

Extending Qualitative Probabilistic Network with Mutual Information Weights

Kun Yue¹, Feng Wang², Mujin Wei¹, Weiyi Liu¹

¹Department of Computer Science and Engineering, School of Information Science and Engineering, Yunnan University, Kunming, China

²Yunnan Computer Technology Application Key Lab, Kunming University of Science and Technology, Kunming, China

Email: kyue@ynu.edu.cn

Received 15 January 2015; accepted 30 January 2015; published 3 February 2015

Academic Editor: Prof. Zhongzhi Shi, Institute of Computing Technology, Chinese Academy of Sciences, China

Copyright © 2015 by authors and Scientific Research Publishing Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

Abstract

Bayesian network (BN) is a well-accepted framework for representing and inferring uncertain knowledge. As the qualitative abstraction of BN, qualitative probabilistic network (QPN) is introduced for probabilistic inferences in a qualitative way. With much higher efficiency of inferences, QPNs are more suitable for real-time applications than BNs. However, the high abstraction level brings some inference conflicts and tends to pose a major obstacle to their applications. In order to eliminate the inference conflicts of QPN, in this paper, we begin by extending the QPN by adding a mutual-information-based weight (MI weight) to each qualitative influence in the QPN. The extended QPN is called MI-QPN. After obtaining the MI weights from the corresponding BN, we discuss the symmetry, transitivity and composition properties of the qualitative influences. Then we extend the general inference algorithm to implement the conflict-free inferences of MI-QPN. The feasibility of our method is verified by the results of the experiment.

Keywords

Qualitative Probabilistic Network (QPN), Inference Conflict, Mutual Information, Influence Weight, Superposition

1. Introduction

Bayesian network (BN) is a well-accepted model to represent a set of random variables and their probabilistic relationships via a directed acyclic graph (DAG) [1] [2]. There is a conditional probability table (CPT) asso-

ciated with each node to represent the quantitative relationships among the nodes. BN has been widely used in different aspects of intelligent applications including representation, prediction, reasoning, and so on [1] [3]. However, exact inference in an arbitrary BN is NP hard [4], and even the approximate inference is also very difficult [5]. As a result, BN's inferences are not very efficient, which is doomed to pose limits to BN-related applications, especially for real-time situations. Fortunately, as the qualitative abstraction of BN, QPN was proposed by Wellman [6] and can be appropriately used to remedy the above drawbacks of a BN. A QPN adopts a DAG to encode random variables and their relationships, which are not quantified by conditional probabilities, but are summarized by qualitative signs [7]. The most important thing is that there has been an efficient QPN inference (reasoning) algorithm, which is based on sign propagation and run in polynomial time [8].

However, the high abstraction level of QPN results in the problem that there is no information left to compare two different qualitative influences in a QPN. Thus, when a node receives two inconsistent signs from its two different neighbor nodes during inferences, it is hard to know which sign is more suitable for the node, so that the “?” (*i.e.*, unknown) sign is obtained as the result sign on the node. This means that there are inference conflicts generated on the node, which lead to less powerful expressiveness and inference capabilities than expected for real-world prediction and decision-making of uncertain knowledge in economics, health care, traffics, etc [9]-[16]. Worst of all, when the ambiguous results caused by inference conflicts are produced, they will be spread to most parts of the network with the reasoning algorithm going on.

To provide a conflict-free inference mechanism of QPN is paid much attention. Various approaches have been proposed from various perspectives. However, in some representative methods [9] [11] [13], the original sample data or threshold values are required or predefined when the influences are weighted quantitatively. This is not consistent with the BN-based applications that focus on probabilistic inferences taking as input BN without the original sample data. Furthermore, a BN is frequently defined by experts instead of learned from data. Even though the original data are available, the efficiency of the process of deriving influence weights is sensitive to the data size. Meanwhile, the uncertainty of the influence weight has not been well incorporated, and the propagation of influence weights lacks solid theoretical foundations during QPN inferences.

Therefore, in this paper we are to consider eliminating the inference conflicts of general QPNs and developing a conflict-free inference method. We derive the quantitative QPN by adding a weight to each qualitative influence from the corresponding BN, while sample data and threshold values are not required. The weight is adopted as the information to compare two different qualitative signs. Therefore, when a node in a general QPN faces an inference conflict, a trade-off will be incorporated to avoid the conflict based on the weights.

Mutual information (MI) is a quantity that measures the mutual dependence of the two random variables [17] [18]. As well, MI is widely applied to test the association degree among the BN nodes [19]-[21]. When node *A* receives two inconsistent signs from its two different neighbor nodes, say *B* and *C*, we can naturally think that the result sign on *A* is more dependent on the sign propagated from *B* than that from *C*, if the mutual information between *A* and *B* is larger than that between *A* and *C*. Based on the above idea, we adopt MI as the weight to quantify the degree of the mutual dependence between two nodes linked by a directed edge.

Generally speaking, the main contributions of this paper can be summarized as follows:

- We define the MI-based weight of a qualitative influence, called MI weight, and extend the traditional QPN by adding a MI weight to each qualitative influence. We call the extended QPN as MI-QPN.
- We propose an efficient algorithm to derive MI weights from the conditional probability tables and prior probability distributions in the corresponding BN instead of the sample data.
- We discuss the symmetry, transitivity and composition properties of qualitative influences in the MI-QPN. Then, we extend the general QPN's inference algorithm to achieve conflict-free inferences with the MI-QPN.
- We give preliminary experiments to verify the feasibility and correctness of our method.

2. Related Work

BN has been successfully established as a framework to describe, manage uncertainty using the probabilistic graphical approach [2] [3]. As the qualitative abstraction of BN, QPN [6] was proposed by Wellman and widely used in decision making, industrial control, forecasting, and so on [22]-[24]. The methods to construct QPNs mainly include deriving QPNs from the corresponding existing BNs [15] or giving QPNs by the domain experts. As for the QPN's inference, an efficient algorithm was proposed by Druzdzel *et al.* [8]. The inference conflict has become a common problem for QPNs and how to resolve the inference conflicts is the subject that was paid much attention in recent years. Various methods were proposed from various perspectives.

