Knowledge Base Merging by Majority

Jinxin Lin Alberto O. Mendelzon

Dept. of Computer Science
University of Toronto
Toronto, M5S 1A4, Canada
{lin,mendel}@db.toronto.edu

Abstract

A common practice for resolving conflicts among a group of agents is to let the majority decide. We formalize this principle and apply it to the problem of merging the knowledge of multiple agents. We then postulate logical properties that all knowledge merging operators should satisfy and give a model-theoretic characterization of all merging operators that satisfy the postulates. It turns out that the operators that satisfy the postulates are precisely those that induce a certain kind of partial pre-order over the set of possible worlds such that the models of the merged knowledge base are the possible worlds that are *minimal* with respect to the pre-order. Then, we review several previous approaches and study them in light of the proposed postulates. We then present a particular method for knowledge merging, *CMerge*, that satisfies all the postulates. We show by example that *CMerge* appears to resolve conflicts among knowledge bases in a plausible way. Finally, we show that *CMerge* can be implemented by an efficient syntactical transformation from the set of knowledge bases to be merged.

1 Introduction

In many fields such as heterogeneous databases [BAP92] or genomic information systems [Fre91] we are often confronted with multiple and conflicting sources of information. Systems organized around reasoning agents [HM85, FH88] face the similar problem of resolving conflicts among contradictory knowledge or beliefs[1] held by different agents. At the same time, these systems want to extract additional knowledge that is not locally held by any agent, but collectively by all of them (called *implicit knowledge* in [HM85]). For example, if an agent knows a and another agent knows $a \rightarrow b$, then combining their knowledge yields b, even though neither one of them individually knows b.

Assume each agent is associated with a knowledge base and the agent reasons about the world according to the knowledge base. Then the problem of how to merge the knowledge of multiple agents is the same as how to merge multiple knowledge bases. One simple method

[1] We do not distinguish between *knowledge* and *belief* in this paper.

for knowledge base merging is to take the disjunction of the maximal consistent subsets of the union of the knowledge bases. This is essentially the approach adopted by Baral *et al.* [BKMS92], who also consider merging knowledge bases in the presence of integrity constraints and priorities over the knowledge bases. However, such an approach does not take into account the "popularity" of a belief among a group of agents. There are many situations (but of course not all) where the number of agents who hold a particular belief is important, and sometimes it may be the only practical way of resolving a conflict. For example, suppose there are three agents with knowledge bases $T_1 = \{a\}, T_2 = \{a\}$ and $T_3 = \{\neg a\}$. The approach of Baral *et al.* yields $a \vee \neg a$, a tautology, which does not support either a or $\neg a$, even though two out of the three agents support a. In contrast, our approach in this paper will produce a, which reflects the view of the majority. Furthermore, our approach is independent of the particular syntax of the knowledge bases to be merged, while the method of Baral *et al.* is syntax-dependent. The method of Revesz [Rev94] also does not take majority into account. For the above example, it returns the same result as the one of Baral *et al.*

Borgida and Imielinski [BI84] proposed several approaches to resolve a conflict in beliefs held by a collection of experts (which they called a "committee"). However, using their protocols, committees may remain indecisive or contradictory on the conflicting issues. One protocol they defined is that a committee supports a sentence α if all members support (imply) α. This protocol is equivalent to taking the disjunction of the knowledge bases. This does not take majority into account, and the committee remains indecisive on conflicting issues. In the example above this method produces the same result as the one of Baral *et al.*, $a \vee \neg a$. Among other protocols they defined, the "democracy" protocol seems closest to the principle of majority. Under the democracy protocol, a committee C is defined to support α if: (1) some subcommittee supports α, and (2) if some other subcommittee S supports $\neg\alpha$ then there is a subcommittee S' such that S' supports α and cardinality$(S') >$ cardinality(S), where a subcommittee is a subset of C and is defined to support a sentence if every member in the subcommittee supports the sentence. Then it is not difficult to verify that the committee consisting of $T_1 = \{a\}, T_2 = \{b\}$ and $T_3 = \{\neg a \vee \neg b\}$ supports a and b and $\neg a \vee \neg b$, which is a contradiction! In the method we describe here, the result of merging is always consistent; in fact this property is one of the basic conditions we postulate for all knowledge merge operators.

Halpern and Moses [HM85] defined the notion of *implicit knowledge* in a group of agents. But their formulation of implicit knowledge applies only to agents who do not have conflicting views; if there is a conflict, the implicit knowledge of the agents is every sentence in the language.

Majority rule is used in many other fields, notably social choice theory (see, e.g., [Kel88]). However, the majority postulate we formalize in this paper is specifically designed for knowledge merging and is different from the majority rule in the other fields. This is due to the fact that our postulate includes a notion of 'partial support' defined in the paper. We show that the simple majority rule commonly used in other fields, i.e., given two alternatives, select the one that has more supporters, will lead to contradiction when applied to knowledge merging, as in the case of the above "democracy" protocol. On the other hand, various voting schemes studied in social choice theory [Kel88] cannot accomplish the purpose of knowledge merging. For example, we would like to achieve new knowledge b by merging two knowledge bases $\{a\}$ and $\{a \to b\}$. But since neither knowledge base supports b, b cannot be obtained through "voting" by the knowledge bases.

The rest of this paper is organized as follows. In Section 2 we define the language and basic notation used in the paper. In Section 3, we formally propose a set of four postulates for knowledge merging by majority. We show that our formalization of majority postulate is reasonable such that a stronger version will lead to contradiction, and a weaker version will not capture the requirement of knowledge base merging. We then present in Section 4 a model-theoretic characterization of the postulates, which demonstrates how the set of postulates captures the 'minimal criterion'. In Section 5 we review several previous approaches and show how those approaches fail to satisfy some of the postulates. Then in Section 6, we propose a majority merging operator *CMerge* and give a syntactic characterization of it. We conclude the paper in Section 7 by pointing out some topics for future research. The proofs of some theorems and lemmas are given in the Appendix.

2 Preliminaries

In this paper, we consider the language L of propositional logic formed from a finite set of atoms $\mathcal{P}$ in the usual way. And we use the standard terminology of propositional logic except for the definitions given below.

A *knowledge base* is a finite set of L sentences. We sometimes refer to a knowledge base as an *agent* and use the two terms interchangeably. A *possible world* is a function from $\mathcal{P}$ to $\{t, f\}$. The set of all possible worlds is denoted $\mathcal{W}$. Logical equivalence is denoted by $\equiv$. A *literal* is an atom or negation of an atom. For convenience, if a literal l denote one of p and $\neg p$ (where p is an atom) then $-l$ denote the other of p and $\neg p$. And $\mathcal{P}(l)$ denotes the atom appearing in l. Given a possible world w, we denote w_l^- to be the possible world that agrees with w on every atom except the atom $\mathcal{P}(l)$, in which case w_l^- maps $\mathcal{P}(l)$ to t if w maps it to f, and maps $\mathcal{P}(l)$ to f if w maps it to t.

A *model* of a propositional formula ψ is a possible world where ψ is true in the usual sense. The set of all the models of ψ is denoted by $[\![\psi]\!]$.

A *pre-order* over the set of possible worlds $\mathcal{W}$ is a reflexive and transitive relation on $\mathcal{W}$. For a pre-order $\leq$, we define $<$ as: $w < w'$ iff $w \leq w'$ and $w' \not\leq w$. Let W be a subset of $\mathcal{W}$. We say a possible world w is *minimal* in W with respect to $\leq$ if $w \in W$ and there is no $w' \in W$ such that $w' < w$. Let $Min(W, \leq)$ denote the set of possible worlds that are minimal in W with respect to $\leq$.

3 Postulates for Majority Merging

In this section, we propose a set of four *postulates* that knowledge merging operators are expected to satisfy. The first three postulates apply to any merge operator; the fourth one is specific to operators that obey majority rule.

Suppose the knowledge bases under consideration are $T_1, \ldots, T_n$. A merging function *Merge* is a mapping from the set of knowledge bases to a new knowledge base $Merge(\{T_1, \ldots, T_n\})$. Our basic postulates for knowledge merging are as follows:

MM1 $Merge(\{T_1, \ldots, T_n\})$ is consistent.

MM2 If $T_1 \wedge \cdots \wedge T_n$ is consistent then $Merge(\{T_1, \ldots, T_n\}) \equiv T_1 \wedge \cdots \wedge T_n$.

MM3 If $T_1 \equiv T_1', \ldots, T_n \equiv T_n'$, then $Merge(\{T_1, \ldots, T_n\}) \equiv Merge(\{T_1', \ldots, T_n'\})$.

Postulate MM1 assures that the result of merging is consistent since we do not want the conflicts among the knowledge bases to remain unresolved. Postulate MM2 takes care of one limiting case, saying that if there is no conflict among the knowledge bases then the result of merging is simply the union of the knowledge bases. This provides "cooperation" among the knowledge bases, meaning that the group obtains the maximum amount of knowledge from each knowledge base and new knowledge is created by this combination. Postulate MM3 requires the merging function to be independent of the syntax of each knowledge base, as we are considering merging the *knowledge* from the knowledge bases and two knowledge bases have the same knowledge if they are logically equivalent. Note also that the order of the knowledge bases is irrelevant since we take the *set* of knowledge bases as the argument of *Merge*.

