# Qualitative Reasoning of Bayesian Belief Using Meta-knowledge

Bon K. Sy

Queens College
of the City University of New York
Department of Computer Science
65-30 Kissena Boulevard
Flushing, NY 11367-0904
BON@QCVAX.BITNET
(718)-520-5100

## Abstract

Ordering composite hypotheses in a Bayesian network based on its associated *a posteriori* probabilities can be exponentially hard. This paper discusses a qualitative reasoning approach which reduces the computational complexity of deriving a partial ordering of composite hypotheses. Such a reasoning makes use of the *meta-knowledge* about the causal relationships among individual hypotheses to deduce qualitative conclusions about the ordering of local composite hypotheses. By doing so, we can employ "divide and conquer" strategy to derive the global ordering of the composite hypotheses from a set of local ordering in which consistencies are guaranteed. We view the contribution of this research is on the integration of qualitative reasoning with the use of local computations to find not only the most likely composite hypotheses, but also the partial ordering of the composite hypotheses.

## I. Introduction

A Bayesian network [Pearl 86,87] is a graphical representation of probabilistic knowledge about the causal relationships among a set of variables (propositions) in an expert system. Each of these variables accounts for a set of possible outcomes, each of which is a hypothesis. A permutation of the outcomes accounted for by different variables is referred to as a <u>composite hypothesis</u>. For example, if the causal relationships among heatstroke (one kind of heat illness) and its pathological states (such as body temperature, level of consciousness, etc.) are represented in terms of a Bayesian network for use in computer aided medical diagnosis, one possible composite hypothesis can be: *not* heatstroke and *high* body temperature and *low* level of consciousness.

The probabilistic inference of a Bayesian network is to derive conclusions about the hypotheses. The conclusions can be the most likely composite hypothesis, or the partial ordering of a set of composite hypotheses, based on its associated Bayesian beliefs which are quantitatively expressed in terms of *a posteriori* (conditional) probabilities. The conclusion about the most likely composite hypothesis is not necessarily sufficient in some applications; e.g, the assessment of design methodologies suggested by a CAD system [Sy, 89] for the development of a nonvocal communication device. In this case, partial ordering of the composite hypotheses is necessary.

Several approaches have been proposed to find the most likely composite hypothesis. These include the use of linear programming [Cooper, 86] and task decomposition [Pearl, 87] based on combining local maximum *causal* and *diagnostic* supports. However, the computational complexities of using linear programming have never been addressed and the consistency of combining local *causal* and *diagnostic* supports to derive an ordering of composite hypotheses cannot be guaranteed. Even worse, finding a partial ordering of all composite hypotheses in a straightforward manner can be exponentially or NP-hard [Cooper, 87][Rege, 88].

In dealing with the class of problems which is NP-hard, three avenues are proposed in a recent research symposium [Sipser, 88]. They are (i) the study of random guess versus rigorous problem-solving algorithms, (ii) computation time versus memory space, and (iii) parallelism. In our previous research [Sy, 88], we explored the potential of random guess and found that the random guess approach can be effective if a set of *good* prediction rules relevant to the constraints of probability theory and the topological structure of a Bayesian network are provided. This finding motivates the study of prediction rules which are encoded as meta — knowledge to indicate the qualitative change of the Bayesian beliefs of the composite hypotheses with respect to the change of the *a posteriori* probability of each individual hypothesis.

In section II the properties of Bayesian networks and their computational problems are described. In section III we will introduce the notion of *conditional influence* which formulates meta-knowledge about the causal relationships among *local* composite hypotheses. The details of qualitative reasoning using meta-knowledge are presented in section IV. In section V the mechanism of qualitative reasoning is demonstrated through an example illustration. Finally, the conclusions are summarized in section VI.

## II. Bayesian Networks and Problem Overview

Bayesian networks are acyclic graphs within which a set of nodes are connected by a set of arcs. The nodes in the graph represent variables (propositions) and the arcs signify causal dependencies among the probabilistic variables. Each variable is denoted by a lower case letter and is quantified by a set of discrete values so that each value corresponds to one hypothesis — or one possible outcome accounted for by the variable; for example, the variable relevant to the body temperature can be *high, normal,* or *low.* For the sake of discussion, each variable is assumed to have only two possible values (e.g. true or false). The uppercase letter, such as $X$ and $\bar{X}$ are used to abbreviate $x = X$ and $x = \bar{X}$ respectively. A simple Bayesian network is shown in Fig. 1.