Lv *et al.* [9] proposed the ambiguity reduction method by associating a qualitative mutual information weight to each qualitative influence in the QPN, but the weight can only be obtained from the given or generated sample data. This is not always feasible, since the BN may be not learned from data, or the sample data cannot be available. Parsons [10] introduced the concept of categorical influence, which is either an influence that serves to increase a probability to 1 or an influence that decreases a probability to 0, regardless of any other influences. By this approach, only some of the inference conflicts can be resolved.

Renooij *et al.* [11] [15] proposed two methods to solve the inference conflicts by using both kappas values and pivots to zoom upon inferences. The kappas value based method distinguishes the strong and weak influences by the interval values that do not have solid theoretical basis for interval value propagation during QPN inferences. Renooij *et al.* [12] presented an enhanced QPN that differs from a regular QPN to distinguish between strong and weak influences. This method is only suitable for the binary-variable situation. A threshold value is required in these methods except that exploits context-specific information. Renooij *et al.* [13] proposed SQPN (semi-QPN) by associating a probability interval value with each qualitative influence to quantify the qualitative influences. De Campos *et al.* [25] discussed the complexity of inferences in polytree-shaped SQPNs. Renooij *et al.* [14] extended QPNs by providing the inclusion of context-specific information about influences and showed that exploiting this information upon inference had the ability to forestall unnecessarily weak results. By this method, a threshold value should be given to distinguish the strong and weak influences.

We extended QPN to solve inference conflicts by adding the weights to qualitative influences based on the rough set theory and interval probability theory respectively in [16] and [26]. By the method in [16], the dependency degree was computed based on the threshold value, but the corresponding sample data was required to obtain the weights. By the method in [26], the interval-probability weights were derived from the corresponding BN's CPT, but the calculation of interval probability during inferences led to approximate results.

3. Preliminaries and Problem Statement

3.1. The Concept of QPN

A QPN has the same graphical structure as the corresponding BN, also represented by a DAG. In a QPN, each node accords with a random variable, and the influence between each pair of the nodes can only be one of the signs including “+”, “−”, “?” and “0”, where “+” (“−”) means the probability of a higher (lower) value for the corresponding variable increases, sign “?” denotes the unknown influence by giving an evidence and “0” represents initial state of the variable without observations. Each edge with a qualitative sign means the qualitative influence between two corresponding variables. The definition of the qualitative influence [6] is given as follows.

Definition 3.1 We say that A positively influences C , written $S^+(A, C)$, if for all values $a_1 > a_2$, c_0 and x , which is the set of all of C 's parents other than A , $P(C > c_0 | a_1 x) \geq P(C > c_0 | a_2 x)$.

This definition means the probability of a higher value of C is increased when given a higher value of A , regardless of any other direct influences on C . A negative qualitative influence S^- and a zero qualitative influence S^0 are defined analogously, by replacing \geq in the above formula by \leq and $=$ respectively. If the qualitative influence between A and C does not belong to the above three kinds, written $S^?(A, C)$.

Example 3.1 Based on **Definition 3.1** and from the BN shown in **Figure 1**, we derive a QPN shown in **Figure 2**.

3.2. QPN Inference and Inference Conflicts

It is known that the qualitative influence of a QPN exhibits various useful properties [6]. The *symmetry* prop-

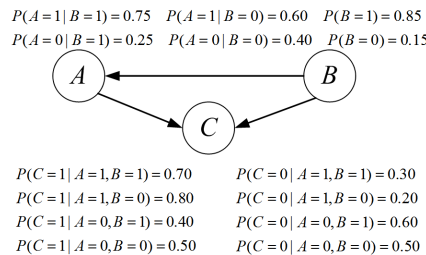


Figure 1. An simple BN.

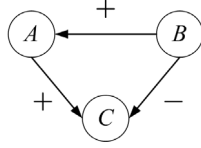


Figure 2. QPN derived from the BN in Figure 1.

erty expresses if there is $S^\delta(A, B)$ in a QPN, then we have $S^\delta(B, A)$. The *transitivity* property can be used to combine qualitative influences along a trail without a head-to-head node into a single influence with the \otimes -operator. The *composition* property asserts that multiple qualitative influences between two nodes along parallel chains combine into a single influence with the \oplus -operator. The ambiguous results (“?”) maybe appear by \oplus -operator during inference, and then will be propagated to other nodes by \otimes -operator.

Building on these three properties and operators, Druzdzel *et al.* proposed an efficient inference algorithm based on sign propagation [8]. The algorithm traces the effect of observing a value for one node on the other nodes by message-passing between neighbors. When the evidences are given to the network, each node receiving a message updates its sign and subsequently sends a message to each (induced) neighbor node that is independent of the observed node. The sign propagation reach all the nodes that are not d-separated with the evidence nodes. During the process of sign propagation, each node has changed its sign at most twice.

The inference conflicts take place when a node receives two different kinds of qualitative signs (“+” and “-”). In fact, the weights of qualitative influences in a QPN are not always equivalent. Thus, in this paper, we will add a weight to each qualitative influence so that a node facing a conflict will take a sign (“+” or “-”) instead of “?” by comparing the corresponding weights.

4. Constructing MI-QPN

4.1. Concepts of Mutual Information

It is well known that information entropy quantifies the information contained in a message and is a measure of the uncertainty associated with a random variable. Now we introduce relevant definition [17] and properties [1].

Definition 4.1 The information entropy H of a discrete random variable X with possible values $\{x_1, x_2, \dots, x_n\}$ is

$$H(X) = -\sum_{i=1}^n P(X = x_i) \log_2 P(X = x_i) \quad (1)$$

Definition 4.2 The conditional information entropy of Y given X is

$$H(Y|X) = -\sum_{i=1}^I \sum_{j=1}^J P(X = x_i) P(Y = y_j | X = x_i) \log_2 P(Y = y_j | X = x_i) \quad (2)$$

where X and Y are two discrete random variables, $x_i \in \{x_1, \dots, x_I\}$, $y_j \in \{y_1, \dots, y_J\}$ and I (or J) denotes the total number of possible values of X (or Y).

