A Syntax-based Framework for Merging Imprecise Probabilistic Logic Programs

Anbu Yue, Weiru Liu
School of Electronics, Electrical Engineering and Computer Science,
Queen's University Belfast, Belfast BT7 1NN, UK
{a.yue, w.liu}@qub.ac.uk

Abstract

In this paper, we address the problem of merging multiple imprecise probabilistic beliefs represented as Probabilistic Logic Programs (PLPs) obtained from multiple sources. Beliefs in each PLP are modeled as conditional events attached with probability bounds. The major task of syntax-based merging is to obtain the most rational probability bound for each conditional event from the original PLPs to form a new PLP. We require the minimal change principle to be followed so that each source gives up its beliefs as little as possible. Some instantiated merging operators are derived from our merging framework. Furthermore, we propose a set of postulates for merging PLPs, some of which extend the postulates for merging classical knowledge bases, whilst others are specific to the merging of probabilistic beliefs.

1 Introduction

The need for dealing with imprecise probabilistic beliefs is present in many real world applications. When multiple sources of probabilistic beliefs are available, one classic objective in statistics is to form consensus beliefs from the sources (e.g., [Nau, 1999; Pate-Cornell, 2002; Kern-Isberner and Rödder, 2004; Chen *et al.*, 2005; Osherson and Vardi, 2006; Bronevich, 2007] among others).

In the literature, many methods have been proposed to aggregate probability distributions where probabilistic beliefs are modeled as multiple probability distributions [Nau, 1999; Bronevich, 2007]. These methods are syntax-irrelevant in the sense that the underlying set of samples is unique and static.

In this paper, we consider situations where probabilistic beliefs are represented as *probabilistic formulas* in the form of $(\psi|\phi)[l, u]$, that is, conditional events are attached with probability bounds. Multiple belief sets may contain overlapped but not entirely the same set of probabilistic conditional events. Probabilistic conditional events explicitly given in an PLP suggest the focus and interests of that PLP in relation to an application. Let us consider a scenario where there are two sources providing relevant information. Source A is more concerned with the relationship between events ϕ and ψ and models its beliefs by estimating a probability bound for $(\psi|\phi)$. Source B is not aware of the relationship between ϕ and ψ and hence does not have this relationship explicitly modeled in its beliefs. Let PLP P_A and P_B be elicited from sources A and B respectively. A rational way to merge P_A and P_B is to generate a PLP in which the bound for $(\psi|\phi)$ is closer to the bound of $(\psi|\phi)$ stated in P_A than that implicitly inferred from P_B. This observation suggests that first a rational merging procedure could be syntax-based in order to preserve any explicitly stated probabilistic beliefs, and second appropriate weights might be attached to sources (or even conditional events) to emphasize their relative importance.

Some syntax-based merging methods were proposed in [Kern-Isberner and Rödder, 2004; Batsell *et al.*, 2002; Osherson and Vardi, 2006] to extract a probability distribution from multiple probabilistic beliefs. However, our main objective of merging multiple PLPs is to obtain a new PLP which has a set of probability distributions as its probabilistic models, so that the impreciseness of the original PLPs can be preserved. Therefore, these methods do not satisfy our requirements.

When beliefs from different sources contradict with each other, each source should give up some of its beliefs in order to get consensus beliefs. Intuitively, we expect that each source gives up its beliefs as little as possible, and this is known as the *minimal change principle*. Another essential requirement is the *mean-value principle*, which states that the resultant probability bound for any conditional event shall not fall beyond the bounds given by the original sources. Constrained by these principles, in this paper we propose a framework to merge imprecise probabilistic beliefs syntactically. Within the framework, we define quantitatively how to measure the amount of beliefs that a source has given up (with respect to a merged result) and what constitute a minimal change (with respect to a set of sources). By instantiating the definitions for measuring the amount that a source shall give up on its beliefs, and for measuring the minimal change, we derive several concrete merging operators that possess different interesting characteristics. To formally regulate the behavior of merging PLPs, we propose a number of postulates in the framework, some of these extend the postulates for merging classical beliefs [Konieczny *et al.*, 2004] whilst others are peculiar to the merging of probabilistic beliefs. It is proved that our instantiated operators satisfy most of these postulates.

The rest of this paper is organized as follows. After a brief review of PLPs in Section 2, we introduce our syntax-based merging framework in Section 3. Then, we provide postulates for merging PLPs in Section 4. Some instantiated operators are defined in Section 5. After comparing our work with related work, we conclude this paper in Section 6.

2 Probabilistic Logic Programs

We consider conditional probabilistic logic programming in this paper [Lukasiewicz, 1998; 2001; 2007]. Let Φ be a finite set of *predicate symbols* and *constant symbols*, and $\mathcal{V}$ be a set of *variables*. An *event* or *formula* can be defined from $\Phi \cup \mathcal{V}$ and connectives $\neg, \wedge, \vee$ as usual. We use Greek letters ϕ, ψ, φ for events. An assignment σ maps each variable to a constant from Φ. The Herbrand semantics can also be canonically defined. We use I to stand for a *possible world*. Notation $I \models_\sigma \phi$ means I is a model of ϕ under σ, and we denote $I \models \phi$ if $I \models_\sigma \phi$ for any assignment σ. A *conditional event* is of the form $\psi|\phi$ with events ψ and ϕ. A *probabilistic formula* is of the form $(\psi|\phi)[l,u]$ which means that the lower and upper probability bounds for conditional event $\psi|\phi$ are l and u, where $l, u \in [0,1]$. In this paper, for simplicity we say $[l,u]$ the probability bound for $(\psi|\phi)$. A *conditional probabilistic logic program (PLP) P* is a set of probabilistic formulas.