We now formalize, in terms of a postulate, the "majority principle" in the context of knowledge base merging. Let us discuss possible ways of doing this.

We say a knowledge base T_i *supports* a sentence α if T_i implies α ($T_i \models \alpha$), and *opposes* α if T_i implies $\neg\alpha$ ($T_i \models \neg\alpha$). One possible formalization of majority rule is to have the merged knowledge base support a sentence α if there are over half (or some fixed percentage) of members in the group of knowledge bases to be merged supporting α. But this postulate is too weak. Consider the following operator that satisfies the postulate: the merged knowledge base is defined to support a sentence α if over half of members support α. Then consider three knowledge bases $T_1 = \{a\}$, $T_2 = \{b, c\}$, and $T_3 = \{\neg c\}$. Intuitively the conflict among the knowledge bases centers on c and we want a and b to be in the merge result, but fewer than half of the knowledge bases support a or b.

A more appealing formalization of majority rule is to require the merged knowledge base to support α whenever there are more members supporting α than $\neg\alpha$. But this can produce contradictory results. For example, suppose $T_1 = \{a\}$, $T_2 = \{\neg a \vee \neg b\}$ and $T_3 = \{b, c\}$. Then the merged knowledge base supports a, b and $\neg a \vee \neg b$.

The problem is that the principle we just proposed applies not only to atomic sentences such as a or b but also to compound ones such as $\neg a \vee \neg b$. In order to avoid inconsistent results, we have to restrict the principle to atomic sentences. We can justify this with a metaphor: in a decision-making committee, members vote on single issues instead of disjunctions or more complex combinations of issues.

Suppose we now restrict our proposal as follows. If there are more members supporting a literal l than its negation $-l$, then the result should support l. Then for the example $T_1 = \{a\}$, $T_2 = \{\neg a \vee \neg b\}$ and $T_3 = \{b, c\}$, the merged knowledge base supports a and b. But there is still a problem with this proposal: it ignores knowledge encoded in compound sentences such as T_2. In fact, merging T_1, T_2 and T_3 has the same result as merging T_1 and T_3, which seems undesirable. On closer examination of the agent T_2, we note that he has some "partial" support for $\neg a$ and for $\neg b$, i.e., he knows that one of $\neg a$ and $\neg b$ must be true but does not know which one. In fact this knowledge $(\neg a \vee \neg b)$ denies the sentence $a \wedge b$. It seems therefore that we should not conclude $a \wedge b$, ignoring T_2's partial support for $\neg a$ and $\neg b$.

To state our final proposal for formalizing the principle of majority, we first define what we mean by a knowledge base "partially supporting" a literal. We say a knowledge base T_i *partially supports* literal l (denoted $T_i \rhd l$) if there is a β, which mentions no atom appearing in l, such that $T_i \models l \vee \beta$ but $T_i \not\models l$ and $T_i \not\models \beta$. That is, the agent believes either l or β is true without knowing which one is. We restrict this β to be a proposition that mentions no atom appearing in l, for the purpose of excluding the cases that $\beta = -l \vee \ldots$. Otherwise, knowledge of the tautologies would be sufficient to have partial knowledge of any literal. This notion of partial support has a simple model-theoretic description.

Proposition 3.1 *A knowledge base ψ partially supports a literal l if and only if $[\![\psi \wedge -l]\!] \neq \emptyset$ and there exists $w \in [\![\psi \wedge l]\!]$ such that $w_l^- \notin [\![\psi]\!]$.*

Proof: 'If' direction: Let ψ be a knowledge base and l be a literal such that $[\![\psi \wedge -l]\!] \neq \emptyset$ and there exists $w \in [\![\psi \wedge l]\!]$ such that $w_l^- \notin [\![\psi]\!]$. Since $w_l^- \notin [\![\psi]\!]$, for each $u \in [\![\psi]\!] - \{w\}$ we can find an atom $p_u \in \mathcal{P}$ such that $p_u \neq \mathcal{P}(l)$ and $u(p_u) \neq w(p_u)$. Define β to be:

$$\beta = \bigvee_{u \in ([\![\psi]\!] - \{w\})} \begin{cases} p_u & \text{if } u(p_u) = t \\ \neg p_u & \text{if } u(p_u) = f. \end{cases}$$

We have for every $u \in [\![\psi]\!] - \{w\}$, $u \models \beta$. Then since $w \models l$, $\psi \models l \vee \beta$. It is clear from the construction of β that β mentions no atom appearing in l and $w \not\models \beta$. Hence $\psi \not\models \beta$. Since $[\![\psi \wedge -l]\!] \neq \emptyset$, we have $\psi \not\models l$. This means that ψ partially supports l.

'Only-If' direction: Let ψ partially support l. Then $[\![\psi \wedge -l]\!] \neq \emptyset$. Otherwise $\psi \models l$ and it follows that $\psi \models l \vee \beta$, for any β, which means that ψ does not partially support l.

From the definition of 'partial-support', we know that there exists β, which mentions no atom appearing in l, such that $\psi \models l \vee \beta$ but $\psi \not\models l$ and $\psi \not\models \beta$. From $\psi \not\models \beta$ we also know that there exists $w \in [\![\psi]\!]$ such that $w \not\models \beta$. Then we have $w_l^- \not\models \beta$ as well since β mentions no atom appearing in l and w and w_l^- assigns the same valuation to every atom except the atom appearing in l. Since $\psi \models l \vee \beta$, we have $w \models l \vee \beta$. But $w \not\models \beta$, so we have $w \models l$. Thus, $w \in [\![\psi \wedge l]\!]$. Since $w \models l$ it is obvious that $w_l^- \not\models l$. It follows from $w_l^- \not\models \beta$ that $w_l^- \not\models l \vee \beta$. Then clearly, $w_l^- \notin [\![\psi]\!]$ since $\psi \models l \vee \beta$, from which the 'Only-If' direction follows. □

The following property about partial support is instructional.

Proposition 3.2 *For a knowledge base ψ and a literal l, if $\psi \models l$ or $\psi \models -l$ or ψ is inconsistent then ψ does not partially supports l.*

Proof: Immediate from the definition of partial support. □

We return to the example $T_1 = \{a\}$, $T_2 = \{\neg a \vee \neg b\}$ and $T_3 = \{b, c\}$. As we mentioned, there seems no reason to conclude a and b since T_2 opposes $a \wedge b$; and obviously no reason as well to conclude $\neg a \vee \neg b$. But there is no question that c should be supported. From these observations, we formalize the majority principle as: a group of knowledge bases supports a literal sentence l if there are more members supporting l than members supporting $-l$ or partially supporting $-l$. We do not count members who partially support l, since they may turn out not to support l, but we count the members who partially support $-l$ since they may turn out to support $-l$. That is, support for l should be stronger than opposition to l, including potential opposition.

Formally, let $|T_i \models l|$ denote the number of T_i's that support l, and $|T_i \rhd l|$ denote the number of T_i's that partially support l, where $i \in [1, n]$. Then the postulate is:

MM4 For a literal sentence l, if

$$|T_i \models l| > |T_i \models -l| + |T_i \rhd -l| \tag{1}$$

then $Merge(\{T_1, \ldots, T_n\})$ implies l.

We note that if a T_i is inconsistent then $T_i \models l$, $T_i \models -l$ and T_i does not partially support $-l$ according to Proposition 3.1. Hence any inconsistent T_i makes no difference to the inequality $|T_i \models l| > |T_i \models -l| + |T_i \rhd -l|$. We can just count those T_i's that are consistent, and rephrase MM4 as:

For a literal l, if there are more consistent T_i's supporting l than supporting or partially supporting $-l$ then $Merge(\{T_1, \ldots, T_n\})$ implies l.

Normally, for a *Merge* that abides by MM4, the condition of $|T_i \models l| > |T_i \models -l|$ does not guarantee the result of *Merge* to support l. But in the following case, it suffices:

Proposition 3.3 *Let Merge be a merging function that abides by MM4. If every member of* $\{T_1, \ldots, T_n\}$ *supports either* l *or* $-l$ *and* $|T_i \models l| > |T_i \models -l|$ *then* $Merge(\{T_1, \ldots, T_n\}) \models l$.

This is due to the fact that there is no member partially supporting $-l$ in this case.

The following is a simple consequence of the definition of MM4, which relates MM4 and the "strict majority rule"; we omit its proof.

Proposition 3.4 *Let Merge be a merging function that abides by MM4. If there are over half of the members supporting* l *but not* $-l$ *then* $Merge(\{T_1, \ldots, T_n\}) \models l$.

In the case that there is no conflict among the knowledge bases, MM2 alone subsumes all the other postulates. This is shown by the following proposition whose proof is trivial and is omitted.

Proposition 3.5 *If* $T_1 \wedge \cdots \wedge T_n$ *is consistent and* $Merge(\{T_1, \ldots, T_n\})$ *satisfies MM2 then* $Merge(\{T_1, \ldots, T_n\})$ *satisfies MM1-4.*

So the postulates MM1 and MM3-4 are particularly useful when there are conflicts among the knowledge bases.

The set of postulates MM1-4 is consistent, since the merge function *CMerge* (presented in a later section) satisfies them for any given set of knowledge bases.

