2008 International Conference on Computer Science and Software Engineering

Deduction from Conditional Knowledge on Bayesian Networks with Interval Probability Parameters

Yong Li[1,2]
[1] Dept. of automation
Kunming university of science and technology
Kunming, P. R. China
leon@people.com.cn

Weiyi Liu[2]
[2] Dept. of computer science and engineering
Yunnan university
Kunming, P. R. China
liuweiyi2000@yahoo.com.cn

***Abstract*—We propose a Bayesian higher-order probability logic reasoning approach with interval probability parameters to the problem of making inference from conditional knowledge, which combines weak conditional probability and conditional event algebra for approximate inferences. We define the bound-limited weak conditional interval probabilities, the corresponding probabilistic description, and the multiplication rules of weak conditional probabilities for joint probability distribution, and use higher-order conditional event to resolve a discrepancy between logic and probability. By extending normal measurable space with conditional event, we bring logic consistent with probability in denoting conditional knowledge, and then transform a higher-order conditional event to normal events and corresponding logical joint events via conditional event algebra. Based on multiplication rules, we compute the quantitative values of the events with interval parameters, and evaluate the value of higher-order conditional event and finish reasoning process. An illustrative example of application of our method shows how we make inferences from conditional knowledge.**

Keywords-Probability logic reasoning; conditional event algebra; Bayesian network; interval probability

I. INTRODUCTION

A common problem of interest in the Artificial Intelligent community and beyond is how to draw conclusion from a knowledge base, such as Γ. As well known, Γ is a set of conditional rules, "b→a", to be read as "if b then a" whose reliability is quantified by conditional probability P(a|b) [1]. The general problem is to estimate the reliability of some rules of interesting that is not in Γ, and the quantitative problem is that given either numerical values or numerical bounds on the reliabilities of the rules, find the estimate of the reliability of the rule of interest. As a graphical representation of probabilistic causal relationships, Bayesian networks (BNs) are effective and widely used frameworks. Probabilistic inferences can be done by computing products of conditional probabilities from BNs [2]. There exists a discrepancy between logic and probability. The "if-then" is also modeled as Boolean element and its reliability is usually denoted as P(b→a). In classical logic, "b→a" is equivalent to "b′∨a", but it can be proved [3]:

$$P(b\rightarrow a)=(P(a|b)+P(b')P(a'|b))\geqslant P(a|b), \qquad (1)$$

with strict inequality holding unless P(b)=1 or P(a|b) =1.

This work is supported by the Chun-Hui Projects of the Educational Department of China under Grant No. Z2005-2-65003, and the National Nature Science Foundation of China under Grant No. 60763007.

On the other hand, how to denote complex conditional rules such as "if if b_1 then a_1 and if b_2 then a_2…then if d then c", i.e., higher-order logic [4], and evaluate its quantitative value? Given a rule "if b then a", the probability of an event c under it may be denoted as P(c|(a|b)). But it can't be evaluated with P(c|(b′∨a)).

And more, reasoning methods have been mainly developed for the analysis of single-valued variables. In the real life, values of variables are always presented with numerical bounds, i.e., interval parameters. BNs have been used in many different intelligent applications, and the representation and inference are comprehensively interpreted in [2]. Starting from conditional independence, Pearl defined the Markov blanket, which describes the direct causes and direct effects given a certain node in the BN. The discovery of Markov blanket is applied in BN's approximate reasoning, such as stochastic simulation [8]. Imprecise probabilities and weak conditional independence are incorporated into BNs [9]. Gyftodimos and Flach discussed how to combine BNs with higher-order data representations [4]. However, existing concepts and definitions of interval probabilities do not satisfy the conditional probability and the multiplication rules for usual scenes [6]. Representation and inference of higher-order conditionals with interval probability parameters on BNs can't be done.