A *probabilistic interpretation Pr* is a probability distribution on $\mathcal{I}_\Phi$, where $\mathcal{I}_\Phi$ is the set of all possible worlds. The *probability* of an event ϕ in Pr under σ is defined as $Pr_\sigma(\phi) = \sum_{I \in \mathcal{I}_\Phi, I \models_\sigma \phi} Pr(I)$. If $Pr_\sigma(\phi) > 0$, we define $Pr_\sigma(\psi|\phi) = Pr_\sigma(\psi \wedge \phi)/Pr_\sigma(\phi)$. We define $Pr \models_\sigma (\psi|\phi)[l,u]$, iff $Pr_\sigma(\phi) = 0$ or $Pr_\sigma(\psi|\phi) \in [l,u]$. Pr is a *probabilistic model* of $(\psi|\phi)[l,u]$, denoted by $Pr \models (\psi|\phi)[l,u]$ iff $Pr \models_\sigma (\psi|\phi)[l,u]$ for all σ. Pr is a *probabilistic model* of a PLP P, denoted by $Pr \models P$, iff Pr is a probabilistic model of all $\mu \in P$. A PLP P is *satisfiable* or *consistent* iff a model of P exists. We define $P \models (\psi|\phi)[l,u]$, iff all probabilistic models of P are also probabilistic models of $(\psi|\phi)[l,u]$. We define $P \models_{tight} (\psi|\phi)[l,u]$, iff $P \models (\psi|\phi)[l,u]$, $P \not\models (\psi|\phi)[l,u']$, $P \not\models (\psi|\phi)[l',u]$ for all $l' > l$ and $u' < u$. Note that, if $P \models (\phi|\top)[0,0]$, then it is canonically defined as $P \models_{tight} (\psi|\phi)[1,0]$, where $[1,0]$ stands for an empty set. We define $P \equiv P'$ iff $Pr \models P \Leftrightarrow Pr \models P'$.

3 A Merging Framework for PLPs

3.1 Problem description

A probabilistic *profile*, denoted as $\mathcal{E}$, is a multi-set of PLPs, i.e. $\mathcal{E} = \{P_1, \ldots, P_n\}$, where $P_1, \ldots, P_n$ are PLPs. For simplicity, we require that each P_i is satisfiable and we define $\Gamma(P_i) = \{(\psi|\phi) \mid \exists l, u, (\psi|\phi)[l,u] \in P_i\}$. $\Gamma(P_i)$ is the set containing all the conditional events $(\psi|\phi)$ that are of interests to the source providing P_i. For a probabilistic profile $\mathcal{E}$, we have $\Gamma(\mathcal{E}) = \bigcup_{i=1}^n \Gamma(P_i)$. We call $\mathcal{E}_1$ and $\mathcal{E}_2$ are equivalent, denoted by $\mathcal{E}_1 \equiv \mathcal{E}_2$, iff there exists a bijection between $\mathcal{E}_1$ and $\mathcal{E}_2$ such that each PLP is equivalent to its image. In this paper, we use $\sqcup$ to denote the multi-set union.

In theory, it is possible for a single conditional event to have two probabilistic bounds. For example, let $P = \{(q(X)|p(X))[0,0.1], (q(X)|p(X))[0.9,1]\}$, then P is satisfiable and $P \models_{tight} (p(X)|\top)[0,0]$. If we have $P' = \{(p(X)|\top)[0,0]\}$, then $P \equiv P'$. However, in practice, a single knowledge base rarely contains two different probability bounds for the same conditional event. This condition is formally defined as follows.

Definition 1 *A PLP P is called* **canonical** *iff for any conditional event $(\psi|\phi)$ in $\Gamma(P)$,*

- $(\psi|\phi)[l,u] \in P \Rightarrow l \leq u$;
- $(\psi|\phi)[l_1,u_1] \in P, (\psi|\phi)[l_2,u_2] \in P \Rightarrow l_1 = l_2, u_1 = u_2$.

In our merging framework defined here, we require that each original PLP P_i is canonical without losing generality since every PLP P_i has a canonical PLP equivalent to it.

Let $\mathcal{E} = \{P_1, \ldots, P_n\}$ be a probabilistic profile, and $P \in \Delta(\mathcal{E})$ be a PLP in the result of merging the PLPs in $\mathcal{E}$ where Δ is a merging operator. We require that P satisfies the following constraints.

- P is canonical.
- For any $P_i \in \mathcal{E}$ and any $(\psi|\phi) \in \Gamma(\mathcal{E})$, suppose that $P_i \models (\psi|\phi)[l_i,u_i]$, then $P \models (\psi|\phi)[l,u]$, s.t., $\min_i l_i \leq l \leq u \leq \max_i u_i$. This requirement is also known as the *mean-value property* which guarantees a probabilistic *Pareto principle* from Social Choice Theory [Sen, 1986] as discussed in [Kern-Isberner and Rödder, 2004].
- Since $\Gamma(P) = \Gamma(\mathcal{E})$, the only difference between P_i and P is the probabilistic bounds for the conditional events in $\Gamma(P_i)$. Therefore, we can measure the differences of these probability bounds. $\Delta(\mathcal{E})$ should contain those PLPs such that each of them is closest to all the PLPs in $\mathcal{E}$ (w.r.t. the probability bounds). This requirement is also known as the *minimal change principle*.
- We can further require that the conditional events in the original PLPs have as similar effects as possible in the merging to get more preferred merging results.

3.2 Strong consistency between PLPs

When two knowledge bases infer contradicting conclusions, we consider them inconsistent. That is true in classical logic, since for two propositional (or first order) knowledge bases K_1 and K_2, $K_1 \cup K_2$ is unsatisfiable iff there exists a ϕ such that $K_1 \models \phi$ and $K_2 \models \neg\phi$. However, two PLPs infer contradicting conclusions (e.g. infer two disjoint bounds for the same conditional event) does not suggest that these two PLPs are inconsistent. Let us demonstrate this with the following example.