Both the arcs and nodes in a Bayesian network are quantified by probability functions. The probability function of a node with variable $x$ is the *a priori* probability distribution $Pr(x)$. The probability function of an arc connecting a node with variable $y$ to a node with variable $x$ is the conditional probability $Pr(x|y)$ [1]. The joint distribution of the variables can be obtained by multiplying appropriate probability functions together. For example, $Pr(abcdef) = Pr(a|bc)Pr(b)Pr(c)Pr(d|c)Pr(e|df)Pr(f)$ in Fig. 1.

A permutation of the variables in a Bayesian network can be a composite hypothesis (H) or an evidence ($S_e$) depending upon whether the variables are observable or not. In general, a composite hypothesis refers to a permutation of a set of quantified variables which include all the variables in a Bayesian network except those appears in $S_e$. However, if this is not the case, such a composite hypothesis is referred to as local composite hypothesis. For example, there are 6 variables (a, b, c, d, e, f) in Fig. 1; if $S_e$ = *EF,* then *ABCD* is a composite hypothesis while *BCD* and *DEF* are local composite hypotheses [2].

The inference process of a Bayesian network is based on computing $Pr(H|S_e)$ to derive the conclusions about the most likely $H^*$ (i.e. $Max_i[Pr(H_i|S_e)]$), or the partial ordering of all $H_i$s, given $S_e$. When the Bayesian beliefs of all $H_i$s (i.e. $Pr(H_i|S_e)$) are computed, a complete partial ordering of all $H_i$s can be generated. In above example, 16 combinations of different values of the variables, a, b, c d (given $S_e = \bar{E}F$) have to be considered in order to find a complete partial ordering of all $H_i$s. When k the number of variables in $H_i$ — increases, the number of Bayesian beliefs to be considered correspondingly increases in the order of $2^k$. It is clear the computational load will soon be a problem when k becomes large. It is unlikely in any application that we need to find a complete partial ordering of the composite hypotheses. However, if we are interested in, for example, the largest four $Pr(abcd|\bar{E}F)$, applying straightforward approach (i.e. exhaustive evaluations of all $Pr(abcd|\bar{E}F)$) will still require the consideration of all combinations which complexity is in the order of $2^k$. In next section, we will discuss a set of meta-rules which reduces the computational complexities of deriving a partial ordering of composite hypotheses.

[1] Whenever necessary, the conditional joint probability will also be stored; for example, $Pr(a|bc)$ and $Pr(e|df)$ in Fig. 1 and table 1.

[2] Note that a permutation of all variables in a Bayesian network is also a local composite hypothesis by definition. However, this special case is useless to our discussions, thus is excluded from our considerations of local composite hypotheses.

## III. Prediction Rule and Conditional Influence

Within the framework of probability theory, Bayesian networks exhibit two properties which lead to a prediction rule for the derivation of a partial ordering of composite hypotheses.

(i) The composite hypotheses, $H_i$s, of a Bayesian networks are mutually exclusive to each other. Mathematically,

$$Pr(\cup_i H_i|S_e) = \sum_i Pr(H_i|S_e) \quad (1)$$

(ii) The set of all $H_i$, $H = \cup_i H_i$, exhaustively covers all the possible combinations of different values of the variables in a Bayesian network. That is:

$$\sum_{i=1}^{2^k} Pr(H_i|S_e) = Pr(H_1 \cup ... \cup H_{2^k}|S_e) = 1 \quad (2)$$

Remark: Consider $Pr(H_i = abcd|EF)$ discussed above, $H_1 \cup ... \cup H_{16}$ (for k = 4) will lead to $A \cup \bar{A}$, $B \cup \bar{B}$, $C \cup \bar{C}$, $D \cup \bar{D}$. The Bayesian belief of any of these (i.e. $Pr(H|S_e)$) is 1.

Based on the equations (1) and (2), we can derive the following prediction rule:

PRR1 (Prediction rule): Given a partial ordering of the probability of m (out of $2^k$) mutually exclusive hypotheses $Pr(H_1|S_e) \geq Pr(H_2|S_e) \geq ... \geq Pr(H_m|S_e)$, there is a $\frac{\alpha-\beta}{C}$ probability that the largest $Pr(H_i|S_e)$ is within the bound $[\beta, \alpha]$ where
$H_i \in \{H_{m+1} ... H_{2^k}\}$,
$C = 1 - Pr(H_1 \cup ... \cup H_m|S_e)$,
$\alpha$ and $\beta$ are some constants such that $Pr(H_{m+1} \cup ... \cup H_{2^k}|S_e) \geq \alpha \geq \beta$.
Remark: The proof of above prediction rule (PRR1) can be found in [Sy, 88].

