

PROBABILISTIC REASONING BASED ON LAYERS OF KNOWLEDGE BASE

TRAN DINH QUE

Abstract. Reasoning in the interval-valued probabilistic logic depends heavily on the basic matrix of truth values of sentences in a knowledge base \mathcal{B} and a target sentence S . However, the problem of determining all such consistent truth value assignments for a set of sentences is NP-complete for propositional logic and undecidable for first-order predicate logic.

This paper first presents a method of approximate reasoning in the interval-valued probabilistic logic by basing on “layers” of a knowledge base. Then, we investigate the method of slightly decreasing the complexity of reasoning via the maximum entropy principle in a point-valued probabilistic knowledge base. Such a method is based on the reduced basic matrix constructed from sentences of the knowledge base without the target sentence.

Tóm tắt. Lập luận trong logic xác suất giá trị khoảng phụ thuộc rất nhiều vào ma trận cơ bản của các giá trị chân lý của các câu trong cơ sở tri thức \mathcal{B} và câu đích S . Tuy nhiên, bài toán xác định tất cả những phép gán giá trị chân lý phi mâu thuẫn cho một tập hợp câu là NP-đầy đủ đối với logic mệnh đề và không quyết định được đối với logic vị từ cấp 1.

Bài báo này trước hết trình bày một phương pháp lập luận xấp xỉ trong logic xác suất giá trị khoảng bằng cách dựa vào “các tầng” của cơ sở tri thức. Sau đó chúng ta sẽ xem xét một phương pháp làm giảm một chút độ phức tạp của lập luận dựa trên nguyên lý entropy tối đại trong cơ sở tri thức xác suất giá trị điểm. Phương pháp lập luận như vậy dựa trên ma trận cơ bản rút gọn được xây dựng từ các câu trong cơ sở tri thức không bao gồm câu đích.

1. INTRODUCTION

In various approaches to handling uncertain information, the paradigm of probabilistic logic has been widely studied in the community of AI researchers (e.g., [1-13]). The interest in probabilistic logic as a research topic for AI was sparked by Nilsson’s paper on probabilistic logic [11].

The probabilistic logic, an integration of logic and the probability theory, determines a probability of a sentence by means of a probability distribution on a sample space composed of *classes of possible worlds*. Each class is defined by means of a tuple of consistent truth values assigned to a set of sentences. The deduction in this logic is then reduced to the linear programming problem. However, the problem of determining all such consistent truth value assignments for a set of sentences is NP-complete for propositional logic and undecidable for first-order logic. There have been a great deal of attempts in the AI community to deal with the drawback (e.g., [1], [8], [10], [13]).

This paper first proposes a method of approximate reasoning based on “layers” of an interval-valued probabilistic knowledge base (iKB). The first layer consists of elements of the iKB such that their sentences have some logical relationship with the target sentence. The second one contains elements of iKB whose sentences have some relationship with sentences in the first layer and so on. Our inference method is based on the idea that the calculation of a value of a sentence is only based directly on its nearest upper layer. Later we consider the deduction of point-valued probabilistic logic via Maximum Entropy (ME) principle. Like the deduction from iKB, ME deduction is also based on the matrix composed of vectors of consistent truth values of the target sentence and sentences in a point-valued knowledge base (pKB). It is possible to build this deduction based on the reduced basic matrix of only sentences in some layers of pKB without the target sentence.

The method of constructing layers from sentences in a knowledge base and a method of approx-

imate reasoning based on them will be presented in the next section. Section 3 presents a method of reducing the size of the basic matrix in the pointed probabilistic reasoning via ME. Our approach is to construct the basic matrix of the sentences in the related layers without referring to the goal sentence. Some conclusions and discussions are presented in Section 4.

2. APPROXIMATE REASONING BASED ON LAYERS OF A KNOWLEDGE BASE

2.1. Entailment problem in probabilistic logic

This section overviews the entailment problem of the interval-valued probabilistic logic [3] and of the point-valued probabilistic logic proposed by Nilsson [11].

