NORTH-HOLLAND

Reasoning in Evidential Networks With Conditional Belief Functions

Hong Xu
IRIDIA and Service d'Automatique, Université Libre de Bruxelles, Belgium

Philippe Smets
IRIDIA, Université Libre de Bruxelles, Belgium

ABSTRACT

In the existing evidential networks applicable to belief functions, the relations among the variables are always represented by joint belief functions on the product space of the variables involved. In this paper, we use conditional belief functions to represent such relations in the network and show some relations between these two kinds of representations. We also present a propagation algorithm for such networks. By analyzing the properties of some special networks with conditional belief functions, called networks with partial dependency, we show that the computation for reasoning can be simplified.

1. INTRODUCTION

Network-based approaches have been widely used for knowledge representation and reasoning with uncertainties. Bayesian networks [3] and valuation-based systems [7] are two well-known frameworks. Bayesian networks are implemented for probabilistic inference, while valuation-based systems can represent several uncertainty formalisms in a unified framework. Graphically, a Bayesian network is a directed acyclic graph, and a valuation-based system is a hypergraph. Nodes in the networks represent random variables; each variable is associated with a finite set of all its possible values, called its frame. In a Bayesian network, arcs represent conditional dependency relations among the variables; in a valuation network, such relations are represented in the form of joint

Address correspondence to Hong Xu, IRIDIA, Université Libre de Bruxelles, 50, Ave. F. Roosevelt, CP 194/6, 1050-Brussels, Belgium.

Received September 1994; accepted March 1995.

International Journal of Approximate Reasoning 1996; 14:155–185

655 Avenue of the Americas, New York, NY 10010
0888-613X/96/$15.00
SSDI 0888-613X(96)00113-1

valuations on the product space of the variables involved. For the case of belief functions, such valuations are the joint belief functions. Recently, Cano et al. [1] have presented an axiomatic system for propagating uncertainty (including belief functions) in Pearl's Bayesian networks, based on Shafer and Shenoy's axiomatic framework [5, 6]. But the belief functions for representing relations of the variables in their system are still represented on the product space. Smets [16] has generalized Bayes' theorem for the case of belief functions and presented the disjunctive rules of combination for two distinct pieces of evidence,[1] which makes it possible to represent knowledge in the form of conditional belief functions and to use them for reasoning in evidential networks. In this paper, we present a network called an evidential network with conditional belief functions and propose a propagation scheme for it. Moreover, we analyze some special properties of this kind of networks, and show that the reasoning process can be simplified in such special cases.

The rest of the paper is organized as follows: In Section 2, we first briefly review belief functions and their rules of combination, both conjunctive and disjunctive. Next, in Section 3, we show some relations between joint belief functions and conditional belief functions which represent the same knowledge. In Section 4, we introduce evidential networks with conditional belief functions and present a propagation scheme for them. In Section 5, we propose some principles for simplifying computation after analyzing the properties of the network with partial dependency, and give an example to show its application. Finally in Section 6, we give some conclusions.

2. DISJUNCTIVE AND CONJUNCTIVE RULES OF COMBINATION

In this section, we introduce the basic concepts of belief functions [4, 11, 17] and summarize the conditioning rules and combination rules for the belief functions. More details can be found in [12, 16].

DEFINITION 2.1 *Let Ω be a finite nonempty set called the frame of discernment (the frame for short). The mapping* $\mathrm{bel}: 2^{\Omega} \to [0,1]$ *is an (unnormalized) belief function if and only if there exists a basic belief assignment (bba)* $m: 2^{\Omega} \to [0,1]$ *such that:*

(i) $\sum_{A \subseteq \Omega} m(A) = 1,$

(ii) $\mathrm{bel}(A) = \sum_{B \subseteq A,\, B \neq \emptyset} m(B),$

(iii) $\mathrm{bel}(\emptyset) = 0.$

[1] Smets [15] has given a definition for the concept of distinct evidence.

Those subsets A such that $m(A) > 0$ are called the *focal elements*. A *vacuous belief function* is a belief function such that $m(\Omega) = 1$ and $m(A) = 0$ for all $A \neq \Omega$, which represents total ignorance.

The value $\text{bel}(A)$ quantifies the strength of the belief that the event A occurs. It measures the same concept as $P(A)$ does in classical probability theory, but bel is not an additive measure. The value $m(A)$ represents the part of belief that supports the fact that A occurs and cannot support any more specific event (due to the lack of information). Note that m is not the counterpart of a probability distribution function p [14]. Both bel and P are defined on 2^{Ω}, but m is defined on 2^{Ω}, whereas p is defined on Ω.

Given a belief function, we can define a *plausibility function* $\text{pl}: 2^{\Omega} \to [0, 1]$ and a *commonality function* $q: 2^{\Omega} \to [0, 1]$ as follows: for $A \subseteq \Omega$,

$$\text{pl}(A) = \text{bel}(\Omega) - \text{bel}(\bar{A}) \text{ and } \text{pl}(\emptyset) = 0,$$

$$q(A) = \sum_{A \subseteq B \subseteq \Omega} m(B),$$

where $\bar{A}$ is the complement of A relative to Ω.

Note that m (basic belief mass), bel (belief function), pl (plausibility function), and q (commonality function) are in one-to-one correspondence with each other.

DEFINITION 2.2 *Let* bel *be our belief on the frame* Ω. *Suppose we learn that* $\bar{A} \subseteq \Omega$ *is false. The resulting conditional belief function*[2] $\text{bel}(\cdot \| A)$ ($\text{bel}(B \| A)$ *can be read as the belief of B given A*) *is obtained through the unnormalized rule of conditioning. For* $B \subseteq \Omega$,

$$m(B \| A) = \begin{cases} \sum_{X \subseteq \bar{A}} m(B \cup X) & \text{if } B \subseteq A \subseteq \Omega, \\ 0 & \text{otherwise}. \end{cases}$$

If a second piece of information concerning the same issue is now available from a different source, we need to integrate it with the first one by the combination of two belief functions defined as follows:

DEFINITION 2.3 *Consider two distinct pieces of evidence on* Ω *represented by* m_1 *and* m_2. *The belief function that quantifies the combined impact of these two pieces of evidence is obtained through the conjunctive rule of combination. We use* $\textcircled{\wedge}$ *to represent the conjunctive combination operator. For all* $A \subseteq \Omega$,

$$\left(m_1 \textcircled{\wedge} m_2\right)(A) = \sum_{A = B \cap C} m_1(B) m_2(C).$$

[2] We use "‖" in place of "|" to emphasize the nonnormalization of our conditioning.

This can also be written in terms of the commonality functions as follows:

$$\left(q_1 \textcircled{\wedge} q_2\right)(A) = q_1(A)q_2(A).$$

DEFINITION 2.4 *Consider two distinct pieces of evidence on* Ω *represented by* m_1 *and* m_2. *The belief function induced by the disjunction of these two pieces of evidence is obtained through the disjunctive rule of combination* [2]. *We use* $\textcircled{\vee}$ *to represent the disjunctive combination operator. For all* $A \subseteq \Omega$,

$$\left(m_1 \textcircled{\vee} m_2\right)(A) = \sum_{A = B \cup C} m_1(B)m_2(C).$$

Let a function $b : 2^{\Omega} \to [0, 1]$ *be defined as* $b(A) = \mathrm{bel}(A) + m(\varnothing)$. *Then the disjunctive combination rule can be written as follows*:

$$\left(b_1 \textcircled{\vee} b_2\right)(A) = b_1(A)b_2(A).$$

The meaning of these two rules is given in [16]. Suppose m_1 and m_2 are the bba's induced by two distinct pieces of evidence $\mathscr{E}_1$ and $\mathscr{E}_2$, respectively. Suppose an agent will hold belief m_1 (m_2) if he knows that $\mathscr{E}_1$ ($\mathscr{E}_2$) prevails. If the agent knows that both $\mathscr{E}_1$ and $\mathscr{E}_2$ prevail, then his belief will be represented by $m_1 \textcircled{\wedge} m_2$. If the agent knows only that at least one of $\mathscr{E}_1$ and $\mathscr{E}_2$ prevails without knowing which one, his belief will be represented by $m_1 \textcircled{\vee} m_2$. The justification of these two rules, called conjunctive and disjunctive rules of combination, can be found in [12] and [16], respectively.

Since m (basic belief mass), bel (belief function), pl (plausibility function), and q (commonality function) are in one-to-one correspondence with each other, the above rules can also be represented by using any of these functions. Details can be found in [16].

Note that all the definitions above are for the nonnormalized case. For the case of normalized belief functions, which means $m(\varnothing) = 0$, the normalized factor $K = 1 - m(\varnothing)$ should be considered in those rules, and the conditioning rule and the conjunctive combination rule turn out to be Dempster's rule of conditioning and of combination. The (unnormalized) $\mathrm{bel}(A \| B)$ turns out to be the (normalized) $\mathrm{bel}(A | B)$, and $\textcircled{\wedge}$ to be $\oplus$ [4, 16]. $\textcircled{\vee}$ doesn't have a counterpart in Shafer's presentation. To be consistent with convention, we will use $\oplus$ instead of $\textcircled{\wedge}$, but the computation is unnormalized.

Let's consider two spaces Θ and X. We use $\mathrm{bel}_X(\cdot \| \theta)$ to represent the belief function induced on the space X given $\theta \subseteq \Theta$. Suppose all we know about X is initially represented by the set $\{\mathrm{bel}_X(\cdot \| \theta_i) : \theta_i \in \Theta\}$. We only know the beliefs on X when we know which element of Θ holds. We do not have further specific knowledge about the belief on X when we only know that the prevailing element of Θ belongs to a given subset θ of Θ. Under very general requirements, Smets [10, 16] has derived the disjunctive rule of combination (DRC) to build $\mathrm{bel}_X(\cdot\|\theta)$ on X for any $\theta \subseteq \Theta$, and the generalized Bayesian theorem (GBT) to build $\mathrm{bel}_\Theta(\cdot\|x)$ on Θ for any $x \subseteq X$.

THEOREM 2.1 (Disjunctive rule of combination [16]) *For all* $\theta \subseteq \Theta$, $x \subseteq X$,

$$m_X(x\|\theta) = \sum_{\bigcup_{i:\theta_i\in\theta} x_i = x} \prod_{i:\theta_i\in\theta} m_X(x_i\|\theta_i),$$

$$\mathrm{pl}_X(x\|\theta) = 1 - \prod_{\theta_i\in\theta} [1 - \mathrm{pl}_X(x\|\theta_i)].$$

THEOREM 2.2 (Generalized Bayesian theorem [16]) *For all* $\theta \subseteq \Theta$, $x \subseteq X$,

$$\mathrm{pl}_\Theta(\theta\|x) = 1 - \prod_{\theta_i\in\theta} [1 - \mathrm{pl}_X(x\|\theta_i)].$$

Note that $\mathrm{pl}_\Theta(\theta\|x) = \mathrm{pl}_X(x\|\theta)$, and this represents the fact that in general $\mathrm{pl}(A\|B) = \mathrm{pl}(B\|A)$, an equality unsatisfied once normalization is introduced. This can be seen from the above two theorems. Now suppose there exists some *a priori* belief bel_0 over Θ. By using Theorems 2.1 and 2.2, we can compute bel on X given bel_0 and $\{\mathrm{bel}_X(\cdot\|\theta_i) : \theta_i \in \Theta\}$:

THEOREM 2.3 [16] *Suppose there exists some* a priori *belief* bel_0 *our* Θ *distinct from the belief induced by the set of conditional belief functions* $\{\mathrm{bel}_X(\cdot\|\theta_i)\} : \theta_i \in \Theta$. *Then* $\forall x \subseteq X$,

$$m_X(x) = \sum_{\theta\subseteq\Theta} m_0(\theta) m_X(x\|\theta),$$

$$\mathrm{pl}_X(x) = \sum_{\theta\subseteq\Theta} m_0(\theta) \mathrm{pl}_X(x\|\theta)$$

$$= \sum_{\theta\subseteq\Theta} m_0(\theta)\left(1 - \prod_{\theta_i\in\theta} [1 - \mathrm{pl}_X(x\|\theta_i)]\right).$$

Note that the above three theorems can also be expressed by using belief functions and commonality functions, which are detailed in Smets [16].