4 A Model-Theoretic Characterization of Merge Operators

The class of all knowledge base merge operators that satisfy MM1-4 can be characterized in model-theoretic terms. Briefly describing, for each merge operator that complies with the postulates and for each set of knowledge bases $\{T_1, \ldots, T_n\}$ there is a metric that orders the possible worlds in $\mathcal{W}$ in terms of their closeness to $\{T_1, \ldots, T_n\}$, so that the result of the merge operation is defined by the set of worlds that are "closest" to the knowledge bases according to this metric. Conversely, each well-behaved metric can be used to define a merge operator that complies with MM1-4.

We say a function that assigns to every set of knowledge bases $\{T_1, \ldots, T_n\}$ a pre-order $\leq_{\{T_1,\ldots,T_n\}}$ over the set of possible worlds $\mathcal{W}$ is a *majority* assignment if the following conditions hold:

1. If $[\![T_1]\!] = [\![T_1']\!], \ldots, [\![T_n]\!] = [\![T_n']\!]$ then $\leq_{\{T_1,\ldots,T_n\}} = \leq_{\{T_1',\ldots,T_n'\}}$.
2. Let $W = [\![T_1]\!] \cap \ldots \cap [\![T_n]\!]$. Then for all $w \in W$ and $w' \in \mathcal{W}$, $w \leq_{\{T_1,\ldots,T_n\}} w'$; In particular, $w <_{\{T_1,\ldots,T_n\}} w'$ if $w' \notin W$.
3. Let l be a literal that satisfies the inequality (1), then for all $w \in [\![l]\!]$, $w <_{\{T_1,\ldots,T_n\}} w_l^-$.

The first condition says that the order for $\{T_1, \ldots, T_n\}$ should be independent of the syntax of each T_i $(i = 1, \ldots, n)$. The second condition says that all models of $T_1 \wedge \ldots \wedge T_n$ are equally close to the knowledge bases and are strictly closer to them than any non-model. The third one says that if there are more knowledge bases supporting l than supporting or partially supporting $-l$, then given any two possible worlds that agree on every atom except the atom $\mathcal{P}(l)$, the world that satisfies l is strictly closer to the knowledge bases than the one that does not satisfy l. We now take $Min(\mathcal{W}, \leq_{\{T_1,\ldots,T_n\}})$ to be the set of possible worlds in $\mathcal{W}$ that are closest to $\{T_1, \ldots, T_n\}$.

Theorem 4.1 *A merging function Merge satisfies postulates MM1-4 iff there exists a majority assignment that maps each set of knowledge base $\{T_1, \ldots, T_n\}$ to a partial pre-order $\leq_{\{T_1,\ldots,T_n\}}$ such that*

$$[\![Merge(\{T_1, \ldots, T_n\})]\!] = Min(\mathcal{W}, \leq_{\{T_1,\ldots,T_n\}}).$$

Proof: For convenience in the proof, let $T = \{T_1, \ldots, T_n\}$.

(If-part) Let *Merge* be a merging function and there exists a majority assignment that maps the set of knowledge base T to a partial pre-order $\leq_T$ such that $[\![Merge(T)]\!] = Min(\mathcal{W}, \leq_T)$. We now prove that *Merge* satisfies MM1-4.

MM1: Assume $Merge(T)$ is inconsistent. This means $Min(\mathcal{W}, \leq_T)$ is empty. It follows that for any $w_1 \in \mathcal{W}$ there is $w_2 \in \mathcal{W}$ such that $w_2 <_T w_1$. Similarly, for w_2, there exists $w_3 \in \mathcal{W}$ such that $w_3 <_T w_2$. Hence there is an infinite sequence such that $\cdots <_T w_i <_T w_{i-1} <_T \cdots <_T w_2 <_T w_1$. We now prove that every world in the sequence is distinct. Assume to the contrary that there are two worlds w_i and w_j $(i > j)$ in the sequence such that $w_i = w_j$. Then $w_i <_T \cdots <_T w_{j+1} <_T w_j$ ($i = j+1$ is allowed). Since $\leq_T$ is transitive it follows that $w_i \leq_T w_{j+1}$. Since $w_i = w_j$ we can substitute w_i with w_j in $w_i \leq_T w_{j+1}$ and we have $w_j \leq_T w_{j+1}$, contradicting the fact that $w_{j+1} <_T w_j$. Hence every world in the sequence is distinct.

As the language we are concerned with is finite, the set of all possible worlds $\mathcal{W}$ is finite. However in the infinite sequence $\cdots <_T w_i <_T w_{i-1} <_T \cdots <_T w_2 <_T w_1$ every world is distinct. This means there are infinite number of possible worlds, a contradiction. Hence $Merge(T)$ is consistent. MM1 is satisfied.

MM2: If $T_1 \wedge \ldots \wedge T_n$ is consistent, $[\![T_1 \wedge \ldots \wedge T_n]\!]$ is not empty. We prove $Min(\mathcal{W}, \leq_T) = [\![T_1 \wedge \ldots \wedge T_n]\!]$. Let $w \in Min(\mathcal{W}, \leq_T)$. Assume $w \notin [\![T_1 \wedge \ldots \wedge T_n]\!]$. Since $[\![T_1 \wedge \ldots \wedge T_n]\!]$ is not empty there is at least a possible world $w' \in [\![T_1 \wedge \ldots \wedge T_n]\!]$. We have $w' <_T w$ according to the second condition of majority assignment. This contradicts the fact that $w \in Min(\mathcal{W}, \leq_T)$.

Let $w \in [\![T_1 \wedge \ldots \wedge T_n]\!]$. Assume $w \notin Min(\mathcal{W}, \leq_T)$. Then there is $w' \in \mathcal{W}$ such that $w' <_T w$. However we have $w \leq_T w'$ by the second condition of majority assignment. This is a contradiction. Hence $Min(\mathcal{W}, \leq_T) = [\![T_1 \wedge \ldots \wedge T_n]\!]$. It follows that $[\![Merge(T)]\!] = [\![T_1 \wedge \ldots \wedge T_n]\!]$. MM2 is satisfied.

MM3: is satisfied by the first condition of majority assignment.

MM4: Let l be a literal such that $|T_i \models l| > |T_i \models -l| + |T_i \rhd -l|$. Assume $Merge(T) \not\models l$. Then there is $w \in [\![Merge(T)]\!]$ such that $w \models -l$. It follows that $w \in Min(\mathcal{W}, \leq_T)$ and $w_l^- \models l$. But then by the third condition of majority assignment, $w_l^- <_T w$, contradicting the fact that $w \in Min(\mathcal{W}, \leq_T)$. Hence $Merge(T) \models l$. MM4 is satisfied.

(Only-if part) Assume that $Merge$ satisfies MM1-4. We define a relation $\leq_T$ as $w \leq_T w'$ if and only if one of the following conditions holds:

C1 $w = w'$.

C2 w is a model of $Merge(T)$.

C3 $w \in [\![l]\!]$ and $w' = w_l^-$ for some literal l such that $|T_i \models l| > |T_i \models -l| + |T_i \rhd -l|$.

C4 There is $w^* \in \mathcal{W}$ such that $w \leq_T w^*$ and $w^* \leq_T w'$.

For the order such defined, we have the following lemma that will be needed later.

Lemma 4.1 *For two possible worlds w and w', if $w \models Merge(T)$ and $w' \not\models Merge(T)$ then $w <_T w'$.*

Proof: Since $w \models Merge(T)$, we have $w \leq_T w'$ according to (C2). Then to prove $w <_T w'$ we need to prove $w' \not\leq_T w$. Assume to the contrary that $w' \leq_T w$. Then there must exist some sequences using (C1-4) to define $w' \leq_T w$. Among them, we assume the shortest to be $w^1 \leq_T w^2 \leq_T \cdots \leq_T w^m$ where $w^1, w^2, \ldots, w^m$, $m \geq 2$ are possible worlds, and $w^1 = w'$ and $w^m = w$. Since $[\![\neg Merge(T)]\!]$ and $[\![Merge(T)]\!]$ are disjoint and $w^1 \in [\![\neg Merge(T)]\!]$ but $w^m \in [\![Merge(T)]\!]$, there must be some w^i and w^{i+1} ($i \in [1, m]$) in the sequence such that $w^i \in [\![\neg Merge(T)]\!]$ and $w^{i+1} \in [\![Merge(T)]\!]$ and $w^i \leq_T w^{i+1}$. Since the sequence is the shortest, $w^i \leq_T w^{i+1}$ is defined using (C1-3) only. It is clear that (C1) does not apply, as one is the model of *Merge*(T) and the other is not. Condition (C2) does not apply either since w^i is not a model of *Merge*(T). It follows that $w^i \leq_T w^{i+1}$ is defined using (C3) only. Then from the definition of (C3), we have $w^i \models l$ and $w^{i+1} = (w^i)^-_l$, for some literal l such that $|T_i \models l| > |T_i \models -l| + |T_i \rhd -l|$. As *Merge* satisfies MM4, $Merge(T) \models l$. But $w^{i+1} \not\models l$, which contradicts the fact that $w^{i+1} \in [\![Merge(T)]\!]$. Hence $w <_T w'$. □

To prove the (only-if) part, we need to show two things: (1) $\leq_T$ is a partial pre-order and the assignment mapping the set of knowledge bases to $\leq_T$ is a majority assignment. (2) $[\![Merge(T)]\!] = Min(\mathcal{W}, \leq_T)$.

