knowledge bases under constraints. *The 21st American National Conference on Artificial Intelligence (AAAI06)*:281-286, 2006
11. Spohn, W. Ordinal conditional functions. a dynamic theory of epistemic states. *Causation in Decision, Belief Change, and Statistics* 2: 105-134, Harper, W., Skyrms, B., eds., Kluwer, Dordrecht, 1988