3. KNOWLEDGE REPRESENTATION USING BELIEF FUNCTIONS

Let $\mathbf{U} = \{X_1, \ldots, X_n\}$ be a finite set of variables representing the universe of discourse for a class of problems, where each X_i represents a relevant aspects of the problem. Associated with X_i, there is a frame Θ_{X_i} which is set of all its possible values. Let X and Y be two disjoint subsets of $\mathbf{U}$; their frames Θ_X and Θ_Y are the product spaces of the frames of the variables they include. For short, we write X, Y for Θ_X, Θ_Y whenever confusion is absent. According to the notation of the previous section, a conditional belief function for Y given X is represented by $\mathrm{bel}_Y(\cdot \| x)$ where $x \subseteq \Theta_X$, which means that we know the belief about Y given that we only know that the actual value of X is in x. Similarly, joint belief functions on X and Y are defined on the space $\Theta_X \times \Theta_Y$ ($\Theta_{X \cup Y}$ or $X \times Y$ for short). Look at the following example:

EXAMPLE 3.1 Let A and B be two variables with frames $\Theta_A = \{a, \bar{a}\}$ and $\Theta_B = \{b, \bar{b}\}$ respectively. To represent a relation between A and B such that if $A = a$ then $B = b$ with $m = 0.9$, by a belief function in joint form, the rule is represented by a belief function on the space $\Theta = \{(a, b), (a, \bar{b}), (\bar{a}, b), (\bar{a}, \bar{b})\}$, with masses 0.9 on the subset $\{(a, b), (\bar{a}, b), (\bar{a}, \bar{b})\}$, and 0.1 on Θ, while with belief functions in conditional form, it is represented by the conditional bba, $m(\{b\}|a) = 0.9$, $m(\Theta_B|a) = 0.1$, $m(\Theta_B|\bar{a}) = 1$, $m(\Theta_B|\Theta_A) = 1$. This is illustrated by Table 1.

From the example, it can be seen that the latter representation is often more natural and easy for the users to provide and to understand. The use of conditional belief functions parallels the use of conditional probabilities in Bayesian networks. Generally, given two disjoint subsets $X, Y \subseteq \mathbf{U}$, to represent conditional belief functions for Y given X by a joint form, one needs $2^{|\Theta_X| \times |\Theta_Y|}$ elements in the worst case, while to represent them by conditional form one only needs $2^{|\Theta_X| + |\Theta_Y|}$ elements in the worst case. Of course, not all belief functions on $\Theta_{X \cup Y}$ admit an equivalent representation by a set of conditional belief functions. But we think that the users' knowledge is encoded in the conditional form and that the joint beliefs they would provide are those based on the known conditional form.

Table 1. A Belief Function in a Conditional Form

	a	$\bar{a}$	Θ_A
b	0.9	0	0
Θ_B	0.1	1	1

Furthermore, in many cases, the users' belief can be represented by the conditional belief functions for Y given $x_i \in \Theta_X$. The conditional belief for Y given $x \subseteq \Theta_X$ is then derived from the DRC. Example 1 is such a case. In the worst case, it needs only $|\Theta_X| \times 2^{|\Theta_Y|}$ elements.

Cano et al. [1] and Shenoy [8, 9] have both introduced the concept of a noninformative belief function.[3] To understand such a concept, we first introduce the concepts of projection, extension, and marginalization.

DEFINITION 3.1 *Projection of configurations simply means dropping the extra coordinates. If X and Y are sets of variables, $Y \subseteq X$, and x_i is a configuration of Θ_X, then let $x_i^{\downarrow Y}$ denote the projection of x_i on Θ_Y. Then $x^{\downarrow Y}$ is a configuration of Θ_Y. If x is a nonempty subset of Θ_X, then the projection of x on Y, denoted by $x^{\downarrow Y}$, is obtained by $x^{\downarrow Y} = \{x_i^{\downarrow Y} | x_i \in x\}$. If y is a subset of Θ_Y, then the extension of y to X, denoted by $y^{\uparrow X}$, is $y \times \Theta_{X-Y}$ (It is also called the cylindric extension of y into X.)*

DEFINITION 3.2 *Suppose m is a bba on B and $A \subseteq B \subseteq \mathbf{U}$, $A \neq \emptyset$. The marginal of m for A, denoted by $m^{\downarrow A}$, is the bba on A defined by*

$$m^{\downarrow A}(a) = \sum_{b \subseteq \Theta_B,\ B^{\downarrow A} = a} m(b) \qquad \textit{for all} \quad a \subseteq \Theta_A.$$

DEFINITION 3.3 *Given two disjoint subsets $X, Y \subseteq \mathbf{U}$, let* bel *be a belief function defined on the space $\Theta_{X \cup Y}$. It is said that* bel *is a noninformative belief function over X if and only if* $\mathrm{bel}^{\downarrow X}$ *is a vacuous belief function over X (Cano et al.* [1]).

Intuitively, the belief function in the above definition gives some information about variables in Y and their relationship with variables in X, but no information about X. This property is easy to verify when the belief is represented in conditional form.

PROPOSITION 1 *Let $\{\mathrm{bel}_Y(\cdot \| x) : x \subseteq \Theta_X\}$ be a family of conditional belief functions for Y given X. It is noninformative over Y iff $\mathrm{bel}_Y(\cdot \| \Theta_X)$ is a vacuous belief function on Y.*

Proof Let $\mathrm{bel}_{X \times Y}$ be the belief function over $X \times Y$ whose conditioning given X is $\{\mathrm{bel}_Y(\cdot \| x) : x \subseteq \Theta_X\}$. It is easy to see that $\mathrm{bel}_{X \times Y}^{\downarrow Y}(y) = \mathrm{bel}_Y(y \| \Theta_X)$. Therefore, the proposition is proved according to the definition of noninformative belief function. ■

[3] Note that Shenoy [8] and Cano et al. [1] called this belief function the "conditional belief function." We change the name to avoid confusion with the classical meaning of "conditional belief function."

Moreover, suppose a belief function bel defined on the space $\Theta_{X \cup Y}$ gives information only on the relation between X and Y, but no information about either X or Y. Then $\text{bel}^{\downarrow X}$ and $\text{bel}^{\downarrow Y}$ are both vacuous on X and Y respectively. That is to say, bel is noninformative over both X and Y. The following shows how to verify such properties when the belief functions are in conditional form. It is based on normalized belief functions, i.e., belief functions such that $m(\emptyset) = 1$.

PROPOSITION 2 $\{\text{bel}_Y(\cdot \| x_i) : x_i \in \Theta_X\}$ *is noninformative over* X *if and only if* $\text{bel}_Y(\cdot \| x_i)$ *is a normalized belief function for each* $x_i \in \Theta_X$, *in which case* $\text{bel}_Y(\cdot \| x)$ *is also normalized for* $x \subseteq \Theta_X$.

Proof From Theorem 2.1, it is easy to prove that $\text{bel}_Y(\cdot \| x)$ for any $x \subseteq \Theta_X$ is normalized if $\text{bel}_Y(\cdot \| x_i)$ is a normalized belief function for each $x_i \in \Theta_X$. That is, $\text{pl}_Y(\Theta_Y \| x) = 1$ for any $x \subseteq \Theta_X$. Thus, for any $x \subseteq \Theta_X$, $\text{pl}_X(x \| \Theta_Y) = \text{pl}_Y(\Theta_Y \| x) = 1$, i.e., $\text{bel}_X(\cdot \| \Theta_Y)$ is a vacuous belief function over X. From the previous proposition, we have that $\{\text{bel}_Y(\cdot \| x_i) : x_i \in \Theta_X\}$ is noninformative over X.

Suppose there exists $x_j \in \Theta_X$ such that $\text{bel}_Y(\cdot \| x_j)$ is unnormalized. Then $\text{pl}_X(\{x_j\} \| \Theta_Y) = \text{pl}_Y(\Theta_Y \| \{x_j\}) < 1$. Thus $\{\text{bel}_Y(\cdot \| x_i) : x_i \in \Theta_X\}$ is not noninformative over X. ■

PROPOSITION 3 *If we only know a family of (normalized) conditional belief functions such as* $\{\text{bel}_Y(\cdot | x_i) : x_i \in \Theta_X\}$ *(see footnote 2), then it is noninformative over* Y *iff for each* $y \subset \Theta_Y$, *there exists* $x_i \in \Theta_X$ *such that* $\text{bel}_Y(y | x_i) = 0$.

Proof For each $y \subset \Theta_Y$, suppose there exists $x_i \in \Theta_X$ such that $\text{bel}_Y(y | x_i) = 0$; then $\text{pl}_Y(\bar{y} | x_i) = 1$. That is, for each $y \subset \Theta_Y$, there exists $x_i \in \Theta_X$ such that $\text{pl}_Y(y | x_i) = 1$. From Theorem 2.1, it is easy to see that $\text{pl}_Y(y | \Theta_X) = 1$ for all $y \subseteq \Theta_Y$. According to Proposition 1, $\{\text{bel}_Y(\cdot | x_i) : x_i \in \Theta_X\}$ is noninformative over Y.

Suppose $\{\text{bel}_Y(\cdot | x_i) : x_i \in \Theta_X\}$ is noninformative over Y. Then we have $\text{pl}_Y(y | \Theta_X) = 1$ for all $y \subseteq \Theta_Y$. From Theorem 2.1, we have that for each $y \subset \Theta_Y$, there exists $x_i \in \Theta_X$ such that $\text{pl}_Y(y | x_i) = 1$; thus $\text{bel}_Y(\bar{y} | x_i) = 0$. That is, for each $y \subset \Theta_Y$, there exists $x_i \in \Theta_X$ such that $\text{bel}_Y(y | x_i) = 0$. ■

In the following, we will show some relations between the belief functions represented in conditional form and in joint form. By using the rules of conditioning, every joint belief function can induce a family of conditional belief functions, but not every family of conditional belief functions is compatible with a joint one. Incompatibility occurs when the set of conditional belief functions cannot be obtained by conditioning some underlying joint belief function. We say that those sets of conditional belief functions that are not compatible with a joint belief function are *invalid*. The joint belief functions that could underlie a family of conditional beliefs are not always unique. Smets [16] has shown that when the conditional belief functions are represented by $\{\text{bel}_Y(\cdot\|x_i) : x_i \in X\}$, we can always construct a joint belief function from it, and the joint belief is unique if the principle of minimal commitment is applied.

