The Belief Calculus and Uncertain Reasoning

Yen-Teh Hsia*

IRIDIA, Université Libre de Bruxelles
50 av. F. Roosevelt, CP 194/6
1050, Brussels, Belgium
r01509@bbrbfu01.bitnet

Abstract

We formulate the Dempster-Shafer formalism of belief functions [Shafer 76] in the spirit of logical inference systems. Our formulation (called the belief calculus) explicitly avoids the use of set-theoretic notations. As such, it serves as an alternative for the use of the Dempster-Shafer formalism for uncertain reasoning.

1. Introduction

Traditionally, the "syntax" of the Dempster-Shafer (D-S) formalism [Shafer 76] has been set-theoretic in nature (e.g., [Gordon and Shortliffe 84; Kong 86; Shafer et al. 87; Yen 89]). In some cases, propositions may be used for belief specifications (e.g., [Smets 88; Zarley et al. 88]). However, to date, there is no purely logic-oriented formulation of this formalism.

Set-theoretic notations are appropriate when we are concerned with general theory rather than applications. But on the other hand, we might also find it difficult to use set-theoretic notations in some application domains. To overcome this notational disadvantage of the D-S formalism, we give an alternative formulation of belief functions in this paper. Our formulation (called the belief calculus) is developed along the lines of natural deduction systems, and it explicitly avoids the use of set-theoretic notations. This differs from the previous research [Ruspini 87; Fagin and Halpern 89] in which the main concern was the "structure" or semantics of the D-S formalism and not its syntax.

To show how the belief calculus may be used for uncertain reasoning, we give three examples. These examples model different real world situations, and they address issues such as independent random variables, belief dependency structures, and "distinct" sources of evidence.

The remainder of this paper is organized as follows. In Section 2, we describe the belief calculus. In Section 3, we show how the belief calculus may be used for uncertain reasoning. In Section 4, we discuss some related issues. Finally, Section 5 concludes.

* This work was supported in part by the DRUMS project funded by the Commission of the European Communities under the ESPRIT II-Program, Basic Research Project 3085.

2. The Belief Calculus

The multivariate formalism. Our formulation of the D-S formalism starts with the multivariate formalism [Kong 86]. That is, we assume that different aspects of the world that are of interests to us are already appropriately formulated as questions or *variables* (e.g., "Is the entity capable of flying?", "Can the object be used to cross the river?" etc). Each of these variables is associated with a set of mutually exclusive and exhaustive values (called the *frame* of the variable) representing all possible answers to the question. A *boolean variable* is one that has an associated frame of {Yes, No}.

Propositions. Primitive propositions (i.e., atoms) are of the form "SomeVariable = SomeValue". From propositions, we build compound propositions using five logical connectives (with the usual semantics): $\neg$ (not), $\vee$ (or), $\wedge$ (and), $\rightarrow$ (if ... then), $\leftrightarrow$ (if and only if).

As a basic requirement of the multivariate formalism, we assume that, for every variable and its associated frame, there is a corresponding *mutual exclusion axiom*. For example, if the frame of the variable A is {h, m, l}, then the mutual exclusion axiom associated with A and its frame is '$((A = h) \wedge \neg(A = m) \wedge \neg(A = l)) \vee (\neg(A = h) \wedge (A = m) \wedge \neg(A = l)) \vee (\neg(A = h) \wedge \neg(A = m) \wedge (A = l))$'. We use $\mathbb{ME}$ to denote the set of all mutual exclusion axioms. For convenience, we also use 'A' as an abbreviation for 'A = Yes' whenever A is a boolean variable (and '$\neg$A' will be logically equivalent to 'A = No' under $\mathbb{ME}$).

Let $A_1, A_2, \ldots, A_N$ be all variables, and let $\Theta_1, \Theta_2, \ldots, \Theta_N$ be their respective frames. A *valuation* is an assignment of an element of Θ_i ($1 \leq i \leq N$) to A_i for every i (i.e., an assignment of a value-vector to the variable-vector $<A_1, A_2, \ldots, A_N>$). A proposition P is said to be true under a valuation V if P is true when all variables occurring in P are replaced with their corresponding values in V; otherwise P is said to be false under V.