Example 1 *Let $P_1 = \{(q(t)|p(t))[0.4,0.5]\}$ and $P_2 = \{(q(t)|p(t))[0.51,0.6]\}$ be two PLPs. Informally, P_1 states that "$q(t)$ looks unlikely to be true when $p(t)$ is true" whilst P_2 says that "$q(t)$ looks likely to be true when $p(t)$ is true". Intuitively, P_1 and P_2 contradict each other. However, $P_1 \cup P_2$ is satisfiable, and $P_1 \cup P_2 \models_{tight} (p(t)|\top)[0,0]$, which says that $p(t)$ cannot be true.*

Obviously, $P_1 \cup P_2$ is not a reasonable candidate for merging P_1 and P_2, since by merging, we want to extract appropriate bounds for the conditional event $(q(t)|p(t))$ rather than simply state that the antecedent $p(t)$ can not be true. In our framework, for any P, if $P \equiv P_1 \cup P_2$ and $\Gamma(P) = \Gamma(P_1 \cup P_2)$, then P is not canonical. Therefore, $P_1 \cup P_2$ does not satisfy the first constraint we outlined above, and thus can not be a candidate of merging result. It is worth noting that the union of PLPs may or may not be canonical.

Definition 2 *Let $P_1, \ldots, P_n$ be PLPs. We define the* **compact union** *of $P_1, \ldots, P_n$, denoted as $\biguplus_{i=1}^n P_i$, as follows*

$$\begin{aligned}\biguplus_{i=1}^n P_i &= \{(\psi|\phi)[l,u] \mid (\psi|\phi) \in \bigcup_{i=1}^n \Gamma(P_i),\\ \text{s.t. } l &= \max\{l_i \mid P_i \models_{tight} (\psi|\phi)[l_i,u_i]\}\\ \text{and } u &= \min\{u_i \mid P_i \models_{tight} (\psi|\phi)[l_i,u_i]\}\end{aligned}$$

In the compact union, only one bound is assigned to a conditional event. The compact union of a set of PLPs is semantically equivalent to the union of these PLPs.

Proposition 1 *Let $P_1, \ldots, P_n$ be PLPs. Then $\biguplus_{i=1}^n P_i \equiv \bigcup_{i=1}^n P_i$.*

Now, we can construct a canonical PLP P' from any PLP P by replacing each $(\psi|\phi)[l,u]$ with $l > u$ in $(P \uplus P)$ by $(\phi|\top)[0,0]$. Obviously, $P' \equiv P$. Two PLPs that infer contradicting conclusions should be considered inconsistent. To achieve this, we have the following definition.

Definition 3 *PLPs $P_1, \ldots, P_n$ are* strongly consistent *iff $\biguplus_{i=1}^n P_i$ is canonical and satisfiable (consistent).*

In Example 1, $P_1 \uplus P_2 = \{(q(t)|p(t))[0.51, 0.5]\}$, which is not canonical, therefore P_1 and P_2 are not strongly consistent.

3.3 Minimal Change

Based on the constraints on merging PLPs given above, we take merging PLPs in $\mathcal{E}$ as a process of constructing a canonical consistent PLP P s.t. $\Gamma(P) = \Gamma(\mathcal{E})$. Then, in addition to the extra conditional events in P from other PLP P_j, the major difference between P and any $P_i \in \mathcal{E}$ is that the probabilistic bounds for the conditional events in $\Gamma(P_i)$ could be different. A quantitative measure about the difference between these probabilistic bounds for conditional events in $\Gamma(P_i)$ can be viewed as a measure about how much belief P_i has given up in order to reach an agreement with other PLPs in $\mathcal{E}$. Ideally, each PLP wants to give up its beliefs as little as possible.

For any l and u, if $0 \le l \le u \le 1$ then we call $[l,u]$ a sub-interval of $[0,1]$.

Definition 4 *Let $[l,u], [l',u']$ be two sub-intervals of $[0,1]$. We define the* **change** *from interval $[l,u]$ to interval $[l',u']$, denoted by $\mathsf{Ch}([l,u],[l',u'])$, as*

$$\mathsf{Ch}([l,u],[l',u']) = |l - l'| + |u - u'|.$$

Definition 5 *Let $[l,u], [l',u']$ be two sub-intervals of $[0,1]$. We define the* **weak change** *from interval $[l,u]$ to interval $[l',u']$, denoted by $\mathsf{wCh}([l,u],[l',u'])$, as*

$$\mathsf{wCh}([l,u],[l',u']) = \max(l - l', 0) + \max(u' - u, 0)$$

Obviously, Ch is symmetric while wCh is asymmetric.

Based on the change of bounds, we can define the change and weak change from one PLP to another.

Definition 6 *Let P_1 and P_2 be two canonical PLPs s.t. $\Gamma(P_1) \subseteq \Gamma(P_2)$. We define the change (resp. weak change) from P_1 to P_2, denoted as $\mathsf{d}(P_1, P_2)$, as*

$$\begin{aligned}\mathsf{d}(P_1,P_2) \;=\; & \Sigma(\{\mathsf{d}([l_1,u_1],[l_2,u_2]) \mid\\ & (\psi|\phi) \in \Gamma(P_1), (\psi|\phi)[l_1,u_1] \in P_1,\\ & \text{and } (\psi|\phi)[l_2,u_2] \in P_2\})\\ \text{where } & \mathsf{d} \text{ is } \mathsf{Ch} (\text{resp. } \mathsf{wCh}).\end{aligned}$$

$\mathsf{Ch}(P_1, P_2)$ is greater than 0 iff some bounds are indeed changed, while $\mathsf{wCh}(P_1, P_2)$ is greater than 0 iff some bounds are loosened.