Consider an extreme case that $m = 1$ and $Pr(H_1|S_e) > 0.5$. This implies that the value of C in PRR1 is less than 0.5. If we set $\beta = 0$, $\alpha = Pr(H_2 \cup ... \cup H_{2^k}|S_e)$, then PRR1 indicates that $H_1$ is the most likely composite hypothesis with certainty. Next suppose we have $m = 3$ and $Pr(H_1|S_e) = 0.35 \geq Pr(H_2|S_e) = 0.2 \geq Pr(H_3|S_e) = 0.15$. Now $Max[Pr(H_4 \cup ... \cup H_{2^k}|S_e)] = C = 0.3$. If we set $\beta = 0$, $\alpha = Max[Pr(H_4 \cup ... \cup H_{2^k}|S_e)] \geq Max_{i \neq 1,2,3}[Pr(H_i|S_e)]$, we will once again find that $Pr(H_1|S_e)$ is the most likely composite hypothesis with certainty.

From the example shown above, we can see that the rate of reaching a conclusion (about finding either the most likely composite hypothesis or the partial ordering of composite hypotheses) depends on how and what composite hypotheses are selected for evaluation. That is, if the composite hypotheses selected for evaluation have relatively large values of Bayesian beliefs, the faster a conclusion can be reached. Referring to the discussion in Section II and Fig. 1, the Bayesian belief of any arbitrary composite hypothesis, for example, $Pr(abcd|S_e = \bar{E}F)$, can be rewritten as a joint probability normalized by *a priori* probability of the variables in $S_e$ (Bayes rule). That is:

$Pr(abcd|S_e)$
$= \frac{1}{Pr(S_e)} Pr(a|bc)Pr(b)Pr(c)Pr(d|c)Pr(\bar{E}|dF)Pr(F)$ (3)

Each term in the right side (except $\frac{1}{Pr(S_e)}$) can be considered as a probability function for a local composite hypothesis. The change of the Bayesian belief of a permutation of variables (i.e. abed) depends on the change of the local probabilities (i.e. the terms in the right hand side). The relative changes and the bound of changes are the meta-knowledge for qualitative reasoning and are encoded in terms of "conditional influence".

The probability function for a local composite hypothesis has a general form $Pr(x|y_1y_2...y_n)$ as is observed in (3). Each variable, $x$ or $y_i$, can have two possible values; i.e. $X$, $\bar{X}$ or $Y_i$, $\bar{Y}_i$ respectively [3]. Given the value of $x$, let's say, $x = X$, we can compare two local probabilities with different values of $y_i$, let's say, $Pr(X|Y_1Y_2)$ and $Pr(X|\bar{Y}_1\bar{Y}_2)$ (for $n = 2$). Suppose $Pr(X|Y_1Y_2) \geq Pr(X|\bar{Y}_1\bar{Y}_2)$, we denote such an inequality by $\langle X|Y_1Y_2, X|\bar{Y}_1\bar{Y}_2\rangle$. Similarly, $\langle Y_1|Z_1, Y_1|\bar{Z}_1\rangle$ denotes $Pr(Y_1|Z_1) \geq Pr(Y_1|\bar{Z}_1)$, and $\langle Z_1, \bar{Z}_1\rangle$ denotes $Pr(Z_1) \geq Pr(\bar{Z}_1)$. To simplify our discussion, we use $\mathbf{Y_i}$ to represent one permutation of the quantified value of $y_1y_2...y_n$. For example, $\mathbf{Y_1} = \bar{Y}_1\bar{Y}_2$, $\mathbf{Y_2} = \bar{Y}_1Y_2$, $\mathbf{Y_3} = Y_1\bar{Y}_2$, and $\mathbf{Y_4} = Y_1Y_2$ for $n = 2$. In addition, we use "I" to denote the collection of all $\langle\bullet\rangle$s. Using these notations, Conditional influence can be defined as a relation over $I$x$I$ as follows:

If there exists $\langle X|\mathbf{Y_i}, X|\mathbf{Y_j}\rangle$ and $\langle Y_k|\mathbf{Z_l}, Y_v|\mathbf{Z_l}\rangle$ such that $Y_k$ occurs in $\mathbf{Y_i}$ and $Y_v$ occurs in $\mathbf{Y_j}$ (for any $k \neq v$), then these two $\langle\bullet\rangle$s are related to each other under conditional influence relation and the relation is represented as an ordered pair shown below [4]:

$$(\langle X|\mathbf{Y_i}, X|\mathbf{Y_j}\rangle, \langle Y_k|\mathbf{Z_l}, Y_v|\mathbf{Z_l}\rangle)$$

Referring to the Bayesian network in Fig. 1 and the numerical values of the probabilistic information shown in table 1, there are 19 such relations. They are:

CI01 : $(\langle A|\bar{B}\bar{C}, A|B\bar{C}\rangle, \langle B, B\rangle)$
CI02 : $(\langle A|\bar{B}C, A|B\bar{C}\rangle, \langle \bar{B}, B\rangle)$
CI03 : $(\langle A|\bar{B}C, A|B\bar{C}\rangle, \langle C, \bar{C}\rangle)$
CI04 : $(\langle \bar{A}|BC, \bar{A}|B\bar{C}\rangle, \langle C, C\rangle)$
CI05 : $(\langle \bar{A}|BC, \bar{A}|\bar{B}\bar{C}\rangle, \langle C, \bar{C}\rangle)$
CI06 : $(\langle \bar{A}|\bar{B}C, \bar{A}|\bar{B}\bar{C}\rangle, \langle C, \bar{C}\rangle)$
CI07 : $(\langle A|\bar{B}, A|B\rangle, \langle \bar{B}, B\rangle)$
CI08 : $(\langle \bar{A}|C, \bar{A}|\bar{C}\rangle, \langle C, \bar{C}\rangle)$
CI09 : $(\langle \bar{D}|C, \bar{D}|\bar{C}\rangle, \langle C, \bar{C}\rangle)$
CI10 : $(\langle E|D\bar{F}, E|\bar{D}\bar{F}\rangle, \langle D|\bar{C}, \bar{D}|\bar{C}\rangle)$
CI11 : $(\langle E|D\bar{F}, E|\bar{D}\bar{F}\rangle, \langle D|\bar{C}, \bar{D}|\bar{C}\rangle)$
CI12 : $(\langle \bar{E}|DF, \bar{E}|\bar{D}F\rangle, \langle D|\bar{C}, \bar{D}|\bar{C}\rangle)$
CI13 : $(\langle \bar{E}|DF, \bar{E}|D\bar{F}\rangle, \langle F, \bar{F}\rangle)$
CI14 : $(\langle \bar{E}|\bar{D}F, \bar{E}|D\bar{F}\rangle, \langle F, \bar{F}\rangle)$
CI15 : $(\langle \bar{E}|DF, \bar{E}|\bar{D}\bar{F}\rangle, \langle D|\bar{C}, \bar{D}|\bar{C}\rangle)$
CI16 : $(\langle \bar{E}|DF, \bar{E}|\bar{D}\bar{F}\rangle, \langle F, \bar{F}\rangle)$
CI17 : $(\langle E|D, E|\bar{D}\rangle, \langle D|\bar{C}, \bar{D}|\bar{C}\rangle)$
CI18 : $(\langle \bar{E}|\bar{D}, \bar{E}|D\rangle, \langle \bar{D}|C, D|C\rangle)$
CI19 : $(\langle \bar{E}|F, E|F\rangle, \langle F, \bar{F}\rangle)$

In next section we will show how the conditional influence relations are used in qualitative reasoning.

## IV. Recognition, Combination, and Propagation Rules

Using the conditional influence relations discussed in Section III, we can examine the qualitative change of the Bayesian belief of a composite hypothesis (i.e. a permutation of quantified variables) with respect to the change of the probabilities of local composite hypotheses. For example, when the composite hypothesis accounted for by abed is changed from ABCD to ABCD, CI04 and CI09 indicate an increase in the probabilities $Pr(a|bc)$, $Pr(d|c)$, and $Pr(c)$, thus the $Pr$ $(abcd \backslash S_e)$.

The selection of composite hypotheses for evaluation during the reasoning process can be based on the conditional influence relations which induce partitions [5] among the local composite hypotheses. For example, the Bayesian belief of the composite hypothesis ABCD (given $S_e$ = EF) in Fig. 1 can be expressed as below:

$$Pr(\bar{A}BC\bar{D}|S_e = \bar{E}F) = \frac{Pr(\bar{A}|BC)Pr(B)Pr(C)Pr(\bar{D}|C)Pr(\bar{E}|\bar{D}F)Pr(F)}{Pr(\bar{E}F)}$$

There are six terms ( e x c l u $\frac{1}{Pr(\bar{E}F)}$)i n the right hand side of above expression. These terms can be partitioned into 4 classes according to the conditional influence relations. They are:

| | |
|---|---|
| Class 1 (according to CI04) : | $Pr(\bar{A}|BC)Pr(C)$ |
| Class 2 (not exist in CI) : | $Pr(B)$ |
| Class 3 (according to CI09) : | $Pr(\bar{D}|C)Pr(C)$ |
| Class 4 (according to CI14) : | $Pr(\bar{E}|\bar{D}F)Pr(F)$ |

Formally, a class refers to the collection of local composite hypotheses which achieve local consistency; i.e., the change of the probability value of one variable (due to the change of its quantified value) will lead to the same qualitative change of the overall probability of the class. For example, when the value of the variable c is changed from C to $\bar{C}$, the values of b o $Pr(c)$ and $Pr(a|c)$ will decrease, thus the overall probability $Pr(a|bc)Pr(c)$ in class 1 will decrease. Similarly, when the value of a is changed from A to A, the probability $Pr(a|bc)$ increases, so as $Pr(a|bc)Pr(c)$.