Formally, a valuation is defined as an assignment of values to all variables. But when the situation permits, we also use the word 'valuation' to mean a (partial) valuation of all variables occurring in some proposition.

Example 1. "(PLACE = Africa) → (TEMP = high)" is a proposition. It is false under the valuation <PLACE, TEMP> = <Africa, medium> and true under the valuation <PLACE, TEMP> = <Europe, medium>.

A *contingent proposition*, then, is a proposition that is true under one valuation but false under another valuation. A *valid proposition* (or tautology), denoted as 'T', is a proposition that is true under every valuation, while an *unsatisfiable proposition*, denoted as 'F', is a proposition that is false under every valuation (and a *satisfiable proposition* is true under at least one valuation).

Beliefs. A *belief* is a formula specified in the following format:

$$P_1\,(m_1) \boldsymbol{\vee} P_2\,(m_2) \boldsymbol{\vee} \ldots \boldsymbol{\vee} P_n\,(m_n)$$
$$\text{where}$$
$$\forall\, i\ (1 \le i \le n),\quad P_i \text{ is a satisfiable proposition and } 0 < m_i \le 1,$$
$$\text{and } \sum m_i = 1. \tag{1}$$

As a convention, if 'T' appears in formula (1), it is usually specified at the end (i.e., as P_n). "T (1)" is called *the vacuous belief.*

Intuitively, each m_i (called the "m-value" of P_i) in formula (1) represents the amount of belief we specifically "allocate" to P_i. That is, formula (1) may be interpreted to mean the following: the world, as we understand it, is such that P_1 holds (with the amount of belief m_1 being allocated to it), *or* P_2 holds (with the amount of belief m_2 being allocated to it), ..., *or* P_n holds (with the remaining amount of belief m_n being allocated to it). This is why we use the symbol '$\boldsymbol{\vee}$' to delimit the '$P_i\,(m_i)$' of a belief. However, this symbol '$\boldsymbol{\vee}$' is not to be confused with the usual logical symbol '$\vee$' that occurs within a proposition.

In the following, we use the term 'intuitive belief' (or 'intuitive beliefs') to denote the intuitive belief(s) we have in mind, and we also use the word 'belief' (or 'beliefs') to denote a formula (or formulas).

Belief sets. We may be able to come up with one single belief (i.e., one formula) which appropriately formalizes our intuitive belief in (almost) every way. However, this is not a very easy task in general, and it may be argued whether the specification of such a joint belief is always necessary.

As human problem solvers, we are often capable of identifying various "independent" aspects of a problem. Once such aspects have been identified, we can then specify a (unique) belief for each of these aspects and use some kind of inference mechanism to "combine" the specified (independent) beliefs. This philosophy is embodied in the D-S formalism. Accordingly, the belief calculus works with sets of beliefs (called belief sets)[1]; when using this calculus for reasoning, we first try to infer a singleton set (containing the combined belief) from the given belief set.

A *belief set* $\mathbb{B}$ is a non-empty set of beliefs. In notation,

$$\mathbb{B} = \{\mathbf{b}_1; \mathbf{b}_2; \ldots ; \mathbf{b}_r\}, \text{ where}$$
$$r \ge 1, \text{ and } \forall\, i\ (1 \le i \le r),\ \mathbf{b}_i \text{ is a belief.}$$

(We use ';' to delimit the specified beliefs. This specification of beliefs does not mean that the beliefs are implicitly ordered.)

We now give the inference rules of the belief calculus[2]. The first three rules are trivial. (Notation: Throughout this paper, we use '|- P' to mean that "the proposition P is provable in the propositional calculus from the set of mutual exclusion axioms $\mathbb{ME}$".)