Example 2 *Let us consider a situation about whether taking drug* a *will reduce the probability of a patient's (with a certain disease* m*) mortality. One* Doctor *thinks that a large proportion of the patient population with this disease will have a highly reduced probability of mortality, whilst another* Doctor *thinks a patient also with condition* d *is unlikely to be benefited from this drug. Obviously, patients having disease* m *and also with condition* d *form a subclass of patients with disease* m.

Let $rMor(X)$ denote "reducing mortality of X" and $dis(X, \mathsf{d})$ denote "X having disease of d*", and let*

$\mu_1 = (rMor(X)|drug(X,\mathsf{a}) \wedge dis(X,\mathsf{m}))$
$\mu_2 = (dis(X,\mathsf{d})|dis(X,\mathsf{m}))$
$\mu_3 = (rMor(X)|drug(X,\mathsf{a}) \wedge dis(X,\mathsf{m}) \wedge dis(X,\mathsf{d}))$

Then, the beliefs from the two doctors can be represented by the two PLPs below:

$$\begin{aligned}P_1 &= \{\mu_1[0.98, 1.0], \mu_2[0.1, 1.0]\}\\ P_2 &= \{\mu_2[0.10, 1.0], \mu_3[0.0, 0.1]\}\end{aligned}$$

Let P and P' be two PLPs s.t.

$$\begin{aligned}P &= \{\mu_1[0.914, 0.994], \mu_2[0.1, 1], \mu_3[0.034, 0.138]\}\\ P' &= \{\mu_1[0.490, 0.955], \mu_2[0.1, 1], \mu_3[0.400, 0.550]\}\end{aligned}$$

We can calculate the (weak) changes from P_1 and P_2 to P and P' respectively, as shown in the following tables, e.g. $\mathsf{Ch}(P_1, P) = 0.083$ *and* $\mathsf{wCh}(P_2, P') = 0.45$.

Ch	P	P'
P_1	0.072	0.535
P_2	0.072	0.850

wCh	P	P'
P_1	0.066	0.490
P_2	0.038	0.450

Definition 7 *An* **aggregation function** Ag *is a total function associating a nonnegative real number to every finite tuple of nonnegative real numbers that satisfies:*

- $x \le y \Rightarrow \mathsf{Ag}(x_1, ..., x, ..., x_n) \le \mathsf{Ag}(x_1, ..., y, ..., x_n)$.
- $\mathsf{Ag}(x_1, ..., x_n) = 0$ *if and only if* $x_1 = \cdots = x_n = 0$.
- *For every nonnegative real number x,* $\mathsf{Ag}(x) = x$.

Functions Max, Sum, and weighted sum wSum are exemplar aggregation functions.

Definition 8 *Let $\mathcal{E}$ be a probabilistic profile, and* Ag *be an aggregation function. Suppose that P is a canonical PLP s.t. $\Gamma(P) = \Gamma(\mathcal{E})$. We define the (weak) change from $\mathcal{E}$ to P, denoted as $\mathsf{d}^{\mathsf{Ag}}(\mathcal{E}, P)$, as $\mathsf{d}^{\mathsf{Ag}}(\mathcal{E}, P) = \mathsf{Ag}(\{\mathsf{d}(P_i, P) \mid P_i \in \mathcal{E}\})$, where* d *is* wCh *or* Ch.

In Example 2, the (weak) changes from $\mathcal{E}$ to P and P' are shown in the following table.

Ag	$\mathsf{Ch}^{\mathsf{Ag}}(\mathcal{E}, \bullet)$		$\mathsf{wCh}^{\mathsf{Ag}}(\mathcal{E}, \bullet)$	
	P	P'	P	P'
Sum	0.144	1.385	0.104	0.940
Max	0.072	0.850	0.066	0.490

To preserve the minimal change principle, P' should not be chosen as the merge result.

Proposition 2 *Let $\mathcal{E}$ be a probabilistic profile, then $\mathsf{wCh}(P_i, \biguplus \mathcal{E}) = 0$ for all $P_i \in \mathcal{E}$.*

Proposition 3 *Let $\mathcal{E}$ be a probabilistic profile and P be a PLP. If $\mathsf{wCh}(P_i, P) = 0$ for all $P_i \in \mathcal{E}$, then $P \models \biguplus \mathcal{E}$.*

The above two propositions state that if we simply union multiple PLPs together, then we get tighter bounds, and thus the weak changes from all source PLPs to the result are 0.

Computing weights

To address the importance of a source for contributing to the calculation of the merged bound for a conditional event during merging, we investigate how weights shall be attached to sources (or even conditional events).

Definition 9 *Let $\mathcal{E} = \{P_1, \ldots, P_n\}$ be a probabilistic profile. Suppose that the set $\Gamma(\mathcal{E})$ can be represented as $\{\mu_1, \ldots, \mu_m\}$. Let $W = (W^l, W^u)$ be a pair of $n \times m$ matrices from $\mathcal{R}^{[0,1]}$, s.t. $\forall k, \sum_i W^l[i][k] = \sum_i W^u[i][k] = 1$. For any $P_i \in \mathcal{E}$ and $\mu_k \in \Gamma(\mathcal{E})$, suppose that $P_i \models_{tight} \mu_k[l_{ik}, u_{ik}]$.* A weighted combination *of $P_1, \ldots, P_n$ w.r.t. W is a PLP, denoted as $\bigotimes^W_{P_i \in \mathcal{E}} P_i$, s.t. $\bigotimes^W_{P_i \in \mathcal{E}} P_i = \{\mu_k[l, u] \mid \mu_k \in \Gamma(\mathcal{E})$, and $l = \sum_{\mu_i \in \Gamma(P_i)} W^l[i][k] \times l_{ik}, u = \sum_i W^u[i][k] \times u_{ik}\}$.*

For simplicity, we use $\bigotimes^W \mathcal{E}$ to denote $\bigotimes^W_{P_i \in \mathcal{E}} P_i$. Weights $W^l[i][j]$ and $W^u[i][j]$ reflect the importance of the lower bound and upper bound for conditional event μ_j contributed by PLP P_i.