The conditional event algebra (CEA) generalizes the logic computations of Boolean algebra to a set of conditional events while maintaining the compatibility of ordinary probability and the conditional probability [5]. Based on CEA, combining the theory of fuzzy logic with random sets, Goodman et al proposed Boolean relational event algebra and Boolean conditional event algebra [6]. The interval-probability theory is applied to the representation of uncertain statistics [7]. Many researchers are interested in representation of uncertain statistics using interval probabilities [6]. Weichselberger gave the system of axioms describing the properties of interval probabilities and mentioned 2 different concepts of the interval conditional probability: the intuitive concept and the canonical concept. It is intuitive to quantify the rules implied in interval sample data by adopting interval probability parameters and CEA instead of traditional conditional probabilities associated with the interval-valued variables.

Based on the concepts of lower and upper approximation of rough set theory [10], in this paper we define the interval

978-0-7695-3336-0/08 $25.00 © 2008 IEEE
DOI 10.1109/CSSE.2008.948

probability, the bound-limited weak conditional probability of interval value samples, corresponding probabilistic description, and the multiplication rule of weak conditional probabilities for joint probability distribution. Further, we propose a Bayesian higher-order probabilistic logic reasoning approach with interval probability parameters, which combines weak conditional probability and CEA.

II. Some Preliminaries

A. Conditional Event Algebra

Let (Ω, B, P) be a standard probability space, Ω be a sample space, B a fixed event domain in the space, and P a certain probability measure. For a, b∈B, P(a) and P(b) defined on the subsets a and b belonging to σ-algebra σ(Ω) of subsets of Ω, that satisfies the following properties [5]:

(1) P(Φ) =0, P(Ω)=1.

(2) $P(b_1 \cup \dots \cup b_j \cup \dots) = P(b_1)+ \cdots +P(b_j)+ \cdots$ for any countable sequence of mutually disjoint b_i of Ω.

The relationship between probability and the Boolean operations ∩, ∪ and ′ on the σ(Ω) is expressed by the familiar relationships P(b′)=1–P(b), P(b∪a)+P(b∩a)=P(b)+P(a).

The basic idea underlying CEA is to find a way if mathematically representing rules of the form "if b then a" as "conditional rules" in such way:

(3) The set of all conditional events is a Boolean algebra with logic connectives ∧, ∨ and ′.

(4) (b|Ω)=b. I.e., ordinary events are a special kind of conditional event.

(5) Any probability measure P defined on the events of Ω can be uniquely extended a probability measure P_0 on conditional events such that $P_0((a|b))=P(a|b)$. That is, conditional probability is a true probability measure on conditional events.

(6) (a|b)∧b=(a∩b). I.e., when a rule (a|b) with antecedent b is "fired" using that antecedent, then what results is knowledge of both the consequent a and the antecedent b of the rule. This is the classical inference rule known as modus ponens.

Given normal events a and b, constructing an extended production probability measurable space (Ω_0, B_0, P_0), $\Omega_0=\Omega\times\Omega\times\dots$, B_0 is a Boolean algebra or σ-algebra extended from B×B…, and P_0 is probability measure in the space. So,

$$(a|b)=((a\wedge b)\times\Omega_0)\vee(b'\times(a\wedge b)\times\Omega_0)\vee(b'\times b'\times(a\wedge b)\times\Omega_0)\vee\dots$$

Given a function f: $B\times B\rightarrow B_0$, for all f are defined as

$$f(a,b)=ab\vee(b'\times ab)\vee(b'\times b'\times ab)\vee\dots \qquad (2)$$

We have

$$P_0(f(a,b))=P_0(ab\vee(b'\times ab)\vee(b'\times b'\times ab)\vee\dots)$$
$$=P_0(ab)+ P_0(b'\times ab)+ P_0(b'\times b'\times ab)+\dots$$
$$=P(ab)\Sigma(P(b'))^j= P(a|b), \quad j=0,1,2,\dots \qquad (3)$$

Together, conditions (1-6) state that if conditionals are mathematically modeled as conditional events, then standard probability theory can be extended to rules and, more generally, to all kinds of contingent events. Thus, it is possible to calculate $P_0((a|b))$ and express the probability weight of the conditional rule "b→a".