We know $H(X)$ is regarded as a measure of uncertainty about a random variable X , and $H(X|Y)$ denotes the amount of uncertainty remaining about X when Y is known. The mutual information is just defined as the difference between $H(X)$ and $H(X|Y)$, which is to denote the amount of uncertainty in X that is removed by knowing Y .

Definition 4.3 The mutual information of two discrete random variables X and Y is defined as

$$I(X, Y) = H(Y) - H(Y|X) \quad (3)$$

4.2. Defining MI-QPN

It is known that the dependency quantity X on Y is the same with that of Y on X . However, the same dependency degree may have different influences on X and Y respectively. This means that the weight of $S^\delta(X, Y)$ does not always equal that of $S^\delta(Y, X)$, although we always have $I(X, Y) = I(Y, X)$. Intuitively, for any node Y with two different neighbor nodes X and Z , if the degree of the mutual dependence between Y and X is greater than that between Y and Z , then the weight of $S^\delta(X, Y)$ should be greater

than that of $S^\delta(Z, Y)$. Consequently, we know the weights of qualitative influences should satisfy the following properties: 1) $MI(X \rightarrow Y)$ does not always equal $MI(Y \rightarrow X)$; 2) For any X, Y and Z , $I(X, Y) > (=) I(Z, Y) \Leftrightarrow MI(X \rightarrow Y) > (=) MI(Z \rightarrow Y)$.

Now, we define the weight of a quantitative influence based on a normalized variant of the mutual information. The weight is called MI weight that satisfies the above properties.

Definition 4.4 The MI weight of a qualitative influence $S^\delta(X, Y)$ is defined as

$$MI(X \rightarrow Y) = \begin{cases} 0, & I(X, Y) = 0; \\ \frac{I(X, Y)}{H(Y)}, & I(X, Y) > 0. \end{cases} \quad (4)$$

From the above definition, we know that the MI weight further satisfies the third property: 3) $MI(X \rightarrow Y) \in [0, 1]$. Especially, $MI(X \rightarrow Y) = 1$ when X depends on Y completely, and $MI(X \rightarrow Y) = 0$ when X and Y are independent.

Then, we consider the concerned computation in each of the three properties of the MI weight as follows:

- For the first property, if $I(X, Y) = I(Y, X) = 0$, we have $MI(X \rightarrow Y) = MI(Y \rightarrow X)$. If $I(X, Y) = I(Y, X) \neq 0$ and $H(X) \neq H(Y)$, we know $H(X) \neq 0$ and $H(Y) \neq 0$. Thus, we have $MI(X \rightarrow Y) \neq MI(Y \rightarrow X)$.
- For the second property, if $I(X, Y) > I(Z, Y)$, we have $I(X, Y) > 0$, since $I(Z, Y) \geq 0$. Then, we have

$$MI(X \rightarrow Y) = I(X, Y)/H(Y), \quad MI(Z \rightarrow Y) = I(Z, Y)/H(Y) \quad (5)$$

since if $I(Z, Y) = 0$, otherwise $MI(Z \rightarrow Y) = 0$. Then, we have

$$MI(X \rightarrow Y) - MI(Z \rightarrow Y) = (I(X, Y) - I(Z, Y))/H(Y) > 0, \quad MI(X \rightarrow Y) > MI(Z \rightarrow Y) \quad (6)$$

If $I(X, Y) = I(Z, Y) = 0$, then we have

$$MI(X \rightarrow Y) - MI(Z \rightarrow Y) = 0 \quad (7)$$

If $I(X, Y) = I(Z, Y) \neq 0$, we have $MI(X \rightarrow Y) = MI(Z \rightarrow Y)$ since

$$MI(X \rightarrow Y) - MI(Z \rightarrow Y) = (I(X, Y) - I(Z, Y))/H(Y) = 0 \quad (8)$$

Therefore, the second property can be derived from **Definition 4.4**.

- For the third property, if $I(X, Y) = 0$, we have

$$MI(X \rightarrow Y) = 0 \quad (9)$$

If $I(X, Y) > 0$, we have

$$MI(X \rightarrow Y) = I(X, Y)/H(Y) = (H(Y) - H(Y|X))/(H(Y)) = 1 - H(Y|X)/H(Y) \quad (10)$$

Thus, we have $I(X, Y) = H(Y) - H(Y|X) \geq 0$, $H(Y) \geq 0$ and $H(Y|X) \geq 0$, and if $H(Y) \neq 0$, we have

$$0 \leq H(Y|X)/H(Y) \leq 1 \quad (11)$$

By Formulae (9), (10) and (11), we can obtain $0 \leq MI(X \rightarrow Y) \leq 1$.

Definition 4.5 If we have $S^\delta(A, B)$, we denote the qualitative influence with an MI weight as $S_{MI(B \rightarrow A)}^\sigma(A, B)$, where $\delta \in \{+, -, ?, 0\}$.

By the symmetry property of the qualitative influence, we know $S^\delta(A, B)$ implies $S^\delta(B, A)$. Based on **Definition 4.5**, we know $S_{MI(B \rightarrow A)}^\sigma(B, A)$. Therefore, two MI weights are added to each QPN edge to repre-

sent bidirectional qualitative influences with the MI weights between the corresponding two nodes.

Definition 4.6 MI-QPN is a DAG $G = (V, Q^w)$, where V is a set of variables (nodes) in the graph, and Q^w is a set of qualitative relations and each qualitative influence in Q^w is associated with a MI weight.

