

## PAPER

## Knowledge Integration by Probabilistic Argumentation

Saung Hnin Pwint OO<sup>†a)</sup>, Member, Nguyen Duy HUNG<sup>†b)</sup>, and Thanaruk THEERAMUNKONG<sup>†c)</sup>, Nonmembers

**SUMMARY** While existing inference engines solved real world problems using probabilistic knowledge representation, one challenging task is to efficiently utilize the representation under a situation of uncertainty during conflict resolution. This paper presents a new approach to straightforwardly combine a rule-based system (RB) with a probabilistic graphical inference framework, i.e., naïve Bayesian network (BN), towards probabilistic argumentation via a so-called probabilistic assumption-based argumentation (PABA) framework. A rule-based system (RB) formalizes its rules into defeasible logic under the assumption-based argumentation (ABA) framework while the Bayesian network (BN) provides probabilistic reasoning. By knowledge integration, while the former provides a solid testbed for inference, the latter helps the former to solve persistent conflicts by setting an acceptance threshold. By experiments, effectiveness of this approach on conflict resolution is shown via an example of liver disorder diagnosis.

**key words:** knowledge integration, probabilistic argumentation, probabilistic graphical models and rules

## 1. Introduction

Knowledge integration (KI) is a process of incorporating new information into a body of existing knowledge [1] when a source is insufficient for making decision itself. In multi-agent systems, agents integrate and share their knowledge to perform a common task [2]–[4]. In ontology integration, several ontologies are integrated for handling semantic heterogeneity and knowledge variation [5]–[8]. In BNs integration [9]–[12], a global BN is constructed by combining local BNs. However, conflict is a common problem when knowledge from multiple sources is combined. For example, a source has knowledge about that  $A$  causes  $B$  while another source has knowledge about that  $B$  causes  $A$ . Unfortunately, conflicts make inconsistent decisions in multi-agent systems and inconsistent semantic information in ontology-based systems, triggering some cycles in BNs.

As a conflict resolution, conflict is assumed as a normalization factor of ignorance in Dempster's rule of combination [13] for merging evidences from independent sources. Unfortunately, Dempster's rule of combination cannot solve the problem of total conflict evidences since it

causes the combination result to be not applicable (N/A). So, [14] extends the Dempster's rule of combination by measuring conflict between two pieces of evidence as a coefficient. Moreover, several approaches [15]–[18] have been explored to investigate combination methods for integrating conflict evidences based on Dempster-Shafer theory (DST). However, the combination results may vary according to their baseline methods. Meanwhile, argumentation is a common sense of reasoning formalism and solves conflict by attacking arguments. It is used in several areas including decision making, knowledge integration, ontology, and multi-agents, for providing accurate final decisions. There are several approaches [19]–[23], which use argumentation to solve the conflict while integrating knowledge. These approaches usually use single representation, such as either DST statements or argumentation statements. However, it is more efficient if we combine straightforwardly different notations, each of which has different strengths in its representation, resulting in improving total process performance. In fact, for combining different representations, it is difficult to detect whether the knowledge of these representations overlap with each other or not due to their different vocabularies in knowledge representation. For example, BN represents knowledge in the form of a directed acyclic graph while RB represents knowledge in the form of rules.

This paper presents a study on integration of two powerful but different presentations with augmentation-based conflict resolution, that is combination of rule-based (RB) and Bayesian network (BN) via probabilistic argumentation. In our approach, defeasible reasoning can be incorporated into the RB system using the ABA framework where any conflict can be solved using the ABA's semantic. Complementarily, the probabilistic inference can be achieved by querying the propositions based on evidences (observations) inside the BN. As one solution, it is possible to set up a rule iff probability of the proposition computed by the BN is greater than 0.5, then the proposition is entailed by the framework. Otherwise, they conflict. To enable the combination of RB and BN, we propose to translate them into a common representation, say the probabilistic assumption-based argumentation (PABA) framework. So, we use the PABA framework as an identical representation since it captures both logical and probabilistic aspects. In the conflict resolution, we formalize them into their respective PABA frameworks and provide a combination way for integrating them into an integrated framework. Then, we infer the proposition into this integrated PABA framework

Manuscript received October 7, 2019.

Manuscript revised February 20, 2020.

Manuscript publicized May 1, 2020.

<sup>†</sup>The authors are with the School of Information, Computer, and Communication Technology, Sirindhorn International Institute of Technology, Thammasat University, Pathum Thani, 12000, Thailand.

a) E-mail: hnin.pwint172004@gmail.com

b) E-mail: hung@siit.tu.ac.th

c) E-mail: thanaruk@siit.tu.ac.th

DOI: 10.1587/transinf.2019EDP7270

for computing the acceptable probability of the proposition. We show that our approach provides more reliable and better explainable results than each single individual representation. As an application area, we apply production rules of a rule-based approach that is implemented by LUCAS [24], [25] and a naïve BN that is generated from 100 patient-records [26]. They are concerned about diagnosis of liver disorders and biliary tracts. We evaluate our approach by inferring common diagnosis such as steatosis hepatitis and primary biliary cirrhosis.

Some related works are shown in Sect. 2 while some theoretical backgrounds are provided in Sect. 3. We describe overview of our approach with its applications, an RB and a naïve BN, in Sect. 4. How to restructure the RB into the ABA framework is described in Sect. 5. In Sect. 6, we show how to detect conflicts between the ABA framework and the naïve BN. In Sect. 7, we resolve the conflict by introducing the PABA framework, translating both RB and BN into this framework, and then conducting threshold-based inference. The experimental results are provided in Sect. 8. Some conclusions and future directions are given in Sect. 9.

## 2. Related Work

So far knowledge integration have been studied in several works including those related to multi-agent systems [3], [4], [27], and Bayesian inference systems [9]–[12], [19]. In [9], a method that handles an ancestral ordering shared by individual BNs, was developed to combine these BNs. The method removes some nodes from two original networks and combines these networks. In [10], a four-step algorithm was proposed to systematically combine the qualitative and the quantitative parts of the different probabilistic graphical models; Bayesian networks (BNs) and influence diagrams (IDs), for heart disease diagnosis. It reduces complexity of structural and quantitative (i.e., conditional distributions) combinations by removing some original dependencies between nodes. In [11], a method to recover the global structure from multiple local Bayesian networks without losing any domain information, was proposed. In [12], the authors presented the probabilistic models as Bayesian Knowledge Bases (BKBs) and proposed an algorithm called Bayesian knowledge fusion that allows easy aggregation and de-aggregation of information from multiple expert sources and facilitates multi-expert decision making by providing a framework in which all opinions can be preserved and reasoned over. In [28], an approach to translate logical knowledge into Bayesian networks was illustrated. In this approach, network composition is applied to build a uniform representation that supports both logical and probabilistic reasoning. In [29], an approach to fuse multiple knowledge sources in geo-spatial decision support systems, was proposed as probabilistic logic. In this work, the probability of a sentence (i.e., a query) was computed when a set of sentences is provided as evidences. Then the probabilistic logic is introduced to represent uncertainties in logical sentences and to reduce the computational complex-

ity of coping with the semantics of the sentence. If multiple sources have conflicts in their knowledge, then it is difficult to be integrated. In propositional bases, conjunctive and disjunctive mergings are two ways to combine these bases depending on whether the bases are conflicting or not [30]. When the bases are not conflicting, the former is preferred to the latter, otherwise, the latter to the former.

Moreover, several approaches [15]–[18] have been explored to investigate hybrid methods for integrating conflict evidences based on Dempster-Shafer theory (DST). In [15], the authors applied Dempster-Shafer Theory (DST) to address the problems of combining information in recommender systems (RSs) such as non-sense combination problem, totally-conflicting combination problem, incomparable rating problem in RS. They analyzed six combination of methods, which are based on DST, and proposed two hybrid combination methods to solve existing conflict in massive functions. The first hybrid method combines the methods of Dubois and Prade's Rule of Combination and the method of Dempster's Rule of Combination. The other method combines Dempster's Rule of Combination and finds the average of rule combination. The results of the paper showed that these hybrid methods outperformed the baseline methods. In [16], the authors proposed a method based on DEMATEL for combining multiple conflicting evidences and reducing computational complexity in the combination. The relation between evidences is determined by constructing a similarity measure matrix. Then conflicting evidence was modified by calculating prominence and importance of evidence. Finally, the method computed the weighted average of evidences as a combination result based on Dempster's rule of combination.

In the past, the argument-based systems (ABSs) helps improve decision making process, particularly when the user provided enough evidences, even the decision was done in a realistic but complicate task [21]. In [19], the WebKIDSS framework was developed as an argumentation-enabled knowledge integration in a decision support system in semantic web in order to solve conflicts which make incomplete and inconsistent semantic web information. In [20], argumentation was used instead of voting to develop a hybrid intelligent system for crop classification. It integrates expert knowledge and decisions of three bases classifiers; decision tree, support vector machine and neural network. Both knowledge and decisions are expressed as rules with defeasible logic programming in order to solve conflicts for providing accurate prediction in crop classification. When ontologies are incomplete and inconsistent, a decision support framework for ontology integration was developed in [22] using argumentation. It determines the membership status of individuals to concepts and provides a global understanding of the knowledge. In [31], preference-based argumentation framework was proposed to solve conflicts by defining merging operators. In [32], a set of argumentation frameworks were combined by merging operators which generate a set of corresponding frameworks rather than a single one. The framework, called Pooling Infor-

mation from Several Agents (for short, PISA), was developed in [23] to enable classification using argumentation. Agents are participated in a dialectical process arguing for a given example to be classified based on their experience. In [33], a scheme of probabilistic argumentation was developed by combining logic and probability in argumentation. Their focus is an annotating logical formulas (rules) of a knowledge base with probability. So far, even the existing methods could solve real world problems using probabilistic knowledge representation, but there have been no methods that efficiently solve conflicts by integrating the representation under a situation of uncertainty. In this paper, we aim to invent a method to solve conflicts by integrating multiple types of knowledge representation under uncertainty.

### 3. Theoretical Backgrounds

This section provides the backgrounds of rule-based system (RB), Bayesian network (BN), and some argumentation frameworks.

#### 3.1 A Rule-Based System

A rule-based system (RB) contains a set of production rules, having a set of conditions and a set of actions by using predicates<sup>†</sup>, functions<sup>††</sup> and their negations. A predicate is an expression of the form either  $t(p_1, \dots, p_{n-1}, v_1)$  or  $t(p_1, \dots, p_{n-1}, [v_1, \dots, v_m])$  where  $t$  is a predicate-name, each  $p_i$  is a parameter and each  $v_i$  is a value. A function is an expression of the form of either  $f(q_1, \dots, q_n) \circ y_1$  or  $f(q_1, \dots, q_n) \circ [y_1..y_m]$  where  $f$  is a function-name, each  $q_i$  is parameter,  $\circ$  is a mathematics operator (i.e.,  $\circ \in \{=, <, \leq, >, \geq\}$ ) and each  $y_j$  is a value. The symbol NOT, which is used in front of a predicate/function, is referred to the negation. The symbol AND (resp. OR) is used as conjunctive (resp. disjunctive) connectors between predicates and functions. A production rule is an expression of the following form:

R: $x$  IF *conditions* THEN *actions* FI

where  $x$  is a rule number. A set of production rules having common conditions are encoded in the following module where each  $a_{st}$  and  $b_{hq}$  are conditions (i.e., functions/predicates/negations), each  $c_i$  is an action (i.e., function/predicate/negation), and  $x$  is a rule's number.