It is possible to change the value of one variable in

[3] Note that the discussion here can be generalized for variables with multi-values other than binary values.

[4] If $Y_k$ and $Y_v$ have no "parents" in the Bayesian network, then the relation will be $(\langle X|\mathbf{Y_i}, X|\mathbf{Y_j}\rangle, \langle Y_k, Y_v\rangle)$.

[5] The partition is not necessarily mutually exclusive.

one class, which simultaneously leads to the change of the overall probabilities of other classes in an inconsistent way. For example, when the value of *d* is changed from D to D, $Pr(\bar{E}|dF)$ will increase as does the overall probability of class 4. However, this will also cause the decrease of the probability *Pr{d\c)* in class 3, thus causing a decrease in the overall probability of class 3.

In order to determine which composite hypotheses should be selected for evaluation, we need to identify those composite hypotheses with which local composite hypotheses are consistent not only within a class, but also among classes. In the selection of composite hypotheses for evaluation, three kinds of rules are required in qualitative reasoning. They are (i) the recognition rule, (ii) the combination rule, and (iii) the propagation rule.

Recognition rule - This rule is for class composition. It is used to recognize the pattern(s) of conditional influence existing in a given composite hypothesis and to partition it into classes, or to combine two classes together if the local composite hypotheses of two classes are consistent with one another. The recognition rule used for class composition can be formulated as below:

RRI (Recognition rule):

(i) If a local composite hypothesis does not exist in any ordered pairs of conditional influence, it is a separate class by itself.

(ii) Otherwise, a class is generated using the composition operation, $\oplus$, defined as below:

$$(\langle V,U\rangle,\langle W,Z\rangle)\oplus(\langle W,L\rangle,\langle M,N\rangle)\ \Rightarrow\ \ll V,W,M\gg$$

where $\ll \bullet \gg$ denotes a collection of local composite hypotheses which are consistent to each other in the same class (i.e. local consistency).

Combination rule This rule is used during the course of reasoning to seek out new hypothesis (i.e. quantified propositional variable) which can be categorized into a given class. The combination rule is defined through the combination operation ® described below:

CR1 (Combination rule):

$$(\langle V,U\rangle,\langle W,Z\rangle)\otimes(\langle V,L\rangle,\langle M,N\rangle)\ \Rightarrow\ \ll V,W,M\gg$$

Propagation rule This rule is for generating conclusion about the potential change of a Bayesian belief due to the change of the probabilities of the partitioned classes. It is used to propagate the conclusions about the combined qualitative changes of the probabilities of two classes; i.e., increase (↑), steady (-), decrease (↓), or unknown (?). The following table summarizes the qualitative change of the overall probability of a combined class with respect to the change of the probability of each individual class.

| $C_1 \backslash C_2$ | ↑ | − | ↓ | ? |
|---|---|---|---|---|
| ↑ | ↑ | ↑ | ? | ? |
| − | ↑ | − | ↓ | ? |
| ↓ | ? | ↓ | ↓ | ? |
| ? | ? | ? | ? | ? |

Qualitative reasoning is envisioned as a process of applying the rules described above in a certain sequence for the selection of composite hypotheses for evaluation, and for the generation of conclusions about the qualitative values of the Bayesian beliefs of the composite hypotheses. Within the scenario of finding the partial ordering of $Pr(H_i|S_e)$, the mechanism of reasoning follows as below:

Step 1:

Randomly select a new composite hypothesis for examination.

Step 2:

Apply the recognition rule to the composite hypothesis for class partition.

Step 3:

Modify the value of a propositional variable in the composite hypothesis and re examine the local consistency of a class using the combination rule.

Step 4:

Apply the propagation rule to generate a conclusion about the overall change of the Bayesian belief of a permutation of propositional variables.

Step 5:

Re-iterate step 2 to 4 until maximum global consistency is obtained.

Step 6:

Evaluate the quantitative value of the Bayesian belief of the selected composite hypothesis.

Step 7:

Use prediction and/or heuristic rules to determine whether the conclusion is reached and whether the reasoning process can be terminated.