Given an iKB

$$\mathcal{B} = \{\langle S_i, I_i \rangle \mid i = 1, \dots, l\},$$

in which S_i ($i = 1, \dots, l$) are sentences, I_i ($i = 1, \dots, l$) are subintervals of the unit interval $[0, 1]$; and a target sentence S . From the set of sentences $\Sigma = \{S_1, \dots, S_l, S_{l+1}\}$, ($S_{l+1} = S$), it is possible to construct a set of classes of possible worlds. Every class is characterized by a vector of consistent truth values of sentences in Σ . In this section, we suppose that $\Omega = \{\omega_1, \dots, \omega_k\}$ is the set of all Σ -classes of possible worlds and $(u_{1j}, \dots, u_{lj}, u_{l+1j})^t$ is a column vector of the truth values of sentences w.r.t. S_1, \dots, S_l, S_{l+1} in the class ω_j .

Let $p = (p_1, \dots, p_k)$ be a probability distribution over the sample space Ω . The truth probability of a sentence S_i is then defined to be the sum of probabilities on possible world classes in which S_i is true, i.e.,

$$\pi(S_i) = u_{i1}p_1 + \dots + u_{ik}p_k$$

or

$$\pi(S_i) = \sum_{\omega_j \models S_i} p_j.$$

We can write these equalities in the form of the following matrix equation

$$\Pi = UP,$$

where $\Pi = (\pi(S_1), \dots, \pi(S_l), \pi(S))^t$, $P = (p_1, \dots, p_k)^t$ and $U = (u_{ij})$ ($i = 1, \dots, l+1$, $j = 1, \dots, k$). The matrix U will be called the *basic matrix* of Σ .

The probabilistic entailment problem is reduced to the linear programming one finding

$$\alpha = \min \pi(S), \quad \beta = \max \pi(S),$$

where

$$\pi(S) = u_{l+1,1}p_1 + \dots + u_{l+1,k}p_k,$$

subject to constraints

$$\begin{cases} \pi_i = u_{i1}p_1 + \dots + u_{ik}p_k \in I_i & (i = 1, \dots, l) \\ \sum_{j=1}^k p_j = 1, \quad p_j \geq 0 & (j = 1, \dots, k). \end{cases}$$

We denote the interval $[\alpha, \beta]$ by $F(S, \mathcal{B})$, and write $\mathcal{B} \vdash \langle S, F(S, \mathcal{B}) \rangle$.

In the special case, when \mathcal{B} is the point-valued probabilistic knowledge base (pKB), i.e., all I_i are points α_i in $[0, 1]$, constraints become equalities

$$\begin{cases} \pi_i = u_{i1}p_1 + \dots + u_{ik}p_k = \alpha_i & (i = 1, \dots, l) \\ \sum_{j=1}^k p_j = 1, \quad p_j \geq 0 & (j = 1, \dots, k). \end{cases}$$

However, in general, $F(S, \mathcal{B})$ is not to be a point value. Some assumption is added to the constraints to derive a point value for a target sentence. The Maximum Entropy (ME) principle is usually used for such a deduction. We will return to this investigation in Section 3.

2.2. Layers of knowledge base

This subsection is devoted to presenting a procedure to produce layers of a knowledge base.

Suppose that $\mathcal{B} = \{\langle S_i, I_i \rangle \mid i = 1, \dots, l\}$ is an iKB, in which S_i are propositional sentences and I_i are interval values of sentences S_i ; S is any target sentence we would like to calculate its probability value.

The reasoning for deriving the probabilistic value of the sentence S from the knowledge base \mathcal{B} depends strongly on the basic matrix of truth values of a subset of sentences in $\Sigma' = \{S_1, \dots, S_l\}$ that have some logical relationship with the target sentence. We will characterise the relationship by layering the set of sentences in the knowledge base.

A subset \mathcal{B}' of \mathcal{B} is *sufficient for S* if the probabilistic values of S deduced from \mathcal{B} and \mathcal{B}' are the same.

It means that if $\mathcal{B} \vdash \langle S, I \rangle$ and $\mathcal{B}' \vdash \langle S, I' \rangle$ then $I = I'$.

Denote $atom(\phi)$ the set of atoms occurring in the sentence ϕ and $atom(\Phi) = \bigcup_{\phi \in \Phi} atom(\phi)$ the set of all atoms in sentences in Φ .