PROVIDED ( $a_{11}$ OR ... OR $a_{1s_1}$ ) AND ... AND ( $a_{t1}$ OR ... OR $a_{ts_t}$ ) THEN
R: $x$
IF ( $b_{11}$ OR ... OR $b_{1h_1}$ ) AND ... AND ( $b_{q1}$ OR ... OR $b_{qh_q}$ ) THEN
$c_1$ AND ... AND $c_m$
FI
...
END;

<sup>†</sup>Predicate returns multiple values at the same time.

<sup>††</sup>Function returns one value at a time.

#### 3.2 A Bayesian Network

**Definition 1:** A Bayesian network is a pair,  $\mathcal{N} = (\mathcal{G}, \mathcal{T})$  where  $\mathcal{G} = (V, E)$  is a directed acyclic graph having a set of nodes;  $V$ , and a set of edges;  $E$ , and  $\mathcal{T}$  is a set of conditional probabilities for all nodes.

The network is a probabilistic graphical model representing uncertainties. A node (random variable)  $V_i$  has a set of exclusive or mutual exclusive values;  $v_1, \dots, v_k$  where  $k > 0$ . The edge from nodes  $V_i$  to  $V_j$  represents the conditional dependency between them and  $V_i$  is a parent of  $V_j$ . Let  $Par(V_j = v_j)$  be a set of parents of  $V_j$  having a value of  $v_j$ . The joint probability distribution of nodes can be computed by the following equation.

$$P(V_1 = v_1, \dots, V_n = v_n) = \prod_{i=1}^n P(V_i = v_i | Par(V_i = v_i)) \quad (1)$$

#### 3.3 The Abstract Argumentation Framework

**Definition 2:** The abstract argumentation framework [34] is a pair,  $\mathcal{AF} = (Ar, Att)$  where  $Ar$  is a set of arguments and  $Att \subseteq Ar \times Ar$  is a set of attacks<sup>†††</sup> between arguments.

A set of arguments,  $S \subseteq Ar$ , is **conflict-free** if it doesn't attack itself. Argument  $A$  is **acceptable** w.r.t  $S$  if any argument attacking the argument  $A$  is attacked by arguments in  $S$ . Conflict-free set  $S$  is **admissible** if each argument in  $S$  is acceptable w.r.t  $S$ .  $S$  is a **complete extension** if it is admissible and every argument in  $S$  is acceptable w.r.t  $S$ .  $S$  is a **preferred extension** if it is admissible and a maximal (w.r.t set inclusion) complete extension.  $S$  is an **ideal extension** if it is admissible and contained in every preferred extension. Let  $f(S)$  be a characteristic function that is denoted as  $f(S) = \{A \in Ar \mid A \text{ is acceptable w.r.t } S\}$ .  $S$  is a **grounded extension** iff it is admissible and the least fix-point of the characteristic function  $f(S)$ . An argument  $A$  is **credulously/grounded/ideally** accepted w.r.t  $\mathcal{AF}$  (i.e.,  $\mathcal{AF} \vdash_x A$  where  $x \in \{cr, gr, id\}$ ) if it is contained in a preferred/grounded/ideal extension of  $\mathcal{AF}$  respectively. It is skeptically accepted by  $\mathcal{AF}$ , (i.e.,  $\mathcal{AF} \vdash_{sk} A$ ), if it belongs to all extensions of  $\mathcal{AF}$ .

#### 3.4 The Assumption-Based Argumentation Framework

Let  $\mathcal{L}$  be a language containing a non-empty set of literals: positive, negative (i.e., classical negation) and negation as failure. Underlying language  $\mathcal{L}$ , the assumption-based argumentation (ABA) framework [35] is an extension of Dung's abstract argumentation framework [34]. The ABA framework identifies sentences by means of inference rules supported by assumptions.

**Definition 3:** The ABA framework is a triple,  $\mathcal{F} = (\mathcal{A}, \mathcal{R}, \neg)$  where  $\mathcal{A}$  is a set of assumptions,  $\mathcal{R}$  is a set of

<sup>†††</sup> $(A, B)$  is denoted that argument  $A$  attacks argument  $B$ .

inference rules and  $\neg^\dagger$  is a total mapping from each assumption to its contrary.

In ABA, the assumption never appears in the head of inference rule  $r \in \mathcal{R}$  of the form  $\alpha_0 \leftarrow \alpha_1, \dots, \alpha_n$  where  $n \geq 0$ . The inference rule is classified into defeasible rule and strict rule. The defeasible rule contains at least one assumption in the body of the rule while the strict rule does not contain assumption in the body of the rule. The strict rule without body of the rule of the form  $\alpha \leftarrow$  is called a fact. An inference rule  $r \in \mathcal{R}$  (resp. a subset of the set of inference rules ( $R \subseteq \mathcal{R}$ )) can generate an argument concluding a proposition  $\pi$ , which is the head of the inference rule  $r$  (resp.  $r' \in R$ ). The argument, which concludes the proposition  $\pi \in \mathcal{L}$  and is supported by a set of assumptions  $Q$ , is denoted as  $(Q, \delta, \pi)$  when there is a deduction  $\delta$  from  $\pi$  to  $Q$ . The detail backward deduction is referred to [35]. Reasonably,  $(Q, \delta, \pi)$  is said to be  $(Q, \pi)$ . An argument  $(Q, \pi)$  attacks argument  $(Q', \pi')$  by attacking some assumptions in  $Q'$ .  $\pi$  is credulously/grounded/ideally accepted w.r.t ABA  $\mathcal{F}$  (i.e.,  $\mathcal{F} \vdash_{x \in \{cr, gr, id\}} \pi$ ) if the argument that concludes  $\pi$  is contained in a credulous/grounded/ideal extension respectively. Also, it is skeptically accepted by  $\mathcal{F}$ , (i.e.,  $\mathcal{F} \vdash_{sk} \pi$ ), if the argument belongs to all extensions of  $\mathcal{F}$ .

### 3.5 The Probabilistic Assumption-Based Argumentation Framework

The Probabilistic assumption-based argumentation (PABA) framework [36], an extension of the ABA framework, is a class of probabilistic argumentation.

**Definition 4:** The PABA framework is a triple,  $\mathcal{P} = (\mathcal{A}_p, \mathcal{R}_p, \mathcal{F})$  where  $\mathcal{A}_p$  is a set of probabilistic assumptions,  $\mathcal{R}_p$  is a set of probabilistic rules and  $\mathcal{F}$  is an ABA framework.

The PABA framework contains probabilistic part consisting of  $\mathcal{A}_p$  and  $\mathcal{R}_p$  and the logical part consisting of ABA  $\mathcal{F}$ . A possible world  $\omega$  of PABA  $\mathcal{P}$  is a maximal consistent subset of  $\mathcal{A}_p \cup \neg\mathcal{A}_p$ . A probabilistic rule in  $\mathcal{R}_p$  is written as  $[\alpha_0 : x] \leftarrow \alpha_1, \dots, \alpha_n$  (i.e.,  $n \geq 0$ ) where  $\alpha_0$  is a probabilistic assumption and  $x$  (i.e.,  $0 \leq x \leq 1$ ) is the probability of  $\alpha_0$ . Its compulsory rule is written as  $[\neg\alpha_0 : 1 - x] \leftarrow \alpha_1, \dots, \alpha_n$ . If the probabilistic rules  $r_p : [\alpha_0 : x] \leftarrow \alpha_1, \dots, \alpha_n$  and  $r'_p : [\alpha_0 : y] \leftarrow \alpha'_1, \dots, \alpha'_m$  (i.e.,  $x \neq y$ ) occur in  $\mathcal{R}_p$  then  $body(r_p) \subseteq body(r'_p)$  or  $body(r'_p) \subseteq body(r_p)$ . If an argument concludes probabilistic assumption  $\alpha$  from a set of assumptions  $Q$ , it is called a probabilistic argument. Otherwise, it is a non-probabilistic argument. Let  $Q$  and  $Q'$  be two sets of assumptions (i.e.,  $Q \neq Q'$ ). We can say that an argument  $(Q, \alpha)$  attacks another argument  $(Q', \alpha')$  satisfying one of three conditions;

1. if  $(Q, \alpha)$  is a non-probabilistic argument and  $\alpha$  is a contrary of assumptions in  $Q'$
2. if both arguments are probabilistic arguments having

$$Q' \subset Q$$

3. if argument  $(Q, \alpha)$  concludes probabilistic assumption  $\alpha$  and argument  $(Q', [\alpha : x])$  concludes compulsory of the probabilistic assumption  $\alpha'$  with its probability.

Let PABA  $\mathcal{P}$  be a probabilistic acyclic in its dependency graph and  $\mathcal{W}$  be a set of all possible worlds of  $\mathcal{P}$ .  $\mathcal{F}_\omega = (\mathcal{A}, \mathcal{R} \cup \mathcal{R}_p \cup \{\alpha \leftarrow \mid \alpha \in \omega\}, \neg)$  is defined as an ABA framework for each possible world  $\omega$  of PABA  $\mathcal{P}$ . ABA  $\mathcal{F}_\omega$  is also called the PABA  $\mathcal{P}_\omega$  condition to  $\omega$ . The probability of an argument, which concludes a proposition  $\pi$  being acceptable w.r.t  $\mathcal{F}_\omega$ , is computed as follows.

$$Prob(\pi) = \sum_{\omega \in \mathcal{W}: ABA\mathcal{F}_\omega \vdash \pi} P(\omega) \quad (2)$$

According to [37], [38], Bayesian PABA framework is a subclass of the PABA framework when PABA  $\mathcal{P}$  subsumes a Bayesian network  $\mathcal{N}$  instead of  $\mathcal{R}_p$  where  $\mathcal{N}$  is developed using a probabilistic logic programming; especially Problog<sup>††</sup>. Therefore, the probability of possible world  $\omega$ ;  $P(\omega)$ , can be computed using the Eq. (1).

**Definition 5:** A Bayesian PABA is a triple,  $\mathcal{P} = (\mathcal{A}_p, \mathcal{N}, \mathcal{F})$  where each  $\alpha \in \mathcal{A}_p$  is a probabilistic assumption that corresponds to a node of Bayesian network  $\mathcal{N}$ ,  $\mathcal{N}$  is subsumed instead of probabilistic rules and  $\mathcal{F} = (\mathcal{A}, \mathcal{R}, \neg)$  is an ABA framework iff it satisfies that  $\alpha$  is not an assumption in  $\mathcal{F}$  and doesn't occur in the head of any inference rule in  $\mathcal{F}$ .

## 4. Overview of Our Approach

This section starts with the application areas of our approach (i.e., an RB and a BN), which are concerned about diagnosis of liver disorders and biliary tracts. Then, this section explains the overview of our approach.

### 4.1 HEPAR-RB

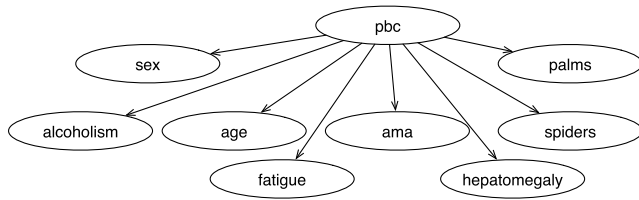
HEPAR-RB is a rule-based system, which is developed by LUCAS [24], [25] regarding the diagnosis of liver disorders. HEPAR-RB has 178 production rules encoding in rule modules. We apply these production rules in our approach.

**Example 1:** Provided that the duration of a patient who suffers from liver disorder is chronic, if the patient abuses alcohol, then diagnoses are alcoholic hepatitis, alcoholic cirrhosis and steatosis hepatitis. If the patient is female, she is over 40, she does not abuse alcohol and she feels either fatigue or generalized pruritus, then the diagnosis is primary biliary cirrhosis. These statements are written in two production rules in a rule module as below.