## V. Example

Qualitative reasoning of the Bayesian belief discussed in this paper can be illustrated through the Bayesian network shown in Fig. 1. The problem is to determine the partial ordering of three most likely composite hypothesis given $S_e$ = *EF*. To start the reasoning process, assume *H* = *ABCD* is selected (according to C101) for consideration. The mechanism of qualitative reasoning is shown below:

**QR1:**

Consider CI01, initial setting of *abcd* is $H = A\bar{B}\bar{C}\bar{D}$

Apply the recognition rule to $H = A\bar{B}\bar{C}\bar{D}$

| | |
|---|---|
| Class 1 (according to CI01) : | $Pr(A\|\bar{B}\bar{C})Pr(\bar{B})$ |
| Class 2 (not exist in CI) : | $Pr(\bar{C})$ |
| Class 3 (not exist in CI) : | $Pr(\bar{D}\|\bar{C})$ |
| Class 4 (according to CI14) : | $Pr(\bar{E}\|\bar{D}F)Pr(F)$ |

Consider CI02, the value of *abcd* is modified to $H = A\bar{B}C\bar{D}$

Apply the combination rule to Classes 1 and 2

| | |
|---|---|
| Class 1 (according to CI02, 03) : | $Pr(A\|\bar{B}C)Pr(\bar{B})Pr(C)$ |
| Class 2 (according to CI09) : | $Pr(\bar{D}\|C)Pr(C)$ |

Class 3 (according to CI14) : $Pr(\bar{E}|\bar{D}F)Pr(F)$

Apply the propagation rule to trace the change of $H$ from $A\bar{B}\bar{C}\bar{D}$ to $A\bar{B}C\bar{D}$, we can conclude $Pr(A\bar{B}C\bar{D}|\bar{E}F) \geq Pr(A\bar{B}\bar{C}\bar{D}|\bar{E}F)$.

Evaluate $Pr(A\bar{B}C\bar{D}|\bar{E}F) \Rightarrow Pr(A\bar{B}C\bar{D}|\bar{E}F) = 0.178$.

Mark $A\bar{B}C\bar{D}$, reinvokes the qualitative reasoning process to search another likely composite hypothesis with the same initial consideration; i.e., $H = A\bar{B}\bar{C}\bar{D}$.

**QR2:**

Consider CI01, we get $H = A\bar{B}\bar{C}\bar{D}$ as is in QR1.

Consider CI15, the value of *abcd* is modified to $H = A\bar{B}\bar{C}D$

Apply the combination rule to classes 3 and 4 shown in QR1.

Class 1 (according to CI01) : $Pr(A|\bar{B}\bar{C})Pr(\bar{B})$

Class 2 (not exist in CI) : $Pr(\bar{C})$

Class 3 (based on CI15, 16) : $Pr(D|\bar{C})Pr(\bar{E}|DF)Pr(F)$

Apply the propagation rule to trace the change of $H$ from $A\bar{B}\bar{C}\bar{D}$ to $A\bar{B}\bar{C}D$, we can conclude $Pr(A\bar{B}\bar{C}D|\bar{E}F) \geq Pr(A\bar{B}\bar{C}\bar{D}|\bar{E}F)$.

Evaluate $Pr(A\bar{B}\bar{C}D|\bar{E}F) \Rightarrow Pr(A\bar{B}\bar{C}D|\bar{E}F) = 0.140$.

**QR3:**

Consider CI04, the initial setting of *abcd* is $H = \bar{A}BC\bar{D}$.

Apply the recognition rule to $H = \bar{A}BC\bar{D}$.

Class 1 (according to CI04) : $Pr(\bar{A}|BC)Pr(C)$

Class 2 (not exist in CI) : $Pr(B)$

Class 3 (according to CI14) : $Pr(\bar{E}|\bar{D}F)Pr(F)$

Evaluate $Pr(\bar{A}BC\bar{D}|\bar{E}F) \Rightarrow Pr(\bar{A}BC\bar{D}|\bar{E}F) = 0.159$.

Combining the results obtained from QR1, QR2, and QR3, the partial ordering of three most likely composite hypotheses is [6]:

$$Pr(A\bar{B}C\bar{D}|\bar{E}F) \geq Pr(\bar{A}BC\bar{D}|\bar{E}F) \geq Pr(A\bar{B}\bar{C}D|\bar{E}F)$$

## VI. Conclusion

This paper presents a qualitative reasoning approach for the derivation of partial ordering of composite hypotheses in a Bayesian network. The mechanism of the reasoning process is based on a set of rules which are used to derive qualitative conclusions about the ordering of the causal relationships of the hypotheses encoded as conditional influence relations. This approach features several distinct significances which make it an attractive alternative to be considered for use in probabilistic reasoning. First, the partitions induced by conditional influence permit local reasoning, thus concurrent processing. Second, reasoning in a partitioned class reduces computational complexities, and yet its consistencies between local and global levels are assured. Third, this approach can be used to infer not only the most likely composite hypothesis, but also the ordering of composite hypotheses. Finally, it is equally important to note that (i) the conclusions generated by qualitative reasoning are usually weaker; e.g., the ordering of the composite hypotheses are known but not their quantitative values, (ii) the efficiency of this approach depends on both the topological structures of a Bayesian network and its probability distribution, and (iii) several other approaches are available and equally attractive if our interest is only on the most likely composite hypothesis.