Example 1. $atom(A \rightarrow B \wedge C) = \{A, B, C\}$. $atom(\{A \wedge B, C \rightarrow \neg D\}) = \{A, B, C, D\}$.

The following note shows us the meaning of introducing the notion of *atom*.

If \mathcal{B}' is a subset of \mathcal{B} such that

$$atom(\mathcal{B}' \cup \{S\}) \cap atom(\mathcal{B} - \mathcal{B}') = \emptyset,$$

then \mathcal{B}' is sufficient for S .

We now consider a procedure to produce layers of a knowledge base based on a logical dependence between its sentences with the sentence S .

Layers of sentences in Σ are constructed recursively as follows:

$$\begin{aligned} L_0^S &= \{S\}, \\ L_1^S &= \{\phi \mid \phi \in \Sigma, \phi \notin L_0^S \text{ and } atom(\phi) \cap atom(L_0^S) \neq \emptyset\}, \\ L_2^S &= \{\phi \mid \phi \in \Sigma, \phi \notin \bigcup_{i=0}^1 L_i^S \text{ and } atom(\phi) \cap atom(L_1^S) \neq \emptyset\} \\ &\vdots \\ L_n^S &= \{\phi \mid \phi \in \Sigma, \phi \notin \bigcup_{i=0}^{n-1} L_i^S \text{ and } atom(\phi) \cap atom(L_{n-1}^S) \neq \emptyset\}, \\ &\vdots \end{aligned}$$

With respect to each L_n^S , let

$$\mathcal{B}_n^S = \{\langle \phi, I_\phi \rangle \mid \langle \phi, I_\phi \rangle \in \mathcal{B} \text{ and } \phi \in L_n^S\}, \quad n \geq 0.$$

Note that if $S \notin \Sigma'$, then $\mathcal{B}_0^S = \{\langle S, [0, 1] \rangle\}$; otherwise $\mathcal{B}_0^S = \{\langle S, I_S \rangle \mid \langle S, I_S \rangle \in \mathcal{B}\}$.

We call the subset \mathcal{B}_n^S to be n^{th} -layer of the knowledge base \mathcal{B} w.r.t. S . If $\phi \in L_i$, the layer \mathcal{B}_{i+1}^S is called the *nearest upper-layer* of the sentence.

It is easy to see that there always exists a number n_0 such that $L_{n_0}^S \neq \emptyset$ but $L_{n_0+1}^S = \emptyset$. We denote

$$\mathcal{B}_{suf(S)} = \bigcup_{i=0}^{n_0} \mathcal{B}_i^S.$$

It is clear that $\mathcal{B}_{suf(S)}$ is a sufficient subset for S .

Consider the following illustrating example.

Example 2. Given a knowledge base

$$\begin{aligned}\mathcal{B} = \{ & B \rightarrow A : [.9, 1], \\ & D \rightarrow B : [.8, .9], \\ & A \wedge C : [.6, .8], \\ & D : [.8, 1], \\ & C : [.2, .7]\}\end{aligned}$$

and a target sentence A .

The knowledge base can be layered into subsets with the target sentence A

$$\begin{aligned}L_0^A &= \{A\}, & \mathcal{B}_0^A &= \{A : [0, 1]\} \\ L_1^A &= \{B \rightarrow A, A \wedge C\}, & \mathcal{B}_1^A &= \{B \rightarrow A : [.9, 1], A \wedge C : [.6, .8]\} \\ L_2^A &= \{D \rightarrow B, C\}, & \mathcal{B}_2^A &= \{D \rightarrow B : [.8, .9], C : [.2, .7]\} \\ L_3^A &= \{D\}, & \mathcal{B}_3^A &= \{D : [.8, 1]\}\end{aligned}$$

Thus, the sufficient subset for A is

$$\mathcal{B}_{suf(A)} = \mathcal{B}.$$

Similarly, layering can be performed for a point-valued probabilistic knowledge base.