(1). It is easy to see that $\leq_T$ is a partial pre-order. We now show that the assignment mapping the set of knowledge bases to $\leq_T$ is a majority assignment. The first condition of majority assignment is satisfied trivially. We now show the remaining conditions.

- For the second condition, if there exists $w \in [\![T_1 \wedge \ldots \wedge T_n]\!]$ then $T_1 \wedge \ldots \wedge T_n$ is consistent. And since *Merge* satisfies MM2, it follows that $w \in [\![Merge(T)]\!]$. By (C2) we have $w \leq_T w'$, for all $w' \in \mathcal{W}$. In particular, if $w' \notin [\![T_1 \wedge \ldots \wedge T_n]\!]$, then $w' \notin [\![Merge(T)]\!]$. By Lemma 4.1 we have $w <_T w'$. Hence the second condition of majority assignment is satisfied.

- For the third condition, suppose there exists l such that $|T_i \models l| > |T_i \models -l| + |T_i \rhd -l|$. Let $w \in [\![l]\!]$. We want to prove $w <_T w^-_l$. From MM4 we have $Merge(T) \models l$. Thus according to (C3), $w \leq_T w^-_l$. It remains to be shown that $w^-_l \not\leq_T w$. Assume to the contrary that $w^-_l \leq_T w$. Then there must exist some sequences using (C1-4) to define $w^-_l \leq_T w$. Among them, we assume the shortest to be $w^1 \leq_T w^2 \leq_T \cdots \leq_T w^m$ where $w^1, w^2, \ldots, w^m$ ($m \geq 2$) are possible worlds, and $w^1 = w^-_l$ and $w^m = w$. Since $[\![-l]\!]$ and $[\![l]\!]$ are disjoint and $w^1 \in [\![-l]\!]$ but $w^m \in [\![l]\!]$, there must be some w^i and w^{i+1} ($i \in [1, m]$) in the sequence such that $w^i \in [\![-l]\!]$ and $w^{i+1} \in [\![l]\!]$ and $w^i \leq_T w^{i+1}$. Since this sequence is the shortest, $w^i \leq_T w^{i+1}$ is defined using (C1-3) only. It is clear that (C1) does not apply, as w^{i+1} satisfies l but the other does not. Assume Condition (C2) is satisfied. Then w^i is a model of *Merge*(T), and since w^-_l is not a model of *Merge*(T), we have $w^i <_T w^-_l$ by Lemma 4.1. But by the transitivity of $\leq_T$, we know $w^1 \leq_T w^i$. That is, $w^-_l \leq_T w^i$ since $w^1 = w^-_l$, contradicting $w^i <_T w^-_l$. Hence $w^i \leq_T w^{i+1}$ is defined using (C3) only. From the definition of (C3), we have $w^i \models l'$ and $w^{i+1} = (w^i)^-_{l'}$ for some literal l' such that $|T_i \models l'| > |T_i \models -l'| + |T_i \rhd -l'|$. Then w^i and w^{i+1} assign the same valuations to all atoms except the one appearing in l'. However, $w^i \models -l$ and $w^{i+1} \models l$. Hence we have $l' = -l$. It follows that $|T_i \models -l| > |T_i \models l| + |T_i \rhd l|$. This contradicts our assumption from the beginning that $|T_i \models l| > |T_i \models -l| + |T_i \rhd -l|$ (note that $|T_i \rhd l|$ and $|T_i \rhd -l|$ are non-negative numbers). Hence $w^-_l \leq_T w$ does not hold. Hence $w_l <_T w^-_l$. The third condition of majority assignment is satisfied.

(2). We now prove $[\![Merge(T)]\!] = Min(\mathcal{W}, \leq_T)$. Let $w \in [\![Merge(T)]\!]$. Assume $w \notin Min(\mathcal{W}, \leq_T)$. Then there exists $w' \in \mathcal{W}$ such that $w' <_T w$. However, since w is a model of *Merge*(T) and according to (C2), we have $w \leq_T w'$. This is a contradiction.

To prove the other direction, let $w \in Min(\mathcal{W}, \leq_T)$ and $w \notin [\![Merge(T)]\!]$. Then by MM1, $Merge(T)$ is consistent, and hence there exists $w' \in \mathcal{W}$ such that $w' \in [\![Merge(T)]\!]$. From Lemma 4.1, we have $w' <_T w$. This contradicts the fact that $w \in Min(\mathcal{W}, \leq_T)$.

Therefore we have $[\![Merge(T)]\!] = Min(\mathcal{W}, \leq_T)$. □

Note the similarity of this theorem to the representation theorems of Katsuno and Mendelzon for theory change operators such as revision and update [KM91b, KM91a]. In theory change, we have a knowledge base T and a sentence α that *must* be incorporated into T; the operators select from all models of α those that are closest to T in a certain sense. In knowledge base merge, we are not restricted to choosing models of any given sentence, but rather we try to compromise among multiple theories to come up with a result that combines as much local knowledge as possible into global knowledge, using majority rule to resolve conflicts.

5 Review of Previous Proposals

We have described in the introductory section several previous approaches for resolving conflicts in a group of agents. We review in this section those approaches in light of postulates MM1-4. In the following, the result of merging is usually a set of sentences. For the first three approaches, the result can be viewed as the set of all sentences that the protocol supports.

5.1 "All Support" Protocol

In [BI84], Borgida and Imielinski discuss a protocol where a group of agents is defined to support a sentence α if every agent in the group support α. Let us call this "All support" protocol.

It is easy to see that the result of "All support" can be described as the disjunction of the knowledge bases. That is, a set of knowledge bases $T_1, \ldots, T_n$ supports a sentence α according to the "All support" protocol iff $T_1 \vee \ldots \vee T_n \models \alpha$. With respect to the postulates MM1-4, we have:

Proposition 5.1 *The "All support" protocol satisfies MM3, but not MM1, MM2, and MM4.*

Proof: MM3 can easily be seen to hold, which shows that the method has a semantic nature rather than syntactic. When all the knowledge bases to be merged are inconsistent, the disjunction of the knowledge bases is also inconsistent. Hence MM1 is violated. However, the disjunction is consistent if at least one of the knowledge bases is consistent. So for most interesting cases, particularly when each knowledge base to be merged is consistent by itself but inconsistent with other knowledge bases, the result of the protocol is consistent. This means that conflict is sure to be resolved by the protocol for most interesting cases.

The protocol violates MM2. Consider two knowledge bases $\{a\}$ and $\{a \rightarrow b\}$. Then the merge result is $a \vee (a \rightarrow b)$, a tautology that does not support anything else. This means that there is no cooperation among the agents and no new knowledge (such as b in this example) is ever achieved, even when there is no conflict among the agents. In other words, the resulting knowledge base of the protocol is more ignorant than any knowledge base before.

For MM4, consider $T_1 = \{a\}, T_2 = \{a\}$ and $T_3 = \{\neg a\}$. Then the result of the protocol is also a tautology $a \vee \neg a$. The conflict on a is not resolved by majority. □

5.2 "Democracy" Protocol

Borgida and Imielinski [BI84] also propose a "Democracy" protocol described as follows. A set of agents is said to support a sentence α iff: (1) some subcommittee supports α, and (2) if some other subcommittee S supports $\neg\alpha$ then there is a subcommittee S' such that S' supports α and cardinality(S') > cardinality(S), where a subcommittee is a subset of agents and is defined to support a sentence if every member in the subcommittee supports the sentence.

Proposition 5.2 *The "Democracy" protocol satisfies MM2-4, but not MM1.*

Proof: See the following "More proponents than opponents" protocol. □

5.3 "More Proponents than Opponents" Protocol

In Section 3, we mentioned a proposal that defines a group of agents to support a sentence if there are more members supporting the sentence than members supporting its negation. Let us call this "More proponents than opponents" protocol. It turns out that the protocol is equivalent to the "Democracy" protocol in the previous subsection.

Proposition 5.3 *The "More proponents than opponents" protocol is equivalent to the "Democracy" protocol and it satisfies MM2-4, but not MM1.*

Proof: The equivalence follows from the fact that in the "Democracy" protocol, a subcommittee is a subset of agents and is defined to support a sentence if every member in the subcommittee supports the sentence.

It is easy to see that the protocol satisfies MM2-3. For MM4, just note that for any literal l, if there are more members supporting l than members supporting or partially supporting $-l$ then there are more members supporting l than members supporting $-l$. From this we also see that the postulate MM4 sets a weaker requirement to merge operators than the "More proponents than opponents" rule applying to literal sentences.

MM1 is violated by the example given in the introductory section. That is, for $T_1 = \{a\}, T_2 = \{b\}$ and $T_3 = \{\neg a \vee \neg b\}$, the protocol supports a and b and $\neg a \vee \neg b$, contradictory. □

5.4 "Maximal Consistent Subsets" Protocol

Baral *et al.* [BKMS92] consider how to combine knowledge bases by manipulating the maximal consistent subsets of the union of the knowledge bases. Although they also consider integrity constraints and priorities, the basic idea is to define the merged result as the disjunction of all the maximal consistent subsets. Let us call this "Maximal consistent subsets" protocol, denoted by *MCS*.

Definition 5.1 *MCS*$(\{T_1, \ldots, T_n\}) = \bigvee\{C \mid C \subseteq U$ *is consistent, and* $\nexists C' \subseteq U$ *such that* C' *is consistent and* $C \subset C'$ $\}$, *where* $U = T_1 \cup \cdots \cup T_n$.