[6] A complete partial ordering of *Pr(abcd\EF)* is shown in table 2.

## VI. References


[Cooper, 86] G.F. Cooper, "A Diagnostic Method that uses Causal Knowledge and Linear Programming in the Application of Bayes' Formula," *Computer Methods and Programs in Biomedicine,* Vol. 22, pp. 223-237, 1986.

[Cooper, 87] G.F. Cooper, "Probabilistic Inference Using Belief Networks is NP-hard," *Knowledge Systems Laboratory, Stanford University, Memo KSL-87-27,* May 1987 (Revised Aug. 1987).

[Pearl, 86] J. Pearl, "Fusion, Propagation, and Structuring in Bayesian Networks," *Artificial Intelligence,* Vol. 28, pp. 241-288, 1986.

[Pearl, 87] J. Pearl, "Distributed Revision of Composite Beliefs," *Artificial Intelligence,* Vol. 33, pp. 173-215, 1987.

[Rege, 88] A. Rege and A.M. Agogino, "Topological Framework for Representing and Solving Probabilistic Inference Problems in Expert Systems," *IEEE Trans, on Systems, Man, and Cybernetics,* May/June 1988, Vol. 18, pp. 402-414.

[Sipser, *88}* M.F. Sipser, "We still don't know if P = NP", *MIT Computer Science Research Symposium,* Oct. 1988.

[Sy, 88] B.K. Sy, "Qualitative Reasoning of Quantitative Uncertainty in Bayesian Networks," *Proc. of the International Computer Science Conf.,* Hong Kong, December 1988.

[Sy, 89] B.K. Sy, "An Al-based CAD/CAM Approach to Assess Design Methodology of a User-specific Nonvocal Communication Device," *Proc. of 15th Northeast Bioengineering Conf.,* Boston, March 1989.