LEMMA 3.1 *Let X and Y be two disjoint subsets of* **U**, *and* $\text{bel}_{X\times Y}$ *be a belief function on the product space* $X \times Y$. *Then its conditional form* $\{\text{bel}_Y(\cdot\|x) : x \subseteq \Theta_X\}$ *is obtained by*

$$m_Y(y\|x) = \sum_{S \subseteq \Theta_{X \cup Y},\, (S \cap x^{\uparrow(X \cup Y)})^{\downarrow Y} = y} m_{X\times Y}(S).$$

Proof This is directly obtained from the conditioning process. ■

LEMMA 3.2 *Suppose a family of normalized conditional belief functions* $\{\text{bel}_Y(\cdot\|x) : x \subseteq \Theta_X\}$ *is compatible with a joint belief. Then it satisfies* $\text{pl}_Y(y\|x_1) \leq \text{pl}_Y(y\|x_2)$ *if* $x_1 \subseteq x_2 \subseteq \Theta_X$.

Proof

$$\begin{aligned}
\text{pl}_Y(y\|x_1) &= \text{pl}_{X\times Y}(y^{\uparrow X\cup Y}\|x_1^{\uparrow X \cup Y}) \\
&= \text{pl}_{X\times Y}(y^{\uparrow X\cup Y} \cap x_1^{\uparrow X \cup Y}) \\
&\leq \text{pl}_{X\times Y}(y^{\uparrow X\cup Y} \cap x_2^{\uparrow X \cup Y}) \\
&= \text{pl}_{X\times Y}(y^{\uparrow X\cup Y}\|x_2^{\uparrow X \cup Y}) = \text{pl}_Y(y\|x_2).
\end{aligned}$$

This concludes the proof. ■

EXAMPLE 3.2 Let A and B be two variables with $\Theta_A = \{a, \bar{a}\}$ and $\Theta_B = \{b, \bar{b}\}$. Let $\Theta = \{ab, a\bar{b}, \bar{a}b, \bar{a}\bar{b}\}$ be briefly denoted by $\{1, 2, 3, 4\}$; then, for example, the subset $\{ab, a\bar{b}\}$ is denoted by 12 for short. Consider bel_1 on Θ: $m(14) = m(23) = 0.1$, $m(123) = m(124) = m(134) = m(234) = 0.1$, and $m(1234) = 0.4$. By applying Lemma 3.1, its corresponding conditional belief function for B given A is shown in Table 2. Consider bel_2 on Θ: $m(23) = 0.2$, $m(134) = m(124) = 0.2$, and $m(1234) = 0.4$. Its corresponding conditional form obtained by applying Lemma 3.1 is shown in Table 3. Comparing the two tables, we find that two different joint belief functions lead to the same conditional form. Therefore, a conditional belief is compatible with more than one joint beliefs.

The ballooning extension of a conditional belief function over X given $y \subset Y$ in the belief function over $X \times Y$ means that the bba $m(x|y)$ $(x \subseteq X)$ is allocated to the set $x^{\uparrow X\times Y} \cup \bar{y}^{\uparrow X \times Y}$, i.e., the largest subset of $X \times Y$ whose intersection with $x^{\uparrow X\times Y}$ is $x^{\uparrow X\times Y}$.

Table 2. Belief Function in Conditional Form for bel_1

	a	$\bar{a}$	Θ_A
b	$m(14) + m(134) = 0.1 + 0.1 = 0.2$	$m(23) + m(123) = 0.1 + 0.1 = 0.2$	0
$\bar{b}$	$m(23) + m(234) = 0.1 + 0.1 = 0.2$	$m(14) + m(124) = 0.1 + 0.1 = 0.2$	0
Θ_B	$m(123) + m(124) + m(1234) = 0.1 + 0.1 + 0.4 = 0.6$	$m(134) + m(234) + m(1234) = 0.1 + 0.1 + 0.4 = 0.6$	$m(14) + m(23) + m(123) + m(124) + m(134) + m(234) + m(1234) = 1$

Table 3. Belief Function in Conditional Form for bel_2

	a	$\bar{a}$	Θ_A
b	$m(134) = 0.2$	$m(23) = 0.2$	0
$\bar{b}$	$m(23) = 0.2$	$m(124) = 0.2$	0
Θ_B	$m(124) + m(1234)$ $= 0.2 + 0.4 = 0.6$	$m(134) + m(1234)$ $= 0.2 + 0.4 = 0.6$	$m(23) + m(124)$ $+m(134) + m(1234)$ $= 1$

The belief function so built is the least committed belief function on $X \times Y$ among all belief functions on $X \times Y$ whose conditioning on $y^{\uparrow X \times Y}$ is equal to $\mathrm{bel}(\cdot|y)$. Formally, we have the following definition:

DEFINITION 3.4 [16] *Let Y and X be two finite spaces, and* $\mathrm{bel}_X(x\|y)$ *be a conditional belief function on X given some* $y \subseteq Y$. *The ballooning extension of* $\mathrm{bel}_X(x\|y)$ *on* $X \times Y$ *is a belief function* $\mathrm{bel}_{X \times Y}$ *computed as follows*:

$$\mathrm{bel}_{X\times Y}((x^{\uparrow X\times Y} \cap y^{\uparrow X\times Y}) \cup \bar{y}^{\uparrow X\times Y}) = \mathrm{bel}_X(x\|y) + m_X(\emptyset\|y).$$

LEMMA 3.3 [10, 16] *Suppose X and Y are two disjoint subsets of* **U**. *If all we know about the relation between X and Y is given by the set of conditional belief functions* $\{\mathrm{bel}_Y(\cdot\|x_i) : x_i \in \Theta_X\}$, *we can construct the belief function on* $X \times Y$ *by first computing the ballooning extension of each* $\mathrm{bel}_Y(\cdot\|x_i)$, *then combining the results using Dempster's rule of combination. The computation can be written as follows: Let* $a \subseteq \Theta_{X \cup Y}$ *and* $y_i = (a \cap \{x_i\}^{\uparrow X \cup Y})^{\downarrow Y}$. *Then*

$$m_{X \cup Y}(a) = \prod_{x_i \in \Theta_X} m_Y(y_i\|x_i).$$

The joint belief function on $X \times Y$ is the one whose conditionings over $x \subseteq X$ and $y \subseteq Y$ result in the DRC and the GBT, respectively.

4. REASONING WITH CONDITIONAL BELIEFS

4.1. Evidential Networks with Conditional Belief Functions

In this section, we use the network proposed by Smets [16] for the propagation of beliefs. Graphically, the network is a directed graph (acyclicity is not required), as shown in Figure 1. The conditional beliefs are defined in a different way from conditional probabilities in the Bayesian

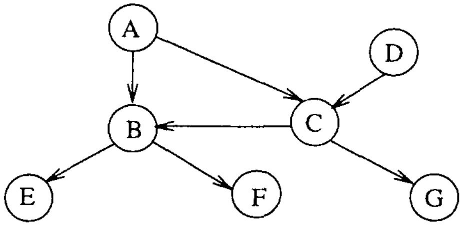


Figure 1. An example of an evidential network with conditional belief functions.

networks (BNs) [3]. In our network, each edge represents a conditional relation between the two nodes it connects. For example, the edges (A, C) and (D, C) mean that we have $\{\text{bel}_C(\cdot\|a_i): a_i \in \Theta_A\}$ and $\{\text{bel}_C(\cdot\|d_i): d_i \in \Theta_D\}$, but *not* $\{\text{bel}_C(\cdot\|a_i, d_i): a_i \in \Theta_A, d_i \in \Theta_D\}$ as is the case in Bayesian networks. In order to distinguish these two kinds of networks, we call ours ENC, which means an evidential network with conditional belief functions. However, if we have a conditional belief such as $\{\text{bel}_C(\cdot\|a_i, d_i): a_i \in \Theta_A, d_i \in \Theta_D\}$, we can build an ENC in which nodes A and D are merged as one node. In this section, we also assume that, for each conditional belief function for Y given X, all we know about Y given X is initially represented by the set $\{\text{bel}_Y(\cdot x_i): x_i \in \Theta_X\}$.

Note that in a BN, there is a joint probability distribution for the network, and the local conditional probabilities can be computed by projecting the global probability on the subset of the variables involved. And for each node in BN, there is only one conditional probability for it, given its parents. In an ENC, we can have knowledge about the relations between two nodes from different sources (we call them local conditional beliefs). Then the global belief for an ENC can be computed from all those local conditional beliefs. However, from the global belief, we can no longer reconstruct each local conditional belief. This distinction results from the different natures of the two models [13]. Also, once an ENC has been constructed, we can add new knowledge, by the conjunctive rule of combination for some variables, at any time, supposing the new knowledge is from an independent source. A BN, in contrast, is always constructed from some underlying joint probability distribution.

4.2. Propagating Beliefs in an ENC

Given any ENC, one can transform each conditional belief function $\text{bel}_Y(\cdot|x)$, $x \subseteq X$, into a joint belief on $X \times Y$ by building its ballooning extension on $X \times Y$. Then the ENC becomes a classical network, to which the VBS algorithm [6, 7] could be directly applied. Such a strategy would be inefficient, and we present hereafter more efficient algorithms that profit from the particular nature of the belief function encountered in the ENC.

It has been shown that one main objective of reasoning processes in an evidential network is to compute the marginals of the global belief functions for some variables. For two disjoint subsets X and Y of $\mathbf{U}$, knowing[4] $\mathrm{bel}_{X\times Y}$ or $\mathrm{pl}_{X\times Y}(x, y)$, $x \subset \Theta_X$, $y \subseteq \Theta_Y$, is equivalent for what concerns propagation to computing the marginal beliefs for the variables. We use BEL_X and bel_{0X} to denote the marginal and *a priori* beliefs for the variable X. Due to Theorem 2.1 and 2.2, given two variables X and Y and the conditional belief $\{\mathrm{bel}_Y(\cdot\|x_i): x_i \in \Theta_X\}$, we could compute and store $\mathrm{bel}_Y(\cdot\|x): x \subseteq \Theta_X$ and $\mathrm{bel}_X(\cdot\|y): y \subseteq \Theta_Y$ in the preprocess, or simply store $\mathrm{pl}(x, y): x \subseteq \Theta_X$, $y \subseteq \Theta_Y$, to save space. Storing pl's takes only half of the space of storing bel's, since $\mathrm{pl}(x, y) = \mathrm{pl}_Y(y\|x) = \mathrm{pl}_X(x\|y)$. Now, we are ready to give the inference algorithm:

Given an ENC represented by $G = (\mathbf{M}, \mathbf{E})$.

Case 1: Suppose G is a polytree, i.e., there is only one (undirected) path between any of two nodes in the network.

The propagation algorithm can be regarded as a message-passing scheme: for each node X in the network, its marginal BEL_X is computed by combining all the messages from its neighbors $\mathrm{ne}(X)$ and its own *a priori* bel_{0X}, i.e.,

$$\mathrm{BEL}_X = \mathrm{bel}_{0X} \oplus (\oplus\{M_{Y\to X} | Y \in \mathrm{ne}(X)\}),$$

where the message $M_{Z\to Y}$ is a belief function on X, so it can be represented by $\mathrm{bel}_{Y\to X}$ (or $m_{Y\to X}$), and is computed as follows: $\forall x \subseteq \Theta_X$,

$$\mathrm{bel}_{Y\to X}(x) = \sum_{y\subseteq\Theta_Y} m_X(x\|y)\,\mathrm{bel}_{\mathrm{ne}(Y)/X\to Y}(y),$$

where

$$\mathrm{bel}_{\mathrm{ne}(Y)/X\to Y} = \mathrm{bel}_{0Y} \oplus (\oplus\{\mathrm{bel}_{Z\to Y} | Z \in \mathrm{ne}(Y), Z \neq X\}).$$

Case 2: If there exist any undirected loops in the network, then some nodes needed to be merged to make the network acyclic, resulting in a new polytree $G' = (\mathbf{M}', \mathbf{E}')$, where some nodes in G' might be subsets of the nodes in G; we call them *merged nodes*. For any merged node v in G', there might be a belief function $\mathcal{B}_v$ obtained by combining the ballooning extension of each conditional belief. Figure 2 illustrates two examples for this process.