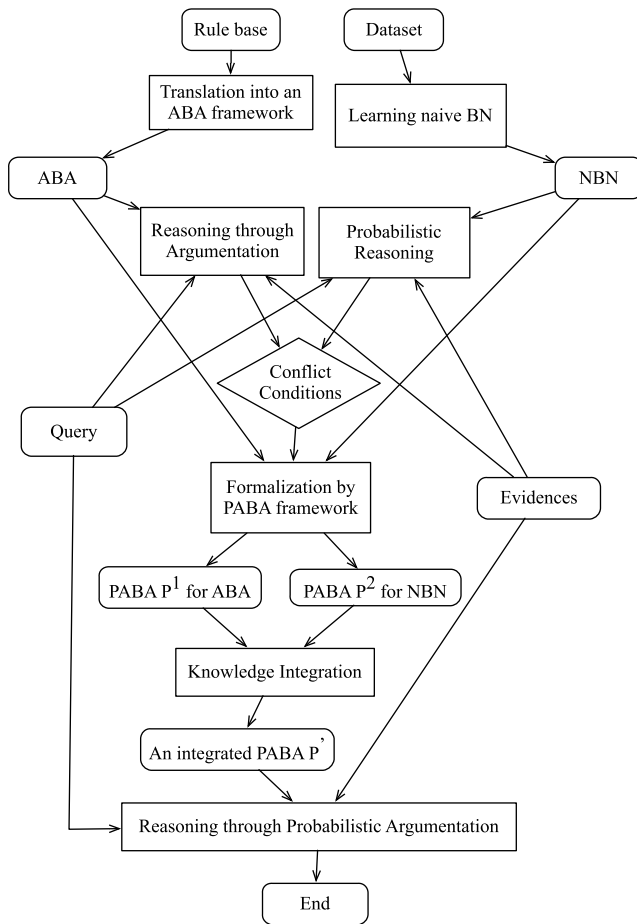
<sup>†</sup> $\bar{x}$  is the contrary of the assumption  $x$  (i.e.,  $\bar{x} = \neg x$ ).

<sup>††</sup><https://dtai.cs.kuleuven.be/problog>





**Fig. 1** The segment of HEPAR-NBN



**Fig. 2** The overview of our approach

PROVIDED duration(patient, complab, chronic) THEN R:3680 IF biochemicalsigns(patient) = "alcohol-abuse" THEN diagnosis(patient, "alcoholic-hepatitis") AND diagnosis(patient, "alcoholic-cirrhosis") AND diagnosis(patient, "steatosis-hepatitis") FI
R:3850 IF sex(patient)=female AND age(patient)>40 AND disease.history(patient, "alcohol- abuse") AND complaint(patient, [fatigue, "generalized-pruritus" ]) THEN diagnosis(patient, "primary-biliary- cirrhosis") FI
... END;

## 4.2 HEPAR Naïve Bayesian Network

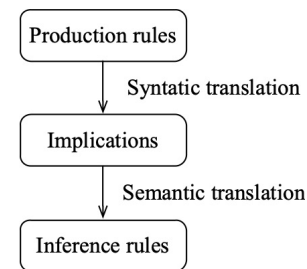
A naïve BN is generated from 100 patients' records using Genie [39] to original Hepar BN developed by [26] and name it as HEPAR-NBN. The Fig. 1 shows the segment of HEPAR-NBN where pbc is referred to primary biliary cirrhosis to be classified.

## 4.3 System Architecture

Our approach takes HEPAR-RB and a dataset, which has 100 patient records, as inputs. The HEPAR-RB is translated into an ABA framework that is named as HEPAR-ABA. Moreover, the dataset is learned to create a naïve BN (i.e., HEPAR-NBN). Given a set of evidences, a proposition is queried by HEPAR-ABA and HEPAR-NBN through argumentation and probabilistic reasoning, respectively. HEPAR-ABA and HEPAR-NBN conflict each other if HEPAR-ABA derives the proposition and HEPAR-NBN provides the probability of the proposition as less than or equal to 0.5. Such conflict is solved by formalizing and integrating HEPAR-ABA and HEPAR-NBN into the integrated one using the PABA framework. Then, the proposition is also evaluated by the integrated PABA framework that calculates the acceptable probability of the proposition. The overview of our approach is shown in Fig. 2.

## 5. Structuring the RB by the ABA Framework

This section includes translation from an RB to an ABA framework for solving conflicts by ABA's semantic. There are two steps in translations: syntactic translation and semantic translation. In the first step, all production rules of the RB is transformed into logical implications. In the second step, these logical implications are transformed into inference rules<sup>†</sup> of the ABA framework by adding assumptions into the bodies of the rules, which become defeasible rules. Figure 3 illustrates the translation steps to structure the RB by the ABA framework. For example, an RB contains an production rule of the form IF b THEN a. In the first step, the production rule is transformed into logical implication of the form  $a \leftarrow b$ . In the second step, this implication



**Fig. 3** Translation steps

<sup>†</sup>The inference rule of the ABA framework is classified into strict rules and defeasible rules.

**Table 1** Translating a RB's syntax to a logical syntax

RB's syntax	Logical syntax	
	<i>head(r)</i>	<i>body(r)</i>
$x$	$x$	
$x_1 \text{ OR } \dots \text{ OR } x_n$	$x_1 \vee \dots \vee x_n$	
$x_1 \text{ AND } \dots \text{ AND } x_n$	$x_1, \dots, x_n$	
NOT $x$	$\neg x$	$\sim x$

**Table 2** Translating a production rule to an inference rule

Production rule	Inference rule
IF $x$ THEN $y$ FI	$y \leftarrow x$
IF $x_1^1 \text{ OR } \dots \text{ OR } x_m^1 \text{ AND } \dots \text{ AND } x_1^n \text{ OR } \dots \text{ OR } x_k^n$ THEN $y_1 \text{ AND } \dots \text{ AND } y_t$ FI	$y_i \leftarrow \bigvee X_1, \dots, \bigvee X_n$ ( $n \geq 0$ and $1 \leq i \leq t$ )
PROVIDED $z$ THEN IF $x$ THEN $y$ FI END;	$y \leftarrow z, x$

where  $X = \{x_1, \dots, x_m\}$  ( $m \geq 0$ ) is a set and  $\bigvee X = x_1 \vee \dots \vee x_m$  is disjunctive between elements of  $X$ .

**Table 3** Translating RB's predicates to literals

Predicates of RB	Literals of logical form
$f(p_1, \dots, p_{n-1}, v)$	$f(p_1, \dots, p_{n-1}, v)$
$f(p_1, \dots, p_{n-1}, [v_1, \dots, v_m])$	$f(p_1, \dots, p_{n-1}, v_1) \vee \dots \vee f(p_1, \dots, p_{n-1}, v_m)$

**Table 4** Translating RB's functions to literals

Functions of RB	Literals of logical form
$f(p_1, \dots, p_n) = v$	$f(p_1, \dots, p_n, v)$
$f(p_1, \dots, p_n) \circ v$	$f(p_1, \dots, p_n, X), X \circ v$
$f(p_1, \dots, p_n) \text{ in } [v_1, \dots, v_m]$	$f(p_1, \dots, p_n, X), X \geq v_1, X \leq v_m$

where  $\circ \in \{<, >, \leq, \geq\}$  and *in* is an operator that checks whether a value is in a range.

is converted into a defeasible rule of the form  $a \leftarrow b, \gamma_a$  where  $\gamma_a$  is an assumption.

### 5.1 Syntactic Translation

Underlying language  $\mathcal{L}$ , the syntax of the RB is transformed into logical syntax as shown in Table 1. Every production rule of the RB is transformed into an inference rule as shown in Table 2. Predicates and functions of the RB are translated into literals as shown in Table 3 and Table 4 respectively.

**Example 2:** Let *du* be “duration”, *pa* be “patient”, *cl* be “complab”, *cr* be “chronic”, *bsigns* be “biochemicalsigns”, *diag* be “diagnosis”, *aabuse* be “alcohol-abuse”, *acirr* be “alcoholic-cirrhosis”, *steatosis* be “steatosis-hepatitis”, *ahepat* be “alcoholic-hepatitis”, *sex* be “sex”, *female* be “female”, *age* be “age”, *dise* be “disease-history”, *comp* be “complaint”, *fatigue* be “fatigue”, *gpru* be “generalized-pruritis” and *pbc* be “primary-biliary-cirrhosis”. The production rules from Example 1 is translated into the following logical implications.

<i>diag</i> ( <i>pa</i> , <i>steatosis</i> )	$\leftarrow$	<i>du</i> ( <i>pa</i> , <i>cl</i> , <i>cr</i> ), <i>bsigns</i> ( <i>pa</i> , <i>aabuse</i> )
<i>diag</i> ( <i>pa</i> , <i>ahepat</i> )	$\leftarrow$	<i>du</i> ( <i>pa</i> , <i>cl</i> , <i>cr</i> ), <i>bsigns</i> ( <i>pa</i> , <i>aabuse</i> )
<i>diag</i> ( <i>pa</i> , <i>acirr</i> )	$\leftarrow$	<i>du</i> ( <i>pa</i> , <i>cl</i> , <i>cr</i> ), <i>bsigns</i> ( <i>pa</i> , <i>aabuse</i> )
<i>diag</i> ( <i>pa</i> , <i>pbc</i> )	$\leftarrow$	<i>du</i> ( <i>pa</i> , <i>cl</i> , <i>cr</i> ), <i>sex</i> ( <i>pa</i> , <i>female</i> ), <i>age</i> ( <i>pa</i> , <i>X</i> ), $X > 40$ , $\sim \text{dise}(\text{pa}, \text{aabuse})$ , ( <i>comp</i> ( <i>pa</i> , <i>fatigue</i> ) $\vee$ <i>comp</i> ( <i>pa</i> , <i>gpru</i> ))

### 5.2 Semantic Translation

Underlying language  $\mathcal{L}$ , the RB is translated into an ABA framework,  $\mathcal{F} = (\mathcal{A}, \mathcal{R}, \neg)$ , for solving conflicts by ABA's semantic. In ABA  $\mathcal{F}$ ,  $\mathcal{A} = \{\gamma_y \mid \gamma_y \text{ is a positive literal for rebutting an argument that concludes } y\} \cup \{\varepsilon \mid \varepsilon \text{ is a positive atom for unknown value}\}$  is a set of assumptions (i.e.,  $\mathcal{A} \subset \mathcal{L}$ ). For each assumption  $\alpha \in \mathcal{A}$ ,  $\neg\alpha$  is a contrary of assumption  $\alpha$  (i.e.,  $\bar{\alpha} = \neg\alpha$ ).  $\mathcal{R}$  consists of three types of inference rules; defeasible, strict and fact. Let  $R_D, R_S$  and  $R_O$  be a set of defeasible rules, a set of strict rules and a set of evidences (i.e., observed facts) respectively.