Proposition 5.4 *MCS satisfies MM1 and MM2, but not MM3 and MM4.*

Proof: MM1 and MM2 are satisfied trivially. For MM3, let $T_1 = \{a, b\}$ and $T_2 = \{\neg b\}$ then the merged result is equivalent to a. But if $T_1 = \{a \wedge b\}$ and $T_2 = \{\neg b\}$ then the result is equivalent to $a \vee \neg b$. Thus the *MCS* depends on the particular syntax of the knowledge bases, violating MM3. The problem with syntactic merging is that an agent cannot contribute to the merge his knowledge by reasoning independently with his own knowledge base.

MCS violates MM4 by the example given in the introductory section. That is, for $T_1 = \{a\}, T_2 = \{a\}$ and $T_3 = \{\neg a\}$, the merged result is $a \vee \neg a$, which does not support a. □

6 A Majority Merging Operator

In this section, we propose a merging operator *CMerge* that satisfies MM1-4. By Theorem 4.1, all we have to do is define the pre-order $\leq_{\{T_1,\ldots,T_n\}}$ which is used to measure the distance between a possible world and the set of knowledge bases $\{T_1, \ldots, T_n\}$. We will define the distance between a possible world and a set of knowledge bases as the sum of the distances between the world and each knowledge base, where the distance to a single knowledge base is defined using the method of Dalal's [Dal88], which is simply the minimum number of atoms on which the world differs from some model of the knowledge base.

6.1 The Definition

Let the distance between two possible worlds w_1 and w_2, $dist(w_1, w_2)$, be the number of atoms whose valuations differ in the two possible worlds; and let the distance between a possible world w and a knowledge base T be

$$dist(w, T) = \min_{w' \in [\![T]\!]} dist(w, w').$$

For the boundary case where T is inconsistent, we define $dist(w, T) = 0$.

There may be some possible worlds that are close to a particular knowledge base but distant from others in the set. Since we want the views of the majority to be reflected, we try to select those worlds that are closest *overall* to the set of knowledge bases.

One simple way of defining *overall distance* between a possible world w and $\{T_1, \ldots, T_n\}$ is to take the sum of the distances between the world and each T_i. Hence we say w is *closer to or at the same distance from* $\{T_1, \ldots, T_n\}$ than w' (denoted by $w \preceq_{\{T_1,\ldots,T_n\}} w'$) if $\sum_{i=1}^{n} dist(w, T_i) \leq \sum_{i=1}^{n} dist(w', T_i)$.

Finally, *CMerge* is defined as:

Definition 6.1 $[\![CMerge(\{T_1, \ldots, T_n\})]\!] = Min(\mathcal{W}, \preceq_{\{T_1,\ldots,T_n\}})$.

That is, a possible world is a model of *CMerge* if and only if its overall distance to the set of knowledge bases is minimum. We have the following property.

Lemma 6.1 *CMerge satisfies MM1-4.*

Proof: In view of Theorem 4.1, all we need is to prove that $\preceq_{\{T_1,\ldots,T_n\}}$ is a pre-order and the assignment that maps $\{T_1, \ldots, T_n\}$ to $\preceq_{\{T_1,\ldots,T_n\}}$ is a majority assignment. It is easy to see $\preceq_{\{T_1,\ldots,T_n\}}$ is a pre-order, and it is also obvious that the assignment which maps $\{T_1, \ldots, T_n\}$ to $\preceq_{\{T_1,\ldots,T_n\}}$ satisfies the first and the second conditions of majority assignment.

We now show the third condition of majority assignment. Let l be a literal such that $|T_i \models l| > |T_i \models -l| + |T_i \rhd -l|$, and $w \in [\![l]\!]$. We want to prove that $w \prec_{\{T_1,\ldots,T_n\}} w_l^-$. That is, we want to prove $\sum_{i=1}^n dist(w, T_i) < \sum_{i=1}^n dist(w_l^-, T_i)$.

First of all, there must be some T_i ($i \in [1, n]$) that is consistent. Otherwise, each T_i ($i \in [1, n]$) supports both l and $-l$ and hence $|T_i \models l| = |T_i \models -l|$, contradicting the assumption that $|T_i \models l| > |T_i \models -l| + |T_i \rhd -l|$.

Without loss of generality, we assume T_i ($i \in [1, m], m \leq n$) to be the consistent T_i. Note that there are more T_i ($i = 1, \ldots, n$) that support l than support or partially support $-l$, and each inconsistent T_i ($i = m + 1, \ldots, n$) supports both l and $-l$ but does not partially support $-l$. Hence, there are more T_i ($i = 1, \ldots, m$) that support l than support or partially support $-l$.

For $i \in [1, m]$, T_i either supports l, or supports $-l$, or partially supports $-l$, or neither of the above three. We consider these cases in the following:

1. For the T_i that supports l, we prove $dist(w, T_i) = dist(w_l^-, T_i) - 1$. Let $u \in [\![T_i]\!]$ such that $dist(w, T_i) = dist(w, u)$, i.e., $dist(w, u) = \min_{v \in [\![T_i]\!]} dist(w, v)$. From $T_i \models l$ we have $u \models l$, and it follows from $w \models l$ that $dist(w, u) = dist(w_l^-, u) - 1$. Then to prove $dist(w, T_i) = dist(w_l^-, T_i) - 1$, it suffices to prove $dist(w_l^-, T_i) = dist(w_l^-, u)$.

   Assume to the contrary that $dist(w_l^-, T_i) \neq dist(w_l^-, u)$. Then there is $v \in [\![T_i]\!]$ such that $dist(w_l^-, v) < dist(w_l^-, u)$. From $T_i \models l$ we have $v \models l$. Then it is easy to see $dist(w, v) = dist(w_l^-, v) - 1$ and $dist(w, u) = dist(w_l^-, u) - 1$ since $w \models l$ and $u \models l$. It follows that $dist(w, v) < dist(w, u)$. But since $v \in [\![T_i]\!]$, this contradicts that $dist(w, u) = \min_{v \in [\![T_i]\!]} dist(w, v)$. Therefore, $dist(w_l^-, T_i) = dist(w_l^-, u)$ and it follows that $dist(w, T_i) = dist(w_l^-, T_i) - 1$.

2. For the T_i that supports $-l$, we have $dist(w, T_i) = dist(w_l^-, T_i) + 1$ accordingly.

3. For the T_i that partially supports $-l$, it is easy to see $dist(w, T_i) \leq dist(w_l^-, T_i) + 1$ since w and w_l^- differs only on the atom $\mathcal{P}(l)$.

4. For the remaining T_i, they do not supports l or $-l$ or partially supports $-l$. We prove $dist(w, T_i) \leq dist(w_l^-, T_i)$. Let $u \in [\![T_i]\!]$ such that $dist(w, u) = dist(w, T_i)$, and $v \in [\![T_i]\!]$ such that $dist(w_l^-, v) = dist(w_l^-, T_i)$. Then to prove $dist(w, T_i) \leq dist(w_l^-, T_i)$, it suffices to prove $dist(w, u) \leq dist(w_l^-, v)$. Assume to the contrary that $dist(w, u) > dist(w_l^-, v)$. We have two cases: $v \models -l$ or $v \models l$.

- $v \models -l$. Since T_i does not supports l, $[\![T_i \wedge l]\!] \neq \emptyset$. Then since T_i does not partially supports $-l$, it follows from the Proposition 3.1 that for all $y \in [\![T_i \wedge -l]\!]$, $y_l^- \in [\![T_i]\!]$. Therefore, $v_l^- \in [\![T_i]\!]$. It is clear that $dist(w, v_l^-) = dist(w_l^-, v)$. Then from the assumption we have $dist(w, u) > dist(w, v_l^-)$. But since $v_l^- \in [\![T_i]\!]$, this contradicts that $dist(w, u) = dist(w, T_i)$.
- $v \models l$. Then since $w \models l$, $dist(w_l^-, v) = dist(w, v) + 1$. It follows from the assumption that $dist(w, u) > dist(w, v)$. But since $v \in [\![T_i]\!]$, this also contradicts that $dist(w, u) = dist(w, T_i)$.

In both cases we derive a contradiction. So $dist(w, u) \leq dist(w_l^-, v)$ and it follows that $dist(w, T_i) \leq dist(w_l^-, T_i)$.

Since there are more T_i $(i = 1, \ldots, m)$ that support l than support or partially support $-l$, it is easy to see that for the T_i $(i = 1, \ldots, m)$ that supports l or supports $-l$ or partially supports $-l$, the sum of $dist(w, T_i)$ is smaller than the sum of $dist(w_l^-, T_i)$. For the T_i $(i = 1, \ldots, m)$ that does not support l or $-l$ or partially support $-l$, the sum of $dist(w, T_i)$ is equal to or smaller than the sum of $dist(w_l^-, T_i)$. Thus, $\sum_{i=1}^{m} dist(w, T_i) < \sum_{i=1}^{m} dist(w_l^-, T_i)$. For the inconsistent T_i $(i \in [m+1, n])$, we have $dist(w, T_i) = dist(w_l^-, T_i) = 0$ by the definition of $dist$. It follows that $\sum_{i=1}^{n} dist(w, T_i) < \sum_{i=1}^{n} dist(w_l^-, T_i)$. Thus, $w \prec_{\{T_1,\ldots,T_n\}} w_l^-$ and the assignment which maps $\{T_1, \ldots, T_n\}$ to $\preceq_{\{T_1,\ldots,T_n\}}$ satisfies the third condition of majority assignment, from which the theorem follows. $\Box$

The following result makes sure that if there are some inconsistent knowledge bases in the group, they do not prevent *CMerge* from obtaining the knowledge from the other consistent knowledge bases.