2.3. Approximate solution based on layers

In the case a knowledge base is large, it is not easy to derive the smallest interval value for a target sentence S from $\mathcal{B}_{suf(S)}$. Layers gives us a method of calculating an approximate value. The idea of approximate reasoning is that the probabilistic value of each sentence is updated by deriving its value based on the nearest upper-layer of this sentence. And when all sentences of the nearest upper-layer of the target sentence are updated, its value is then calculated. We now formalise the above presentation.

Without loss of generality, we suppose that \mathcal{B} is a sufficient knowledge base and S is a target sentence. It is layered into subsets $\mathcal{B}_0^S, \mathcal{B}_1^S, \dots, \mathcal{B}_{n_0}^S$, where $\mathcal{B}_{n_0}^S$ is the highest layer in the knowledge base. Remind that L_i^S ($i = 1, \dots, n_0$) are subsets of sentences w.r.t. \mathcal{B}_i^S .

Update of a sentence ϕ is recursively defined as follows:

(i) For all $\phi \in L_{n_0}^S$, ϕ is updated;

(ii) $\phi \in L_i^S$, ($i < n_0$), is updated if all $\psi \in L_{i+1}^S$ are updated and $\mathcal{B}_{(i+1,u)}^S \vdash \langle \phi, I_\phi \rangle$, where

$\mathcal{B}_{(i+1,u)}^S$ is the updated layer of \mathcal{B}_{i+1}^S .

If \mathcal{B}_1^S is updated into $\mathcal{B}_{(1,u)}^S$ and $\mathcal{B}_{(1,u)}^S \vdash \langle S, I_S \rangle$, then I_S is the approximate value for S .

Thus, the approximate calculation of interval value for a sentence consists of three steps:

1. Divide the knowledge base into layers with the lowest layer being the target sentence S .

2. Update the values for sentences of \mathcal{B}_{i-1} from the nearest upper-layer \mathcal{B}_i . This process starts from $i = n_0$ till \mathcal{B}_1 is updated into $\mathcal{B}_{(1,u)}^S$.

3. Calculate the value for S from $\mathcal{B}_{(1,u)}^S$.

Example 3. (continued) In Example 2, we have constructed the layers of the knowledge base. If we base on the whole $\mathcal{B}_{suf(A)}$, it is necessary to build a 6×14 -basic matrix of 6 rows and 14 columns. It is possible to calculate the value for A according to the above approximate method.

In the process of updating, $D \rightarrow B$ and $B \rightarrow A$ are stable, i.e., their values are $[.8, .9]$ and $[.9, 1]$, respectively. Since the value of C is $[.2, .7]$, $A \wedge C$ is updated to $[.6, .7]$. Thus, a value of A is deduced from the 1th updated layer

$$\mathcal{B}_{(1,u)}^A = \{B \rightarrow A : [.9, 1], A \wedge C : [.6, .7]\}.$$

The basic matrix for sentences $\Sigma = \{B \rightarrow A, A \wedge C, A\}$ is

$$\begin{pmatrix} 1 & 1 & 1 & 0 \\ 1 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 \end{pmatrix}$$

We need to compute

$$\min(p_1 + p_3) \quad \max(p_1 + p_3)$$

on the domain determined by

$$\begin{cases} .9 \leq p_1 + p_2 + p_3 \leq 1 \\ .6 \leq p_1 \leq .7 \\ p_1 + p_2 + p_3 + p_4 = 1 \end{cases}$$

The value of A is then $[.6, 1]$.

We compare now the computable value with a value derived from the anytime deduction proposed by Frish and Haddawy [8]. Anytime deduction is based on a set of thirty two rules enumerated from (i) to (xxxii). In the above example, applying (xx) first to $D : [.8, 1]$ and $D \rightarrow B : [.8, .9]$ yields $B : [.6, .9]$; then combining it with $B \rightarrow A : [.9, 1]$ via the rule (xx) results to $A : [.5, 1]$. In the same way, combining $C : [.2, .7]$ and $A : [0, 1]$ via the rule (xxv) gives $A \wedge C : [0, .7]$ and then with $A \wedge C : [.6, .8]$ via (xvii) gives $A \wedge C : [.6, .7]$; applying (xxvi) to this result yields $A : [.6, 1]$. Applying (xvii) to two ways of computation of A , we have $A : [.6, 1]$. The derived interval equals to the interval value of A deduced by our method of approximate reasoning.