1. Commutation:

$$\frac{\{\ldots \boldsymbol{\vee} P_i\,(m_i) \boldsymbol{\vee} \ldots \boldsymbol{\vee} P_j\,(m_j) \boldsymbol{\vee} \ldots;\ \mathbf{b}_2;\ \mathbf{b}_3;\ \ldots ;\ \mathbf{b}_r\}}{\{\ldots \boldsymbol{\vee} P_j\,(m_j) \boldsymbol{\vee} \ldots \boldsymbol{\vee} P_i\,(m_i) \boldsymbol{\vee} \ldots;\ \mathbf{b}_2;\ \mathbf{b}_3;\ \ldots ;\ \mathbf{b}_r\}}$$

2. Addition:

$$\frac{\begin{array}{l}\vdash (P \leftrightarrow R),\\ \{\ldots \boldsymbol{\vee} P\,(m_i) \boldsymbol{\vee} \ldots \boldsymbol{\vee} R\,(m_j) \boldsymbol{\vee} \ldots;\ \mathbf{b}_2;\ \mathbf{b}_3;\ \ldots ;\ \mathbf{b}_r\}\end{array}}{\{\ldots \boldsymbol{\vee} P\,(m_i + m_j) \boldsymbol{\vee} \ldots \boldsymbol{\vee} \ldots;\ \mathbf{b}_2;\ \mathbf{b}_3;\ \ldots ;\ \mathbf{b}_r\}}$$

3. Substitution:

$$\frac{\vdash (P \leftrightarrow R),\ \{\ldots \boldsymbol{\vee} P\,(m) \boldsymbol{\vee} \ldots ;\ \mathbf{b}_2;\ \mathbf{b}_3;\ \ldots ;\ \mathbf{b}_r\}}{\{\ldots \boldsymbol{\vee} R\,(m) \boldsymbol{\vee} \ldots ;\ \mathbf{b}_2;\ \mathbf{b}_3;\ \ldots ;\ \mathbf{b}_r\}}$$

We also need a fourth rule for inferring combined beliefs (⊕ is defined below).

4. Combination[3]:

$$\frac{\{\mathbf{b}_1;\ \mathbf{b}_2;\ \mathbf{b}_3;\ \ldots ;\ \mathbf{b}_r\},\ \mathbf{b}_1 \oplus \mathbf{b}_2 = \mathbf{b}_c}{\{\mathbf{b}_c;\ \mathbf{b}_3;\ \ldots ;\ \mathbf{b}_r\}}$$

The combination operator ⊕. Let $\mathfrak{B}$ denote the set of all beliefs. The combination operator ⊕ (read as "Dempster's combination") is a partial function that maps from $\mathfrak{B} \times \mathfrak{B}$ to $\mathfrak{B}$.

Intuitively, ⊕ (e.g., $\big(A\,(.8) \boldsymbol{\vee} (\neg A)\,(.2)\big) \oplus \big(A\,(.5) \boldsymbol{\vee} B\,(.3) \boldsymbol{\vee} T\,(.2)\big)$) may be thought of as a two step process.

The first step is to apply the (*independence*) assumption that allocating p_i to P_i in the first belief *and* allocating r_j to R_j in the second belief *should mean* allocating $p_i{*}r_j$ to

[1] with singleton sets as special cases.

[2] The belief calculus is formulated along the lines of natural deduction systems. However, due to space limitations, we only describe the inference mechanism of this system.

[3] This is the only inference rule (in the context of the belief calculus) that can be used to reduce the number of beliefs in a belief set.

$P_i \wedge R_j$ in the combined belief (e.g., $((A \wedge A)\ (.4) \vee (A \wedge B)\ (.24) \vee (A \wedge T)\ (.16) \vee (\neg A \wedge A)\ (.1) \vee (\neg A \wedge B)\ (.06) \vee (\neg A \wedge T)\ (.04))$).

The second step is to apply the (*coherence*) assumption that the two beliefs that are being combined are meant to be coherent; this is done by taking away all "pairs" containing unsatisfiable propositions (e.g., $(\neg A \wedge A)\ (.1)$) and redistributing their m-values (e.g., .1) to the remaining propositions by proportions (e.g., $((A \wedge A)\ (.444) \vee (A \wedge B)\ (.267) \vee (A \wedge T)\ (.178) \vee (\neg A \wedge B)\ (.067) \vee (\neg A \wedge T)\ (.044))$). This step is also known as "renormalization".

Formally, $\oplus$ is defined as follows. (Notation: Let S be an ordered set of "pairs" $(P_1\ (m_1), P_2\ (m_2), \ldots, P_n\ (m_n))$, then by '$\vee$ S', we mean the formula '$P_1\ (m_1) \vee P_2\ (m_2) \vee \ldots \vee P_n\ (m_n)$'.)