Proposition 6.1 *$CMerge(\{T_1, \ldots, T_n\}) \equiv CMerge(\{T'_1, \ldots, T'_m\})$, where $T'_1, \ldots, T'_m$ are the consistent knowledge bases in $\{T_1, \ldots, T_n\}$.*

Proof: It suffices to prove $\preceq_{\{T_1,\ldots,T_n\}} = \preceq_{\{T'_1,\ldots,T'_m\}}$. This means to prove $w \preceq_{\{T_1,\ldots,T_n\}} w'$ iff $w \preceq_{\{T'_1,\ldots,T'_m\}} w'$, where $w, w' \in \mathcal{W}$. That is, to prove $\sum_{i=1}^{n} dist(w, T_i) \leq \sum_{i=1}^{n} dist(w', T_i)$ iff $\sum_{i=1}^{m} dist(w, T'_i) \leq \sum_{i=1}^{m} dist(w', T'_i)$. Note that if T_i is inconsistent, $dist(w, T_i) = 0$. Hence $\sum_{i=1}^{n} dist(w, T_i) = \sum_{i=1}^{m} dist(w, T'_i)$ and $\sum_{i=1}^{n} dist(w', T_i) = \sum_{i=1}^{m} dist(w', T'_i)$. It follows that $w \preceq_{\{T_1,\ldots,T_n\}} w'$ iff $w \preceq_{\{T'_1,\ldots,T'_m\}} w'$. $\Box$

When T_i is inconsistent for every $i \in [1, n]$, all worlds in $\mathcal{W}$ are equally close to $\{T_1, \ldots, T_n\}$ and consequently $[\![CMerge(\{T_1, \ldots, T_n\})]\!] = \mathcal{W}$, i.e., the result of *CMerge* is a tautology.

6.2 A Syntactic Characterization of *CMerge*

The result of *CMerge* can be obtained by a syntactic transformation of the knowledge bases to be merged, without recourse to any model-theoretic construction.

First, we assume the knowledge base is put in the disjunctive normal form known as *dual clause form*, where a *dual clause* is a set of literals among which $\wedge$ is implied, and the knowledge base is equivalent to the disjunction of all the clauses. We use $DNF(\alpha)$ to denote the dual clause form of α. We require all the dual clauses in $DNF(\alpha)$ to be satisfiable if α is satisfiable. If α is unsatisfiable, we define $DNF(\alpha) = \{\emptyset\}$.

We will use the structure called *multiset* in the syntactic characterization. A *multiset* is a data structure similar to *set* except that duplicate elements are allowed in a multiset. We express a multiset like a set, e.g., a multiset $C = \{a, b, a\}$. We define the operation *combination* ($\sqcup$) between two sets A and B as: $A \sqcup B$ results in a multiset containing all the elements of A and all the elements of B. *Combination* is similar to the operation *union* of sets, except that during the operation of combination duplicate elements are not removed. The membership predicate 'in' ($\in$) between an element and a multiset is as in set theory.

We say a multiset C of literals contains an *inconsistent pair* $\langle l, -l \rangle$ if both l and $-l$ are in C. Given a multiset C, we can break it into two sub-multisets: one is $f(C)$, which contains all the inconsistent pairs that could be found in C; the other one is $t(C)$, which contains the remaining elements of C after $f(C)$ is removed. For example, suppose $C = \{a, a, a, \neg a, \neg a, b, \neg b, c\}$. Then $f(C) = \{a, \neg a, a, \neg a, b, \neg b\}$ and $t(C) = \{a, c\}$. It is useful to notice that $t(C)$ is always satisfiable, and $f(C)$ is unsatisfiable unless it is empty. We can call $t(C)$ the *consistent part* of C and $f(C)$ the *inconsistent part* of C. Let $|C|$ denote the cardinality of C, where C is a set or multiset.

Theorem 6.1 *(Syntactic equivalence of CMerge)*

$$\begin{aligned} &\mathit{CMerge}(\{T_1, \ldots, T_n\}) \equiv \\ &\quad \bigvee\{\bigwedge t(C) \mid C \in S, \textit{and } \nexists C' \in S \textit{ such that } |f(C')| < |f(C)|\}, \end{aligned}$$

where $S = \{D_1 \sqcup \ldots \sqcup D_n \mid D_i \in DNF(T_i), i = 1, \ldots, n\}$.

Proof: For a possible world w and a set of literals D, let $\mathit{diff}(w, D)$ denote the number of literals in D that are not supported by w, i.e., $\mathit{diff}(w, D) = |\{l \in D \mid w \not\models l\}|$. Then we have the following lemma:

Lemma 6.2 *Let w be a possible world and T be a knowledge base. Then*

$$\mathit{dist}(w, T) = \min_{D \in DNF(T)} \mathit{diff}(w, D).$$

Proof: If T is inconsistent (unsatisfiable) then $DNF(T) = \{\emptyset\}$ and $\mathit{dist}(w, T) = 0$. It is easy to see that the lemma holds. Hence we assume T is satisfiable.

Assume $\mathit{dist}(w, T) > \min_{D \in DNF(T)} \mathit{diff}(w, D)$. Let $D' \in DNF(T)$ such that $\mathit{diff}(w, D') = \min_{D \in DNF(T)} \mathit{diff}(w, D)$. Then $\mathit{dist}(w, T) > \mathit{diff}(w, D')$. From the definition of DNF, we know that all the dual clauses in $DNF(T)$ is satisfiable and hence D' is satisfiable. Thus there is not $p \in D'$ such that $\neg p \in D'$. We let w' be a possible world such that $w'(p) = t$ if $p \in D'$, $w'(p) = f$ if $\neg p \in D'$, and $w'(p) = w(p)$ otherwise. Then it is obvious that $w' \models D'$ and therefore $w' \in [\![T]\!]$. It is also easy to see $\mathit{dist}(w, w') = \mathit{diff}(w, D')$. Then from $\mathit{dist}(w, T) > \mathit{diff}(w, D')$ it follows that $\mathit{dist}(w, T) > \mathit{dist}(w, w')$. But since $w' \in [\![T]\!]$, this contradicts the fact that $\mathit{dist}(w, T) = \min_{u \in [\![T]\!]} \mathit{dist}(w, u)$.

Conversely, assume $\mathit{dist}(w, T) < \min_{D \in DNF(T)} \mathit{diff}(w, D)$. Let $w' \in [\![T]\!]$ such that $\mathit{dist}(w, w') = \mathit{dist}(w, T)$. Then $\mathit{dist}(w, w') < \min_{D \in DNF(T)} \mathit{diff}(w, D)$. Since $w' \in [\![T]\!]$,

$w' \models D'$ for some $D' \in DNF(T)$. Then $w' \models l$ for all $l \in D'$. Hence we have $\mathit{diff}(w, D') \leq \mathit{dist}(w, w')$. Then from $\mathit{dist}(w, w') < \min_{D \in DNF(T)} \mathit{diff}(w, D)$ it follows that $\mathit{diff}(w, D') < \min_{D \in DNF(T)} \mathit{diff}(w, D)$. But since $D' \in DNF(T)$, this is a contradiction.

Therefore $\mathit{dist}(w, T) = \min_{D \in DNF(T)} \mathit{diff}(w, D)$. □

Lemma 6.3 $\displaystyle\sum_{i=1}^{n} \mathit{diff}(w, D_i) = \mathit{diff}(w, t(D_1 \sqcup \ldots \sqcup D_n)) + |f(D_1 \sqcup \ldots \sqcup D_n)|/2,$ *where w is any possible world and $D_1, \ldots, D_n$ are any sets of literals.*

Proof: Follows from the definitions of *diff* and $t(D_1 \sqcup \ldots \sqcup D_n)$ and $f(D_1 \sqcup \ldots \sqcup D_n)$. □

We now proceed to prove the main theorem.