On the other hand, it is easy to prove that if a PLP P satisfies the mean-value property w.r.t. $\mathcal{E}$ then there must exist a W s.t. $P = \bigotimes^W \mathcal{E}$.

Example 3 (Cont. of Example 2) *Assume that weights assigned to the three conditional events in P_1 and P_2 are*

$$W^l = W^u = \begin{array}{c} \mu_1 \quad\; \mu_2 \quad\; \mu_3 \\ \left(\begin{array}{ccc} 0.9325 & 0.5 & 0.0425 \\ 0.0675 & 0.5 & 0.9575 \end{array} \right) \end{array} \begin{array}{l} \\ \dashrightarrow P_1 \\ \dashrightarrow P_2 \end{array}$$

Then, the weighted combination of P_1 and P_2 produces P.

4 Postulates

In classical merging, some postulates are provided [Konieczny *et al.*, 2004]. In this paper, we adapt the postulates from [Konieczny *et al.*, 2004] for merging PLPs. The merging operator Δ maps a probabilistic profile to a set of PLPs. Suppose that $P \in \Delta(\mathcal{E})$, then ideally P should satisfy the following:

IC1 P is satisfiable.

IC2 If $\mathcal{E}$ is strongly consistent then $P \models P_i$ for all $P_i \in \mathcal{E}$.

IC3 If $\mathcal{E}_1 \equiv \mathcal{E}_2$, then $\Delta(\mathcal{E}_1) \equiv \Delta(\mathcal{E}_2)$.

IC4 If $\mathcal{E} = \{P_1, P_2\}$, then $\exists P_1' \in \Delta(\{P_1, P_2\})$ $P_1' \uplus P_1$ is satisfiable iff $\exists P_2' \in \Delta(\{P_1, P_2\})$ $P_2' \uplus P_2$ is satisfiable.

IC5 If $\mathcal{E} = \mathcal{E}_1 \sqcup \mathcal{E}_2$, $P_1 \in \Delta(\mathcal{E}_1)$, and $P_2 \in \Delta(\mathcal{E}_2)$, then $P_1 \uplus P_2 \models P$.

IC6 Let $\mathcal{E} = \mathcal{E}_1 \sqcup \mathcal{E}_2$, $P_1 \in \Delta(\mathcal{E}_1)$, $P_2 \in \Delta(\mathcal{E}_2)$. If $P_1 \uplus P_2$ is satisfiable, then $\exists P' \in \Delta(\mathcal{E})$, $P' \models P_1 \uplus P_2$.

IC7 Let $P_1 \in \Delta(\mathcal{E} \sqcup \{P_i\})$ and $P_2 \in \Delta(\mathcal{E} \sqcup \{P_i'\})$, where P_i' is obtained by replacing a probabilistic formula $(\psi_i|\varphi_i)[l_i, u_i]$ in P_i with $(\psi_i|\varphi_i)[l_i', u_i']$. Then $\mathsf{Ch}(P_1, P_2) \leq \mathsf{Ch}(P_i, P_i')$.

IC8 If $P_i \models (\psi|\phi)[l, u]$ for all i, then $P \models (\psi|\phi)[l, u]$.

Maj For any $\mathcal{E}$ and $P \in \mathcal{E}$. Suppose that $P_1 = \cdots = P_n = P$ and $P' \in \Delta(\mathcal{E} \sqcup \{P_1, \ldots, P_n\})$, then $\lim_{n\to\infty} \mathsf{wCh}(P, P') = 0$.

Arb For any $\mathcal{E}$ and $P \in \mathcal{E}$. Suppose that $P_1 = \cdots = P_n = P$, then $\Delta(\mathcal{E}) = \Delta(\mathcal{E} \sqcup \{P_1, \ldots, P_n\})$.

In these postulates, IC1-IC6, and **Maj** are essentially equivalent to those provided in [Konieczny *et al.*, 2004] with trivial integrity constraint $\top$ and **Arb** is to that provided in [Meyer, 2001]. IC7 says that a slightly change of the sources will not affect the result too much, and the change of the results will always be smaller than the change of the sources. IC8, also known as *consensus* postulate, means that the common beliefs of the original PLPs should be preserved after merging.

5 Merging Operators

5.1 Merging by dilation of the probability bounds

A straightforward method to merge multiple PLPs is to weaken all of them until they become strongly consistent and then generate their compact union to get a new PLP.

Definition 10 *Let P be a canonical PLP. A PLP P' is called a dilation of P iff $P \models P'$ and $\Gamma(P) = \Gamma(P')$.*

A dilation of a PLP is obtained by loosing the probability bounds for some conditional events in the original PLP.

Definition 11 *Let $\mathcal{E}$ be a probabilistic profile, we define* dilation *merging operators, denoted by $\Delta^{(\mathsf{wCh},\mathsf{Ag})}(\mathcal{E})$ where* Ag *is an aggregation function, as $P \in \Delta^{(\mathsf{wCh},\mathsf{Ag})}(\mathcal{E})$ iff $P = \arg\min_P \{\mathsf{wCh}^{\mathsf{Ag}}(\mathcal{E}, P) \mid P$ is satisfiable and $P = \bigotimes^W(\mathcal{E})$ for some pair of weight matrices $W\}$.*

Proposition 4 *Let $\mathcal{E} = \{P_1, \ldots, P_n\}$ be a probabilistic profile, and $P \in \Delta^{(\mathsf{wCh},\mathsf{Max})}(\mathcal{E})$. Then there exits $\mathcal{E}' = \{P_1', \ldots, P_n'\}$, s.t. P_i' is a dilation of P_i, and $P \equiv \biguplus \mathcal{E}'$.*

The above proposition indicates that the merging result can also be obtained by first weakening the beliefs of every P_i such that these weakened PLPs are strongly consistent with others, and the merging result is the compact union of the weakened PLPs.