For each logical implication of the form  $y \leftarrow \bigvee X_1, \dots, \bigvee X_n$  that is syntactically transformed from the production rule of the RB, it is converted into the following defeasible rule by adding an assumption  $\gamma_y$  in the body of the rule. Then the rule is added into  $R_D$ .

$$y \leftarrow \bigvee X_1, \dots, \bigvee X_n, \gamma_y$$

If there is another defeasible rule  $r' \in R_D$  of the form  $y' \leftarrow \bigvee Y_1, \dots, \bigvee Y_m, \gamma_{y'}$  where  $y'$  contradicts  $y$  (i.e.,  $\neg y = y'$ ) then the following strict rule, which claims  $\neg\gamma_y$  (i.e., contrary of the assumption  $\gamma_y$ ), is added into  $R_S$ .

$$\neg\gamma_y \leftarrow \gamma_{y'}$$

For each assumption  $\varepsilon \in \mathcal{A}$  of the form  $f(p_1, \dots, \text{unknown})$  for unknown value, the following strict rule, which claims  $\neg f(p_1, \dots, \text{unknown})$  (i.e., contrary of assumption  $f(p_1, \dots, \text{unknown})$ ), is added into  $R_S$ .

$$\neg f(p_1, \dots, \text{unknown}) \leftarrow \bigvee_1^m f_i(p_1, \dots, v_i) \quad (1 \leq i \leq m)$$

Let  $O = \{o_1, \dots, o_n\}$  is a set of literals for evidences (i.e., observed facts). For each  $o \in O$ , the following inference rule (i.e., a fact) is added into  $R_O$ .

$$o \leftarrow$$

**Example 3:** Suppose that HEPAR-ABA  $\mathcal{F} = (\mathcal{A}, \mathcal{R}, \neg)$  is an ABA framework that is transformed from HEPAR-RB.  $\mathcal{R}$  consists of the following rules from  $R_D$  and  $R_S$  and the inference rules from  $R_O$  of Example 4.

$R_D$ :		
<i>diag</i> ( <i>pa</i> , <i>steatosis</i> )	$\leftarrow$	<i>du</i> ( <i>pa</i> , <i>cl</i> , <i>cr</i> ), <i>bsigns</i> ( <i>pa</i> , <i>aabuse</i> ), $\gamma_{\text{diag}(\text{pa}, \text{steatosis})}$
<i>diag</i> ( <i>pa</i> , <i>ahepat</i> )	$\leftarrow$	<i>du</i> ( <i>pa</i> , <i>cl</i> , <i>cr</i> ), <i>bsigns</i> ( <i>pa</i> , <i>aabuse</i> ), $\gamma_{\text{diag}(\text{pa}, \text{ahepat})}$
<i>diag</i> ( <i>pa</i> , <i>acirr</i> )	$\leftarrow$	<i>du</i> ( <i>pa</i> , <i>cl</i> , <i>cr</i> ), <i>bsigns</i> ( <i>pa</i> , <i>aabuse</i> ), $\gamma_{\text{diag}(\text{pa}, \text{acirr})}$
<i>diag</i> ( <i>pa</i> , <i>pbc</i> )	$\leftarrow$	<i>du</i> ( <i>pa</i> , <i>cl</i> , <i>cr</i> ), <i>sex</i> ( <i>pa</i> , <i>female</i> ), <i>age</i> ( <i>pa</i> , <i>X</i> ), $X > 40$ , $\sim \text{dise}(\text{pa}, \text{aabuse})$ ,

		$(\text{comp}(\text{pa}, \text{fatigue}) \vee \text{comp}(\text{pa}, \text{gpru})),$
		$\gamma_{\text{diag}}(\text{pa}, \text{pbc})$
$\text{wl}(\text{pa}, \text{sig})$	$\leftarrow$	$\text{comp}(\text{pa}, \text{wl}), \text{al}(\text{pa}), \gamma_{\text{wl}}(\text{pa}, \text{sig})$
$\neg \text{wl}(\text{pa}, \text{sig})$	$\leftarrow$	$\text{comp}(\text{pa}, \text{wl}), \sim \text{al}(\text{pa}), \gamma_{\neg \text{wl}}(\text{pa}, \text{sig})$
$\text{disor}(\text{pa}, \text{begin})$	$\leftarrow$	$\text{comp}(\text{pa}, \text{col}), \text{rad}(\text{pa}, \text{unknown}),$
		$\text{comp}(\text{pa}, \text{vom}), \gamma_{\text{disor}}(\text{pa}, \text{begin})$
$R_S :$		
$\neg \gamma_{\text{wl}}(\text{pa}, \text{sig})$	$\leftarrow$	$\sim \text{wl}(\text{pa}, \text{sig})$
$\neg \gamma_{\neg \text{wl}}(\text{pa}, \text{sig})$	$\leftarrow$	$\text{wl}(\text{pa}, \text{sig})$
$\neg \text{rad}(\text{pa}, \text{unknown})$	$\leftarrow$	$\text{rad}(\text{pa}, \text{back}) \vee \text{rad}(\text{pa}, \text{araq})$

$\mathcal{A}$  is the set of assumptions having  $\{\gamma_{\text{diag}}(\text{pa}, \text{steatosis}), \gamma_{\text{diag}}(\text{pa}, \text{ahepat}), \gamma_{\text{diag}}(\text{pa}, \text{acirr}), \gamma_{\text{diag}}(\text{pa}, \text{pbc}), \gamma_{\text{wl}}(\text{pa}, \text{sig}), \gamma_{\neg \text{wl}}(\text{pa}, \text{sig}), \gamma_{\text{disor}}(\text{pa}, \text{begin}), \text{rad}(\text{pa}, \text{unknown})\}$ . Their contraries are as follows.

$$\begin{aligned} \overline{\gamma_{\text{diag}}(\text{pa}, \text{steatosis})} &= \neg \gamma_{\text{diag}}(\text{pa}, \text{steatosis}) \\ \overline{\gamma_{\text{diag}}(\text{pa}, \text{ahepat})} &= \neg \gamma_{\text{diag}}(\text{pa}, \text{ahepat}) \\ \overline{\gamma_{\text{diag}}(\text{pa}, \text{acirr})} &= \neg \gamma_{\text{diag}}(\text{pa}, \text{acirr}) \\ \overline{\gamma_{\text{diag}}(\text{pa}, \text{pbc})} &= \neg \gamma_{\text{diag}}(\text{pa}, \text{pbc}) \\ \overline{\gamma_{\text{wl}}(\text{pa}, \text{sig})} &= \neg \gamma_{\text{wl}}(\text{pa}, \text{sig}) \\ \overline{\gamma_{\neg \text{wl}}(\text{pa}, \text{sig})} &= \neg \gamma_{\neg \text{wl}}(\text{pa}, \text{sig}) \\ \overline{\text{rad}(\text{pa}, \text{unknown})} &= \neg \text{rad}(\text{pa}, \text{unknown}). \end{aligned}$$

For reasoning through the ABA framework, let  $Q$  be a subset of  $\mathcal{A}$  that supports an argument  $A$  for concluding a proposition  $\pi$ . Argument  $A$  is written as  $(Q, \pi)$ . If no argument attacks argument  $A$  then argument  $A$  is grounded acceptable w.r.t ABA  $\mathcal{F}$  so the proposition  $\pi$  is concluded.

**Example 4:** (Continue Example 3) Suppose that a set of evidences for HEPAR-ABA  $\mathcal{F}$  is  $O_{\mathcal{F}} = \{\text{sex}(\text{pa}, \text{female}), \text{age}(\text{pa}, 45), \neg \text{dise}(\text{pa}, \text{aabuse}), \text{comp}(\text{pa}, \text{fatigue})\}$ . Then the following rules of  $R_O$  are added into  $\mathcal{R}$  as facts.

$R_O :$
$\text{sex}(\text{pa}, \text{female}) \leftarrow$
$\text{age}(\text{pa}, 45) \leftarrow$
$\neg \text{dise}(\text{pa}, \text{aabuse}) \leftarrow$
$\text{comp}(\text{pa}, \text{fatigue}) \leftarrow$

According to HEPAR-ABA  $\mathcal{F}$ , there is no argument that concludes  $\text{diag}(\text{pa}, \text{pbc})$ . So, the argument isn't entailed by HEPAR-ABA  $\mathcal{F}$  (i.e.,  $\mathcal{F} \not\models_{x \in \{cr, gr, id\}} \text{diag}(\text{pa}, \text{pbc})$ ).

## 6. Unifying the ABA $\mathcal{F}$ and the Naïve BN $\mathcal{N}$

Let  $O_{\mathcal{F}}$  be a set of evidences for the ABA  $\mathcal{F}$ ,  $O_{\mathcal{N}}$  be a set of evidences for the naïve BN  $\mathcal{N}$ , which is implemented using Problog.  $O = O_{\mathcal{F}} \cup O_{\mathcal{N}}$  contains all evidences from  $\mathcal{F}$  and  $\mathcal{N}$ . Underlying a language  $\mathcal{L}$ , vocabulary of  $\mathcal{F}$  has literals of the form  $f(p_1, \dots, p_n)$  while vocabulary of  $\mathcal{N}$  has node-value pairs of the form  $v(x)$  where  $v$  is a node of  $\mathcal{N}$  and  $x$  is a value of  $v$ . The problem is to determine correspondence between  $\mathcal{F}$  and  $\mathcal{N}$  based on specific domain. Querying a proposition  $\pi$  to  $\mathcal{F}$  and  $\mathcal{N}$ ,  $\pi$  is transformed into a literal denoted as  $\pi_{\mathcal{F}}$  and a node-value pair denoted as  $\pi_{\mathcal{N}}$ . Intuitively,  $\pi_{\mathcal{F}}$  and  $\pi_{\mathcal{N}}$  are correspondence.

### 6.1 Interface Specification

This section describes interface specification as a solution to the previous problem. We apply online terminologies (SNOMED-CT and WordNet) to determine their correspondence by mapping from node-value pairs of the naïve BN  $\mathcal{N}$  to literals of the ABA  $\mathcal{F}$ .

**Definition 6:** Interface specification is triple,  $\oplus = (F^{\oplus}, N^{\oplus}, M^{\oplus})$  where  $F^{\oplus}$  is a set of literals of an ABA  $\mathcal{F}$ ,  $N^{\oplus}$  is a set of node-value pairs of a naïve BN  $\mathcal{N}$  and  $M^{\oplus} \subseteq N^{\oplus} \times F^{\oplus}$  is a total mapping from  $N^{\oplus}$  to  $F^{\oplus}$ .

The mapping proceeds according to the following steps of method MapFromBNtoABA.

- (i) Assign an empty set into  $M^{\oplus}$ .
- (ii) Take a node-pair  $\delta \in N^{\oplus}$ .
- (iii) According to SNOMED-CT and WorldNet terminologies, determine whether literal  $\theta \in F^{\oplus}$  corresponds node-value pair  $\delta$  or not.
- (iv) If they correspond then  $(\delta, \theta)$  is added into  $M^{\oplus}$ .
- (v) Until there is no literal in  $F^{\oplus}$ , go to step (iii).
- (vi) Until there is no node-value pair in  $N^{\oplus}$ , go to step (ii).

**Example 5:** Suppose that we have an HEPAR-ABA  $\mathcal{F}$  from Example 3 and Example 4 and HEPAR-NBN  $\mathcal{N}$  as shown in Fig. 1. The total mapping from  $\mathcal{N}$  to  $\mathcal{F}$  is a set  $M^{\oplus} = \{(\text{steatosis}(\text{present}), \text{diag}(\text{pa}, \text{steatosis})), (\text{alcoholism}(\text{present}), \text{bsigns}(\text{pa}, \text{aabuse})), (\text{alcoholism}(\text{present}), \text{dise}(\text{pa}, \text{aabuse})), (\text{pbc}(\text{present}), \text{diag}(\text{pa}, \text{pbc})), (\text{age}(\text{age}31\_50), \text{age}(\text{pa}, 45)), (\text{sex}(\text{female}), \text{sex}(\text{pa}, \text{female})), (\text{fatigue}(\text{present}), \text{comp}(\text{pa}, \text{fatigue}))\}$ .

So the ABA  $\mathcal{F}$  and the naïve BN  $\mathcal{N}$  are correspondence.

### 6.2 Inferring Proposition from Evidences

When the propositions  $\pi_{\mathcal{F}}$  and  $\pi_{\mathcal{N}}$  are inferred by the ABA  $\mathcal{F}$  and the naïve BN  $\mathcal{N}$  respectively,  $\pi_{\mathcal{F}}$  and  $\pi_{\mathcal{N}}$  are correspondence according to the mapping MapFromBNtoABA. There are four conditions to check whether  $\mathcal{F}$  and  $\mathcal{N}$  agree or conflict as shown in Table 5 where  $x \in \{cr, gr, sk, id\}$  is a semantic of the ABA framework.