B. Interval Probabilities of Events with Interval Data

Let the universe of discourse U be a continuous one-dimensional space. A subset S of U is called an interval value, if all $v_1,v_2,v_3\in U$ and $v_1\leqslant v_2\leqslant v_3$, $v_1,v_3\in S$ implies $v_2\in S$. We write an interval value ā=[a–,a+], where a– and a+ are the lower and upper bounds of ā respectively. Based on the concepts of lower and upper approximation of rough set theory [10], now we introduce the interval probability for describing events with interval-valued samples.

Definition 1 Let b be a random event with interval value samples. The lower and upper probabilities, written L(b) and U(b) respectively, are defined as follows:

$$L(b)=\Sigma P(x_i),\ x_i \text{ in } b;\ U(b)=\Sigma P(x_i),\ x_i\cap b\neq\Phi, \qquad (4)$$

where x_i is a sample in the sample space Ω.

Definition 2[7] Given a sample space Ω and a σ-field F of random events in it. A set function P(.) defined on F is called Kolmogrov-function, if it satisfies:

(1) For A∈F, P(a)⩾0;

(2) P(Ω)=1;

(3) For a_i, a_j∈F , if a_i and a_j are mutually independent, then $P(a_1\cup\dots\cup a_k)=\Sigma P(a_i)$, i=1,…,k.

Interval-valued set functions [L(.),U(.)] on F are called I-function, if they satisfy:

(1) (Weak condition) For all a∈F, L(a)⩽P(a)⩽U(a);

(2) (Strong condition) For all a∈F, U(a)=1–L(a′).

We only discuss the weak condition restriction.

The classical concept of the conditional probability can be generalized to the interval probability in different ways. Under the intuitive concept, we can convey L(a|b) and U(a|b) by inf(P(a∩b)/P(b)), P ∈ {P(.)|P(b)>0}; sup(P(a∩b)/P(b)), P ∈ {P(.)|P(b)>0}, respectively, in which P (.) is defined on F [7].

For making inferences on a BN with interval probabilities, bound-limited weak conditional probability is advanced.

Definition 3 The bound-limited weak conditional probabilities BL(a|b) and BU(a|b) are defined according to various cases of the relationships among L(a), U(a), L(a|b) and U(a|b):

(1) If L(a) ⩽ L(a|b) ⩽ U(a|b) ⩽ U(a), then BL(a|b)=max{EL(a|b), L(a)} and BU(a|b)=min{EU(a|b), U(a)}.

(2) If L(a|b) ⩽ L(a) ⩽ U(a) ⩽ U(a|b), then BL(a|b)=min{EL(a|b), L(a)} and BU(a|b)=max{EU(a|b), U(a)}.

(3) If L(a|b) ⩽ L(a) ⩽ U(a|b) ⩽ U(a), then BL(a|b)=min{EL(a|b), L(a)} and BU(a|b)=min{EU(a|b), U(a)}.

(4) If L(a) ⩽ L(a|b) ⩽ U(a) ⩽ U(a|b), then BL(a|b)=max{EL(a|b), L(a)} and BU(a|b)=max{EU(a|b), U(a)}.

(5) If $L(a) \leqslant U(a) \leqslant L(a|b) \leqslant U(a|b)$, then $BL(a|b)=\max\{EL(a|b), U(a)\}$ and $BU(a|b) = \min\{EU(a|b), 1\}$.

(6) If $L(a|b)\leqslant U(a|b)\leqslant L(a)\leqslant U(a)$, then $BL(a|b)=EL(a|b)$ and $BU(a|b) = \min\{EU(a|b),L(a)\}$.