---

[4] We use $\mathrm{pl}_{X\times Y}(x, y)$ to denote $\mathrm{pl}_{X\times Y}(x^{\uparrow X\cup Y} \cap y^{\uparrow X\cup Y})$ for the sake of simplicity. It has been shown that $\mathrm{pl}_{X\times Y}(x, y) = \mathrm{pl}_X(x\|y) = \mathrm{pl}_Y(y\|y)$ for $x \subseteq X$, $y \subseteq Y$ [16].

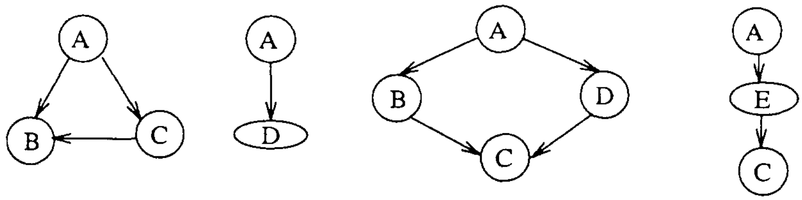


Figure 2. Examples of absorbing loops in the ENCs.

In Figure 2(a), the loop is absorbed by merging nodes B and C; the resulting graph is shown in Figure 2(b), where $D = \{B, C\}$, and the new conditional belief function $\text{bel}_D(\cdot\|a_i)$ is obtained by combining $\text{bel}_B(\cdot\|a_i)$ and $\text{bel}_C(\cdot\|a_i)$ on the space $\Theta_D = \Theta_{B \cup C}$: for all $a_i \in \Theta_A$, $d \subseteq \Theta_D$,

$$m_D(d\|a_i) = \sum_{b^{\uparrow\{B,C\}} \cap c^{\uparrow\{B,C\}} = d} m_B(b\|a_i) m_C(c\|a_i). \tag{4.1}$$

Obviously, $\text{bel}_D(d\|a_i)$ is normalized iff $\text{bel}_B(\cdot\|a_i)$ and $\text{bel}_C(\cdot\|a_i)$ are normalized, since the subset $b^{\uparrow\{B,C\}} \cap c^{\uparrow\{B,C\}}$ can never be an empty set. Moreover, the conditional belief function between B and C becomes $\mathscr{B}_D$ in Figure 2(b), obtained by applying Lemma 3.3. Thus $\mathscr{B}_D$ is a belief function on Θ_D.

Figure 2(c) is another example of an ENC with a loop. In this case, we merge B and D, resulting in the graph shown in Figure 2(d), where $E = \{B, D\}$. Here $\text{bel}_E(\cdot\|a_i)$ is obtained by combining $\text{bel}_B(\cdot\|a_i)$ and $\text{bel}_D(\cdot\|a_i)$ on the space $\Theta_E = \Theta_{B \cup D}$ using Equation (4.1). As for $\text{bel}_C(\cdot\|e)$, $e \subseteq \Theta_E$, we compute it for three cases:

1. $\forall e_t = (b_i, d_j) \in \Theta_E$,

$$m_C(c\|e_t) = \sum_{s_1 \cap s_2 = c} m_C(s_1\|b_i) m_C\left(s_2\|d_j\right).$$

2. For $e \subseteq \Theta_E$, if e can be represented by $b \times d$, where $b \subseteq \Theta_B$, $d \subseteq \Theta_D$, then

$$m_C(c\|e) = \sum_{s_1 \cap s_2 = c} m_C(s_1\|b) m_C(s_2\|d),$$

where $m_C(\cdot\|b)$ and $m_C(\cdot\|d)$ are obtained from $m_C(\cdot\|b_i)$ and $m_C(\cdot\|d_j)$ respectively by applying the DRC (Theorem 2.1).

3. For any other $e \subseteq \Theta_E$, we first construct a conditional belief function $\mathrm{bel}_{C \cup D}(\cdot \| b_i)$ from $m_C(\cdot \| b_i)$ such that

$$m_{C \cup D}(s \| b_i) = m_C(c \| b_i),$$

where $s = c^{\uparrow \{C, D\}} \cap ((e \cap b_i^{\uparrow E})^{\downarrow \{D\}})^{\uparrow \{C, D\}}$. Let $\mathrm{bel}^b_{C \cup D}$ be the belief function resulting from combining ballooning extension of $m_C(\cdot \| d_i)$; then

$$m_C(c \| e) = \left(\mathrm{bel}^b_{C \cup D} \oplus \left(\textcircled{\vee}\{\mathrm{bel}_{C \cup D}(\cdot \| b_i) | b_i \in \Theta_B\}\right)\right)^{\downarrow \{C\}}.$$

Alternatively, $\mathrm{bel}_C(\cdot \| e)$, $e \subseteq \Theta_E$ can be computed by first combining the ballooning extensions of the two conditional beliefs $\mathrm{bel}_C(\cdot \| b_i)$ and $\mathrm{bel}_C(\cdot \| d_j)$ on the space $\Theta_{B \cup C}$ and $\Theta_{C \cup D}$, then applying Lemma 3.1 to transform the resulting belief to a conditional form $\mathrm{bel}_C(\cdot \| e)$, $e \subseteq \Theta_E$, and $\mathrm{bel}_E(\cdot \| c)$, $c \subseteq \Theta_C$. However, this takes more space for the computation. Since there is no direct relation between B and D, $\mathscr{B}_E$ is a vacuous belief function.

After rearranging the network to an acyclic one, we then use a similar algorithm in case 1 for the propagation: Suppose each node X in G' is a subset and has a $\mathscr{B}_x$. Thus, for any nonmerged node, it is a singleton, and $\mathscr{B}_x$ is a vacuous belief function. Then the computation is as follows: for any node $A = \{X_1, \ldots, X_t\}$ in G',

$$\mathrm{Bel}_A = \mathscr{B}_A \oplus (\oplus \{M_{Y \to A} | Y \in \mathrm{ne}(A)\})$$

$$\mathrm{BEL}_{X_i} = \mathrm{bel}_{0X_i} \oplus \left(\mathrm{Bel}_A \oplus \left(\oplus \left\{\mathrm{bel}_{0X_j} | X_j \in A,\ X_j \neq X_i\right\}\right)\right)^{\downarrow X_i}.$$

The message $M_{Y \to A}$ from Y to A is computed as follows: for $Y = \{Y_1, \ldots, Y_n\}$,

$$\mathrm{bel}_{Y \to A}(a) = \sum_{y \subseteq \Theta_Y} m_A(a \| y) \mathrm{bel}_{\mathrm{ne}(Y)/A \to Y},$$

where

$$\mathrm{bel}_{\mathrm{ne}(Y)/A \to Y} = \mathscr{B}_Y \oplus \left(\oplus \left\{\mathrm{bel}_{0Y_j} | Y_j \in Y\right\}\right) \oplus (\oplus \{M_{Z \to Y} | Z \in \mathrm{ne}(Y), Z \neq A\}).$$

From the above propagation scheme, it can be found that, in an ENC, any computations involving two connected variables (or merged nodes), say X and Y, are proposed on the space Θ_X or Θ_Y, while in the network with joint beliefs, such computations are always on the product space $\Theta_{X \cup Y}$. Thus the computation in the former needs fewer set comparisons and multiplications than that in the latter. Although the above representation and propagation algorithm are for networks which only have binary relations between the nodes, it could be generalized to the case where relations are for any number of nodes by using a graphical representation such as directed valuation networks [8].

5. EFFICIENT COMPUTATION FOR THE ENC WITH PARTIAL DEPENDENCY

In the previous section, we have proposed some ideas to solve the problem where there are loops in the network. For the case where there are very complicated loops, the computation is not quite obvious. In this section, we show that for some ENCs with complicated structure but with some special properties called partial dependency, we can reduce the computation by simplifying the structure.

5.1. ENC with Partial Dependency

In this subsection, we give the definition of an ENC with partial dependency and show some properties of such networks.

DEFINITION 5.1 *Given two variables A, X in an ENC, an edge (A, X) represents* $\{\text{bel}_X(\cdot|a_i) : a_i \in \Theta_A\}$. *The set* $\{a_i \setminus m_X(\Theta_X|a_i) = 1,\ a_i \in \Theta_A\}$ *is called an irrelevant set to X, and denoted by $\mathcal{I}_A^X$. If X is a set of variables, then the irrelevant set to X is defined as* $\mathcal{I}_A^X = \{a_j \setminus m_{X_i}(\Theta_{X_i}|a_j) = 1,\ X_i \in X\}$.

From the definition, it's easy to see that $\forall a_j \in \bar{\mathcal{I}}_A^X$, $m_X(\Theta_X|a_j) < 1$, where $\bar{\mathcal{I}}_A^X$ is the complement of $\mathcal{I}_A^X$ relative to Θ_A. Thus we say that $\bar{\mathcal{I}}_A^X$ is *relevant to X*. Such relations occur commonly in diagnosis problems and rule-based systems. In Example 1, we say that a is relevant to B, but $\bar{a}$ irrelevant to B. Intuitively, it means that given some knowledge about a, we can induce knowledge about B, but no matter what we know about $\bar{a}$, we can't induce any new knowledge about B. Thus we say $\bar{a}$ is irrelevant to B.

DEFINITION 5.2 *An ENC is called an ENC with partial dependency if there exists a variable A such that for any $X \in \mathbf{U}$ and $(A, X) \in \mathbf{E}$, $\mathcal{I}_A^X \neq \emptyset$ and $\mathcal{I}_A^X \neq \Theta_A$.*

The following two lemmas state the properties of the simplest ENC with partial dependency, where there are only two variables in the networks.

LEMMA 5.1 *Given two variables* X, Y *and* $\{\mathrm{bel}_Y(\cdot|x_i): x_i \in \Theta_X\}$, *let* $\mathscr{I}_X^Y$ *be as defined above. Then for any* $S \subseteq \Theta_X$, $m_Y(\Theta_Y|S) = 1$ *if* $S \cap \mathscr{I}_X^Y \neq \emptyset$.

Proof The result can be directly derived by using the GBT. ■

LEMMA 5.2 *Given two variables* X, Y *and* $\{\mathrm{bel}_Y(\cdot|x_i): x_i \in \Theta_X\}$, *let* $\mathscr{I}_X^Y$ *be as defined above. Suppose that we have some belief* bel_{0Y} *on* Y. *By Theorems* 2.1–2.3, *we can compute the belief of* X. *If* $m_X(S) \neq 0$, *then* $S \supseteq \mathscr{I}_X^Y$.