Example 4 *Let P_1 and P_2 be two PLPs as given in Example 2. We have $\Delta^{(\mathsf{wCh},\mathsf{Max})}$ as a singleton set. Suppose that $\Delta^{(\mathsf{wCh},\mathsf{Max})} = \{P''\}$, then we have $P'' = \{\mu_1[0.916, 0.916], \mu_2[0.1, 1], \mu_3[0.164, 0.164]\}$.*

Proposition 5 *Let $\mathcal{E} = \{P_1, P_2\}$ be a probabilistic profile, and $P \in \Delta^{(\mathsf{wCh},\mathsf{Max})}(\mathcal{E})$, then $\mathsf{wCh}(P_1, P) = \mathsf{wCh}(P_2, P)$.*

5.2 Merging by Keeping Impreciseness

Decreasing impreciseness VS. keeping impreciseness

Merging by dilation may drastically decrease the impreciseness of beliefs in that the bounds for some conditional events are much tighter than their bounds stated in the original PLPs. However, sometimes we still want the merging result be imprecise to maintain the reliability. For instance, let P_1 and P_2 be as given in Example 1. By dilation merging, we get $\Delta^{(\mathsf{wCh},\mathsf{Max})}(\{P_1, P_2\}) = \{P\}$ where $P = \{(q(t)|p(t))[0.505, 0.505]\}$. Probability 0.505 is chosen because it is the only value that is closest to both of the upper bound for $(q(t)|p(t))$ given in P_1 and the lower bound for $(q(t)|p(t))$ given in P_2. However, when the imprecise bounds given in P_1 and P_2 suggest that both of them agree that imprecise probability bounds are more suitable for the conditional event due to lack of information, the bound $[0.505, 0.505]$ is too tight as the merging result. This motivates us to propose the following operator to keep impreciseness.

The merging operator

In general, we do not require that $W^l = W^u$. In fact, in the dilation operator, the corresponding weights do not satisfy the above condition.

However, there are some intuition to require $W^l = W^u$ if we regard the lower and the upper bound are equally important (or equally possible) for any conditional event in a PLP.

When requiring $W^l = W^u$, it is more nature to use Ch to measure the difference of a probabilistic profile to a PLP, since in Ch, the differences of the lower bounds and upper bounds are equally treated. Therefore, for the operators $\Delta^{(\mathsf{Ch},\mathsf{Ag})}$, we require that $W^l = W^u$:

Definition 12 *Let $\mathcal{E}$ be a probabilistic profile, we define merging operators $\Delta^{(\mathsf{Ch},\mathsf{Ag})}(\mathcal{E})$ as $P \in \Delta^{(\mathsf{Ch},\mathsf{Ag})}(\mathcal{E})$ iff $P = \arg\min_P\{\mathsf{Ch}^{\mathsf{Ag}}(\mathcal{E}, P) \mid P$ is satisfiable and $P = \bigotimes^W \mathcal{E}$ for some pair $W = (W^l, W^u)$ and $W^l = W^u\}$, where Ag is an aggregation function.*

Example 5 *Let P_1 and P_2 be as given in Example 1. By merging, we get that $\Delta^{(\mathsf{Ch},\mathsf{Max})}(\{P_1, P_2\}) = \{P\}$ where $P = \{(\psi|\phi)[0.451, 0.550]\}$.*

Example 6 *Let PLP P_1 and P_2 be as given in Example 2. Then $\Delta^{(\mathsf{Ch},\mathsf{Max})}(\{P_1, P_2\})$ is a singleton set, which contains the only one PLP P as given in Example 2.*

Comparing to Example 4, in the above example, the belief about the probability of μ_3 is imprecise. The upper bound 0.134 for μ_3 is lower than 0.164 given in Example 4, this is because the lower bound 0 for μ_3 in P_2 has more effect in this operator.

5.3 Preferred Candidates

However, PLPs that satisfy Definition 12 are not unique in general, as shown by the following example.

Example 7 *Let $P_1 = \{(p|\top)[1, 1], (q|\top)[1, 1]\}$ and $P_2 = \{(p \wedge q|\top)[0, 0]\}$. Then P_1 and P_2 are unsatisfiable. Let P and P' be given as*

$$P = \{(p|\top)[0.75, 1], (q|\top)[0.75, 1], (p \wedge q)[0.5, 1]\}$$
$$P' = \{(p|\top)[0.8, 1], (q|\top)[0.7, 1], (p \wedge q)[0.5, 1]\}$$

We have that $P, P' \in \Delta^{(\mathsf{Ch},\mathsf{Max})}(\{P_1, P_2\})$.

In this example, P is more reasonable than P', since in P, p and q are symmetrically defined, just as that they are symmetrically defined in P_1 and P_2. In other words, the weights for $(p|\top)$ and $(q|\top)$ w.r.t. P_1 and P_2 are equivalent when obtaining P and inequivalent when obtaining P'. Intuitively, we should choose the PLPs from the merging results such that the weights for the conditional events in the original PLPs are as close as possible.

Let $S = \langle s_1, \ldots, s_n\rangle$ be a finite sequence of real numbers in descending order. We define $\trianglelefteq$ as a lexicographic order among descending ordered sequences, i.e. $S_1 \trianglelefteq S_2$ iff

- $|S_1| \leq |S_2|$ and $s_{1i} = s_{2i}$ for all $i \leq |S_1|$, or
- $\exists i$ s.t. $s_{1i} < s_{2i}$ and $\forall j < i$, $s_{1j} = s_{2j}$.