**Example 6:** According to Example 4, there is no argument that concludes  $\text{diag}(\text{pa}, \text{pbc})$  so HEPAR-ABA  $\mathcal{F}$  doesn't entail  $\text{diag}(\text{pa}, \text{pbc})$ . Assume that a set of evidences for HEPAR-NBN  $\mathcal{N}$  (see Fig. 1) is  $O_{\mathcal{N}} = \{\text{age}(\text{age}31\_50), \text{sex}(\text{female}), \text{alcoholism}(\text{absent}), \text{fatigue}(\text{present})\}$ . When we infer  $\text{pbc}(\text{present})$  by HEPAR-NBN  $\mathcal{N}$ , the probability of  $\text{pbc}(\text{present})$  given

**Table 5** Reasoning by the ABA  $\mathcal{F}$  and the naïve BN  $\mathcal{N}$

Argumentation	Probabilistic Reasoning	Statement
$\mathcal{F} \vdash_x \pi_{\mathcal{F}}$	$P(\pi_{\mathcal{N}}) > 0.5$	Agree
$\mathcal{F} \not\vdash_x \pi_{\mathcal{F}}$	$P(\pi_{\mathcal{N}}) \leq 0.5$	Agree
$\mathcal{F} \vdash_x \pi_{\mathcal{F}}$	$P(\pi_{\mathcal{N}}) \leq 0.5$	Conflict
$\mathcal{F} \not\vdash_x \pi_{\mathcal{F}}$	$P(\pi_{\mathcal{N}}) > 0.5$	Conflict

$O_N$  is 0.7 using Eq.(1). Therefore HEPAR-ABA  $\mathcal{F}$  and HEPAR-NBN  $\mathcal{N}$  conflict.

Suppose that we add  $\text{du}(\text{pa}, \text{cl}, \text{cr})$  as an evidence into  $O_{\mathcal{F}}$  (i.e.,  $O_{\mathcal{F}} = O_{\mathcal{F}} \cup \{\text{du}(\text{pa}, \text{cl}, \text{cr})\}$ ). Firstly, an inference rule of the form  $\text{du}(\text{pa}, \text{cl}, \text{cr}) \leftarrow$  is added into  $\mathcal{R}$ . Secondly, an argument  $(\{\gamma_{\text{diag}(\text{pa}, \text{pbc})}\}, \text{diag}(\text{pa}, \text{pbc}))$  is grounded acceptable by HEPAR-ABA  $\mathcal{F}$  (i.e.,  $\mathcal{F} \vdash_{gr} \text{diag}(\text{pa}, \text{pbc})$ ). Thus, HEPAR-ABA  $\mathcal{F}$  and HEPAR-BN  $\mathcal{N}$  agree.

## 7. Conflict Resolution

Given the set of evidences  $O = O_{\mathcal{F}} \cup O_N$ , the ABA  $\mathcal{F}$  and the naïve BN  $\mathcal{N}$  conflict iff  $\mathcal{F}$  entails (resp. doesn't entail)  $\pi_{\mathcal{F}}$  while the probability of  $\pi_N$ ;  $P(\pi_N | O_N)$ , is less than or equal to 0.5 (resp. greater than 0.5). We solve the conflicts by formalizing and integrating  $\mathcal{F}$  and  $\mathcal{N}$  using the PABA framework.

### 7.1 Structuring the ABA $\mathcal{F}$ and the Naïve BN $\mathcal{N}$ by the PABA Framework

The PABA framework includes probabilistic part (i.e., a set of probabilistic assumptions  $\mathcal{A}_p$  and a set of probabilistic rules  $\mathcal{R}_p$ ) and logical part (i.e., an ABA  $\mathcal{F}$ ). At this point, we can formalize the ABA  $\mathcal{F} = (\mathcal{R}, \mathcal{A}, \neg)$  and the naïve BN  $\mathcal{N}$  in their respective PABAs.

For the ABA  $\mathcal{F}$ , it does not contain uncertainties so there is no  $\mathcal{A}_p$  and  $\mathcal{R}_p$  (i.e.,  $\mathcal{A}_p = \emptyset$  and  $\mathcal{R}_p = \emptyset$ ). A PABA framework that structures the ABA  $\mathcal{F}$  is  $\mathcal{P}^1 = (\emptyset, \emptyset, \mathcal{F})$ .

**Example 7:** The PABA framework, which structures HEPAR-ABA  $\mathcal{F}$  from Example 4, is denoted as HEPAR-PABA  $\mathcal{P}^1$  in which its  $\mathcal{A}_p$  and  $\mathcal{R}_p$  are empties. So, the HEPAR-PABA is  $\mathcal{P}^1 = (\emptyset, \emptyset, \mathcal{F})$ .

For the naïve BN  $\mathcal{N}$ , it is concerned about uncertainties so a PABA framework that structures the naïve BN  $\mathcal{N}$  is  $\mathcal{P}^2 = (\mathcal{A}_p, \mathcal{N}, (\emptyset, \emptyset, \emptyset))$ . Here, a node of  $\mathcal{N}$  may has multiple values (i.e., mutual exclusive) so all node-value pairs<sup>†</sup> of  $\mathcal{N}^{\oplus}$  (instead of nodes) is taken as probabilistic assumptions of  $\mathcal{A}_p$ .  $\mathcal{N}^{\dagger\dagger}$  is subsumed in  $\mathcal{R}_p$ . There is no ABA framework in  $\mathcal{P}^2$  so its ABA framework is  $(\emptyset, \emptyset, \emptyset)$ . According to Definition 5,  $\mathcal{P}^2 = (\mathcal{A}_p, \mathcal{N}, (\emptyset, \emptyset, \emptyset))$  is a Bayesian PABA.

**Example 8:** Using the PABA framework, it structures HEPAR-NBN  $\mathcal{N}$  of Fig.1, and it is denoted as HEPAR-PABA  $\mathcal{P}^2$ . Let  $\text{pbc}$  be  $\text{pbc}(\text{present})$ ,  $\text{pbcab}$  be  $\text{pbc}(\text{absent})$ ,  $\text{sf}$  be  $\text{sex}(\text{female})$ ,  $\text{sm}$  be  $\text{sex}(\text{male})$ ,  $\text{alcop}$  be  $\text{alcoholism}(\text{present})$ ,  $\text{aab}$  be  $\text{alcoholism}(\text{absent})$ ,  $\text{age0}$  be  $\text{age}(\text{age0\_age30})$ ,  $\text{age31}$  be  $\text{age}(\text{age31\_age50})$ ,  $\text{age51}$  be  $\text{age}(\text{age51\_age65})$ ,  $\text{age65}$  be  $\text{age}(\text{age65\_100})$ ,  $\text{fp}$  be  $\text{fatigue}(\text{present})$ ,  $\text{fab}$  be  $\text{fatigue}(\text{absent})$ ,  $\text{amap}$  be  $\text{ama}(\text{present})$ ,  $\text{amaab}$  be  $\text{ama}(\text{absent})$ ,  $\text{spp}$  be  $\text{spiders}(\text{present})$ ,  $\text{spab}$  be  $\text{spider}(\text{absent})$ ,  $\text{pap}$  be  $\text{palms}(\text{present})$ ,  $\text{paab}$  be

<sup>†</sup>Node-value pair is written as  $v(x)$  where  $v$  is a node of  $\mathcal{N}$  and  $x$  is a value of  $v$ .

<sup>††</sup> $\mathcal{N}$  is implemented using Prolog.

$\text{palms}(\text{absent})$ ,  $\text{hepp}$  be  $\text{hepatomegaly}(\text{present})$  and  $\text{hepab}$  be  $\text{hepatomegaly}(\text{absent})$ . And, they are taken as probabilistic assumptions of  $\mathcal{A}_p$ . HEPAR-NBN  $\mathcal{N}$  is subsumed in  $\mathcal{R}_p$  but the logical part of PABA  $\mathcal{P}^2$  is empty. So, HEPAR-PABA  $\mathcal{P}^2 = (\mathcal{A}_p, \mathcal{N}, (\emptyset, \emptyset, \emptyset))$  is a Bayesian PABA framework.

### 7.2 Integrating Two PABA Frameworks

This section provides some extra notations.  $O_N = \{o_1, \dots, o_k\}$  and  $O_{\mathcal{F}} = \{o_{k+1}, \dots, o_n\}$  are sets of evidences for the naïve BN  $\mathcal{N}$  and the ABA  $\mathcal{F}$  respectively.  $\pi$  is a proposition that is converted into  $\pi_{\mathcal{F}}$  and  $\pi_N$  for inferring the ABA  $\mathcal{F}$  and the naïve BN  $\mathcal{N}$  respectively. According to the method `MapFromBNtoABA`,  $\pi_{\mathcal{F}}$  and  $\pi_N$  are correspondence and  $O_{\mathcal{F}}$  and  $O_N$  are correspondence.  $p(\pi_N, o_1, \dots, o_k)$  and  $p(o_1, \dots, o_k)$  are literals that is referred to the joint probability of  $\pi_N, o_1, \dots, o_k$  and the marginal probability of  $o_1, \dots, o_k$  respectively. For instance,  $\text{p}(\text{pbc}, \text{age31}, \text{sf}, \text{aab}, \text{fp})$  and  $\text{p}(\text{age31}, \text{sf}, \text{aab}, \text{fp})$  are literals, which is referred to joint and marginal probabilities computed by BN's Eq. (1).  $\text{probid}(X, Y)^{\dagger\dagger\dagger}$  is a literal that returns a probability  $Y$  of  $X$  being acceptable by a PABA framework.  $\text{pr}(\pi_N, o_1, \dots, o_k, Z)$  is a literal that is referred to the posterior probability of  $\pi_N$  given  $o_1, \dots, o_k$  where  $Z$  is a probability between 0 and 1.  $\text{predict}(\pi_N, o_1, \dots, o_k, X)$  is a literal denoting that  $\pi_N$  is predicted to be true (resp. false) if  $Z$  is greater than 0.5 (resp. if  $Z$  is either less than or equal to 0.5) where  $X \in \{\text{true}, \text{false}\}$ .  $\text{isHasEvidence}(o_1, \dots, o_k)$  is an atom for checking whether evidences,  $o_1, \dots, o_k$ , are included in  $O_N$ .

Now, we combine PABAs  $\mathcal{P}^1 = (\emptyset, \emptyset, \mathcal{F})$  and  $\mathcal{P}^2 = (\mathcal{A}_p, \mathcal{N}, (\emptyset, \emptyset, \emptyset))$  into  $\mathcal{P}' = (\mathcal{A}'_p, \mathcal{R}'_p, \mathcal{F}')$ , where  $\mathcal{F} = (\mathcal{R}, \mathcal{A}, \neg)$  and  $\mathcal{F}' = (\mathcal{R}', \mathcal{A}', \neg)$  are ABA frameworks.

**Definition 7:**  $\mathcal{P}' = (\mathcal{A}'_p, \mathcal{R}'_p, \mathcal{F}')$  is a PABA framework, which integrates PABAs  $\mathcal{P}^1 = (\emptyset, \emptyset, \mathcal{F})$  and  $\mathcal{P}^2 = (\mathcal{A}_p, \mathcal{N}, (\emptyset, \emptyset, \emptyset))$ .