Since both RHS and LHS are consistent, we let $w \models RHS$ and $w' \models LHS$. Then $w \models \bigwedge t(D_1 \sqcup \ldots \sqcup D_n)$, for some $D_1 \in DNF(T_1), \ldots, D_n \in DNF(T_n)$ such that $|f(D_1 \sqcup \ldots \sqcup D_n)| \leq |f(C)|$ for all $C \in S$. Let $Q_i \in DNF(T_i)$ ($i = 1, \ldots, n$ respectively) such that $\mathit{diff}(w', Q_i) = \min_{D \in DNF(T_i)} \mathit{diff}(w', D)$. Then the first thing we get is:

$$|f(D_1 \sqcup \ldots \sqcup D_n)| \leq |f(Q_1 \sqcup \ldots \sqcup Q_n)|. \tag{2}$$

For the later development, we prove two inequalities, one concerning w:

$$\sum_{i=1}^{n} \mathit{dist}(w, T_i) \leq |f(D_1 \sqcup \ldots \sqcup D_n)|/2, \tag{3}$$

and another concerning w':

$$\sum_{i=1}^{n} \mathit{dist}(w', T_i) \geq |f(Q_1 \sqcup \ldots \sqcup Q_n)|/2. \tag{4}$$

We first prove (3). From the fact that $w \models \bigwedge t(D_1 \sqcup \ldots \sqcup D_n)$ we have $\mathit{diff}(w, t(D_1 \sqcup \ldots \sqcup D_n)) = 0$. It follows from Lemma 6.3 that $\sum_{i=1}^{n} \mathit{diff}(w, D_i) = |f(D_1 \sqcup \ldots \sqcup D_n)|/2$. By Lemma 6.2 we have $\sum_{i=1}^{n} \mathit{dist}(w, T_i) = \sum_{i=1}^{n} \min_{D \in DNF(T_i)} \mathit{diff}(w, D)$. So to prove (3), it suffices to prove $\sum_{i=1}^{n} \min_{D \in DNF(T_i)} \mathit{diff}(w, D) \leq \sum_{i=1}^{n} \mathit{diff}(w, D_i)$. Assume to the contrary that $\sum_{i=1}^{n} \min_{D \in DNF(T_i)} \mathit{diff}(w, D) > \sum_{i=1}^{n} \mathit{diff}(w, D_i)$. Then there must be at least an $i \in [1, n]$ such that $\min_{D \in DNF(T_i)} \mathit{diff}(w, D) > \mathit{diff}(w, D_i)$. But since $D_i \in DNF(T_i)$, this is a contradiction. Therefore $\sum_{i=1}^{n} \min_{D \in DNF(T_i)} \mathit{diff}(w, D) \leq \sum_{i=1}^{n} \mathit{diff}(w, D_i)$, from which (3) is established.

Next we prove (4). By Lemma 6.2 and the definition of Q_i ($i = 1, \ldots, n$) we have

$$\sum_{i=1}^{n} \mathit{dist}(w', T_i) = \sum_{i=1}^{n} \min_{D \in DNF(T_i)} \mathit{diff}(w', D) = \sum_{i=1}^{n} \mathit{diff}(w', Q_i). \tag{5}$$

By Lemma 6.3 we know $\sum_{i=1}^{n} \mathit{diff}(w', Q_i) \geq |f(Q_1 \sqcup \ldots \sqcup Q_n)|/2$, from which together with the equation (5) we obtain (4).

Then the proof of the main theorem is by contradiction and is carried out in the following two steps:

(a). Assume $w \not\models LHS$. Then since $w' \models LHS$, we have

$$\sum_{i=1}^{n} dist(w', T_i) < \sum_{i=1}^{n} dist(w, T_i).$$

Then using (3) and (4), we get $|f(Q_1 \sqcup \ldots \sqcup Q_n)| < |f(D_1 \sqcup \ldots \sqcup D_n)|$. This contradicts (2). Hence, $w \models LHS$.

(b). Assume $w' \not\models RHS$. Our goal is to derive a contradiction by proving

$$\sum_{i=1}^{n} dist(w, T_i) < \sum_{i=1}^{n} dist(w', T_i).$$

As $|f(D_1 \sqcup \ldots \sqcup D_n)| \leq |f(Q_1 \sqcup \ldots \sqcup Q_n)|$, there are two cases:

- $|f(D_1 \sqcup \ldots \sqcup D_n)| < |f(Q_1 \sqcup \ldots \sqcup Q_n)|$. Then using (3) and (4), we obtain $\sum_{i=1}^{n} dist(w, T_i) < \sum_{i=1}^{n} dist(w', T_i)$.

- $|f(D_1 \sqcup \ldots \sqcup D_n)| = |f(Q_1 \sqcup \ldots \sqcup Q_n)|$. Then $|f(Q_1 \sqcup \ldots \sqcup Q_n)| \leq |f(C)|$ for all $C \in S$. It must be the case that $w' \not\models t(Q_1 \sqcup \ldots \sqcup Q_n)$ otherwise $w' \models RHS$ which contradicts our assumption. This means $diff(w', t(Q_1 \sqcup \ldots \sqcup Q_n)) > 0$. Then by Lemma 6.3, we have $\sum_{i=1}^{n} diff(w', Q_i) > |f(Q_1 \sqcup \ldots \sqcup Q_n)|/2$. Then from equation (5) we obtain $\sum_{i=1}^{n} dist(w', T_i) > |f(Q_1 \sqcup \ldots \sqcup Q_n)|/2$.

  Then using (3) and the fact that $|f(D_1 \sqcup \ldots \sqcup D_n)| = |f(Q_1 \sqcup \ldots \sqcup Q_n)|$, we establish $\sum_{i=1}^{n} dist(w', T_i) > \sum_{i=1}^{n} dist(w, T_i)$.

In both cases we have $\sum_{i=1}^{n} dist(w', T_i) > \sum_{i=1}^{n} dist(w, T_i)$. This contradicts the fact that $w' \models LHS$.

The theorem follows from (a) and (b). $\Box$

The idea of the syntactic equivalence is to select from each DNF knowledge base a dual clause such that the combination of the dual clauses has minimum number of inconsistent pairs; and the result of *CMerge* is the disjunction of the consistent parts of these combinations. Notice that any direct manipulation or storage of complete models of the knowledge bases is not required for the syntactic equivalence of *CMerge*. Based on this syntactic equivalence, efficient algorithms can be developed to compute the result of merging multiple DNF knowledge bases with *CMerge*. Of course, converting the knowledge base to DNF may itself require exponential time, and the problem of computing *CMerge* on arbitrary knowledge bases is easily seen to be NP-hard.

6.3 Examples

We now look at some examples of *CMerge*. This first one shows an example of transformation as described in Theorem 6.1.

Example 6.1 Given three knowledge bases $T_1 = \{a, c\}$, $T_2 = \{a \rightarrow b, \neg c\}$ and $T_3 = \{c\}$. Their disjunctive normal forms $DNF(T_1) = \{\{a, c\}\}$, $DNF(T_2) = \{\{\neg a, \neg c\}, \{b, \neg c\}\}$ and $DNF(T_3) = \{\{c\}\}$.

Then $S = \{C_1, C_2\}$ where $C_1 = \{a, c, \neg a, \neg c, c\}$ and $C_2 = \{a, c, b, \neg c, c\}$.

We have $f(C_1) = \{a, \neg a, c, \neg c\}$ and $t(C_1) = \{c\}$, while $f(C_2) = \{c, \neg c\}$ and $t(C_2) = \{a, b, c\}$. Since C_1 has two inconsistent pairs while C_2 has only one inconsistent pair,

$$\begin{aligned} \mathit{CMerge}(\{T_1, T_2, T_3\}) &\equiv t(C_2) \\ &= \{a, b, c\} \\ &\equiv \{a, a \rightarrow b, c\}. \end{aligned}$$

Example 6.2 Let $T_1 = \{a, c\}$, $T_2 = \{a \rightarrow b, \neg c\}$ and $T_3 = \{b \rightarrow e, c\}$.
Then $\mathit{CMerge}(\{T_1, T_2, T_3\}) \equiv \{a, a \rightarrow b, b \rightarrow e, c\}$.

We observe that the conflict among the knowledge bases centers on c. Since there are two knowledge bases who support c while only one opposes c, the result supports c, reflecting the opinion of the majority. The result also implies the other propositions irrelevant to c, from the three knowledge bases allowing the derivation of the implicit knowledge b and e.

Example 6.3 Suppose in the last example, a fourth knowledge base $T_4 = \{\neg c\}$ is added into the group. Then $\mathit{CMerge}(\{T_1, T_2, T_3, T_4\}) \equiv \{a, a \rightarrow b, b \rightarrow e\}$.

Now in the group there are two knowledge bases supporting c and two supporting $\neg c$. Therefore, c is left undecided. But as in the last example, the propositions irrelevant to c are preserved.

Suppose another knowledge base $T_5 = \{\neg c\}$ is added into the group. Then $\mathit{CMerge}(\{T_1, T_2, T_3, T_4, T_5\}) \equiv \{a, a \rightarrow b, b \rightarrow e, \neg c\}$. The result supports $\neg c$ since now $\neg c$ has a majority of supporters.

Example 6.4 Suppose $T_1 = \{b\}$
$T_2 = \{a, a \rightarrow b\}$
$T_3 = \{\neg b\}$.
Then $\mathit{CMerge}(\{T_1, T_2, T_3\}) \equiv \{a, a \rightarrow b\} \equiv a \wedge b$.

The result supports b since the majority of knowledge bases support b. The result also supports a since T_2 supports a and the other knowledge bases have no opinion about a. Note if we change T_2 to $\{a, b\}$, the result is the same. *CMerge* is independent of the syntax of the knowledge bases.

Example 6.5 Suppose $T_1 = \{a\}$
$T_2 = \{a \rightarrow b\}$
$T_3 = \{a, \neg b\}$.
Then $\mathit{CMerge}(\{T_1, T_2, T_3\}) \equiv \{a, a \rightarrow b\} \vee \{a, \neg b\} \equiv a$.

One way to interpret the result is as follows. The merged knowledge base keeps T_1 and eliminates T_2 and T_3 since T_2 and T_3 are mutually exclusive – one of T_2 and T_3 is logically equivalent to the negation of the other (i.e., $\neg(a \rightarrow b) \equiv a \wedge \neg b$).