Let $\mathcal{S}$ be a finite multi-set of real numbers, we define $sq(\mathcal{S})$ as a descending ordered sequence of the elements from $\mathcal{S}$, then we can define a partial order $\preccurlyeq$ as $\mathcal{S}_1 \preccurlyeq \mathcal{S}_2$ iff $sq(\mathcal{S}_1) \trianglelefteq sq(\mathcal{S}_2)$. For instance, if $\mathcal{S}_1 = \{0.8, 0.9, 0.8, 0.6\}$ and $\mathcal{S}_2 = \{0.8, 0.9, 0.7, 0.7\}$, then $\mathcal{S}_2 \preccurlyeq \mathcal{S}_1$.

Let P and P' be PLPs s.t. $\Gamma(P_1) = \Gamma(P)$. We define $\mathcal{C}^d(P', P)$ as $\mathcal{C}^d(P', P) = \{d([l', u'], [l, u]) \mid \mu[l', u'] \in P'$ and $\mu[l, u] \in P\}$, where d is wCh or Ch.

Set $\mathcal{C}^d(P', P)$ contains the (weak) change values from the bounds given in an original PLP (P') to those given in a candidate (P) of merging result for the conditional events in P'.

Definition 13 *Let $\mathcal{E} = \{P_1, \ldots, P_n\}$ be a probabilistic profile. Define $\Delta_{\preccurlyeq}^{(\mathsf{d},\mathsf{Max})}(\mathcal{E})$ as a set of PLPs s.t.*

$$P \in \Delta_{\preccurlyeq}^{(\mathsf{d},\mathsf{Max})}(\mathcal{E}) \quad \text{iff} \quad \forall P' \in \Delta^{(\mathsf{d},\mathsf{Max})}(\mathcal{E}), \nexists P_i \in \mathcal{E}, \text{ and } \mathcal{C}^{\mathsf{d}}(P_i, P) \preccurlyeq \mathcal{C}^{\mathsf{d}}(P_i, P')$$

where d is wCh or Ch.

The above definition applies the min-max principle to require that the maximum (weak) change from the probabilistic formulas in the original PLPs be minimized.

Example 8 (Cont. of Example 7) *Now, we have that $\Delta_{\preccurlyeq}^{(\mathsf{Ch},\mathsf{Max})}(\mathcal{E}) = \{P\}$. So, the less reasonable PLP P' is eliminated from the candidates of merging results.*

Similarly, for operators $\Delta^{(\mathsf{d},\mathsf{Sum})}$, we can obtain $\Delta_{\preccurlyeq}^{(\mathsf{d},\mathsf{Sum})}$ by requiring the (weak) changes from the original PLPs to the merging result be as close as possible, but we omit the details due to space limitation.

5.4 Properties

Proposition 6 *Let Ag be an aggregation function.*
$\Delta^{(\mathsf{wCh},\mathsf{Ag})}$ satisfies postulates IC1, IC2, and IC4-IC8;
$\Delta^{(\mathsf{Ch},\mathsf{Ag})}$ satisfies postulates IC1, IC4, IC5, IC7, and IC8;
$\Delta_{\preccurlyeq}^{(\mathsf{wCh},\mathsf{Ag})}$ satisfies postulates IC1, IC2 and IC4-IC8;
$\Delta_{\preccurlyeq}^{(\mathsf{Ch},\mathsf{Ag})}$ satisfies postulates IC1, IC4, IC5, IC7, and IC8.

Since postulate IC3 means that the merging operator is syntax-irrelevant and our framework is syntax-based, our operators can not satisfy IC3. We can extend our framework so that two semantically equivalent but syntactically different conditional events (denoted like $(\psi_1|\phi_1) \equiv (\psi_2|\phi_2)$) can be taken as the same conditional event. As a consequence, our merging framework is beyond syntax-based in which how conditional events are expressed is irrelevant.

Proposition 7 *Our merging operators $\Delta^{(\mathsf{d},\mathsf{Sum})}$ and $\Delta_{\preccurlyeq}^{(\mathsf{d},\mathsf{Sum})}$ satisfy postulate* **Maj***, while operators $\Delta^{(\mathsf{d},\mathsf{Max})}$ and $\Delta_{\preccurlyeq}^{(\mathsf{d},\mathsf{Max})}$ satisfy postulate* **Arb***, where d is wCh or Ch.*

It is worth noting that in propositional logic, no merging operators can simultaneously satisfy IC2, IC4, IC6 and **Arb**.

6 Related Work and Conclusion

Related work In the literature, many methods were proposed to merge or aggregate probability distributions [Clemen and Winkler, 1993; Jacobs, 1995; Pate-Cornell, 2002; Chen *et al.*, 2005]. In these methods, each source represents its beliefs by a single probability distribution, and the merging result is also a probability distribution. Therefore, these methods can be classified as model-based. There are some methods [Nau, 1999; Bronevich, 2007] that can aggregate imprecise probabilities (the lower or upper bounds for a conditional event), but these methods are also syntax-irrelevant, in the sense that they require the original beliefs to be expressed on the same set of (conditional) events. In [Clemen and Winkler, 1993; Jacobs, 1995; Chen *et al.*, 2005; Nau, 1999; Bronevich, 2007], aggregation functions are used to calculate the probabilities of the events, and the (weighted) sum function is a common choice. In contrast, our framework starts with a set of PLPs, and each PLP has a set of probability distributions associated with it. Second, our method is syntax-based; explicitly stated conditional events in the original PLPs are treated as of most relevant beliefs of the sources and our merging framework preserves these beliefs as much as possible. On the technical aspect, the probability bounds from the merged result are also the aggregation of the bounds from the original PLPs. Different from the methods mentioned in this paragraph, the weights for the conditional events reflect the relevant importance of conditional events to in the original PLPs.