We have BL(a|b) and BU(a|b) satisfy the weak condition of I-functions, since $BL(a|b) \leqslant P(a|b) \leqslant BU(a|b)$. Here we substitute BL(a|b) and BU(a|b) for L(a|b) and U(a|b) respectively. In line with the concept of canonical conditional probabilities, we give the multiplication rules as follows:

$$L(abc)\approx BL(abc)=L(a)BL(b|a)BL(c|ab),$$
$$U(abc)\approx BU(abc)=U(a)BU(b|a)BU(c|ab). \quad (5)$$

From Definition 3, we know L(abc) is not exactly equal to BL(abc), but a certain approximation.

III. Probability logic reasoning on Bayesian networks with interval probability parameters

In symbol, the reasoning question about higher-order conditional event "if if b_1 then a_1 and if b_2 then a_2…then if d then c" can be expressed as the scheme:

$$G=[(a|b)_J;(c|d)]. \quad (6)$$

The a_J, b_J, c, d are events under any fixed Boolean algebra B, $(a|b)_J$ expresses $(a_j|b_j)_{j \text{ in } J}$, P is a function denoted as P: $B\rightarrow[0,1]$ (unit interval). $(a|b)_J$ is a set of given rules, (c|d) is inferred conclusion from$(a|b)_J$, and $P(a|b)_J$ in [0,1]. According to CEA, G can be denoted as a set of conjunction events [11]:

$$A(G)=\cap\{a_jb_j,a_j'b_j,b_j'\}\cap\{cd,c'd,d'\}=\{\omega_1,\ldots,\omega_{m+1}\}. \quad (7)$$

If no conjunction were null then $m=3^{card(J)+1}-1$, card(J) is dimension of J. According to (7), G can be denoted as

$$f(a_j,b_j,c,d)=\vee(\omega_j), \text{ j in I(f)}. \quad (8)$$

From (8), for any event f in Boolean (A(G)), there is a uniquely determined index set I(f) in {1,…,m,m+1} determining it and correspondingly for any probability measure P:

$$P((a|b)_J;(c|d))=P(A(G))=P_0(f(a_j,b_j,c,d))=\Sigma P(\omega_j). \quad (9)$$

According to what mentioned above in Definition 3, when the conditional information $(a|b)_J$ is presented with interval-valued, we can obtain the bound-limited estimation of ω_j,

$$BL(\omega_j)\leqslant P(\omega_j)\leqslant BU(\omega_j). \quad (10)$$

Correspondingly, we have

$$\Sigma BL(\omega_j)\leqslant\Sigma P(\omega_j)\leqslant\Sigma BU(\omega_j), \text{ for all j in I(f)}. \quad (11)$$

So, estimating the quantitative value of G is equivalent to estimating the value of normal probability events and correspond joint events ω_1, …, ω_{m+1}. By calculating the probabilities of the events, we can obtain the quantitative value of G. But the intractability of exact inference of interval probability is obvious. Based on the property of Markov blankets, we consider adopting the bound-limited weak interval conditional probabilities for approximate higher-order probability logic reasoning.

A Markov blanket MB(X) of a node X in a BN is any subset S (X not in S) of nodes for which X is independent of U-S-X given S. In any BN, the union of the following 3 types of neighbors is sufficient for forming a Markov blanket of a node X: the direct parents of X, the direct successors of X and all direct parents of X's direct successors [2]. Now we give the basic idea immediately as follows. First, according to weak conditional probability, the probability of G with interval parameters can expressed as [BL(G), BU(G)]. According to (10) and (11), it can be expressed as $[\Sigma BL(\omega_i), \Sigma BU(\omega_i)]$, i=1,…,m+1. The ω_i is a joint event and its corresponding $BL(\omega_i)$ and $BU(\omega_i)$ can be calculated using (5) with the Markov blanket of each event in the joint event. By summing $BL(\omega_i)$ and $BU(\omega_i)$ up respectively, we can estimate the final result of G.