**Example 9:** (Continue to Example 6) Given a set of evidences  $O = O_{\mathcal{F}} \cup O_N$ , HEPAR-ABA  $\mathcal{F}$  from Example 3 to Example 4 and HEPAR-NBN (see Fig. 1) conflicts in querying whether a patient suffers primary biliary cirrhosis or not because  $\mathcal{F}$  doesn't entail  $\text{diag}(\text{pa}, \text{pbc})$  but probability of  $\text{pbc}$  given a set of evidences  $O_N$  is 0.7 (see Example 6). To solve such conflict, we combine HEPAR-PABAs  $\mathcal{P}^1 = (\emptyset, \emptyset, \mathcal{F})$  from Example 7 and  $\mathcal{P}^2 = (\mathcal{A}_p, \mathcal{N}, (\emptyset, \emptyset, \emptyset))$  from Example 8 into the integrated HEPAR-PABA  $\mathcal{P}' = (\mathcal{A}'_p, \mathcal{R}'_p, \mathcal{F}')$  where  $\mathcal{F} = (\mathcal{R}, \mathcal{A}, \neg)$  and  $\mathcal{F}' = (\mathcal{R}', \mathcal{A}', \neg)$  are ABA frameworks.

$\mathcal{A}_p$  of  $\mathcal{P}^2$  is assigned into  $\mathcal{A}'_p$  of  $\mathcal{P}'$  so  $\mathcal{A}'_p$  consists of all probabilistic assumptions of  $\mathcal{A}_p$ .  $\mathcal{N}$  of  $\mathcal{P}^2$  is subsumed in place of  $\mathcal{R}'_p$  of  $\mathcal{P}'$ . And  $\mathcal{F}$  of  $\mathcal{P}^1$  is assigned into  $\mathcal{F}'$  of  $\mathcal{P}'$ . In ABA  $\mathcal{F}' = (\mathcal{R}', \mathcal{A}', \neg)$ ,  $\mathcal{R}'$  consists of all rules of  $\mathcal{R}$  of  $\mathcal{F}$

<sup>†††</sup> $\text{probid}(X, Y)$  of PRENGINE [37], [38] is used for computing an acceptable probability  $Y$  of  $X$  by the PABA framework.



and the following additional rules.

- For literal  $p(\pi_N, o_1, \dots, o_k)$ , we add the following inference rule is added into  $\mathcal{R}'$ .

$$p(\pi_N, o_1, \dots, o_k) \leftarrow \pi_N, o_1, \dots, o_k$$

- For literal  $p(o_1, \dots, o_k)$ , we add the following inference rule into  $\mathcal{R}'$ .

$$p(o_1, \dots, o_k) \leftarrow o_1, \dots, o_k$$

- For literal  $pr(\pi_N, o_1, \dots, o_k, Z)$ , we add the following inference rule into  $\mathcal{R}'$ .

$$\begin{aligned} pr(\pi_N, o_1, \dots, o_k, Z) &\leftarrow \text{probid}(p(\pi_N, o_1, \dots, o_k), X), \\ &\quad \text{probid}(p(o_1, \dots, o_k), Y), \\ &\quad Z \text{ is } X/Y \end{aligned}$$

- For each literal  $\text{predict}(\pi_N, o_1, \dots, o_k, X)$ , we add the following inference rules are added into  $\mathcal{R}'$ .

$$\begin{aligned} \text{predict}(\pi_N, o_1, \dots, o_k, X) &\leftarrow pr(\pi_N, o_1, \dots, o_k, Z), \\ &\quad Z > 0.5, X \text{ is true} \\ \text{predict}(\pi_N, o_1, \dots, o_k, X) &\leftarrow pr(\pi_N, o_1, \dots, o_k, Z), \\ &\quad Z \leq 0.5, X \text{ is false} \end{aligned}$$

- Since the naïve BN  $\mathcal{N}$  predicts either  $\pi_N$  or  $\neg\pi_N$  based on the set of evidences  $O_N$ , we create the following defeasible rules by putting two assumptions  $\gamma_{bn\_predict(\pi_N)}$  and  $\gamma_{bn\_predict(\neg\pi_N)}$  in the bodies of the rules respectively. And they are added into  $\mathcal{R}'$ .

$$\begin{aligned} bn\_predict(\pi_N) &\leftarrow \text{predict}(\pi_N, o_1, \dots, o_k, \text{true}), \\ &\quad \gamma_{bn\_predict(\pi_N)} \\ bn\_predict(\neg\pi_N) &\leftarrow \text{predict}(\pi_N, o_1, \dots, o_k, \text{false}), \\ &\quad \gamma_{bn\_predict(\neg\pi_N)} \end{aligned}$$

- Similarly, since ABA  $\mathcal{F}$  entails either  $\pi_{\mathcal{F}}$  or  $\neg\pi_{\mathcal{F}}$  based on the set of evidences  $O_{\mathcal{F}}$ , we create the following defeasible rules by putting two assumptions  $\gamma_{aba\_predict(\pi_{\mathcal{F}})}$  and  $\gamma_{aba\_predict(\neg\pi_{\mathcal{F}})}$  in the bodies of the rules respectively. And they are added into  $\mathcal{R}'$ .

$$\begin{aligned} aba\_predict(\pi_{\mathcal{F}}) &\leftarrow \pi_{\mathcal{F}}, \gamma_{aba\_predict(\pi_{\mathcal{F}})} \\ aba\_predict(\neg\pi_{\mathcal{F}}) &\leftarrow \neg\pi_{\mathcal{F}}, \gamma_{aba\_predict(\neg\pi_{\mathcal{F}})} \end{aligned}$$

- Actually, we can say that the naïve BN  $\mathcal{N}$  predicts  $\pi_N$  (resp.  $\neg\pi_N$ ) is wrong if ABA  $\mathcal{F}$  entails  $\neg\pi_{\mathcal{F}}$  (resp.  $\pi_{\mathcal{F}}$ ) and  $\mathcal{N}$  is insufficient. The following inference rules are added into  $\mathcal{R}'$ .

$$\begin{aligned} \neg\gamma_{bn\_predict(\pi_N)} &\leftarrow \neg\pi_{\mathcal{F}}, \sim bn\_sufficient \\ \neg\gamma_{bn\_predict(\neg\pi_N)} &\leftarrow \pi_{\mathcal{F}}, \sim bn\_sufficient \end{aligned}$$

- Similarly, ABA  $\mathcal{F}$  predicts  $\pi_{\mathcal{F}}$  (resp.  $\neg\pi_{\mathcal{F}}$ ) is wrong if the naïve BN  $\mathcal{N}$  predicts  $\neg\pi_N$  (resp.  $\pi_N$ ) and  $\mathcal{N}$  is sufficient. So we add the following defeasible rules into  $\mathcal{R}'$  in which we assume  $bn\_sufficient$  as an assumption.

$$\neg\gamma_{aba\_predict(\pi_{\mathcal{F}})} \leftarrow \text{predict}(\pi_N, o_1, \dots, o_k, \text{false}), \\ bn\_sufficient$$

$$\neg\gamma_{aba\_predict(\neg\pi_{\mathcal{F}})} \leftarrow \text{predict}(\pi_N, o_1, \dots, o_k, \text{true}), \\ bn\_sufficient$$

- The naïve BN  $\mathcal{N}$  is insufficient if evidences of  $O_N$  aren't related to node-value pairs of  $\mathcal{N}^{\oplus}$  or there is no evidences in  $O_N$ . Then we assume that evidences of  $O_N$  are unknown. So, we add a defeasible rule of the form  $\neg bn\_sufficient \leftarrow bn\_evidence(\text{unknown})$  into  $\mathcal{R}'$  where  $bn\_evidence(\text{unknown})$  is taken as an assumption.
- If  $O_N$  contains node-value pairs of  $\mathcal{N}^{\oplus}$  as a set of evidences then we can say that evidences of  $O_N$  are known. So we add the following inference rule into  $\mathcal{R}'$ .

$$\neg bn\_evidence(\text{unknown}) \leftarrow \text{isHasEvidence}(o_1, \dots, o_k)$$

- Since  $\pi$  is the original proposition that is transformed into  $\pi_{\mathcal{F}}$  and  $\pi_N$ , we infer  $\pi$  as a query into PABA  $\mathcal{P}'$ . We added the following four inference rules into  $\mathcal{R}'$ .

$$\begin{aligned} \pi &\leftarrow bn\_predict(\pi_N) & \neg\pi &\leftarrow bn\_predict(\neg\pi_N) \\ \pi &\leftarrow aba\_predict(\pi_{\mathcal{F}}) & \neg\pi &\leftarrow aba\_predict(\neg\pi_{\mathcal{F}}) \end{aligned}$$

$\mathcal{A}'$  consists of all assumptions of  $\mathcal{A}$  of  $\mathcal{F}$  and new assumptions;  $\gamma_{bn\_predict(\pi_N)}$ ,  $\gamma_{bn\_predict(\neg\pi_N)}$ ,  $\gamma_{aba\_predict(\pi_{\mathcal{F}})}$ ,  $\gamma_{aba\_predict(\neg\pi_{\mathcal{F}})}$ ,  $bn\_sufficient$  and  $bn\_evidence(\text{unknown})$ .  $\mathcal{F}'$  contains contraries of  $\mathcal{F}$  and additional contraries of new assumptions. The additional contraries are

$$\begin{aligned} \overline{\gamma_{bn\_predict(\pi_N)}} &= \neg\gamma_{bn\_predict(\pi_N)} \\ \overline{\gamma_{bn\_predict(\neg\pi_N)}} &= \neg\gamma_{bn\_predict(\neg\pi_N)} \\ \overline{\gamma_{aba\_predict(\pi_{\mathcal{F}})}} &= \neg\gamma_{aba\_predict(\pi_{\mathcal{F}})} \\ \overline{\gamma_{aba\_predict(\neg\pi_{\mathcal{F}})}} &= \neg\gamma_{aba\_predict(\neg\pi_{\mathcal{F}})} \\ \overline{bn\_sufficient} &= \neg bn\_sufficient \text{ and} \\ \overline{bn\_evidence(\text{unknown})} &= \neg bn\_evidence(\text{unknown}). \end{aligned}$$

**Example 10:** (Continue Example 9) In the integrated HEPAR-PABA  $\mathcal{P}' = (\mathcal{A}', \mathcal{R}', \mathcal{F}')$ ,  $\mathcal{A}'_p$  consists of all assumption of  $\mathcal{A}_p$  (see Example 8).  $\mathcal{R}'_p$  subsumes  $\mathcal{N}$  (see Fig. 1). Let  $pa\_pbc$  be a proposition to be queried by  $\mathcal{P}'$ .  $\mathcal{R}'$  contains all rules of  $\mathcal{R}$  of HEPAR-ABA  $\mathcal{F}$  (see. Example 3 and Example 4) and additional rules, which are as follows.