4.3. Deriving MI Weights from BN

We can obtain a prior probability distribution from each orphans node and a CPT from each non-orphans one taking as input the BN directly. If B_0 is a orphans node, we denote the prior probability distribution as set $PT(B_0) = \{P(B_0 = s_0) | 1 \leq s_0 \leq S_0\}$. Otherwise, we denote the CPT of node B_0 with parents B_1, \dots, B_n as the probability set $CT(B_0 | B_1, \dots, B_n) = \{P(B_0 = s_0 | B_1 = s_1, \dots, B_n = s_n) | 1 \leq i \leq n, 1 \leq s_i \leq S_i, 1 \leq s_0 \leq S_0\}$, where S_i denotes the total number of possible states of B_i , and s_i means the s_i -th state of B_i . It is worth noting that B_1, \dots, B_n are independent given the value s_0 of their child node B_0 .

In order to derive the MI weights between B_0 and its parent B_i ($1 \leq i \leq n$), it is necessary to obtain the conditional probability sets associated with B_0 and B_i , denoted as

$$CT(B_0 | B_i) = \{P(B_0 = s_0 | B_i = s_i) | 1 \leq s_i \leq S_i, 1 \leq s_0 \leq S_0\}.$$

It is also necessary to compute the prior probability distribution of B_0 and B_i , denoted as

$$PT(B_0) = \{P(B_0 = s_0) | 1 \leq s_0 \leq S_0\}, \quad PT(B_i) = \{P(B_i = s_i) | 1 \leq s_i \leq S_i\}.$$

Then, we can use $CT(B_0 | B_i)$, $PT(B_0)$ and $PT(B_i)$ to compute the MI weights of the qualitative influences associated with B_0 and B_i according to **Definition 4.4**. Let s_i and S_i be the s_i -th state and the total number of possible states of random variable B_i ($0 \leq i \leq n$), respectively, $s_i \in \{1, \dots, S_i\}$. Let $D_i = \{1, \dots, S_i\}$, if $P(B_1 = s_1, \dots, B_n = s_n) > 0$, for all s_i and for any $m \in \{1, \dots, n\}$, we have

$$P(B_0 = s_0 | B_1 = s_1, \dots, B_m = s_m) = \sum_{s_{m+1}=1}^{S_{m+1}} \sum_{s_{m+2}=1}^{S_{m+2}} \dots \sum_{s_n=1}^{S_n} P(B_0 = s_0 | B_1 = s_1, \dots, B_m = s_m) \prod_{j=m+1}^n P(B_j = s_j)$$

where $P(B_0 = s_0 | B_1 = s_1, \dots, B_m = s_m)$ denotes the probability of B_0 taking the s_0 -th state, given B_i taking the s_i -th state for each i ($i \in \{1, \dots, m\}$), i.e., $P(B_0 = s_0 | B_1 = s_1 \cap \dots \cap B_m = s_m)$.

It can be seen that we need traverse the CPT once to derive the conditional probability set associated with one node and one of its parents. With the CPTs and the prior probability distributions of a BN, we propose **Algorithm 1** to compute the MI weights for each qualitative influence in the MI-QPN.

As the basic operation in **Algorithm 1**, the multiplication operation for computing relevant probabilities in *get conditional probability* is the most time-consuming operation. Let m and n be the number of non-orphans nodes and the maximal in-degree of these nodes respectively. For convenience, we suppose the number of each node's possible values, denoted as s , is the same as that in the corresponding BN. Thus, the total number of rows of each orphans node's CPT is s^{m+1} . For each orphans node, the multiplication operation can be expressed for $f(m, s) = 2 * f(m/s) + s^{m+1}$ times, which leads to the time complexity of **Algorithm 1** in the worst case is $n * f(m, s) < \log_2^{(m)} * s^{m+1} * n$. This means that to compute the MI weights for a QPN, we need scan each conditional probability table in the relevant BN for $\log_2 m$ times at most.

Example 4.3 We consider the MI weights for the QPN in **Figure 2**. By **Algorithm 1**, we can obtain the corresponding MI-QPN shown in **Figure 3**.

5. Conflict-Free Inference with the MI-QPN

It is known that $S^\delta(A, B)$ implies $S^\delta(B, A)$ according to the *symmetry* property. Then, we know there exist $S_{MI(B,A)}^\delta(B, A)$ in the MI-QPN by **Definition 4.4**, since $S^\delta(B, A)$ exists in the corresponding QPN. We can obtain

$$S_{MI(A \rightarrow B)}^\delta(A, B) \Rightarrow S_{MI(B \rightarrow A)}^\delta(B, A) \quad (12)$$

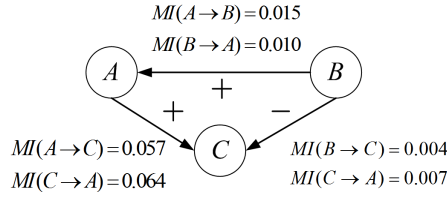


Figure 3. MI-QPN.

Algorithm 1. Deriving MI weights from a BN.**Input:** A BN**Output:** All the MI weights of the MI-QPN**Symbol Description:**

- B_0 : an orphans node in a BN or a non-orphans one with parents B_1, \dots, B_n
- $TP(B_k, \dots, B_m) = \{PT(B_k), \dots, PT(B_m)\}$: the prior probability distribution set of nodes B_k, \dots, B_m
- $CT(B_0|B_k, \dots, B_m) = \{P(B_0 = s_0|B_k = s_k, \dots, B_m = s_m) | 1 \leq s_0 \leq S_0, k \leq i \leq m, 1 \leq s_i \leq S_i\}$: the CPT of B_0 with some of its parents B_k, \dots, B_m
- $TC(B_0|B_k, \dots, B_m) = \{CT(B_0|B_k), \dots, CT(B_0|B_m)\}$: the set of conditional probability sets

Steps:

1. Obtain a topological sequence of all the nodes in the BN
2. **for** each node B_0 in the topological sequence **do**
3. **if** B_0 is a non-orphans node **then**
4. $CT(B_0|B_1, \dots, B_n) \leftarrow \text{get conditional probability}(B_0, TP(B_1, \dots, B_n), CT(B_0|B_1, \dots, B_n))$
5. Compute $PT(B_0)$ with $TC(B_0|B_1)$ and $PT(B_1)$
6. **for** $i \leftarrow 1$ to n **do**
7. Compute $MI(B_0 \rightarrow B_i)$ and $MI(B_i \rightarrow B_0)$ with $TP(B_0|B_1, \dots, B_n)$ and $CT(B_0|B_1, \dots, B_n)$.
8. **end for**
9. **end if**
10. **end for**