Algorithm 1 Approximate higher-order conditional event probability logic reasoning on a BN with interval probability parameters.

Input:

G: the higher-order conditional event $[(a|b)_J;(c|d)]$; Events a, b, c , d are the nodes of a BN,

BN: a BN with interval probability parameters.

Output: the estimates of BL(G) and BU(G).

Variables:

ω_i: a joint event,

ni: the total number of events in ω_i,

ω_{ij}: a normal event in ω_i, j=1,…, ni.

Step 1. Use CEA to change G to be corresponding joint events $\{\omega_1,\ldots,\omega_{m+1}\}=\cap\{a_ib_i,a_i'b_i,b_i'\}\cap\{cd,c'd,d'\}$, $i\leq3^{card(J)+1}$; ω_i is composed of $\omega_{i,1},\ldots,\omega_{i,ni}$.

Step 2. For i←1 to m+1 do

According to the properties of the Markov blanket, Compute the interval probability values of $BL(\omega_i)$ and $BU(\omega_i)$ with multiplication rules:

$$BL(\omega_i)=L(\omega_{i,1})BL(\omega_{i,2}|\omega_{i,1})\ldots BL(\omega_{i,ni}|\omega_{i,1}\ldots\omega_{i,ni-1});$$

$$BU(\omega_i)=U(\omega_{i,1})BU(\omega_{i,2}|\omega_{i,1})\ldots BU(\omega_{i,ni}|\omega_{i,1}\ldots\omega_{i,ni-1});$$

Step 3. Sum $BL(\omega_i)$, $BU(\omega_i)$ up respectively and quit.

The validation of Algorithm 1 is discussed briefly as following: According to CEA, a higher-order conditional event G can be changed to be a set of joint events. From the canonical conditional probabilities, the joint event can be calculated with the multiplication rules as (5). That means, by summing all values of joint events up, the corresponding bound-limited weak interval conditional probability of G can be obtained.

IV. Illustrative example

Considering the corresponding BN of Fig. 1, with contains knowledge a, b, c, d and e nodes with interval-valued. The question of estimating the reliability of knowledge d can be the posterior distribution of (d|(b|a),(c|a)). It can be described as

$G=[(b|a),(c|a);\ (d|\Omega)]$. Its measure weight is $[BL((d|\Omega)|(b|a),(c|a)),\ BU((d|\Omega)|(b|a),(c|a))]$.

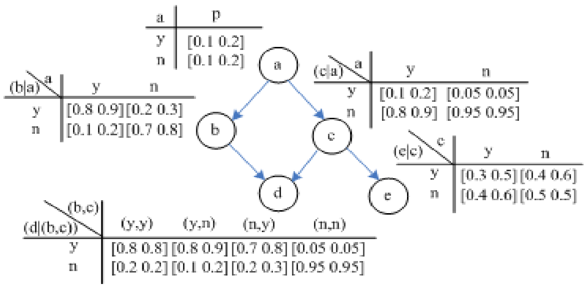

Figure 1. A BN with interval probability parameters.

Now we use Algorithm 1 to estimate G. According to (5), there exists

$$A(G)=\{ba,b'a,a'\}\wedge\{ca,c'a,a'\}\wedge\{d\times\Omega,d'\times\Omega,\Omega'\}. \quad (12)$$

For $\Omega'=\Phi$, we have

$A(G)=\{ba, b'a, a'\}\wedge\{ca, c'a, a'\}\wedge\{d\times\Omega,d'\times\Omega\}$
$=\{abcd,ab'cd,abc'd,ab'c'd,a'd,abcd',ab'cd',abc'd',a'b'c'd',a'd'\}$
$=\{\omega_1,\ldots,\omega_{10}\}$.