3. MAXIMUM ENTROPY DEDUCTION BASED ON THE REDUCED BASIC MATRIX

In this section, we investigate a method of reducing the complexity of computation in applying the Maximum Entropy Principle for deriving a point value for a sentence from a point-valued probabilistic knowledge base.

3.1. Maximum Entropy Deduction

We first review a technique named Maximum Entropy Principle [11] to select a probability distribution among distributions holding some initial conditions given by a knowledge base.

Suppose that

$$\mathcal{B} = \{ \langle S_i, \alpha_i \rangle \mid i = 1, \dots, l \}$$

is pKB and S is a sentence ($S \neq S_i, i = 1, \dots, l$). As presented in Section 2, denote $F(S, \mathcal{B})$ the set of values of $\pi(S) = \sum_{\omega_i \models S} p_i = u_{l+1,1}p_1 + \dots + u_{l+1,k}p_k$, where $p = (p_1, \dots, p_k)$ varies in the domain defined by conditional equation

$$\Pi = U^+ P, \quad (1)$$

where $\Pi = (1, \alpha_1, \dots, \alpha_l)^t$ and U^+ is the basic matrix composing of columns of truth values of sentences S_1, \dots, S_l, S_{l+1} ($S_{l+1} = S$) with the first row being units.

According to Maximum Entropy Principle, in order to obtain a single value for S , we must select a distribution p such that the following optimization problem holds

$$H(p) = - \sum_{j=1}^k p_j \log p_j \rightarrow \max, \quad (2)$$

where p subjects to constraints determined by the conditional equation (1).

Suppose that (p_1, \dots, p_k) is a solution of the above problem. Then the probability of S is denoted by

$$F(S, \mathcal{B}) = u_{11}p_1 + \dots + u_{l+1,k}p_k.$$

Let a_0, a_1, \dots, a_l be parameters for rows of U^+ . Each p_i is defined according to a_j by means of i^{th} -column of U^+

$$p_i = a_0 \prod_{u_{i,j}=1, 1 \leq j \leq l} a_j \quad (i = 1, \dots, k). \quad (3)$$

From the initial conditions of the knowledge base, we can compute a_i and then p_i . Thus the point probability value of S is then derived. We call the deduction based on the Maximum Entropy Principle to be the *Maximum Entropy deduction* or shortly ME deduction.

3.2. Maximum Entropy Deduction with the Reduced Basic Matrix

As presented above, the ME deduction is based on the basic matrix constructed from the target sentence and all sentences in the initial knowledge base. The larger the basic matrix is, the more complex the computation is. In fact, coefficients a_i in (3) are only related to the matrix of truth values of sentences in the knowledge base. The complexity is slightly decreased if ME deduction is based on the basic matrix constructed only from sentences of the knowledge base without the target sentence.

As presented in Subsection 2.2, the probabilistic inference only depends on the sufficient subset for the target sentence. Without loss of generality, we suppose that $\mathcal{B} = \mathcal{B}_{\text{suf}(S)}$, $\Omega = \{\omega_1, \dots, \omega_k\}$ is a set of possible world classes determined by $\Sigma = \{S_1, \dots, S_l\}$ and U^+ is the *reduced basic matrix* of sentences in Σ with the first row being units.

In each class ω_i , S can have either one truth value true/false or both truth values true and false. For ease of presentation, we suppose that on classes $\omega_1, \dots, \omega_m$, the sentence S gets one truth value and on $\omega_{m+1}, \dots, \omega_k$, S has both values true and false. Thus, the extended set of possible world classes w.r.t. $\Sigma \cup \{S\}$ has the form

$$\Omega^+ = F \cup E,$$

where $F = \{\omega_1, \dots, \omega_m\}$ and $E = \{\omega_{m+1}^+, \omega_{m+1}^-, \dots, \omega_k^+, \omega_k^-\}$. We have the following proposition.