Get conditional probability ($B_0, TP(B_k, \dots, B_m), CT(B_0|B_1, \dots, B_n)$)

1. **if** $m < n$ **then**
2. $mid \leftarrow \lceil m-k/2 \rceil$
3. $CT(B_0|B_k, \dots, B_{mid}) \leftarrow \{P(B_0|B_k = s_k, \dots, B_{mid} = s_{mid}) | 1 \leq s_0 \leq S_0, k \leq i \leq mid, 1 \leq s_i \leq S_i\}$
4. $CT(B_0|B_{mid+1}, \dots, B_m) \leftarrow \{P(B_0|B_{mid+1} = s_{mid+1}, \dots, B_m = s_m) | 1 \leq s_0 \leq S_0, mid+1 \leq i \leq n, 1 \leq s_i \leq S_i\}$
5. $TC(B_0|B_k, \dots, B_{mid}) \leftarrow \text{get conditional probability}(B_0, TP(B_k, \dots, B_{mid}), CT(B_0|B_1, \dots, B_{mid}))$
6. $TC(B_0|B_{mid+1}, \dots, B_m) \leftarrow \text{get conditional probability}(B_0, TP(B_{mid+1}, \dots, B_m), CT(B_0|B_{mid+1}, \dots, B_m))$
7. **return** $TC(B_0|B_k, \dots, B_{mid}) \cup TC(B_0|B_{mid+1}, \dots, B_m)$
8. **end if**
9. **return** $CT(B_0|B_k)$

In order to address the *transitivity* property, we consider the MI-QPN fragment in Figure 4, where X denotes the predecessors of B other than A , and Y represents the predecessors of C other than B . From this network, we have $S_{MI(A \rightarrow B)}^{\delta_1}(A, B)$ and $S_{MI(B \rightarrow C)}^{\delta_2}(B, C)$. Based on $S_{MI(A \rightarrow B)}^{\delta_1}(A, B)$ and $S_{MI(B \rightarrow C)}^{\delta_2}(B, C)$, we are to derive the qualitative influence of A on C with an MI weight, denoted as $S_{MI(A \rightarrow C)}^{\delta}(A, C)$.

First, based on QPN's *transitivity* property, we can obtain $\delta = \delta_1 \otimes \delta_2$. Then, we derive the MI weight of qualitative influence A on C , $MI(A \rightarrow C)$. The transitivity property is denoted as

$$S_{MI(A \rightarrow B)}^{\delta_1}(A, B) \& S_{MI(B \rightarrow C)}^{\delta_2}(B, C) \Rightarrow S_{MI(A \rightarrow C)}^{\delta = \delta_1 \otimes \delta_2}(A, C) \quad (13)$$

In order to discuss the *composition* property, we consider the MI-QPN fragment in Figure 5, where X denotes the predecessors of C other than A and B . From the MI-QPN, we have $S_{MI(A \rightarrow C)}^{\delta_1}(A, C)$ and $S_{MI(B \rightarrow C)}^{\delta_2}(B, C)$. Then, we are to derive the combination influence of $S_{MI(A \rightarrow C)}^{\delta_1}(A, C)$ and $S_{MI(B \rightarrow C)}^{\delta_2}(B, C)$, denoted as

$$S_{MI(X \rightarrow C)}^{\delta}(X, C) = S_{MI(A \rightarrow C)}^{\delta_1}(A, C) \vee S_{MI(B \rightarrow C)}^{\delta_2}(B, C) \quad (14)$$

where $X = AB$, $MI(A \rightarrow C)$ denotes the composition weight and " \vee " is the composition operator.

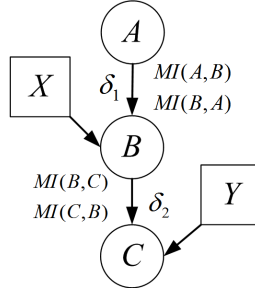


Figure 4. MI-QPN fragment.

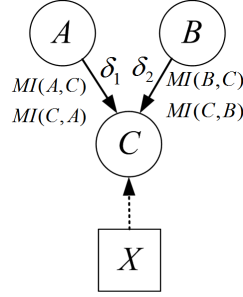


Figure 5. MI-QPN MI-QPN fragment.

Intuitively, the composition operator “ \vee ” ought to satisfy the following properties: 1) The composition weight belongs to $[0, 1]$; 2) The composition operation is commutative; 3) The composition operation is associative; 4) Combining two influences with the same qualitative signs (e.g., two “+” signs or two “-” signs) will result in an influence with a greater MI weight; 5) Combining two influences with different qualitative signs will result in an influence dependent on but less than the larger one. Inspired by the evidence theory and the basic idea of evidence combination [27], we define the composition operator “ \vee ” based on evidence superposition.

Definition 5.1 $S_{MI(X \rightarrow C)}^\delta(X, C) = S_{MI(A \rightarrow C)}^{\delta_1}(A, C) \vee S_{MI(B \rightarrow C)}^{\delta_2}(B, C)$, defined as follows:

- if $\delta_1 = \delta_2$, then

$$\delta = \delta_1 \text{ and } MI(X \rightarrow C) = MI(A \rightarrow C) + MI(B \rightarrow C) - MI(A \rightarrow C) * MI(B \rightarrow C) \quad (15)$$

- if $\delta_1 \neq \delta_2$ and $MI(A \rightarrow C) \geq MI(B \rightarrow C)$, then

$$\delta = \delta_2 \text{ and } MI(X \rightarrow C) = MI(A \rightarrow C) - MI(B \rightarrow C) + MI(A \rightarrow C) * MI(B \rightarrow C) \quad (16)$$

- if $\delta_1 \neq \delta_2$ and $MI(A \rightarrow C) < MI(B \rightarrow C)$, then

$$\delta = \delta_2 \text{ and } MI(X \rightarrow C) = MI(B \rightarrow C) - MI(A \rightarrow C) + MI(A \rightarrow C) * MI(B \rightarrow C) \quad (17)$$

- if $\delta_1 \neq \delta_2$ and $MI(A \rightarrow C) = MI(B \rightarrow C)$, then

$$\delta = \delta_1 \oplus \delta_2 \text{ and } MI(X \rightarrow C) = MI(A \rightarrow C) \quad (18)$$

Based the above properties, we give **Algorithm 2** for conflict-free inferences with the MI-QPN.