$(P_1\ (p_1) \vee P_2\ (p_2) \vee \ldots \vee P_M\ (p_M)) \oplus$
$(R_1\ (r_1) \vee R_2\ (r_2) \vee \ldots \vee R_N\ (r_N))$
$=$

if $\exists\ (i, j) \in \{1, \ldots, M\} \times \{1, \ldots, N\}$
such that $P_i \wedge R_j$ is satisfiable,

then

$\vee\ \{(P_i \wedge R_j)\ (p_i{*}r_j / (1 - \Sigma_{(h,k)\in S}\, p_h{*}r_k))\ |$
S is the maximum subset of
$\{1, \ldots, M\} \times \{1, \ldots, N\}$ such that
$\forall\ (h,k) \in S, \vdash ((P_h \wedge R_k) \leftrightarrow F)$,
and
$(i, j) \in (\{1, \ldots, M\} \times \{1, \ldots, N\}) \setminus S$
$\}$[4],

otherwise undefined.

$\oplus$ is an associative operation (i.e., $(\mathbf{b}_1 \oplus \mathbf{b}_2) \oplus \mathbf{b}_3$, if defined, is the same as $\mathbf{b}_1 \oplus (\mathbf{b}_2 \oplus \mathbf{b}_3)$). If we use $\underline{\oplus}$ to denote "$\oplus$ without renormalization", then $\underline{\oplus}$ is obviously associative; but more importantly, $\mathbf{b}_1 \oplus \mathbf{b}_2 \oplus \ldots \oplus \mathbf{b}_r$ (whatever the order of combinations is), if defined, is the same as $\mathbf{b}_1\ \underline{\oplus}\ \mathbf{b}_2\ \underline{\oplus} \ldots \underline{\oplus}\ \mathbf{b}_r$ followed by one single renormalization.

Also, r beliefs $\mathbf{b}_1, \mathbf{b}_2, \ldots, \mathbf{b}_r$ are said to be *incompatible* whenever $\mathbf{b}_1 \oplus \mathbf{b}_2 \oplus \ldots \oplus \mathbf{b}_r$ (whatever the order of combinations is) is undefined.

Example 2. $(A\ (1)) \oplus ((\neg A)\ (1))$ is undefined.

The calculus. The belief set $\mathbb{B}_2$ is *Dempster-Shafer provable from* the belief set $\mathbb{B}_1$, denoted as '$\mathbb{B}_1 \vdash_{DS} \mathbb{B}_2$', if $\mathbb{B}_2$ can be inferred from $\mathbb{B}_1$ after a finite number of applications of the (four) inference rules.

We are now ready to define the overall belief BEL in an arbitrary proposition. Let $\mathfrak{BS}$ be the set of all belief sets and $\mathfrak{P}$ be the set of all propositions. BEL is a partial function from $\mathfrak{BS} \times \mathfrak{P}$ to [0, 1], defined as follows (let $\mathbb{B}$ be a belief set and R be a proposition.)

$BEL(\mathbb{B}, R) = \Sigma_{\vdash (P_i \rightarrow R)}\, m_i$

where $\mathbb{B} \vdash_{DS} \{P_1\ (m_1) \vee P_2\ (m_2) \vee \ldots \vee P_n\ (m_n)\}$.

/* i.e., we must first deduce a singleton from $\mathbb{B}$ */

Clearly, $BEL(\mathbb{B}, R)$ will be undefined (for every proposition R) whenever the beliefs contained in $\mathbb{B}$ are incompatible.

Note that, even if $BEL(\mathbb{B}, R)$ is defined for some contingent proposition R, it can still be zero. This simply means that we have no idea whether R holds in the world (because we are not aware of anything that logically supports it). But having no (intuitive) belief in R does not necessarily mean that we have any (intuitive) belief in $\neg R$ (i.e., $BEL(\mathbb{B}, \neg R) > 0$), because our degree of belief in a proposition ($\neg R$ in this case), as defined by BEL, is always dependent on whether we are aware of anything that logically supports it and not on whether we are ignorant of anything supporting its negation (i.e., R). This is one of the characteristics of the D-S formalism.

Example 3. $BEL(\{(A \rightarrow B)\ (1)\}, B) = 0$.

Example 4.