Proposition 1. *Suppose that p is a probability distribution satisfying ME principle on Ω . We have*

$$\pi(S) = \sum_{\omega_i \models S, \omega_i, 1 \leq i \leq m} p_i + \frac{1}{2} \sum_{\omega_i \models S, m+1 \leq i \leq k} p_i. \quad (4)$$

Proof. Suppose $p^+ = (p'_1, \dots, p'_m, p_{m+1}^+, p_{m+1}^-, \dots, p_k^+, p_k^-)$ is the probability distribution on Ω^+ satisfying ME and (1). According to the method of constructing this distribution, we have

$$p_{m+1}^+ = p_{m+1}^-, \dots, p_k^+ = p_k^-.$$

Therefore, if $p = (p_1, \dots, p_m, p_{m+1}, \dots, p_k)$ is the probabilistic distribution on Ω satisfying (1) and ME, then

$$\begin{aligned} p_i &= p'_i \quad (i = 1, \dots, m), \\ p_i &= 2p_i^+ \quad (i \geq m+1). \end{aligned}$$

It is easy to derive (4) from these equalities. The proposition is proved.

In summary, the computation of the point value for a sentence S via ME consists of three steps:

1. Construct the sufficient subset for S to eliminate unnecessary information.
2. Find an entropy-maximizing p based on the reduced basic matrix U^+ of the sentences in the sufficient subset.
3. Calculate $\pi(S)$ via the equality (4).

Example 4. Given a knowledge base

$$\begin{aligned} \mathcal{B} &= \{A : \alpha_1, \\ &A \rightarrow B : \alpha_2, \\ &B \rightarrow C : \alpha_3\} \end{aligned}$$

and a target sentence C .

It is clear that $\mathcal{B} = \mathcal{B}_{\text{suf}(C)}$. The reduced basic matrix for the set of sentences in \mathcal{B} with the first

row of units is

$$U^+ = \begin{pmatrix} 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 0 & 1 & 0 \\ 1 & 1 & 1 & 0 & 1 \\ 1 & 0 & 0 & 1 & 1 \end{pmatrix}$$

in which the second row is the truth values of A , the third and fourth ones are of $A \rightarrow B$ and $B \rightarrow C$, respectively. Thus, there are five classes of possible world $\omega_1, \dots, \omega_5$ corresponding to five column vectors (eliminating the first row):

$$v_1 = (1, 1, 1)^t, v_2 = (1, 1, 0)^t, v_3 = (0, 1, 0)^t, v_4 = (1, 0, 1)^t, v_5 = (0, 1, 1)^t.$$

Components of p_i are written in the form

$$p_1 = a_0 a_1 a_2 a_3, p_2 = a_0 a_1 a_2, p_3 = a_0 a_2, p_4 = a_0 a_1 a_3, p_5 = a_0 a_2 a_3$$

with (a_0, a_1, a_2, a_3) satisfying the system of equations

$$\begin{cases} a_0 a_1 a_2 a_3 + a_0 a_1 a_2 + a_0 a_1 a_3 + a_0 a_2 a_3 = \alpha_1 \\ a_0 a_1 a_2 a_3 + a_0 a_1 a_2 + a_0 a_2 + a_0 a_2 a_3 = \alpha_2 \\ a_0 a_1 a_2 a_3 + a_0 a_1 a_3 + a_0 a_2 a_3 = \alpha_3 \\ a_0 a_1 a_2 a_3 + a_0 a_1 a_2 + a_0 a_2 + a_0 a_1 a_3 + a_0 a_2 a_3 = 1 \end{cases}$$

Solving yields

$$\begin{aligned} a_0 &= (1 - \alpha_1)(1 - \alpha_2)(1 - \alpha_1 + \alpha_2 - \alpha_3) / (\alpha_1 + \alpha_3 - 1)(\alpha_2 - \alpha_3), \\ a_1 &= (\alpha_2 - \alpha_3) / (1 - \alpha_1), \\ a_2 &= (\alpha_1 + \alpha_3 - 1)(\alpha_2 - \alpha_3) / (1 - \alpha_2)(1 - \alpha_1 + \alpha_2 - \alpha_3), \\ a_3 &= (1 + \alpha_1 + \alpha_3) / (1 - \alpha_1 + \alpha_2 - \alpha_3). \end{aligned}$$