Now, we discuss the time complexity of **Algorithm 2** for MI-QPN inferences. First, whether a node will be visited is determined by its node sign whose changes are specified in QPN’s general inference algorithm [8], while the influence weights are just used to decide the node sign during inferences. Second, **Algorithm 2** is the same as the general QPN’s inference algorithm if there does not exist “?” during inferences [8]. However, **Algorithm 2** can be used to avoid the “?” results as possible by means of weighting influences quantitatively, where each time of weight computation w.r.t. the composition property can be fuelled by Step 5 in Merge-Sign in $O(1)$ time. Therefore, the time complexity of **Algorithm 2** will be the same as that for the general QPN’s inferences.

Algorithm 2. Conflict-free inference with an MI-QPN.

Input: an MI-QPN MQ
Output: Sign of the influence of evidence node e on each node in MQ
Steps:

1. **for** each node $n \in MQ$ **do**
2. $sign[n] \leftarrow 0, MIweight[n] \leftarrow 0, \text{Propagate-Sign}(\phi, e, e, sign, 1)$
3. **end for**
- Propagate-Sign**(trail, from, to, msgsign, msgMIweight)
 1. Merge-Sign ($tosign, toMIweight, msgsign, msgMIweight$)
 2. $sign[to] \leftarrow tosign, MIweight[to] \leftarrow toMIweight, trail \leftarrow trail \cup \{to\}$
 3. **for** each (induced) neighbor of to **do**
 4. $linksign \leftarrow$ sign of influence between to and n
 5. $msgsign \leftarrow linksign \otimes sign[to]$
 6. $msgMIweight \leftarrow MIweight[to] * MI(to \rightarrow n)$
 7. Merge-Sign($nsign \leftarrow sign[n], nMIweight \leftarrow MIweight[n], msgsign, msgMIweight$)
 8. **if** $n \notin trail$ and $sign[n] \neq nsign$ **then**
 9. Propagate-Sign(trail, to, n, msgsign, msgMIweight)
 10. **end if**
 11. **end for**
- Merge-Sign**($tosign, toMIweight, msgsign, msgMIweight$)
 /* $tosign$ and $toMIweight$ denotes both input parameters and return values*/
 1. **if** $tosign \neq msgsign$ **then**
 2. **if** $toMIweight \geq msgMIweight$ **then**
 3. $toMIweight \leftarrow toMIweight - msgMIweight + toMIweight * msgMIweight$
 4. **else if** $toMIweight < msgMIweight$ **then**
 5. $tosign \leftarrow msgsign, toMIweight \leftarrow msgMIweight - toMIweight + toMIweight * msgMIweight$
 6. **else**
 7. $tosign \leftarrow tosign \oplus msgsign$
 8. **end if**
 9. **else**
 10. $toMIweight \leftarrow msgMIweight + toMIweight - toMIweight * msgMIweight$
 11. **end if**

6. Experimental Results

To test the performance of our method in this paper, we implemented our algorithms for constructing and inferring MI-QPN. We take BN's inference results obtained from Netica [28] as the criteria to decide whether the increase (or decrease) trends of the nodes indicated by the inference results of the corresponding MI-QPN are correct.

It is well known that Wet-Grass network is a classic BN for whether the *Our Grass* is wet, related to *Our Wall*, *Rain*, *Our Sprinkler* and *Neighbor's Grass* status, containing 5 binary variables and 6 edges [2] [28]. From the Wet-Grass BN, we know the qualitative influences between *node Rain* and *Neighbor's Grass* and the ones between *Rain* and *Our Grass* are the same. To illustrate our methods and show the results straightforwardly, we modify the knowledge in the network by removing node *Neighbor's Grass* and the corresponding edge, which will not affect the rest part of the original network and we denote the possible values of the nodes *wet*, *was on*, *rained* as 1, and *dry*, *was off*, *didn't rain* as 0. Moreover, we rename *Our Wall* as *Wall* or *W*, *Our Grass* as *Grass* or *G* and *Our Sprinkler* as *Sprinkler* or *S* for short. Then, the modified network is shown in Figure 6.

First, we derive the corresponding QPN and MI-QPN shown in Figure 7 and Figure 8 respectively. Then, we compare the inference results of the QPN and those of the MI-QPN to verify the feasibility of our algorithms. By using the general QPN's inference algorithm [8] and our inference algorithm for an MI-QPN, we take each node in the QPN and the MI-QPN as evidence and record the inference results of other nodes. The comparisons are shown in Table 1, from which we can see all possible inference conflicts have been eliminated in the MI-QPN.

Then, we compare the inference results on the modified Wet-Grass BN and those on the derived MI-QPN. We take each node of the BN as evidence, and then record the inference results of other nodes shown in Table 2, where E-Node and E-Sign denotes the evidence and corresponding sign respectively. The E-sign $0 \rightarrow 1$ represents that the state of the evidence node changes from 0 to 1, indicating that the probability of state 1 is changed from 0 to 1. If a node in the MI-QPN takes sign "+" ("−") as the final inference result and the increase (decrease) of the probability of state 1 of the corresponding node in the BN, we can conclude that the inference result is correct.