In [Batsell *et al.*, 2002; Osherson and Vardi, 2006], a syntax-based method was proposed for aggregating probabilistic beliefs, where each source gives a set of probabilistic formulas. In their method, all statements from the sources are put together (union) to form a (possibly inconsistent) set and the aggregation procedure is to find a probability distribution that is closest to all the statements in the set. Their method did not explicitly consider the change from a source to the merging result. On the contrary, our method uses the changes to guide the merging procedure.

In [Kern-Isberner and Rödder, 2004], a method based on the maximum entropy principle is proposed. In their method, a single probability distribution is obtained as the merging result even when the sources are imprecise. On the contrary, in our framework, imprecise PLPs are returned as the result of merging.

In addition, we provided postulates for merging probabilistic beliefs and proved our instantiated operators satisfy most of them. Our postulates generalize the postulates for merging classical knowledge bases provided in [Konieczny *et al.*, 2004], and satisfy properties that should be considered in the view of probability theory. For example, IC8 is the consequence of mean-value property.

On the other hand, IC8 can be reduced to propositional logic:

IC8' Let $K_1, \ldots, K_n$ be propositional knowledge bases, and $K_i \models \phi$ for all $i \in [1, n]$, then the merging result $\Delta(K_1, \ldots, K_n) \models \phi$.

However, not all propositional merging operators satisfy IC8'. For example. Let operators $\Delta_{IC}^{d_D,\mathsf{Sum},\mathsf{Sum}}$ and $\Delta_{IC}^{d_D,\mathsf{Sum},\mathsf{Max}}$ be as given in [Konieczny *et al.*, 2004]. Let $IC = \top$, $K_1 = \{p, p \to q\}$, and $K_2 = \{p, p \to \neg q\}$, then $\Delta_{IC}^{d_D,\mathsf{Sum},\mathsf{Sum}}(K_1, K_2) \models p$ but $\Delta_{IC}^{d_D,\mathsf{Sum},\mathsf{Max}}(K_1, K_2) \not\models p$.

In [Kern-Isberner and Rödder, 2004], a simple probabilistic *Pareto principle* was proposed stating that the merging result should assign probability value x to $(B|A)$ for merging $K_1 = \cdots = K_n = \{(B|A)[x]\}$. Actually, it is essentially a special case of postulate IC2 when all original PLPs are the same, and all of our operators satisfy this probabilistic Pareto principle even for those which do not generally satisfy IC2. In addition, we do not require the original PLPs to contain only one probabilistic formula.

Conclusion In this paper, we proposed a syntax-based framework for merging PLPs. Explicit conditional events in original PLPs suggest in what aspect an original PLP is relevant to the scenario, and a merging procedure should respect this and therefore should be syntax-based. From our framework, concrete merging operators can be defined and they possess different interesting properties.

We also provided postulates for merging imprecise probabilistic beliefs, and our postulates extend the postulates for merging propositional knowledge bases.

Acknowledgments

We would like to thank Anthony Hunter and Jianbing Ma for their comments on an earlier version of this paper.

References

[Batsell *et al.*, 2002] R. Batsell, L. Brenner, D. N. Osherson, S. Tsavachidis, and M. Y. Vardi. Eliminating incoherence from subjective estimates of chance. In *KR*, pages 353–364, 2002.

[Bronevich, 2007] A. G. Bronevich. Necessary and sufficient consensus conditions for the eventwise aggregation of lower probabilities. *Fuzzy Sets and Systems*, 158(8):881–894, 2007.

[Chen *et al.*, 2005] Y. Chen, C.-H. Chu, T. Mullen, and D. M. Pennock. Information markets vs. opinion pools: an empirical comparison. In *ACM Conference on Electronic Commerce*, pages 58–67, 2005.

[Clemen and Winkler, 1993] R. T. Clemen and R. L. Winkler. Aggregating point estimates: a flexible modeling approach. *Manage. Sci.*, 39(4):501–515, 1993.

[Jacobs, 1995] R. A. Jacobs. Methods for combining experts' probability assessments. *Neural Comput.*, 7(5):867–888, 1995.

[Kern-Isberner and Rödder, 2004] G. Kern-Isberner and W. Rödder. Belief revision and information fusion on optimum entropy. *Int. J. Intell. Syst.*, 19(9):837–857, 2004.

[Konieczny *et al.*, 2004] S. Konieczny, J. Lang, and P. Marquis. DA^2 merging operators. *Artif. Intel.*, 157(1-2):49–79, 2004.

[Lukasiewicz, 1998] T. Lukasiewicz. Probabilistic logic programming. In *ECAI*, pages 388–392, 1998.

[Lukasiewicz, 2001] T. Lukasiewicz. Probabilistic logic programming with conditional constraints. *ACM Trans. Comput. Log.*, 2(3):289–339, 2001.

[Lukasiewicz, 2007] T. Lukasiewicz. Nonmonotonic probabilistic logics under variable-strength inheritance with overriding: Complexity, algorithms, and implementation. *Int. J. Approx. Reasoning*, 44(3):301–321, 2007.

[Meyer, 2001] T. A. Meyer. On the semantics of combination operations. *J. of Applied Non-Classical Logics*, 11(1-2):59–84, 2001.

[Nau, 1999] R. Nau. The aggregation of imprecise probabilities. In *ISIPTA*, pages 286–295, 1999.

[Osherson and Vardi, 2006] D. Osherson and M. Y. Vardi. Aggregating disparate estimates of chance. *Games and Economic Behavior.*, 56(1):148–173, 2006.

[Pate-Cornell, 2002] E. Pate-Cornell. Fusion of intelligence information: A bayesian approach. *Risk Analysis*, 22(3):445–454, 2002.

[Sen, 1986] A. Sen. Social choice theory. In K. J. Arrow and M. D. Intriligator, editors, *Handbook of Mathematical economics*, volume III, pages 1073–1181. Elsevier Science B. V., 1986.