Proof From Lemma 5.1 we have that, $\forall x \subseteq \Theta_X$, if $x \cap \mathscr{I}_X^Y \neq \emptyset$, then $m_Y(\Theta_Y|x) = 1$, i.e., $\mathrm{pl}_Y(y|x) = 1$ for all $y \subseteq \Theta_Y$. Then, by Theorems 2.2 and 2.3, $\mathrm{pl}_X(x|y) = \mathrm{pl}(y|x) = 1$ for such x. Thus by Lemma 3.1,

$$\mathrm{pl}_X(x) = \sum_{y \subseteq \Theta_Y} m_0(y)\,\mathrm{pl}_X(x|y) = \sum_{y \subseteq \Theta_Y} m_0(y) = 1.$$

Therefore, for all $S \subseteq \Theta_X$, if $m_X(S) \neq 0$, S must contain all the elements of $\mathscr{I}_X^Y$, i.e., $S \supseteq \mathscr{I}_X^Y$. ■

5.2. Computation in ENC with Partial Dependency

From the properties of an ENC with partial dependency (Lemmas 5.1 and 5.2) described above, we can simplify the computation for some cases of such ENCs. Consider the network shown in Figure 3, where in (a), G_i is a set of variables, and suppose $\mathscr{I}_A^{G_i}$ is irrelevant to G_i. Figure 3(b) shows the details in each G_i. To describe the computation, let's begin by recalling the concept of partition.

DEFINITION 5.3 *Let* $\Theta = \{\theta_1, \ldots, \theta_p\}$ *be a frame of discernment. A set* $\mathscr{P}_\Theta$ *of subsets of* Θ *is a partition of* Θ *if the elements in* $\mathscr{P}_\Theta$ *are all nonempty and disjoint, and their union is* Θ. *We also call* $\mathscr{P}_\Theta$ *a coarsening of* Θ, *and* Θ *a refinement of* $\mathscr{P}_\Theta$.

From the definition, we have that for all $\theta_i \in \Theta$, there exists $x_j \in \mathscr{P}_\Theta$

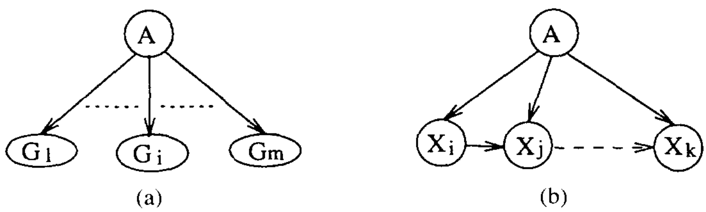


Figure 3. An ENC with partial dependency, where (b) illustrates the details of G_i of (a).

such that $\theta_i \in x_j$. We denote such a relation by $\Lambda(\theta_i) = x_j$. Then, $\forall\theta \subseteq \Theta$, $\Lambda(\theta) = \{\Lambda(\theta_i) \setminus \theta_i \in \Theta\}$. Let bel_1 be a belief function on Θ; then the belief bel_2 on $\mathscr{P}_\Theta$ induced by bel_1, say, by coarsening, is obtained as follows: $\forall x \subseteq \mathscr{P}_\Theta$,

$$m_2(x) = \sum_{\Lambda(\theta)=x} m_1(\theta).$$

Let bel_2 be a belief function on $\mathscr{P}_\Theta$; then the belief bel_1 on Θ induced by bel_2, say, by refinement, is obtained as follows: $\theta \subseteq \Theta$, $x \subseteq \mathscr{P}_\Theta$,

$$m_1(\theta) = m_2(x),$$

where $\theta = \bigcup\{\theta' \setminus \Lambda(\theta') = x\}$.

Given a network as shown in Figure 3, we can represent it as shown in Figure 4. Each A_i has a frame Θ_{A_i} which is a partition of Θ_A such that $\forall a_k \in \Theta_A$, $\Lambda(a_k) = \mathscr{I}_A^{G_i}$ if $a_k \in \mathscr{I}_A^{G_i}$ otherwise $\Lambda(a_k) = \{a_k\} = \hat{a}_k$. Each $A_i \to G_i$ part can be regarded as a subnetwork, and the belief functions passed between A and A_i are performed by refinement and coarsening between the two frames. Let's look at the following example:

Example 5.1 Suppose we have four variables in the network shown in Figure 5(a): A, X, Y, and Z. Their frames are $\Theta_A = \{a_1, a_2, a_3, a_4, a_5\}$, $\Theta_X = \Theta_Y = \Theta_Z = \{+, -\}$. The relations among them are represented by conditional belief functions in Table 4.

Now suppose we have some prior beliefs on X and Z: $m_{0X}(+) = 0.8$, $m_{0X}(\Theta_X) = 0.2$, $m_{0Z}(-) = 1$. To compute the marginal for A, if we use the joint belief for the relation between A and X, then the combination is performed on the product space $\Theta_{A \cup X}$ and $\Theta_{A \cup Z}$; if we use the condi-

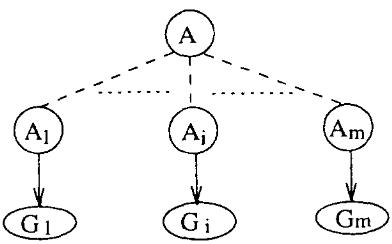


Figure 4. Alternative structure of the network in Figure 3 for simplified computation.

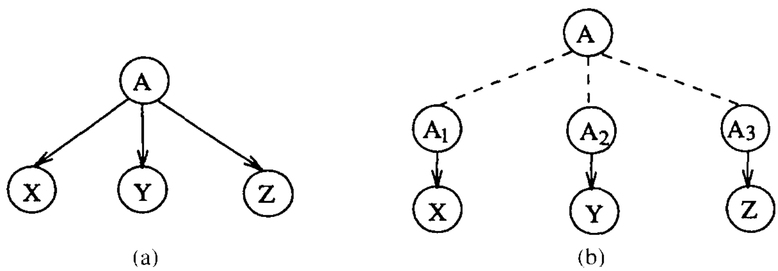
Table 4. Conditional Belief Functions for Example 1

	$m_X(x \vert a_i), i = 1,\ldots,5$					$m_Y(y \vert a_i), i = 1,\ldots,5$				
	a_1	a_2	a_3	a_4	a_5	a_1	a_2	a_3	a_4	a_5
+	.9	.7	0	0	0	0	.7	.2	.4	0
−	.0	.3	0	0	0	0	.3	.6	.1	0
Θ	.1	.0	1	1	1	1	.0	.2	.5	1

	$m_Z(z \vert a_i), i = 1,\ldots,5$				
	a_1	a_2	a_3	a_4	a_5
+	0	0	0	.6	.9
−	0	0	0	.3	.0
Θ	1	1	1	.1	.1

tional beliefs represented in Table 4 and the propagation scheme described above, the computation is performed on the frame Θ_A, which is more efficient. Moreover, if we use the result of Lemma 5.2, the computation can be simplified further. The following steps illustrate such a computation:

1. Transform the network in Figure 5(a) to the network shown in Figure 5(b), where each Θ_{A_i} is a partition of Θ_A: $\Theta_{A_1} = \{\hat{a}_1, \hat{a}_2, \mathscr{I}_1\}$, where $\mathscr{I}_1 = \mathscr{I}_A^X = \{a_3, a_4, a_4\}$; $\Theta_{A_2} = \{\hat{a}_2, \hat{a}_3, \hat{a}_4, \mathscr{I}_2\}$, where $\mathscr{I}_2 = \mathscr{I}_A^Y = \{a_1, a_5\}$; and $\Theta_{A_3} = \{\hat{a}_4, \hat{a}_5, \mathscr{I}_3\}$, where $\mathscr{I}_3 = \mathscr{I}_A^Z = \{a_1, a_2, a_3\}$. Then $\text{bel}_X(\cdot|\hat{a}_j)$, $\hat{a}_j \in \Theta_{A_1}$, is obtained from $\text{bel}_X(\cdot|a_j)$, $a_j \in \Theta_A$, and $\text{bel}_X(\cdot|\mathscr{I}_1)$, $\mathscr{I}_1 \in \Theta_{A_1}$, is obtained by applying the DRC. Symmetrically, we can get the other two conditional beliefs. The resulting conditional beliefs are shown in Table 5 (in the table, we use A_i for Θ_{A_i}, for short).
2. Use the DRC to compute $\text{bel}_{A_1}(\cdot|x)$ and $\text{bel}_{A_3}(\cdot|z)$.



(a) (b)

Figure 5. A simple example of an ENC (a) and its structure for computation (b).

Table 5. Conditional Beliefs Induced from Table 4 for the Partition of Θ_A

	$m_X(x\|a_i), a_i \in A_1$			$m_Y(y\|a_i), a_i \in A_2$				$m_Z(z\|a_i), a_i \in A_3$		
	$\hat{a}_1$	$\hat{a}_2$	$\mathscr{I}_1$	$\hat{a}_2$	$\hat{a}_3$	$\hat{a}_4$	$\mathscr{I}_2$	$\hat{a}_4$	$\hat{a}_5$	$\mathscr{I}_3$
+	.9	.7	0	.7	.2	.4	0	.6	.9	0
−	.0	.3	0	.3	.6	.1	0	.3	.0	0
Θ	.1	.0	1	.0	.2	.5	1	.1	.1	1

3. Use Theorems 2.2 and 2.3 to compute bel_{A_i}, $i = 1,2,3$. In particular, bel_{A_2} is vacuous by Proposition 3; $m_{A_1}(\{\hat{a}_1, \mathscr{I}_1\}) = 0.24$, $m_{A_1}(\Theta_{A_1}) = 0.76$; and

$$m_{A_3}(\{\mathscr{I}_3\}) = 0.54, \qquad m_{A_3}(\{\hat{a}_4, \mathscr{I}_3\}) = 0.36,$$

$$m_{A_3}(\{\hat{a}_5, \mathscr{I}_3\}) = 0.06, \qquad m_{A_3}(\Theta_{A_3}) = 0.04.$$

4. Compute the above two beliefs on the frame Θ_A by refinement, and combine them to get our desired result.

Obviously, this computation is more efficient, since in steps 2 and 3, the computation is on the frame Θ_{A_i}, which is smaller than Θ_A.

Moreover, if the network has the properties defined below, we can also simplify the computation for each subnetwork shown in Figure 3(b) if the network has some unrelated variables defined as follows.

Definition 5.4 *Let X, Y, and A be three variables in an ENC with partial dependency. Suppose we have* $\text{bel}_X(\cdot|a_i)$ *and* $\text{bel}_Y(\cdot|a_i)$ *for* $a_i \in \Theta_A$. *Let* $\mathscr{I}_A^X, \mathscr{I}_A^Y$ *be defined as in Definition* 5.1. *We say that X and Y are unrelated through A, denoted by $\mu(X,Y,A)$, if one of the following conditions is satisfied*:

1. $\mathscr{I}_A^X \cap \mathscr{I}_A^Y \neq \emptyset$, *or*
2. $\mathscr{I}_A^X \cap \mathscr{I}_A^Y = \emptyset$, $\mathscr{I}_A^X \cup \mathscr{I}_A^Y = \Theta_A$, *and* $\text{bel}_X(\cdot|\overline{\mathscr{I}}_A^X)$ *or* $\text{bel}_Y(\cdot|\overline{\mathscr{I}}_A^Y)$ *obtained from* $\text{bel}_X(\cdot|a_i)$ *or* $\text{bel}_Y(\cdot|a_i)$ *by the DRC is vacuous.*

This relation can also be extended to two disjoint subsets, where $\mathscr{I}_A^B = \bigcap_{X \in G} \mathscr{I}_A^X$. The following theorems and their corollaries give solutions for simplifying the computation.

Theorem 5.1 *Let X, Y, and A be defined as in Definition* 5.4, *and* $\{\text{bel}_Y(\cdot|x_i) : x_i \in \Theta_X\}$ *be a family of conditional belief functions for Y given X. Suppose $\mu(X,Y,A)$ and that we have no* a priori *information on A or Y, but on X we have* bel_{0X}. *Then* BEL_Y *is only dependent on* $\text{bel}_Y(\cdot|x_i)$ *and* bel_{0X}.