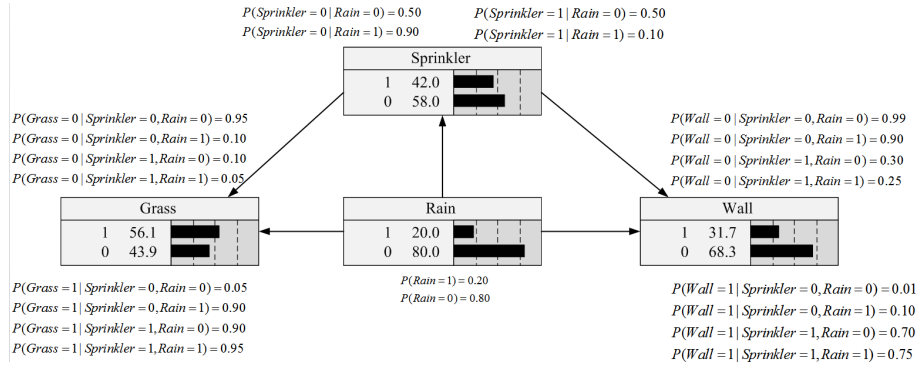


Figure 6. The modified Wet-Grass BN.

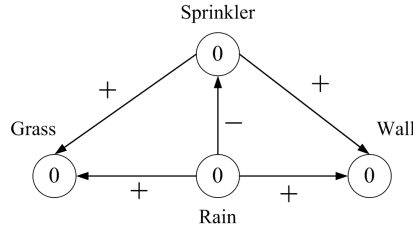


Figure 7. Wet-Grass QPN.

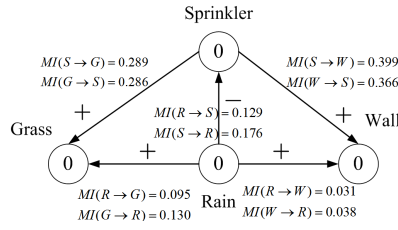


Figure 8. Wet-Grass MI-QPN.

Table 1. Comparisons of inference results between QPN and MI-QPN.

Evidence		QPN					MI-QPN			
Node	Sign	Rain	Sprinkler	Grass	Wall		Rain	Sprinkler	Grass	Wall
Rain	+	+	−	?	?		+	−	+	−
Sprinkler	+	−	+	?	?		−	+	+	+
Grass	+	?	?	+	?		+	+	+	+
Wall	+	?	?	?	+		−	+	+	+

Table 2. Comparisons of inference results between the modified BN and the derived MI-QPN.

	E-Node	E-Sign	Rain	Sprinkler	Grass	Wall
BN	Rain	0 → 1	0 → 1	0.5 → 0.10	0.475 → 0.905	0.355 → 0.165
	Sprinkler	0 → 1	0.31 → 0.0476	0 → 1	0.314 → 0.902	0.0379 → 0.702
	Grass	0 → 1	0.0433 → 0.323	0.0934 → 0.676	0 → 1	0.0782 → 0.504
	Wall	0 → 1	0.245 → 0.104	0.183 → 0.931	0.408 → 0.892	0 → 1
MI-QPN	Rain	+	+	−	+	−
	Sprinkler	+	−	+	+	+
	Grass	+	+	+	+	+
	Wall	+	−	+	+	+

7. Conclusions and Future Work

In this paper, we introduced mutual-information based weights (MI weights) to qualitative influences in QPNs to resolve conflicts during the inferences. We first defined the MI weights based on mutual information and MI-QPN by extending the traditional QPN. Then, we proposed the method to derive the MI weights for the MI-QPN from the corresponding BN without sample data or threshold values. By theoretic analysis, we know the method for deriving MI weights is effective. Furthermore, we discussed the symmetry, transitivity and composition properties in the MI-QPN, and extended the general influence algorithm to implement the conflict-free inferences of MI-QPN. The feasibility of our method was verified by the results of the preliminary experiment.

Our work in this paper also leaves open some other interesting research issues. We are to further consider adding the MI weights to the qualitative synergies and discussing the method to resolve the inference conflicts caused by two inconsistent signs with the same MI weight. As well, we will further resolve the conflicts that take place during the fusion or integration of multiple QPNs by adding the MI weights to the qualitative influences in the QPNs. These are exactly our future work.

Acknowledgements

This work was supported by the National Natural Science Foundation of China (61472345), the Natural Science Foundation of Yunnan Province (2014FA023, 2013FA013), the Yunnan Provincial Foundation for Leaders of Disciplines in Science and Technology (2012HB004), and the Program for Innovative Research Team in Yunnan University (XT412011).