Based on multiplication rules and the properties of Markov blanket, we have

$BL(\omega_1)=BL(abcd)=L(a)BL(b|a)BL(c|a)BL(d|bc)=0.1\times0.8\times0.1\times0.8=0.0064$,
$BU(\omega_1)=BU(abcd)=U(a)BU(b|a)BU(c|a)BU(d|bc)=0.2\times0.9\times0.2\times0.8=0.0288$,
$BL(\omega_2)=0.0007$, $BU(\omega_2)=0.0064$, $BL(\omega_3)=0.0512$,
$BU(\omega_3)=0.1458$, $BL(\omega_4)=0.0004$, $BU(\omega_4)=0.0018$,
$BL(\omega_5)=0.0217$, $BU(\omega_5)=0.0677$, $BL(\omega_6)=0.0016$,
$BU(\omega_6)=0.0072$, $BL(\omega_7)=0.002$, $BU(\omega_7)=0.0024$,
$BL(\omega_8)=0.0064$, $BU(\omega_8)=0.0324$, $BL(\omega_9)=0.0076$,
$BU(\omega_9)=0.0342$, $BL(\omega_{10})=0.066$, $BU(\omega_{10})= 0.1588$.

The $[BL((d|\Omega)|(b|a),(c|a)),\ BU((d|\Omega)|(b|a),(c|a))]$, i.e. $[\Sigma BL(\omega_i),\ \Sigma BU(\omega_i)]$ $(i=1,\ldots,10)$, is $[0.1766,0.4855]$. So, the probability of G is estimated.

V. SUMMARY

In making inferences from conditional knowledge, we use CEA to denote conditional rules for resolving the problem of the discrepancy between probability and classic logic, and transform higher-order conditional event to correspond joint events. By using weak conditional probability, we calculate the lower and upper bounds of these events respectively. At last, the quantitative value of higher-order conditional event is estimated. Based on the method proposed in this paper, the integration of a BN with interval probability parameters has been studied.

REFERENCES

[1] D. Bamber, I. R. Goodman, and H. T. Nguyen, "Deduction from conditional knowledge", Soft Computing, vol. 8, No. 4, 2004, pp. 247–255.

[2] J. Pearl, Probabilistic Reasoning in Intelligent Systems: Network of Plausible Inference. San Mates, California, Morgen Kaufmann Publisher, Inc., 1998.

[3] I. R. Goodman, R. P. S. Mahler, and H. T. Nguyen, Mathematics of Data Fusion. Kluwer Academic Publisher, 1997.

[4] Eilas Gyftodimos and Peter Flach, "Combining Bayesian networks with higher-order data representations", A. F. Famili et al. (EDs.): IDA 2005, LNCS 3646, 2005, pp.145–156.

[5] C. T. William, D. Bamber, I. R. Goodman, and H. T. Nguyen, "A new method for representing linguistic quantifications by random sets with applications to tracking and data Fusion", ISIF, vol. 2, 2002, pp.1308–1315.

[6] I. R. Goodman, H. T. Nguyen, and E. A. Walker, Conditional Inference and Logic for Intelligent systems: A Theory of Measure-Free Conditioning. Amsterdam: North-Holland, 1991.

[7] K. Weichselberger, "The theory of interval probability as unifying concept for uncertainty", Int. J. Approximate Reasoning, vol. 24, 2000, pp.149–170.

[8] J. Pearl, "Evidential reasoning using stochastic simulation of casual models", Int. J. Artificial Intelligence, vol. 32, 1987, pp.245–257.

[9] S. K. M. Wong and C. J. Buta, "Contextual weak independence in Bayesian networks", Proc. Of the 15th Conf. UAI, 1999, pp. 670–690.

[10] Z. Pawlak, "Rough sets", Int. J. Computer and Information Science, vol. 11, 1982, pp.341–356.

[11] D. Bamber and I. R. Goodman, "New uses of second order probability techniques in estimating critical probabilities in command & control decision-making", Proc. of the 2000 Command & Control Research & Technology Symp., June, 2000.