THEOREM 5.2 *Let X, Y, and A be three variables in an ENC with partial dependency as shown in Figure 6(a). Let* $\mathscr{I}_A^X, \mathscr{I}_A^Y$ *be defined as in Definition* 5.1 *and* $\bar{\mathscr{I}}_A^X \cap \bar{\mathscr{I}}_A^Y = \emptyset$. *Suppose we have the conditional belief function* $\{\mathrm{bel}_Y(\cdot|x_i) : x_i \in \Theta_X\}$ *for Y given X, and that we have no* a priori *information on A or Y, but on X we have* bel_{0X}. *Then if there is only one focal element in* bel_{0X}, *i.e.,* $m_X(x_0) = 1$ *where* $x_0 \subseteq \Theta_X$, *then* BEL_A *can be computed by Figure* 6(*b*).

COROLLARY 1 *Let X, Y, and A be as in Definition* 5.4, *and* $\{\mathrm{bel}_Y(\cdot|x_i) : x_i \in \Theta_X\}$ *be a family of conditional belief functions for Y given X as shown in Figure* 6(*a*). *Suppose* $\mu(X, Y, A)$ *and that we have no* a priori *belief on A, X, and Y. If* $\mathrm{bel}_Y(\cdot|x_i)$ *is noninformative over both X and Y, then the marginals for all variables are vacuous.*

Proof This result can be directly obtained from Theorems 5.1 and 5.2 by considering that there is a prior belief on X such that $m(\Theta_X) = 1$. ■

COROLLARY 2 *Let X, Y, and A be as in Definition* 5.4, *and* $\{\mathrm{bel}_Y(\cdot|x_i) : x_i \in \Theta_X\}$ *be a family of conditional belief functions for Y given X such that it is noninformative over both X and Y (Figure* 6(*a*)). *Suppose* $\mu(X, Y, A)$ *and that we have no prior beliefs on A or Y. Let* bel_{0X} *be* a priori *on X. Then* BEL_Y *is computed from* bel_{0X} *and* $\mathrm{bel}_Y(\cdot|x_i)$, *and* BEL_A *is computed as follows: Let* $m_{A|x}$ $(x \subseteq \Theta_X)$ *denote the resulting belief for A when* $m_{0X}(x) = 1$; *then* $\forall a \subseteq \Theta_A$,

$$m(a) = \sum_{x \subseteq \Theta_X} m_{0X}(x) m_{A|x}(a),$$

Proof This can be obtained by applying Theorems 2.3 and 5.2. ■

COROLLARY 3 *Let X, Y, and A be defined as in Definition* 5.4 *Suppose we have no prior beliefs on A or Y. Now suppose we have observations about X. Then* BEL_Y *is vacuous if* $\mu(X, Y, A)$.

Proof This can be obtained directly from Theorem 5.1 by considering that there is no relation between X and Y. ■

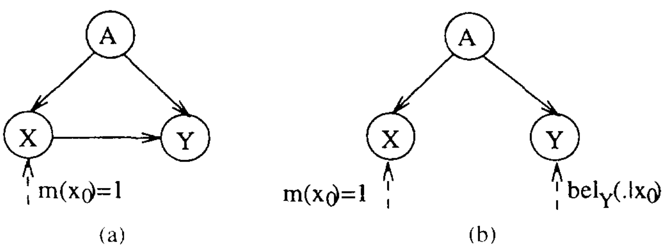


Figure 6. A simple case of ENC where the computation can be simplified.

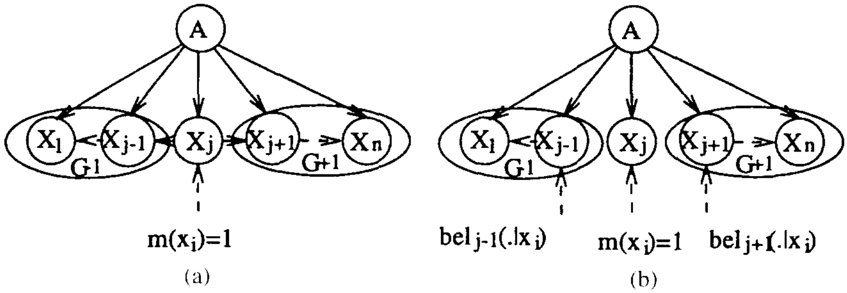


Figure 7. A general case of an ENC whose computation can be simplified.

The above theorems and the corollaries are only for a network with three variables. However, the results can be extended to more general cases.

THEOREM 5.3 *Consider an ENC as shown in Figure 7(a). Suppose we have* a priori *belief* bel_{0j} *for* X_j *such that* $m_{0j}(x_i) = 1$ *where* $x_i \in \Theta_{X_j}$. *Then the network in Figure 7(a) is equivalent to the one in Figure 7(b).*

THEOREM 5.4 *Let* G_1 *and* G_2 *be two sets of variables as shown in Figure 8(a). Suppose some elements* $\bar{\mathcal{F}}_A^i$ $(i = 1, 2)$ *of* A *are irrelevant to* G_i, *and* $\bar{\mathcal{F}}_A^1 \cap \bar{\mathcal{F}}_A^2 = \emptyset$. *Suppose there are some prior beliefs for the variables in* G_1, *say on* X_1, *and from the chain* $X_1 \ldots X_n$, *we get that the belief on* X_{j+1} *is vacuous. Then* BEL_A *can be computed by the network shown in Figure 8(b).*

It's easy to see that Figures 7(b) and 8(b) have similar structure to Figure 3(a). Thus they can be simplified further.

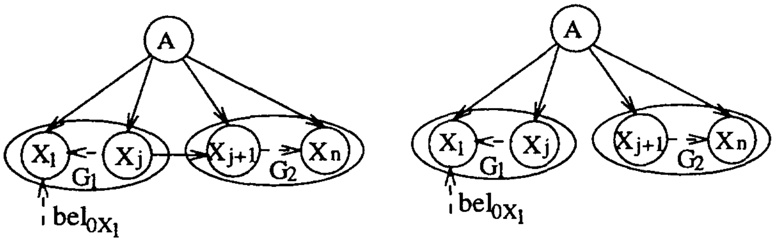


Figure 8. Another case of an ENC whose computation can be simplified.

5.3. An Example of Efficient Computation in ENC

The following example shows how to use the theorems in the previous subsection to reduce the computation for a complicated ENC with partial dependency.

EXAMPLE 5.2 Figure 9 shows an ENC representation for an example of nuclear waste disposal in [18]. In the network, we have 43 variables: 1 diagnosis, 21 tests, and 11 symptoms, where $\Theta_{\text{diagnosis}} = \{a, b, c, d, e, f, g, \omega\}$, $\Theta_{s_i} = \{\text{yes}, \text{no}\} : i = 1, \ldots, 11$, $\Theta t_j = \{+, -\} : j = 1, \ldots, 21$.

The conditional belief functions among the variables are as follows:

- Table 6 shows the conditional belief for s_i $(i = 1, \ldots, 11)$ given the diagnosis. The belief function for s_i given x_i $(\in \Theta_{\text{diagnosis}})$ is a simple support function. For example, the relation of diagnosis and s_1 is represented as $m_{s_1}(\{\text{yes}\}|f) = 0.4$, $m_{s_1}(\Theta_{s_1}|f) = 0.6$.
- The conditional beliefs for t_i $(i = 15, \ldots, 21)$ given the diagnosis are $m_{t_i}(+|d_i) = 0.99$, $m_{t_i}(\Theta_{t_i}|d_i) = 0.01$, $m_{t_i}(\Theta_{t_i}|\omega) = 1$, where $d_{15}, \ldots, d_{21}$ represent $a, \ldots, g$, respectively. The relation between diagnosis and t_{12} is $m_{t_{12}}(+|\omega) = 1$.
- The condition beliefs for t_i given s_i $(i = 1, \ldots, 11)$ are $m_{t_i}(+|\text{yes}) = 0.99$, $m_{t_i}(\Theta_{t_i}|\text{yes}) = 0.01$, $m(\Theta_{t_i}|\text{no}) = 1$.
- The conditional beliefs for t_i given t_j if there is an arrow connecting them $(i, j = 1, \ldots, 10, 12, 13, 14)$ are $m_{t_i}(+|+) = 0.9$, $m_{t_i}(\Theta_{t_i}|+) = 0.1$, $m_{t_i}(\Theta_{t_i}|-) = 1$.

In order to solve the problem, we need the following computations:

1. The belief on diagnosis, given that we know the result of each test.
2. Suppose we have the result of some test. We need to compute the belief on diagnosis given that we also know the result of another test.

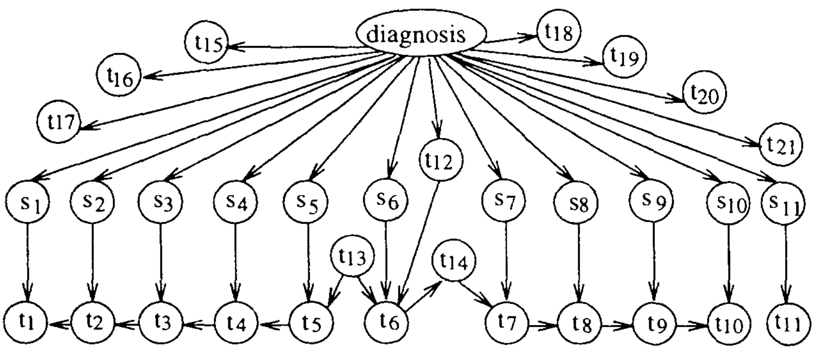


Figure 9. An ENC for an example nuclear waste disposal problem.

Table 6. Conditional Belief Functions between Diagnoses and s_i's

	a	b	c	d	e	f	g	ω
$s_1(+)$						.4		
$s_2(+)$	.95							
$s_3(+)$						.5		
$s_4(+)$						.6		
$s_5(+)$		.9						
$s_6(+)$							.5	
$s_7(+)$		.6						
$s_8(+)$				.2				
$s_9(+)$					.8			
$s_{10}(+)$			.6					
$s_{11}(+)$			.7	.3				

3. We may also compute the belief on diagnosis given that we know the results of several tests.

Here we only show how to simplify the network for the computation according to the above theorems.

1. Suppose we have that the result of t_4 is "$-$". According to Theorem 5.3, we remove the edges (t_3, t_4) and (t_5, t_4), by assigning belief on t_5 as $m(-) = 0.9$, $m(\Theta_{t_5}) = 0.1$. According to Theorem 5.4, we remove the edge (t_7, t_8). The network is thus separated into 11 subnetworks as shown in Figure 10, where the frame of diagnosis can be reduced in each subnetwork shown in the variable D_i. From Corollary 1, it is not difficult to see that the subnetworks with $D_i (i = 1, 4, \ldots, 11)$ will have vacuous beliefs on D_i after the propagation. Thus the computa-

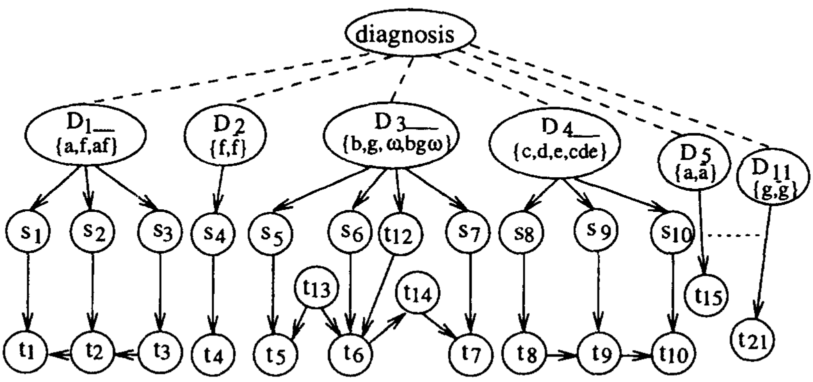


Figure 10. Simplified graph for Figure 9.