References

- [1] Liu, W.Y., Li, W.H. and Yue, K. (2007) Intelligent Data Analysis. Science Press, Beijing.
- [2] Pearl, J. (1988) Probabilistic Reasoning in Intelligent Systems: Network of Plausible Inference. Morgan Kaufmann, San Mateo.
- [3] Li, W.H., Liu, W.Y. and Yue, K. (2008) Recovering the Global Structure from Multiple Local Bayesian Networks. *International Journal on Artificial Intelligence Tools*, **17**, 1067-1088. <http://dx.doi.org/10.1142/S0218213008004308>
- [4] Cooper, G.F. (1990) The Computational Complexity of Probabilistic Inference Using Bayesian Belief Networks. *Artificial Intelligence*, **42**, 393-405. [http://dx.doi.org/10.1016/0004-3702\(90\)90060-D](http://dx.doi.org/10.1016/0004-3702(90)90060-D)
- [5] Dagum, P. and Luby, M. (1993) Approximating Probabilistic Inference in Bayesian Belief Networks Is NP-Hard. *Artificial Intelligence*, **60**, 141-153. [http://dx.doi.org/10.1016/0004-3702\(93\)90036-B](http://dx.doi.org/10.1016/0004-3702(93)90036-B)
- [6] Wellman, M.P. (1990) Fundamental Concepts of Qualitative Probabilistic Networks. *Artificial Intelligence*, **44**, 257-303. [http://dx.doi.org/10.1016/0004-3702\(90\)90026-V](http://dx.doi.org/10.1016/0004-3702(90)90026-V)
- [7] Bolt, J.H., Renooij, S. and Van der Gaag, L.C. (2003) Upgrading Ambiguous Signs in QPNs. *Proceedings of the 19th Conference in Uncertainty in Artificial Intelligence*, Acapulco, 7-10 August 2003, 73-80.
- [8] Druzdzel, M.J. and Henrion, M. (1993) Efficient Reasoning in Qualitative Probabilistic Networks. *Proceedings of the 11th National Conference on Artificial Intelligence*, Washington DC, 11-15 July 1993, 548-553.
- [9] Lv, Y.L. and Liao, S.Z. (2011) Ambiguity Reduction Based on Qualitative Mutual Information in Qualitative Probabilistic Networks. *Pattern Recognition and Artificial Intelligence*, **24**, 123-129.
- [10] Parsons, S. (1995) Refining Reasoning in Qualitative Probabilistic Networks. *Proceedings of the 11th Conference on Uncertainty in Artificial Intelligence*, Montreal, 18-20 August 1995, 427-434.
- [11] Renooij, S., Parsons, S. and Pardieck, P. (2003) Using Kappas as Indicators of Strength in Qualitative Probabilistic Networks. *Proceedings of European Conferences on Symbolic and Quantitative Approaches to Reasoning with Uncertainty*, Aalborg, 2-5 July 2003, 87-99.
- [12] Renooij, S. and Van der Gaag, L.C. (1999) Enhancing QPNs for Trade-Off Resolution. *Proceedings of the 15th Conference on Uncertainty in Artificial Intelligence*, Stockholm, 30 July-1 August 1999, 559-566.
- [13] Renooij, S., Van der Gaag, L.C. and Parsons, S. (2002) Context-Specific Sign-Propagation in Qualitative Probabilistic Networks. *Artificial Intelligence*, **140**, 207-230. [http://dx.doi.org/10.1016/S0004-3702\(02\)00247-3](http://dx.doi.org/10.1016/S0004-3702(02)00247-3)
- [14] Renooij, S., Van der Gaag, L.C., Parsons, S. and Green, S. (2000) Pivotal Pruning of Trade-Off in QPNs. *Proceedings of the 16th Conference in Uncertainty in Artificial Intelligence*, Stanford, 30 June-3 July 2000, 515-522.
- [15] Renooij, S. and Van der Gaag, L.C. (2002) From Qualitative to Quantitative Probabilistic Network. *Proceedings of the 18th Conference on Uncertainty in Artificial Intelligence*, Edmonton, 1-4 August 2002, 422-429.

- [16] Yue, K., Yao, Y., Li, J. and Liu, W.Y. (2010) Qualitative Probabilistic Networks with Reduced Ambiguities. *Applied Intelligence*, **33**, 159-178. <http://dx.doi.org/10.1007/s10489-008-0156-5>
- [17] Cover, T.M. and Thomas, J.A. (1993) Elements of Information Theory. John Wiley & Sons, Inc., Hoboken.
- [18] Shannon, C.E. and Weaver, W. (1949) The Mathematical Theory of Communication. University of Illinois Press, Champaign.
- [19] Chen, X.W., Anantha, G. and Lin, X.T. (2008) Improving Bayesian Network Structure Learning with Mutual Information-Based Node Ordering in the K2 Algorithm. *IEEE Transactions on Knowledge and Data Engineering*, **20**, 628-640. <http://dx.doi.org/10.1109/TKDE.2007.190732>
- [20] De Campos, L.M. (2006) A Scoring Function for Learning Bayesian Networks Based on Mutual Information and Conditional Independence Tests. *Journal of Machine Learning Research*, **7**, 2149-2187.
- [21] Nicholson, A.E. and Jitnah, N. (1998) Using Mutual Information to Determine Relevance in Bayesian Networks. *Proceedings of the 5th Pacific Rim International Conference on Artificial Intelligence*, Singapore, 22-27 November 1998, 399-410.
- [22] Ibrahim, Z.M., Ngom, A. and Tawk, A.Y. (2011) Using Qualitative Probability in Reverse-Engineering Gene Regulatory Networks. *IEEE/ACM Transactions on Computational Biology and Bioinformatics*, **8**, 326-334. <http://dx.doi.org/10.1109/TCBB.2010.98>
- [23] Liu, W.Y., Yue, K., Liu, S.X. and Sun, Y.B. (2008) Qualitative-Probabilistic-Network-Based Modeling of Temporal Causalities and Its Application to Feedback Loop Identification. *Information Sciences*, **178**, 1803-1824. <http://dx.doi.org/10.1016/j.ins.2007.11.021>
- [24] Yue, K., Qian, W.H., Fu, X.D., Li, J. and Liu, W.Y. (2014) Qualitative-Probabilistic-Network-Based Fusion of Time-Series Uncertain Knowledge. *Soft Computing*, Published Online. <http://dx.doi.org/10.1007/s00500-014-1381-y>.
- [25] De Campos, C.P. and Cozman, F.G. (2013) Complexity of Inferences in Polytree-Shaped Semi-Qualitative Probabilistic Networks. *Proceedings of the 27th AAAI Conference on Artificial Intelligence*, Bellevue, 14-18 July 2013, 217-223.
- [26] Yue, K., Liu, W. and Yue, M. (2011) Quantifying Influences in the Qualitative Probabilistic Network with Interval Probability Parameters. *Applied Soft Computing*, **11**, 1135-1143. <http://dx.doi.org/10.1016/j.asoc.2010.02.013>
- [27] Shafer, G. (1986) The Combination of Evidence. *International Journal of Intelligent Systems*, **1**, 155-179. <http://dx.doi.org/10.1002/int.4550010302>
- [28] Norsys Software Corp. (2007) Netica 3.17 Bayesian Network Software from Norsys. <http://www.norsys.com>

Scientific Research Publishing (SCIRP) is one of the largest Open Access journal publishers. It is currently publishing more than 200 open access, online, peer-reviewed journals covering a wide range of academic disciplines. SCIRP serves the worldwide academic communities and contributes to the progress and application of science with its publication.

Other selected journals from SCIRP are listed as below. Submit your manuscript to us via either submit@scirp.org or **Online Submission Portal**.