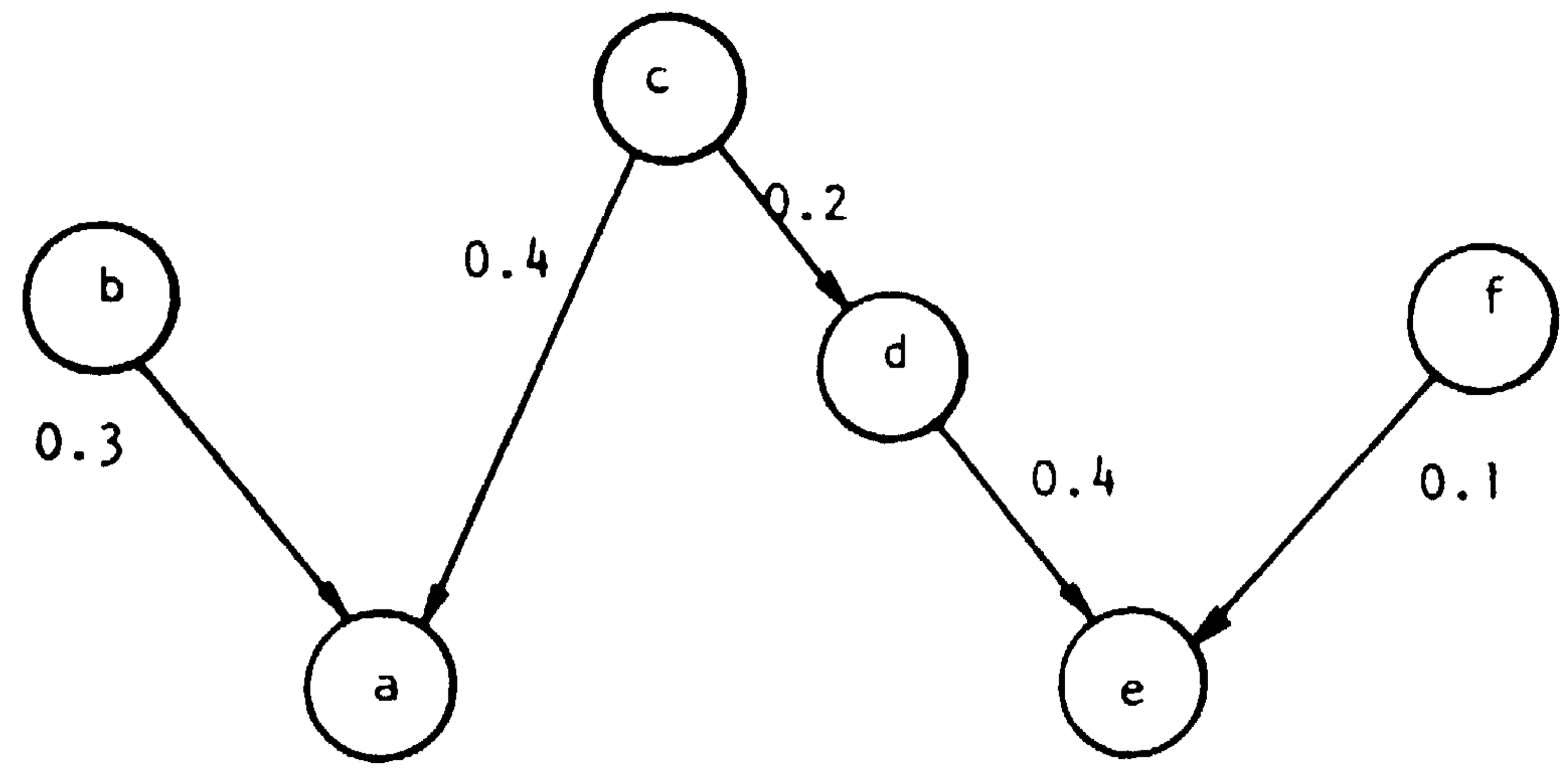


Fig. 1 : Bayesian Network

| | | | |
|---|---|---|---|
| $Pr(B) = 0.4$ | $Pr(C) = 0.65$ | $Pr(F) = 0.55$ | |
| $Pr(D\|C) = 0.2$ | $Pr(D\|\bar{C}) = 0.8$ | | |
| $Pr(A\|B) = 0.3$ | $Pr(A\|\bar{B}) = 0.65$ | $Pr(A\|C) = 0.4$ | $Pr(A\|\bar{C}) = 0.714$ |
| $Pr(E\|D) = 0.4$ | $Pr(E\|\bar{D}) = 0.17$ | $Pr(E\|F) = 0.1$ | $Pr(E\|\bar{F}) = 0.465$ |
| $Pr(A\|\bar{B}\bar{C}) = 0.8$ | $Pr(A\|\bar{B}C) = 0.6$ | $Pr(A\|B\bar{C}) = 0.4$ | $Pr(A\|BC) = 0.2$ |
| $Pr(E\|\bar{D}\bar{F}) = 0.26$ | $Pr(E\|\bar{D}F) = 0.14$ | $Pr(E\|D\bar{F}) = 0.74$ | $Pr(E\|DF) = 0.06$ |
| $Pr(abcdef) = Pr(a\|bc)Pr(b)Pr(c)Pr(d\|c)Pr(e\|df)Pr(f)$ | | | |

Table 1 : Probabilistic Knowledge of Fig. 1

$$
\begin{aligned}
&Pr(A\bar{B}C\bar{D}|\bar{E}F) \geq Pr(\bar{A}BC\bar{D}|\bar{E}F) \geq Pr(A\bar{B}\bar{C}D|\bar{E}F) \geq Pr(\bar{A}\bar{B}C\bar{D}|\bar{E}F) \geq \\
&\quad Pr(\bar{A}B\bar{C}D|\bar{E}F) \geq Pr(A\bar{B}CD|\bar{E}F) \geq Pr(AB\bar{C}D|\bar{E}F) \geq Pr(\bar{A}BCD|\bar{E}F) \geq \\
&\quad\quad Pr(ABC\bar{D}|\bar{E}F) \geq Pr(\bar{A}\bar{B}\bar{C}D|\bar{E}F) \geq Pr(\bar{A}\bar{B}CD|\bar{E}F) \geq Pr(A\bar{B}\bar{C}\bar{D}|\bar{E}F) \geq \\
&\quad\quad\quad Pr(\bar{A}B\bar{C}\bar{D}|\bar{E}F) \geq Pr(ABCD|\bar{E}F) \geq Pr(AB\bar{C}\bar{D}|\bar{E}F) \geq Pr(\bar{A}\bar{B}\bar{C}\bar{D}|\bar{E}F)
\end{aligned}
$$

Table 2 : Complete Partial Ordering of $Pr(abcd|\bar{E}F)$