$\{A\ (.8) \vee (\neg A)\ (.2);\ A\ (.5) \vee B\ (.3) \vee T\ (.2)\}$

/* i.e., $\mathbb{B}$ */

$\vdash_{DS} \{(A \wedge A)\ (.444) \vee (A \wedge B)\ (.267) \vee (A \wedge T)\ (.178) \vee (\neg A \wedge B)\ (.067) \vee (\neg A \wedge T)\ (.044)\}$

Combination

$\vdash_{DS} \{A\ (.444) \vee (A \wedge B)\ (.267) \vee A\ (.178) \vee (\neg A \wedge B)\ (.067) \vee \neg A\ (.044)\}$

Substitution

$\vdash_{DS} \{A\ (.622) \vee (A \wedge B)\ (.267) \vee (\neg A \wedge B)\ (.067) \vee \neg A\ (.044)\}$

Addition

Therefore

$BEL(\mathbb{B}, A) = .622 + .267 = .889$, and $BEL(\mathbb{B}, \neg A) = .067 + .044 = .111$.

Relating to the "usual" formulation. To see the relation between the belief calculus and the usual set-theoretic formulation of belief functions [Shafer 76], consider the following mapping: let $A_1, A_2, \ldots, A_N$ be all the variables, and let $\Theta_1, \Theta_2, \ldots, \Theta_N$ be their respective frames. Then each proposition P corresponds to exactly one subset of the joint frame $\Theta_1 \times \Theta_2 \times \ldots \times \Theta_N$ (i.e., the set of all "total" valuations that make P true). Similarly, for each subset of the joint frame, there is a corresponding set of logically equivalent propositions. Let S_P be the subset of the joint frame that corresponds to the proposition P and let S_R be the subset of the joint frame that corresponds to the proposition R, then: $\vdash (P \rightarrow R)$ if and only if $S_P \subseteq S_R$; $S_P \cup S_R$ corresponds to $P \vee R$; $S_P \cap S_R$ corresponds to $P \wedge R$; and $\Theta_1 \times \Theta_2 \times \ldots \times \Theta_N \setminus S_P$ corresponds to $\neg P$. This

4We assume these '$P_i \wedge R_j$' are lexicographically ordered in (i, j).

provides a straightforward translation between the language of the belief calculus and the language of the usual set-theoretic formulation of belief functions.

Example 5. We now use the belief calculus notations to describe the idea of combining the ATMS [de Kleer 86a] with belief functions (e.g., [Laskey and Lehner 89]). Let $\mathfrak{J} = \{P_1, P_2, \ldots, P_n\}$ be the set of (boolean) propositional clauses that have been transmitted to the ATMS (premises are specified in $\mathfrak{J}$ as "$\rightarrow$ C" or "$\rightarrow \neg$C")[5]. Let $\{A_1, A_2, \ldots, A_k\}$ be a distinguished set of primitives (i.e., assumptions) such that either A_i or $\neg A_i$ (or both) occurs in $\mathfrak{J}$. Furthermore, let $\mathbb{B}$ contain the following (and only the following) beliefs: (a) for each P_i, "P_i (1)" is in $\mathbb{B}$, (b) for each A_i, "A_i (m) $\vee$ $\neg A_i$ (1-m)" is in $\mathbb{B}$[6]. Then for any literal B (i.e., a primitive Q or its negation $\neg$Q) occurring in $\mathfrak{J}$, we can compute BEL($\mathbb{B}$, B) using the nogoods and the label associated with B as a basis. For more details, see [D'Ambrosio 87, 88; Laskey and Lehner 88, 89; Provan 89a, 89b].

3. Uncertain Reasoning

The specification of independent beliefs. The D-S formalism encourages the use of the following methodology: we first identify the "independent" aspects of the problem at hand; and then we specify a belief for each of the identified aspects. Therefore, the purpose of this section is to show how we can use the belief calculus for uncertain reasoning *once the independent aspects of the problem at hand have been identified.*