$p(pbc, \text{age31}, \text{sf}, \text{aab}, \text{fp})$	$\leftarrow$	$pbc, \text{age31}, \text{sf}, \text{aab}, \text{fp}$
$p(\text{age31}, \text{sf}, \text{aab}, \text{fp})$	$\leftarrow$	$\text{age31}, \text{sf}, \text{aab}, \text{fp}$
$pr(pbc, \text{age31}, \text{sf}, \text{aab}, \text{fp}, Z)$	$\leftarrow$	$\text{probid}(p(pbc, \text{age31}, \text{sf}, \text{aab}, \text{fp}), X), \text{probid}(p(\text{age31}, \text{sf}, \text{aab}, \text{fp}), Y), Z \text{ is } X/Y$
$\text{predict}(pbc, \text{age31}, \text{sf}, \text{aab}, \text{fp}, X)$	$\leftarrow$	$pr(pbc, \text{age31}, \text{sf}, \text{aab}, \text{fp}, Z), Z > 0.5, X \text{ is true}$
$\text{predict}(pbc, \text{age31}, \text{sf}, \text{aab}, \text{fp}, X)$	$\leftarrow$	$pr(pbc, \text{age31}, \text{sf}, \text{aab}, \text{fp}, Z), Z \leq 0.5, X \text{ is false}$
$bn\_predict(pbc)$	$\leftarrow$	$\text{predict}(pbc, \text{age31}, \text{sf}, \text{aab}, \text{fp}, \text{true}), \gamma_{bn\_predict(pbc)}$

$\text{bn\_predict}(\neg \text{pbc})$	$\leftarrow$	$\text{predict}(\text{pbc}, \text{age31}, \text{sf}, \text{aab}, \text{fp}, \text{false}), \gamma_{\text{bn\_predict}}(\neg \text{pbc})$
$\text{aba\_predict}(\text{diag}(\text{pa}, \text{pbc}))$	$\leftarrow$	$\text{diag}(\text{pa}, \text{pbc}), \gamma_{\text{aba\_predict}}(\text{diag}(\text{pa}, \text{pbc}))$
$\text{aba\_predict}(\neg \text{diag}(\text{pa}, \text{pbc}))$	$\leftarrow$	$\neg \text{diag}(\text{pa}, \text{pbc}), \gamma_{\text{aba\_predict}}(\neg \text{diag}(\text{pa}, \text{pbc}))$
$\neg \gamma_{\text{bn\_predict}}(\text{pbc})$	$\leftarrow$	$\neg \text{diag}(\text{pa}, \text{pbc}), \sim \text{bn\_sufficient}$
$\neg \gamma_{\text{bn\_predict}}(\neg \text{pbc})$	$\leftarrow$	$\text{diag}(\text{pa}, \text{pbc}), \sim \text{bn\_sufficient}$
$\neg \gamma_{\text{aba\_predict}}(\text{diag}(\text{pa}, \text{pbc}))$	$\leftarrow$	$\text{predict}(\text{pbc}, \text{age31}, \text{sf}, \text{aab}, \text{fp}, \text{false}), \text{bn\_sufficient}$
$\neg \gamma_{\text{aba\_predict}}(\neg \text{diag}(\text{pa}, \text{pbc}))$	$\leftarrow$	$\text{predict}(\text{pbc}, \text{age31}, \text{sf}, \text{aab}, \text{fp}, \text{true}), \text{bn\_sufficient}$
$\neg \text{bn\_sufficient}$	$\leftarrow$	$\text{bn\_evidence}(\text{unknown})$
$\neg \text{bn\_evidence}(\text{unknown})$	$\leftarrow$	$\text{isHasEvidence}(\text{age31}, \text{sf}, \text{aab}, \text{fp})$
$\text{pa\_pbc}$	$\leftarrow$	$\text{bn\_predict}(\text{pbc})$
$\neg \text{pa\_pbc}$	$\leftarrow$	$\text{bn\_predict}(\neg \text{pbc})$
$\text{pa\_pbc}$	$\leftarrow$	$\text{aba\_predict}(\text{diag}(\text{pa}, \text{pbc}))$
$\neg \text{pa\_pbc}$	$\leftarrow$	$\text{aba\_predict}(\neg \text{diag}(\text{pa}, \text{pbc}))$

$\mathcal{A}$  contains all assumptions of  $\mathcal{A}$  of HEPAR-ABA  $\mathcal{F}$  and six new assumptions, which are  $\gamma_{\text{bn\_predict}}(\text{pbc})$ ,  $\gamma_{\text{bn\_predict}}(\neg \text{pbc})$ ,  $\gamma_{\text{aba\_predict}}(\text{diag}(\text{pa}, \text{pbc}))$ ,  $\text{bn\_sufficient}$  and  $\gamma_{\text{aba\_predict}}(\neg \text{diag}(\text{pa}, \text{pbc}))$  and  $\text{bn\_evidence}(\text{unknown})$ .

$\mathcal{F}'$  contains contraries from  $\mathcal{F}$  and the following contraries of new assumptions.

$$\begin{aligned} \overline{\gamma_{\text{bn\_predict}}(\text{pbc})} &= \neg \gamma_{\text{bn\_predict}}(\text{pbc}) \\ \overline{\gamma_{\text{bn\_predict}}(\neg \text{pbc})} &= \neg \gamma_{\text{bn\_predict}}(\neg \text{pbc}) \\ \overline{\gamma_{\text{aba\_predict}}(\text{diag}(\text{pa}, \text{pbc}))} &= \neg \gamma_{\text{aba\_predict}}(\text{diag}(\text{pa}, \text{pbc})) \\ \overline{\gamma_{\text{aba\_predict}}(\neg \text{diag}(\text{pa}, \text{pbc}))} &= \neg \gamma_{\text{aba\_predict}}(\neg \text{diag}(\text{pa}, \text{pbc})) \\ \overline{\text{bn\_sufficient}} &= \neg \text{bn\_sufficient} \\ \overline{\text{bn\_evidence}(\text{unknown})} &= \neg \text{bn\_evidence}(\text{unknown}) \end{aligned}$$

Using PRENGINE [37], [38], our approach is implemented to investigate a conflict resolution when ABA  $\mathcal{F}$  and the naïve BN  $\mathcal{N}$  conflict each other. After PABAs  $\mathcal{P}^1$  for ABA  $\mathcal{F}$  and  $\mathcal{P}^2$  for the naïve BN  $\mathcal{N}$  have been combined into the integrated PABA  $\mathcal{P}'$ , the conflicts between ABA  $\mathcal{F}$  and the naïve BN  $\mathcal{N}$  are solved. Assume that  $\text{probidIntegrate}(\pi, \pi_{\mathcal{F}}, \pi_{\mathcal{N}}, o_1, \dots, o_k, R)$  is a literal that returns the probability  $R$  of proposition  $\pi$  which is entailed by the integrated PABA  $\mathcal{P}'$ .

**Example 11:** (Continue Example 10) Using PRENGINE [37], [38], proposition  $\text{pa\_pbc}$  is being acceptable by the integrated HEPAR-PABA  $\mathcal{P}'$  with acceptable probability 1. When atom  $\text{probidIntegrate}(\text{pa\_pbc}, \text{diag}(\text{pa}, \text{pbc}), [\text{pbc}], [\text{age31}, \text{sf}, \text{aab}, \text{fp}], R)$  is called, it results 0.73. So, our approach determines that the diagnosis of the patient is PBC.

According to Definition 5, the integrated PABA  $\mathcal{P}'$  is a Bayesian PABA because it subsumes the naïve BN  $\mathcal{N}$  in  $\mathcal{R}_p$  and probabilistic assumptions correspond to all nodes of  $\mathcal{N}$  as node-value pairs. And PABA  $\mathcal{P}'$  satisfies that each probabilistic assumption of  $\mathcal{A}'_p$  is not an assumption in  $\mathcal{F}'$  and doesn't occur in the head of any inference rule in  $\mathcal{F}'$ .

## 8. Experimental Result

We implement our approach, which combines HEPAR-ABA and HEPAR-NBN (i.e., inputs of our approach), regarding to the diagnosis of liver disorders and biliary tract. HEPAR-ABA is an ABA framework, which structures HEPAR-RB. HEPAR-RB is a rule-based system, which is developed by LUCAS [24], [25], having 178 rules about the diagnosis of liver disorders and biliary tracts. Meanwhile, we construct HEPAR-NBN (i.e., a naïve BN) by learning one hundred records<sup>†</sup> of liver disorders patients.

The testing dataset of our approach contains forty-two patients' records for primary biliary cirrhosis and seven patients' records for steatosis hepatitis in which two patients suffer both disorders. The dataset also includes fifty-three patients' records for other liver disorders. We evaluate our approach by comparing its performance with the performance of HEPAR-NBN and HEPAR-ABA testing on the diagnosis of primary biliary cirrhosis and steatosis hepatitis. We compute confusion matrix, recall, precision, f-measure and accuracy for each HEPAR-NBN, HEPAR-ABA and our approach. For the diagnosis of primary biliary cirrhosis, Table 6 and Table 7 show the confusion matrix and the results of recall, precision, f-measure and accuracy of HEPAR-NBN, respectively. Table 8 and Table 9 show the confusion matrix and the results of recall, precision, f-measure

**Table 6** Confusion matrix of HEPAR-NBN for primary biliary cirrhosis

		Actual		
Prediction		Present	Absent	Total
	Present	39	5	44
	Absent	3	53	56
	Total	42	58	100

**Table 7** Recall/precision/f-measure/accuracy of HEPAR-NBN for primary biliary cirrhosis

	Recall	Precision	F-measure	Accuracy
Present	92.86%	86.64%	90.70%	92%
Absent	91.38%	94.64%	92.98%	

**Table 8** Confusion matrix of HEPAR-ABA for primary biliary cirrhosis

		Actual		
Entailment		$\text{diag}(\text{pa}, \text{pbc})$	$\neg \text{diag}(\text{pa}, \text{pbc})$	Total
	$\text{diag}(\text{pa}, \text{pbc})$	35	11	46
	$\neg \text{diag}(\text{pa}, \text{pbc})$	7	47	54
	Total	42	58	100

**Table 9** Recall/precision/f-measure/accuracy of HEPAR-ABA for primary biliary cirrhosis

	Recall	Precision	F-measure	Accuracy
$\text{diag}(\text{pa}, \text{pbc})$	83.33%	76.09%	79.54%	82%
$\neg \text{diag}(\text{pa}, \text{pbc})$	81.03%	87.04%	83.93%	

<sup>†</sup>Using GENIE/SMILE software, one hundred patients' records are extracted from the original HEPAR-BN, which is developed by ONIESKO [26].

**Table 10** Confusion matrix of our approach for primary biliary cirrhosis

Entailment	Actual			
		pa_pbc	$\neg$ pa_pbc	Total
	pa_pbc	39	5	44
	$\neg$ pa_pbc	3	53	56
	Total	42	58	100

**Table 11** Recall/precision/f-measure/accuracy of our approach for primary biliary cirrhosis

	Recall	Precision	F-measure	Accuracy
pa_pbc	92.86%	86.64%	90.70%	92%
$\neg$ pa_pbc	91.38%	94.64%	92.98%	

**Table 12** Confusion matrix of HEPAR-NBN for steatosis hepatitis

		Actual		
		Present	Absent	Total
Prediction	Present	6	1	7
	Absent	1	92	93
	Total	7	93	100

**Table 13** Recall/precision/f-measure/accuracy of HEPAR-NBN for steatosis hepatitis

	Recall	Precision	F-measure	Accuracy
Present	85.71%	85.71%	85.71%	98%
Absent	98.92%	98.92%	98.92%	

**Table 14** Confusion matrix of HEPAR-ABA for steatosis hepatitis

Entailment	Actual			
		diag(pa, steat)	$\neg$ diag(pa, steat)	Total
	diag(pa, steat)	4	17	21
	$\neg$ diag(pa, steat)	3	76	79
	Total	7	93	100

**Table 15** Recall/precision/f-measure/accuracy of HEPAR-ABA for steatosis hepatitis

	Recall	Precision	F-measure	Accuracy
diag(pa, steat)	57.14%	19.05%	28.57%	80%
$\neg$ diag(pa, steat)	81.72%	96.20%	88.37%	

and accuracy of HEPAR-ABA, respectively. Table 10 and Table 11 show the confusion matrix and the results of recall, precision, f-measure and accuracy of HEPAR-ABA, respectively. For the diagnosis of steatosis hepatitis, Let  $\text{diag}(\text{pa}, \text{steat})$  be a proposition to be entailed by HEPAR-ABA and  $\text{pa\_steat}$  be a proposition to be entailed by our approach. Table 12 and Table 13 show the confusion matrix and the results of recall, precision, f-measure and accuracy of HEPAR-NBN, respectively. Table 14 and Table 15 show the confusion matrix and the results of recall, precision, f-measure and accuracy of HEPAR-ABA, respectively. Table 16 and Table 17 show the confusion matrix and the results of recall, precision, f-measure and accuracy of HEPAR-ABA, respectively.