Another way to look at the result is as the following. The merged knowledge base supports a since the majority support a. For b it is a striking of balance between $a \rightarrow b$ and b: one knowledge base (T_3) opposes b, but because the majority support a, T_2 generates support for b using its rule $a \rightarrow b$. The opposing and supporting forces are in balance, and hence the issue of b is undecided – the result of merging supports neither b nor $\neg b$.

Example 6.6 Suppose in the last example, a fourth knowledge base $T_4 = \{\neg b\}$ is added into the group. Then $CMerge(\{T_1, T_2, T_3, T_4\}) \equiv \{a, \neg b\}$. The resulting knowledge base opposes b since now the opposing force is stronger than the supporting force for b.

Suppose instead that $T_4 = \{a \rightarrow b\}$. Then $CMerge(\{T_1, T_2, T_3, T_4\}) \equiv \{a, a \rightarrow b\} \equiv a \wedge b$. The merged knowledge base supports b since now the supporting force for b is stronger than the opposing force.

Table 1 compares the results of *CMerge* and the results produced by other methods.

7 Conclusion

We have proposed a set of four postulates that knowledge merging operators are expected to satisfy, within which the fourth postulate formalizes the principle of majority in the context of knowledge base merging. We then investigated the model-theoretic characterization of the set of postulates and proposed the particular knowledge merging operator, *CMerge*, that satisfies the postulates and that appears to resolve conflicts among the knowledge bases in a plausible way.

We notice that it is possible to generalize our result to merging knowledge bases with weights. In that case the merge result should reflect the views of the knowledge bases with higher weights, but for the knowledge bases with the same weights, majority principle should still govern the merging process. We will report the result in a future paper.

Another obvious research direction is to extend the result to first-order case, which is practically appealing. It is also interesting to consider merging the knowledge including 'meta-knowledge' of the agents, e.g., an agent's knowledge about some other agents. The situation can quickly become complicated as 'meta-knowledge' is added in.

Acknowledgments

We thank Anthony Bonner, Hector Levesque, Fangzhen Lin and Ray Reiter for helpful discussions on the subject. We are grateful to the Institute for Robotics and Intelligent Systems of Canada for providing funding to this research.

References

[BAP92] M.W. Bright, A.R.Hurson, and S.H. Pakzad. A taxonomy and current issues in multidatabase systems. *Computer*, 25(3):50–59, 1992.

[BI84] A. Borgida and T. Imielinski. Decision making in committees-a framework for dealing with inconsistency and non-monotonicity. In *Proceedings Nonmonotonic Reasoning Workshop*, pages 21–32, 1984.

[BKMS92] C. Baral, S. Kraus, J. Minker, and V. S. Subrahmanian. Combining knowledge bases consisting of first-order theories. *Computational Intelligence*, 8:45–71, 1992.

Table 1: A Comparison of *CMerge* and Other Methods

	The knowledge bases to be merged	All support	More proponents than opponents (Democracy)	Maximal consistent subsets	*CMerge*
1	$\{a\}$, $\{a\}$, $\{\neg a\}$	$a \vee \neg a$. a tautology	supports $a, \neg a \vee p, \neg a \vee \neg p$. Contradictory	$a \vee \neg a$. A tautology	a
2	$\{a, c\}$,$\{a \rightarrow b, \neg c\}$, $\{b \rightarrow d, c\}$	$\{a, c\} \vee \{a \rightarrow b, \neg c\} \vee \{b \rightarrow d, c\}$. Does not imply b, d, or c.	supports c, $\neg c \vee p$, $\neg c \vee \neg p$. Contradictory	$\{a, a \rightarrow b, b \rightarrow d\}$. Implies b and d, but does not imply c.	$\{a, a \rightarrow b, b \rightarrow d, c\}$. Implies majority view on c, and the implicit knowledge b and d.
3	$\{a, c\}$, $\{a \rightarrow b, \neg c\}$,$\{b \rightarrow d, c\}$, $\{\neg c\}$	$\{a, c\} \vee \{b \rightarrow d, c\} \vee \{\neg c\}$. Does not imply b, or d.	supports $c \vee p, c \vee \neg p, \neg c \vee p, \neg c \vee \neg p$. Contradictory	same as (2)	$\{a, a \rightarrow b, b \rightarrow d\}$
4	$\{b\}$, $\{a$,$a \rightarrow b\}$, $\{\neg b\}$	$b \vee (a \wedge b) \vee \neg b$. A tautology	supports b, $\neg b \vee p$, $\neg b \vee \neg p$. Contradictory	$\{a,\ a \rightarrow b,\ b\} \vee \{a, \neg b\} \vee \{\neg b, a \rightarrow b\}$. Does not imply a, or b	$\{a, a \rightarrow b\}$
5	$\{b\}$, $\{a, b\}$, $\{\neg b\}$	same as (4)	same as (4)	$\{a, b\} \vee \{a, \neg b\}$. Differs from (4). Syntax dependent	same as (4)
6	$\{a\}$,$\{a \rightarrow b\}$, $\{a, \neg b\}$	$a \vee (a \rightarrow b) \vee (a \wedge \neg b)$. A tautology	supports $a, \neg b, \neg a \vee b \vee p$, $\neg a \vee b \vee \neg p$. Contradictory	$\{a,\ a \rightarrow b\} \vee \{a, \neg b\} \vee \{\neg b, a \rightarrow b\}$.	a
7	$\{a\}$,$\{a \rightarrow b\}$, $\{a, \neg b\}$, $\{\neg b\}$	$a \vee (a \rightarrow b) \vee (a \wedge \neg b) \vee \neg b$. A tautology	supports the sentences listed in (6). Contradictory	same as (6)	$a \wedge \neg b$
8	$\{a\}$,$\{a \rightarrow b\}$, $\{a, \neg b\}$, $\{a \rightarrow b\}$	$a \vee (a \rightarrow b) \vee (a \wedge \neg b)$. A tautology	supports the sentences listed in (6). Contradictory	same as (6)	$\{a$,$a \rightarrow b\}$

Note: in the result of "More proponents than opponents", p is any symbol that does not appear in any of the given knowledge bases.

[Bou93] C. Boutilier. Revision sequences and nested conditionals. In *Proceedings of International Joint Conference on Artificial Intelligence (IJCAI)*, 1993.

[Dal88] Mukesh Dalal. Investigations into a theory of knowledge base revision: preliminary report. In *Proceedings of the 7th National Conference of the American Association for Artificial Intelligence*, pages 475–479, 1988.

[FH88] R. Fagin and J. Halpern. Belief, awareness, and limited reasoning. *Artificial Intelligence*, 34:39–76, 1988.

[FHV91] R. Fagin, J. Halpern, and M. Vardi. A model-theoretic analysis of knowledge. *Journal of the ACM*, 38(2):382–428, 1991.

[FKUV86] R. Fagin, G. M. Kuper, J. D. Ullman, and M. Vardi. Updating logical databases. *Advances in Computing Research*, 3:1–18, 1986.

[Fre91] K. A. Frenkel. The human genome project and informatics. *Communications of the Association for Computing Machinery*, 34:42, 1991.

[FUV83] R. Fagin, J. D. Ullman, and M. Vardi. On the semantics of updates in databases. In *2nd ACM SIGACT-SIGMOD Symposium on Principles of Database Systems*, pages 352–365, 1983.

[Hal86] J. Halpern. Reasoning about knowledge: an overview. In *Proceedings of the Conference on Theoretical Aspects of Reasoning about Knowledge*, 1986.

[HM85] J. Halpern and Y. Moses. A guide to the modal logics of knowledge and belief. In *Proceedings IJCAI-85*, pages 480–490, 1985. A complete version appears in *Artificial Intelligence*, 54(3):319-379, 1992, under the title "A guide to completeness and complexity for modal logics of knowledge and belief".

[HM90] J. Halpern and Y. Moses. Knowledge and common knowledge in a distributed environment. *Journal of the ACM*, 37(3):549–587, 1990.

[Kel88] Jerry S. Kelly. *Social Choice Theory: An Introduction.* Springer-Verlag, 1988.

[KM91a] H. Katsuno and A. O. Mendelzon. On the difference between updating a knowledge base and revising it. In *Proceedings of the 2nd International Conference on Principles of Knowledge Representation and Reasoning*, pages 387–394, 1991.

[KM91b] H. Katsuno and A. O. Mendelzon. Propositional knowledgebase revision and minimal change. *Artificial Intelligence*, 52:263–294, 1991.

[Lev84] Hector J. Levesque. A logic of implicit and explicit belief. FLAIR Texh. Rept. 32, Fairchld Lab. for AI Research, Palp Alto. A preliminary version appears in *Proc. of the 4th National conference of the American Association for Artificial Intelligence*, pages 198-202,1984, 1984.

[Lin94] Jinxin Lin. Consistent belief reasoning in the presence of inconsistency. In *Proceedings of Fifth Conference on Theoretical Aspects of Reasoning about Knowledge*, pages 80–94. Morgan Kaufmann, San Francisco, CA, 1994.

[Rev92] P. Revesz. On the semantics of theory change: arbitration between old and new information. In *Proceedings of the Twelfth ACM SIGACT-SIGMOD-SIGART Symposium on Principles of Database Systems*, pages 71–82, 1992.

[Rev94] P. Revesz. On the semantics of arbitration. Manuscript, 1994.