Example 6. (adapted from [Kong 86]) Two sites A and B are connected by a one-way valve which, when working, allows water to flow from A to B. The probability that this valve is working (i.e., not blocked) is p_1. Similarly, sites B and C are connected by a one-way valve (with a working probability of p_2). These two valves work independently. We have no information as to whether there is any water going into A or B or C, but we are interested in whether there is any water in each site. Therefore, we formulate the working of the two valves as two independent random variables, and $\mathbb{B}$ = {V1Working (p_1) $\vee$ $\neg$V1Working (1-p_1); (V1Working $\rightarrow$ (WaterA $\rightarrow$ WaterB)) (1); V2Working (p_2) $\vee$ $\neg$V2Working (1-p_2); (V2Working $\rightarrow$ (WaterB $\rightarrow$ WaterC)) (1)}, and the values of BEL($\mathbb{B}$, WaterA), BEL($\mathbb{B}$, WaterB) and BEL($\mathbb{B}$, WaterC) are zero at the moment.

Suppose we just learned that <WaterB> = <Yes>. Then BEL($\mathbb{B}$ $\cup$ {WaterB (1)}, WaterB) = 1, BEL($\mathbb{B}$ $\cup$ {WaterB (1)}, WaterC) = p_2, and BEL($\mathbb{B}$ $\cup$ {WaterB (1)}, WaterA) remains zero.

Belief dependency structures. In general, the uncertainties we want to specify may be intuitively related. When this is the case, we can no longer formulate these uncertainties as independent random variables. Nevertheless, we can try to work out a dependency structure (in a sense similar to the idea of the Bayesian causal trees [Pearl 86]) among the variables, and we make sure that the way a variable (e.g., A) depends on a valuation of other variables (e.g., (B = Yes) $\wedge$ (C = h)) is independent of the ways this same variable (i.e., A) depends on other valuations of these other variables (i.e., (B = Yes) $\wedge$ (C = m); (B = No) $\wedge$ (C = l); etc.). This is the rational behind the following technique which uses a method described in [Smets 78] for specifying a belief set from independent conditional beliefs:

Let $\mathcal{A}_0 = \{A_1, A_2, \ldots, A_N\}$ (e.g., {Bird, Penguin, Fly}) be a set of variables. We first specify a set $\mathfrak{C}$ of categorical beliefs about these variables (e.g., $\mathfrak{C}$ = {(Penguin $\rightarrow$ Bird $\wedge$ $\neg$Fly) (1)}). Then, we recursively apply the following three steps until the variables contained in $\mathcal{A}_i$ ($i \geq 0$) do not directly "depend on" each other.

Step 1: From $\mathcal{A}_i$ (e.g., $\mathcal{A}_0$), we identify exactly one variable[7] A (e.g., Fly) and also a subset $\mathfrak{B}_i$ of $\mathcal{A}_i \backslash \{A\}$ so that A directly "depends on" the valuation of the elements of $\mathfrak{B}_i$ (e.g., $\mathfrak{B}_0$ = {Bird, Penguin} is a subset of {Bird, Penguin, Fly}\{Fly} so that Fly directly depends on the valuation of the two variables Bird and Penguin).

Step 2: For *each and every* logically possible valuation of the elements of $\mathfrak{B}_i$ (e.g., <Bird, Penguin> = <Yes, No>; <Bird, Penguin> = <Yes, Yes>; <Bird, Penguin> = <No, No>), we assess an independent belief about the valuation of A (e.g., we assess the belief "(Fly = Yes) (.9) $\vee$ ((Fly = Yes) $\vee$ (Fly = No)) (.1)" for the valuation <Bird, Penguin> = <Yes, No>). If this assessed belief is non-vacuous and non-categorical, we (need to) translate it into the following belief:[8]

("the valuation" $\rightarrow$ A's value is in ValueSet1) (m_1) $\vee$
("the valuation" $\rightarrow$ A's value is in ValueSet2) (m_2) $\vee$
...
("the valuation" $\rightarrow$ A's value is in ValueSetM) (m_M),
(e.g., "(Bird $\wedge$ $\neg$Penguin $\rightarrow$ (Fly = Yes)) (.9) $\vee$ (Bird $\wedge$ $\neg$Penguin $\rightarrow$ (Fly = Yes) $\vee$ (Fly = No)) (.1)",
or simply, "(Bird $\wedge$ $\neg$Penguin $\rightarrow$ Fly) (.9) $\vee$ T (.1)").