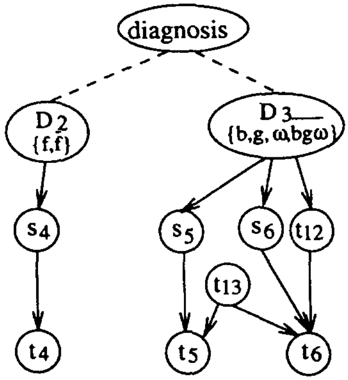


Figure 11. Simplified graph for case 1.

tions are only needed in the network with D_2 and D_3. Furthermore, again by Theorem 5.4, the variables t_{14}, t_7, and s_7 can be removed. Therefore, the biggest subnetwork has only seven variables, and the size of the biggest frame is 4 (shown in Figure 11).

2. Suppose we have that the results for test_{12} are $-$ and for test_2 are $+$. Using the same strategy, the network will be simplified to Figure 12.
3. Similarly, we can reduce the computation in case 3.

From Example 2, we find that, for the ENC with partial dependency, the network can be simplified by cutting some loops and separating the network into smaller subnetworks. Obviously, this makes the computation more efficient in the simplified structure than in the original one.

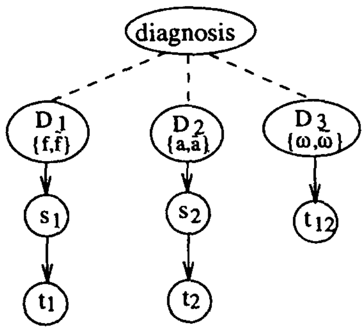


Figure 12. Simplified graph for case 2.

6. CONCLUSIONS

We have presented an evidential network (ENC) which uses conditional belief functions for the knowledge representation and reasoning. By comparing some relations between the representations by joint belief and by conditional belief, it is found that the conditional form is more natural and it takes less space. We also provided an algorithm for reasoning in ENCs. The presented algorithm of reasoning is only for a network where all the relations are binary; the extension of the algorithm to a general case will be studied in future work. We have shown that the computation of an ENC can be simplified due to the property of partial dependency. Further studies are still needed to systematize this simplification process, and it may be conjectured that one possible solution is to represent the knowledge in several networks instead of in one global network.

APPENDIX: PROOFS OF THE THEOREMS

Proof of Theorem 5.1

Before proving the theorem, we first state the following lemma which will be needed in the proof.

LEMMA 6.1 *Let X, Y, and A be three variables in an ENC with partial dependency. Let $\mathscr{I}_A^X, \mathscr{I}_A^Y$ be defined as in Definition 5.1. Suppose $\bar{\mathscr{I}}_A^X \cap \bar{\mathscr{I}}_A^Y = \emptyset$. If we merge X and Y as one node, say S, then $\forall \omega \subseteq \Theta_{X \cup Y}$, $a \subseteq \Theta_A$,*

$$\mathrm{pl}_{X \times Y}(\omega | a) = 1 - \left[1 - \mathrm{pl}_X\left(\omega^{\downarrow X} | a \cap \bar{\mathscr{I}}_A^X\right)\right]\left[1 - \mathrm{pl}_Y(\omega^{\downarrow Y} | a \cap \mathscr{I}_A^X)\right]. \tag{6.1}$$

Proof Let $A_1 = \bar{\mathscr{I}}_A^X$, $A_2 = \bar{\mathscr{I}}_A^Y$, $A_3 = \mathscr{I}_A^X \cap \mathscr{I}_A^Y$. It is easy to see that A_1, A_2, and A_3 constitute a partition of Θ_A, and that A_1 contains the elements relevant to X, but irrelevant to Y; A_2 those relevant to Y, but irrelevant to X; and A_3 those irrelevant to either of X and Y.

From the inference algorithm described in the above section, we merge X and Y as one node, by applying Equation (4.1), for all $\omega \subseteq \Theta_{X \cup Y}$:

$$m_{X \times Y}(\omega | a_i) = \sum_{x^{\uparrow X \cup Y} \cap y^{\uparrow X \cup Y} = \omega} m_X(x | a_i) m_Y(y | a_i).$$

Since A_1 is irrelevant to Y, we have that $\forall a_i \in A_1$, $m_{X \times Y}(\omega | a_i) = m_X(\omega^{\downarrow X} | a_i)$; thus $\mathrm{pl}_{X \times Y}(\omega | a_i) = \mathrm{pl}_X(\omega^{\downarrow X} | a_i)$. Symmetrically, $\forall a_j \in A_2$, $\mathrm{pl}_{X \times Y}(\omega | a_j) = \mathrm{pl}_Y(\omega^{\downarrow Y} | a_j)$. Since A_3 is irrelevant to either of X or Y,

$\forall a_k \in A_3$, $\mathrm{pl}_{X\times Y}(\omega|a_k) = \mathrm{pl}_X(\omega^{\downarrow X}|a_k) = \mathrm{pl}_Y(\omega^{\downarrow Y}|a_k) = 1$. By applying the DRC, we have the $\forall a \subseteq \Theta_A$,

$$\begin{aligned}
\mathrm{pl}_{X\times Y}(\omega|a) &= 1 - \prod_{a_i \in a} [1 - \mathrm{pl}_{X\times Y}(\omega|a_i)] \\
&= 1 - \prod_{a_i \in a\cap A_1} [1 - \mathrm{pl}_{X\times Y}(\omega|a_i)] \\
&\quad \times \prod_{a_j \in a\cap(A_2\cup A_3)} \left[1 - \mathrm{pl}_{X\times Y}\left(\omega|a_j\right)\right] \\
&= 1 - \prod_{a_i \in a\cap A_1} \left[1 - \mathrm{pl}_X(\omega^{\downarrow X}|a_i)\right] \\
&\quad \times \prod_{a_j \in a\cap(A_2\cup A_3)} \left[1 - \mathrm{pl}_Y\left(\omega^{\downarrow Y}|a_j\right)\right] \\
&= 1 - \left[1 - \mathrm{pl}_X\left(\omega^{\downarrow X}|a \cap \bar{\mathcal{I}}_A^X\right)\right]\left[1 - \mathrm{pl}_Y(\omega^{\downarrow Y}|a \cap \mathcal{I}_A^X)\right],
\end{aligned}$$

which concludes the proof. ■

Now let's look at the proof of Theorem 5.1:

Proof Let $\mathcal{I}_A^X, \mathcal{I}_A^Y \subseteq \Theta_A$ be the sets of irrelevant elements to X and Y, respectively. By merging X and Y as one node, say S, and by applying Lemma 6.1, we have that, $\forall \omega \subseteq \Theta_{X\cup Y}$,

$$\mathrm{pl}_{X\times Y}(\omega|\Theta_A) = 1 - \left[1 - \mathrm{pl}_X\left(\omega^{\downarrow X}|\bar{\mathcal{I}}_A^X\right)\right]\left[1 - \mathrm{pl}_Y(\omega^{\downarrow Y}|\mathcal{I}_A^X)\right]. \quad (6.2)$$

Since $\mu(X, Y, A)$, suppose $\mathcal{I}_A^X \cap \mathcal{I}_A^Y \neq \emptyset$; then for all $a \subseteq \Theta_A$ such that $a \cap \mathcal{I}_A^X \cap \mathcal{I}_A^Y \neq \emptyset$, we have $\mathrm{pl}_Y(y|a \cap \mathcal{I}_A^X) = 1$ for all $y \subseteq \Theta_Y$. Thus,

$$\mathrm{pl}_{X\times Y}(\omega|\Theta_A) = 1 - \left[1 - \mathrm{pl}_X\left(\omega^{\downarrow X}|\bar{\mathcal{I}}_A^X\right)\right](1 - 1) = 1. \quad (6.3)$$

If $\mathcal{I}_A^X \cap \mathcal{I}_A^Y = \emptyset$, $\mathcal{I}_A^X \cup \mathcal{I}_A^Y = \Theta_A$, then Equation (6.2) can be written as

$$\mathrm{pl}_{X\times Y}(\omega|\Theta_A) = 1 - \left[1 - \mathrm{pl}_X\left(\omega^{\downarrow X}|\bar{\mathcal{I}}_A^X\right)\right]\left[1 - \mathrm{pl}_Y\left(\omega^{\downarrow Y}|\bar{\mathcal{I}}_A^Y\right)\right]. \quad (6.4)$$

Suppose $\mathrm{bel}_Y(\cdot|\bar{\mathcal{I}}_A^Y)$ obtained from $\mathrm{bel}_Y(\cdot|a_i)$ by the DRC is vacuous; then $\forall y \subseteq \Theta_Y$, $\mathrm{pl}_Y(y|\bar{\mathcal{I}}_A^Y) = 1$; thus

$$\mathrm{pl}_{X\times Y}(\omega|\Theta_A) = 1 - \left[1 - \mathrm{pl}_X\left(\omega^{\downarrow X}|\bar{\mathcal{I}}_A^X\right)\right](1 - 1) = 1. \quad (6.5)$$

Suppose $\mathrm{bel}_X(\cdot|\bar{\mathcal{I}}_A^X)$ obtained from $\mathrm{bel}_X(\cdot|a_i)$ by the DRC is vacuous; then $\forall x \subseteq \Theta_X$, $\mathrm{pl}_X(x|\bar{\mathcal{I}}_A^X) = 1$; thus

$$\mathrm{pl}_{X\times Y}(\omega|\Theta_A) = 1 - (1 - 1)\left[1 - \mathrm{pl}_Y\left(\omega^{\downarrow Y}|\bar{\mathcal{I}}_A^Y\right)\right] = 1. \quad (6.6)$$

From Equations (6.3), (6.5), and (6.6), we have that $\mathrm{bel}_{X\times Y}(\cdot|\Theta_A)$ is vacuous. As there is no *a priori* belief on A, $\mathrm{bel}_{A\to S}$ is vacuous during the propagation. Therefore BEL_Y is dependent only the $\mathrm{bel}_Y(\cdot|x_i)$ and bel_{0Y}. ■

Proof of Theorem 5.2

Proof Since $\mathcal{I}_A^X, \mathcal{I}_A^Y$ are irrelevant to X and Y, respectively, by merging X and Y as one node, say S, and by applying Lemma 6.1, we have $\forall \omega \subseteq \Theta_{X \cup Y}$, $a \subseteq \Theta_A$,

$$\begin{aligned} \mathrm{pl}_A(a|\omega) &= \mathrm{pl}_{X \times Y}(\omega|a) \\ &= 1 - \left[1 - \mathrm{pl}_X\left(\omega^{\downarrow X}|a \cap \bar{\mathcal{I}}_A^X\right)\right]\left[1 - \mathrm{pl}_Y(\omega^{\downarrow Y}|a \cap \mathcal{I}_A^X)\right]. \end{aligned}$$