Thus, the entropy-maximizing p is given by:

$$p = \begin{pmatrix} (\alpha_2 - \alpha_3) / (\alpha_1 + \alpha_3 - 1) \\ \alpha_2 - \alpha_3 \\ 1 - \alpha_1 \\ 1 - \alpha_2 \\ (1 - \alpha_1)(\alpha_1 + \alpha_3 - 1) / (1 - \alpha_1 + \alpha_2 - \alpha_3) \end{pmatrix}$$

Since C has one true value on ω_1 , two truth values in classes ω_4 and ω_5 (false value on ω_2, ω_3), the probability of A is then

$$\begin{aligned} \pi(A) &= p_1 + \frac{1}{2}p_4 + \frac{1}{2}p_5 \\ &= \frac{1}{2}(1 - \alpha_2) + \frac{(\alpha_1 + \alpha_3 - 1)(1 - \alpha_1 + 2\alpha_2 - 2\alpha_3)}{2(1 - \alpha_1 + \alpha_2 - \alpha_3)}. \end{aligned}$$

4. CONCLUSION

This paper has presented a method of layering a knowledge base based on the logical relationship between sentences of the knowledge base with a target sentence. By means of layers, we can perform approximate reasoning in order to derive an interval value for the sentence. Our approximate method is different from the anytime deduction proposed by Frish and Haddawy [8]. While our one is based on the process of updating of all sentences before deriving an interval value for the target sentence, their anytime deduction is based on a set of rules.

We have also presented a method of calculating the point probabilistic value of a sentence via the Maximum Entropy Principle by not referring to the target sentence when constructing the basic matrix. This method slightly decreases the size of the matrix in the computation process.

We have presented a comparative example between our approximate method and the anytime deduction proposed by Frish and Haddawy. A complete comparison of this approximate method with the other ones will be a topic of our further work.

Acknowledgement. I am greatly indebted to my supervisor, Prof. Phan Dinh Dieu, for invaluable suggestions.

REFERENCES

- [1] K. A. Anderson. Characterizing consistency in probabilistic logic for a class of Horn clauses. *Mathematical Programming* **66** (1994) 257–271.
- [2] F. Bacchus, A. J. Grove, J. Y. Halpern and D. Koller. From statistical knowledge bases to degrees of belief, *Artificial Intelligence* **87** (1–2) (1996) 75–143.
- [3] P. D. Dieu, On a theory of interval-valued probabilistic logic, *Research Report, NCSR Vietnam, Hanoi*, 1991.
- [4] P. D. Dieu and P. H. Giang, Interval-valued probabilistic logic for logic programs, *Journal of Computer Science and Cybernetics* **10** (3) (1994) 1–8.
- [5] P. D. Dieu and T. D. Que, From a convergence to a reasoning with interval-valued probability, *Journal of Computer Science and Cybernetics* **13** (3) (1997) 1–9.
- [6] R. Fagin, J. Y. Halpern, and N. Megiddo, A logic for reasoning about probabilities, *Information and Computation* **87** (1990) 78–128.
- [7] R. Fagin and J. Y. Halpern, Uncertainty, Belief and Probability, *Computational Intelligence* **7** (1991) 160–173.
- [8] A. M. Frish and P. Haddawy, Anytime deduction for probabilistic logic, *Artificial Intelligence* **69** (1994) 93–122.
- [9] R. Kruse, E. Schwecke, and J. Heinsohn, *Uncertainty and Vagueness in Knowledge Based Systems*, Springer-Verlag, Berlin - Heidelberg, 1991.
- [10] R. T. Ng and V. S. Subrahmanian, Probabilistic logic programming. *Information and Computation* **101** (1992) 150–201.
- [11] N. J. Nilsson, Probabilistic logic, *Artificial Intelligence* **28** (1986) 71–78.
- [12] T. D. Que, About semantics of probabilistic logic, Submitted to *Computer Science and Cybernetics*.
- [13] P. Snow, Compressed constraints in probabilistic logic and their revision, *Uncertainty in Artificial Intelligence* (1991) 386–391.

Received November 15, 1999

*Department of Information Technology,
Posts and Telecommunications Institute of Technology,
Hanoi, Vietnam.*