The main objective of our approach is to combine total

**Table 16** Confusion matrix of our approach for steatosis hepatitis

Entailment	Actual			
	pa_steats	$\neg$ pa_steats	Total	
	pa_steats	6	1	7
	$\neg$ pa_steats	3	53	56
	Total	7	93	100

**Table 17** Recall/precision/f-measure/accuracy of our approach for steatosis hepatitis

	Recall	Precision	F-measure	Accuracy
pa_steats	85.71%	85.71%	85.71%	98%
$\neg$ pa_steats	98.92%	98.92%	98.92%	

conflict and different knowledge representations efficiently, especially an RB and a naïve BN. So, our approach outperforms the other methods in the combination as a conflict resolution for these representations. One of the advantages of our approach is that the performance of HEPAR-NBN can be improved when the evidences for HEPAR-NBN are insufficient to predict liver disorders. In the base case, we assume that HEPAR-NBN does not have evidence to predict a diagnosis. Then the accuracy of HEPAR-NBN decreases to 0. However, our approach predicts the liver disorder from HEPAR-ABA. So, the accuracy of our approach gets 82% and 80% for the diagnosis of primary biliary cirrhosis and steatosis hepatitis, respectively. We approve that our approach is comparative to the accuracy of the BN and more accurate than the result of the ABA framework.

## 9. Conclusion

Even there exist some frameworks on handling probability in logic programming, there are very few works on knowledge integration for probabilistic argumentation. We combine an ABA framework, which is converted from an RB, with a naïve BN using a probabilistic Argumentation, especially the PABA framework. Based on a specific domain, the ABA framework is certain, while the naïve BN covers uncertainties so we investigate an approach to integrate them. The conflict between the RB and the naïve BN is solved by the semantics of the PABA framework using PRENGINE [37], [38]. As an application area, we integrate HEPAR-RB and HEPAR-NBN, which are related to the diagnosis of liver disorders. However, our approach has two limitations. Our approach requires either a naïve BN or a dataset to generate the naïve BN. One limitation is that unreliable naïve BN or dataset may impact the performance of our approach. Another limitation is that mapping from the naïve Bayesian network to the ABA framework may vary based on the terminologies.

## Acknowledgments

We would like to offer our special thanks to Dr. Peter Lucas, Extraordinary professor of Artificial intelligence, Leiden Institute for Advanced Computer Science, Leiden University for his kindness and providing the HEPAR rule-based sys-

tem to be applied in this research.

## References

- [1] K.S. Murray, "Ki: A tool for knowledge integration," AAAI/IAAI, vol.1, pp.835–842, 1996.
- [2] M. Hernes and J. Sobieska-Karpińska, "Knowledge integration in multi-agent decision support system for financial e-services," 2016 Federated Conference on Computer Science and Information Systems (FedCSIS), pp.1283–1287, IEEE, 2016.
- [3] K.C. Lee, N. Lee, and H. Lee, "Multi-agent knowledge integration mechanism using particle swarm optimization," Technological Forecasting and Social Change, vol.79, no.3, pp.469–484, 2012.
- [4] T.H. Tran and N.T. Nguyen, "An algorithm for agent knowledge integration using conjunctive and disjunctive structures," KES International Symposium on Agent and Multi-Agent Systems: Technologies and Applications, pp.703–712, Springer, 2008.
- [5] Y.M. Chen, Y.J. Chen, C.C. Wen, H.C. Chu, et al., "Ontology-based knowledge integration for distributed product knowledge service," Proc. World Congress on Engineering and Computer Science, 2009.
- [6] N. Huang and S. Diao, "Ontology-based enterprise knowledge integration," Robotics and Computer-Integrated Manufacturing, vol.24, no.4, pp.562–571, 2008.
- [7] L. Ling, Y. Hu, X. Wang, and C. Li, "An ontology-based method for knowledge integration in a collaborative design environment," The International Journal of Advanced Manufacturing Technology, vol.34, no.9, pp.843–856, Oct. 2007.
- [8] Y.y. Wang, L.k. Song, F.l. Wang, H.s. Ji, and X.m. Dai, "Ontology-based knowledge integration of collaborative design for aviation complex products," 2011 IEEE 18Th International Conference on Industrial Engineering and Engineering Management (IE&EM), pp.1894–1898, IEEE, 2011.
- [9] G. Feng, J.-D. Zhang, and S.S. Liao, "A novel method for combining bayesian networks, theoretical analysis, and its applications," Pattern Recognition, vol.47, no.5, pp.2057–2069, 2014.
- [10] C.a. Jiang, T.Y. Leong, and P. Kim-Leng, "A framework for probabilistic graphical model combination," AMIA Annual Symposium Proceedings, p.370, American Medical Informatics Association, 2005.
- [11] W. Li, W. Liu, and K. Yue, "Recovering the global structure from multiple local bayesian networks," International Journal on Artificial Intelligence Tools, vol.17, no.06, pp.1067–1088, 2008.
- [12] E. Santos, J.T. Wilkerson, and E.E. Santos, "Fusing multiple bayesian knowledge sources," International Journal of Approximate Reasoning, vol.52, no.7, pp.935–947, 2011.
- [13] G. Shafer, "Dempster's rule of combination," International Journal of Approximate Reasoning, vol.79, pp.26–40, 2016.
- [14] Y. Zhao, R. Jia, and P. Shi, "A novel combination method for conflicting evidence based on inconsistent measurements," Information Sciences, vol.367, pp.125–142, 2016.
- [15] V.-D. Nguyen and V.-N. Huynh, "On information fusion in recommender systems based on dempster-shafer theory," 2016 IEEE 28th International Conference on Tools with Artificial Intelligence (ICTAI), pp.78–85, IEEE, 2016.
- [16] W. Zhang and Y. Deng, "Combining conflicting evidence using the dematel method," Soft computing, vol.23, no.17, pp.8207–8216, 2019.
- [17] F. Xiao, "An improved method for combining conflicting evidences based on the similarity measure and belief function entropy," International Journal of Fuzzy Systems, vol.20, no.4, pp.1256–1266, 2018.
- [18] U. Mönks, H. Dörksen, V. Lohweg, and M. Hübner, "Information fusion of conflicting input data," Sensors, vol.16, no.11, p.1798, 2016.
- [19] N.K. Janjua, F.K. Hussain, and O.K. Hussain, "Semantic information and knowledge integration through argumentative reasoning to support intelligent decision making," Information Systems Frontiers, vol.15, no.2, pp.167–192, 2013.
- [20] Ș. Conțiu and A. Groza, "Improving remote sensing crop classification by argumentation-based conflict resolution in ensemble learning," Expert Systems with Applications, vol.64, pp.269–286, 2016.
- [21] J. Introne and L. Iandoli, "Improving decision-making performance through argumentation: An argument-based decision support system to compute with evidence," Decision Support Systems, vol.64, pp.79–89, 2014.
- [22] S.A. Gómez, C.I. Chesñevar, and G.R. Simari, "Ontoarg: A decision support framework for ontology integration based on argumentation," Expert Systems with Applications, vol.40, no.5, pp.1858–1870, 2013.
- [23] M. Wardeh, F. Coenen, and T.B. Capon, "Pisa: A framework for multiagent classification using argumentation," Data & Knowledge Engineering, vol.75, pp.34–57, 2012.
- [24] P. Lucas and A. Janssens, "Development and validation of hepar, an expert system for the diagnosis of disorders of the liver and biliary tract," Medical Informatics, vol.16, no.3, pp.259–270, 1991.
- [25] P. Lucas, "Refinement of the hepar expert system: tools and techniques," Artificial intelligence in medicine, vol.6, no.2, pp.175–188, 1994.
- [26] A. Oniésko, P. Lucas, and M.J. Druzdzel, "Comparison of rule-based and bayesian network approaches in medical diagnostic systems," Conference on Artificial Intelligence in Medicine in Europe, pp.283–292, Springer, 2001.
- [27] A. Zygmunt, J. Koźlak, and L. Siwik, "Agent-based environment for knowledge integration," International Conference on Computational Science, pp.885–894, Springer, 2009.
- [28] J. Wang and M. Valtorta, "A framework for integration of logical and probabilistic knowledge," AAAI, 2011.
- [29] T.C. Henderson, R. Simmons, D. Sacharny, A. Mitiche, and X. Fan, "A probabilistic logic for multi-source heterogeneous information fusion," 2017 IEEE International Conference on Multisensor Fusion and Integration for Intelligent Systems (MFI), pp.530–535, IEEE, 2017.
- [30] S. Benferhat, D. Dubois, S. Kaci, and H. Prade, "Possibilistic merging and distance-based fusion of propositional information," Annals of Mathematics and Artificial Intelligence, vol.34, no.1–3, pp.217–252, 2002.
- [31] L. Amgoud and S. Kaci, "An argumentation framework for merging conflicting knowledge bases," International Journal of Approximate Reasoning, vol.45, no.2, pp.321–340, 2007.
- [32] J. Delobelle, A. Haret, S. Konieczny, J.G. Mailly, J. Rossit, and S. Woltran, "Merging of abstract argumentation frameworks," KR, vol.2016, pp.33–42, 2016.
- [33] F. Cerutti and M. Thimm, "Probabilistic augmentations for knowledge representation formalisms," Proceedings of the Workshop on Hybrid Reasoning and Learning (HRL '18), 2018.
- [34] P.M. Dung, "On the acceptability of arguments and its fundamental role in nonmonotonic reasoning, logic programming and n-person games," Artificial intelligence, vol.77, no.2, pp.321–357, 1995.
- [35] P.M. Dung, R.A. Kowalski, and F. Toni, "Assumption-based argumentation," in Argumentation in Artificial Intelligence, pp.199–218, Springer, 2009.
- [36] P.M. Dung and P.M. Thang, "Towards (probabilistic) argumentation for jury-based dispute resolution," COMMA, vol.216, pp.171–182, 2010.
- [37] N.D. Hung, "A probabilistic argumentation engine," ICTAI 2016: 28th International Conference on Tools with Artificial Intelligence, 2016.
- [38] N.D. Hung, "Probabilistic argumentation for decision making a toolbox and application," Journal of Intelligent Informatics and Smart Technology, vol.1, no.1, Sept. 2016.
- [39] M.J. Druzdzel, "Smile: Structural modeling, inference, and learning engine and genie: a development environment for graphical decision-theoretic models," AAAI/IAAI, pp.902–903, 1999.



**Saung Hnin Pwint Oo** received B.C.Sc, B.C.Sc (Hons), M.C.Sc in Computer Science from University of Computer Studies (Mandalay). Now, she is a Ph.D (IT) student of SIIT, Thammasat University. She is doing research in knowledge integration using Probabilistic argumentation.



**Nguyen Duy Hung** received a B.Eng. in Electronics and Telecommunication from Hanoi University of Technology and an M.Eng. in Computer Science and Ph.D. in Remote Sensing and GIS from Asian Institute of Technology. Now, he is serving as an Assistant Professor at SIIT, Thammasat University. His research interests include artificial intelligence, multi-agents, argumentation, proof procedures, dialogues and contract dispute resolution systems.



**Thanaruk Theeramunkong** received a B.Eng. in Electric and Electronics Engineering, an M.Eng. and Ph.D. in Computer Science from the Tokyo Institute of Technology in 1990, 1992 and 1995, respectively. Now, he is serving as a Professor at SIIT, Thammasat University, and Associate Fellow at the Royal Society of Thailand. His research interests include data mining, machine learning, natural language processing, and knowledge engineering.