Let bel_{XY} be the joint belief over $X \times Y$ of $\{\mathrm{bel}_Y(\cdot|x_i): x_i \in \Theta_X\}$, by applying Lemma 3.3. Then each focal element s of bel_{XY} must satisfy $s^{\downarrow X} = \Theta_X$. Since $m_{0X}(x) = 1$, let bel_{0XY} be the resulting belief function on combining bel_{0X} and bel_{XY}. It's easy to see that each focal element ω of bel_{0XY} must satisfy $\omega^{\downarrow X} = x$ and $\mathrm{bel}_{0XY}^{\downarrow Y} = \mathrm{bel}_Y(\cdot|x)$. To compute BEL_A, we use Theorem 2.3:

$$\begin{aligned} \mathrm{pl}_A(a) &= \sum_{\omega \subseteq \Theta_{X \cup Y}} m_{0XY}(\omega)\, \mathrm{pl}_A(a|\omega) \\ &= \sum_{\omega \subseteq \Theta_{X \cup Y}} m_{0XY}(\omega) \\ &\quad \times \left\{1 - \left[1 - \mathrm{pl}_X\left(\omega^{\downarrow X}|a \cap \bar{\mathcal{I}}_A^X\right)\right]\left[1 - \mathrm{pl}_Y(\omega^{\downarrow Y}|a \cap \mathcal{I}_A^X)\right]\right\} \\ &= \sum_{\omega \subseteq \Theta_{X \cup Y}} m_{0XY}(\omega) \\ &\quad \times \left\{1 - \left[1 - \mathrm{pl}_X\left(x|a \cap \bar{\mathcal{I}}_A^X\right)\right]\left[1 - \mathrm{pl}_Y(\omega^{\downarrow Y}|a \cap \mathcal{I}_A^X)\right]\right\} \\ &= \sum_{\omega^{\downarrow Y} \subseteq \Theta_Y} m_{0XY}^{\downarrow Y}(\omega) \\ &\quad \times \left\{1 - \left[1 - \mathrm{pl}_X\left(x|a \cap \bar{\mathcal{I}}_A^X\right)\right]\left[1 - \mathrm{pl}_Y(\omega^{\downarrow Y}|a \cap \mathcal{I}_A^X)\right]\right\} \\ &= \sum_{\omega^{\downarrow Y} \subseteq \Theta_Y} \left[m_{0X} \oplus m_Y(\omega^{\downarrow Y}|x)\right] \\ &\quad \times \left\{1 - \left[1 - \mathrm{pl}_X\left(x|a \cap \bar{\mathcal{I}}_A^X\right)\right]\left[1 - \mathrm{pl}_Y(\omega^{\downarrow Y}|a \cap \mathcal{I}_A^X)\right]\right\}. \end{aligned} \tag{6.7}$$

Equation (6.7) implies that BEL_A can be computed from the network in Figure 6(b).

Proof of Theorem 5.3

Proof Let BEL be the joint belief function of the network, and BEL′ the belief function obtained by combining all the beliefs except $\mathrm{bel}_{j,j-1}$, $\mathrm{bel}_{j,j+1}$, and bel_{0j}, where $\mathrm{bel}_{j,j-1}, \mathrm{bel}_{j,j+1}$ are the joint beliefs of the corresponding conditional beliefs obtained by applying Lemma 3.3. Then, in Figure 7(a) we have

$$\mathrm{BEL} = \mathrm{BEL}' \oplus (\mathrm{bel}_{j,j-1} \oplus \mathrm{bel}_{j,j+1} \oplus \mathrm{bel}_{0j}). \tag{6.8}$$

Since bel_{0j} is such that $m(x_i) = 1$, it's easy to see that

$$\text{bel}_{j,j-1} \oplus \text{bel}_{0j} = \text{bel}_{0j} \oplus (\text{bel}_{j,j-1} \oplus \text{bel}_{0j})^{\downarrow X_{j-1}}.$$

and

$$\text{bel}_{j,j+1} \oplus \text{bel}_{0j} = \text{bel}_{0j} \oplus (\text{bel}_{j,j+1} \oplus \text{bel}_{0j})^{\downarrow X_{j+1}}.$$

From Theorem 2.3, we can find that

$$(\text{bel}_{j,j\pm 1} \oplus \text{bel}_{0j})^{\downarrow X_{j,j\pm 1}} = \text{bel}_{j\pm 1}(\cdot|x).$$

Then Equation (6.8) can be rewritten as

$$\text{BEL} = \text{BEL}' \oplus \left[\text{bel}_{j+1}(\cdot|x) \oplus \text{bel}_{j-1}(\cdot|x) \oplus \text{bel}_{0j}\right].$$

Thus, BEL is also the joint belief function of the network in Figure 7(b). As the variables in the two networks are in one-to-one correspondence, the marginal distributions of the two networks will be the same. Therefore, they are equivalent. ■

Proof of Theorem 5.4

The following lemma is needed for the proof of Theorem 5.4:

LEMMA 6.2 *Let* $\text{bel}_0, \text{bel}_1, \text{bel}_2$, *and* bel_3 *be four belief functions on* Θ_A. *Suppose* bel_3 *is such that* $\forall a \subseteq \Theta_A$, $m_3(a) = x_1 m_1(a) + x_2 m_2(a)$, *where* $x_1, x_2 \geq 0$, $x_1 + x_2 = 1$. *Let* bel_{ij} *denote the belief function resulted from the combination of* bel_i *and* bel_j. *Then* $\forall a \subseteq \Theta_A$,

$$m_{03}(a) = x_1 m_{01}(a) + x_2 m_{02}(a).$$

Proof

$$\begin{aligned} m_{03}(a) &= \sum_{b \cap c = a} m_0(b) m_3(c) = \sum_{b \cap c = a} m_0(b)[x_1 m_1(c) + x_2 m_2(c)] \\ &= \sum_{b \cap c = a} [x_1 m_0(b) m_1(c) + x_2 m_0(b) m_2(c)] \\ &= x_1 \sum_{b \cap c = a} m_0(b) m_3(c) + x_2 \sum_{b \cap c = a} m_0(b) m_3(c) \\ &= x_1 m_{01}(a) + x_2 m_{02}(a). \end{aligned}$$

This concludes the proof. ■

Proof of Theorem 5.4 Let Y and Z be the merged nodes of variables in G_1 and G_2 respectively. From Theorem 5.2, we have that, for any *a priori* belief such that $m_Y(y) = 1$, the BEL_A can be computed as follows:

$$m_{A_1}(a) = m_Y(y)m_A(a|y),$$

$$m_{A_2}(a) = \sum_{z \subseteq \Theta_Z} m_Z(z|y)m_A(a|z),$$

$$\mathrm{BEL}_{A|y} = \mathrm{BEL}_{A_1} \oplus (\mathrm{BEL}_{A_2} \oplus \mathrm{bel}_{0A}). \tag{6.9}$$

Suppose there are prior beliefs on some variables in G_1, and we get that, from the chain $X_1 \ldots X_n$, the belief on X_{j+1} is vacuous. Let bel_{0Y} be the joint belief for Y got from the *a priori* beliefs through the chain. It's easy to see that, for each focal element y of bel_{0Y}, $\mathrm{bel}_Z(\cdot|y)$ is always the same. Thus BEL_{A_2} is always the same. Then, from Lemma 6.2 and Equation (6.9), we have

$$\begin{aligned} m_A(a) &= \sum_{y \subseteq \Theta_Y} m_{0Y}(y)m_{A|y}(a) \\ &= \sum_{y \subseteq \Theta_Y} m_{0Y}(y)m_A(a|y) \oplus (\mathrm{BEL}_{A_2} \oplus \mathrm{bel}_{0A}) \sum_{y \subseteq \Theta_Y} m_{0Y}(y) \\ &= (\mathrm{BEL}_{A_2} \oplus \mathrm{bel}_{0A}) \oplus \sum_{y \subseteq \Theta_Y} m_{0Y}(y)m_A(a|y). \end{aligned}$$

The above equation is the solution for computing BEL_A in Figure 8(b). ■

ACKNOWLEDGMENT

The authors would like to thank the anonymous referees for their comments on the paper. This research work has been partially supported by the ESPRIT III, Basic Research Project Action 6156 (DRUMS II) funded by a grant from the Commission of the European Communities. The first author is supported by a grant of IRIDIA, Université Libre de Bruxelles.

References

1. Cano, J., Delgado, M., and Moral, S., An axiomatic framework for propagating uncertainty in directed acyclic networks, *Internat. J. Approx. Reason.* 8, 253–280, 1993.

2. Dubois, D., and Prade, H., A set theoretical view of belief functions, *Internat. J. Gen. Systems* 12, 193–226, 1986.

3. Pearl, J., *Probabilistic Reasoning in Intelligence Systems: Networks of Plausible Inference*, Morgan Kaufmann, Los Altos, Calif., 1988.

4. Shafer, G., *A Mathematical Theory of Evidence*, Princeton U.P., 1976.

5. Shafer, G., and Shenoy, P. P., Local computation in hypertrees, Working Paper No. 201, School of Business, Univ. of Kansas, Lawrence, 1988.

6. Shenoy, P. P., and Shafer, G., Axioms for probability and belief functions propagation, in *Uncertainty in Artificial Intelligence 4* (R. D. Shachter, T. S. Levitt, L. N. Kanal, and J. F. Lemmer, Eds.), North-Holland, Amsterdam, 159–198, 1990.

7. Shenoy, P. P., Valuation-based systems: A framework for managing uncertainty in expert systems, in *Fuzzy Logic for the Management of Uncertainty* (L. A. Zadeh and J. Kacprzyk, Eds.), 83–104, 1992.

8. Shenoy, P. P., Valuation networks and conditional independence, in *Proceedings of the 9th Conference on Uncertainty in Artificial Intelligence* (D. Heckerman and A. Mamdani, Eds.), Morgan Kaufmann, San Mateo, Calif., 191–199, 1993.

9. Shenoy, P. P., Conditional independence in valuation-based systems, *Internat. J. Approx. Reason.* 10, 203–234, 1994.

10. Smets, Ph., Un modèle mathématico-statistique simulant le processus du diagnostic médical, Doctoral Dissertation, Univ. Libre de Bruxelles, 1978.

11. Smets, Ph., Belief functions, in *Nonstandard Logics for Automated Reasoning* (Ph. Smets, A. Mamdani, D. Dubois, and H. Prade, Eds.), Academic, London, 253–286, 1988.

12. Smets, Ph., The combination of evidence of the transferable belief model, *IEEE Trans. Pattern Anal. and Machine Intell.* 12, 447–458, 1990.

13. Smets, Ph., Patterns of reasoning with belief functions, *J. Appl. Non-Classical Logics* 1(2), 166–170, 1991.

14. Smets, Ph., The transferable belief model and random sets, *Internat. J. Intell. Systems* 7, 37–46, 1992.

15. Smets, Ph., The concept of distinct evidence, *Proceedings of the 4th International Conference on Information Processing and Management of Uncertainty in Knowledge-Based Systems*, 789–794, 1992.

16. Smets, Ph., Belief functions: The disjunctive rule of combination and the generalized Bayesian theorem, *Internat. Approx. Reason.* 9, 1–35, 1993.

17. Smets, Ph., and Kennes, R., The transferable belief model, *Artificial Intelligence* 66, 191–234, 1994.

18. Xu, J., Hsia, T.-T., and Smets, Ph., Transferable belief model for decision making in valuation-based systems, *IEEE Trans. Systems Man Cybernet.*, to appear.