---

[5]The ATMS actually uses a (positive) primitive to represent the negation of another (positive) primitive [de Kleer 86b]. However, for simplicity, we can think of the ATMS as if it accepted a negated primitive directly.

[6]A more general specification will be to specify exactly one of the following for A_i : "A_i (m_1) $\vee$ T (1-m_1)", "$\neg A_i$ (m_1) $\vee$ T (1-m_1)", or "A_i (m_1) $\vee$ $\neg A_i$ (m_2) $\vee$ T (1-m_1-m_2)" (the specification of "T (1-m_1-m_2)" is optional).

[7]Actually, we can identify more than one variable if we want. Here, just for simplicity, we restrict it to be one.

[8]This translation is based on the principle of minimum specificity [Dubois and Prade 86].

Step 3: Let $\mathcal{A}_{i+1}$ be $\mathcal{A}_i \backslash \{A\}$ (e.g., $\mathcal{A}_1$ = {Bird, Penguin, Fly}\{Fly} = {Bird, Penguin}).

Once we get to $\mathcal{A}_{final}$, we can, if we want, specify a (non-vacuous) belief for each of the variables remaining in $\mathcal{A}_{final}$ (e.g., $\mathcal{A}_{final} = \mathcal{A}_2$ = {Bird}, and we have the option of specifying a belief such as "Bird (.7) ∨ T (.3)" for the variable Bird; however, we prefer to have "T (1)" for Bird in this case).

Example 8. The belief set

$\mathbb{B}$ = {(Penguin → Bird ∧ ¬Fly) (1);
(Bird ∧ ¬Penguin → Fly) (.9) ∨ T (.1);
(Bird → ¬Penguin) (.95) ∨ T (.05);}

is obtained from the following "constraints":

valuation	belief about some variable
Bird ∧ ¬Penguin	Fly (.9) ∨ T (.1)
Bird ∧ Penguin	¬Fly (1)
¬Bird ∧ ¬Penguin	T (1)
¬Bird ∧ Penguin	(logically impossible)
Bird	¬Penguin (.95) ∨ T (.05)
¬Bird	¬Penguin (1)

Therefore, BEL($\mathbb{B}$ ∪ {Bird (1)}, ¬Penguin) = .95, BEL($\mathbb{B}$ ∪ {Bird ∧ ¬Penguin (1)}, Fly) = .9, BEL($\mathbb{B}$ ∪ {Bird ∧ ¬Penguin (1)}, ¬Fly) = 0, etc.

In addition, BEL($\mathbb{B}$, ¬Penguin) = .95, BEL($\mathbb{B}$, Bird) = 0, BEL($\mathbb{B}$ ∪ {Bird (1)}, Fly) = .855, BEL($\mathbb{B}$ ∪ {Bird (1)}, ¬Fly) = 0, BEL($\mathbb{B}$ ∪ {Fly (1)}, ¬Penguin) = 1.0, BEL($\mathbb{B}$ ∪ {Fly (1)}, Bird) = 0, BEL($\mathbb{B}$ ∪ {Fly (1)}, ¬Bird) = 0, etc.

"Distinct" sources of evidence. We sometimes encounter the following situation[9]: (1) there are one or more sources that provide us with information, and each source has full confidence in the information it provides; (2) the information provided by each source directly "indicts" some elements of the frame of the "main variable" (i.e., the one we are interested in); (3) we can make a reliability estimation for each of these sources[10]; (4) the reliabilities of the sources are independent.

When we are in this kind of situation, we can put our evidence about each particular source into a unique group, and we specify a belief dependency structure according to each group of (related) evidence. The resulting belief set, then, consists of several belief dependency structures intersecting on the main variable.

Example 7. Our friend is ill, and doctors can not pinpoint the problem. Since it may involve life and death, we bring our friend to two doctors B and C that are famous in this area. It is reasonable to assume that these two doctors are independent in making their diagnoses (because they received their trainings in different medical doctrines, they live in different cities, they do not confer to each other, etc.). We also did some background study about the two doctors. Therefore we know that B is extremely busy, B has more authority in this area than C does, and C has a reputation of always doing his best for his patients. Also, our actual experience with the two doctors seems to confirm this background information. We are interested in the reliabilities of the two doctors. We are also concerned that B's being busy may mean that B does not spend enough time examining our friend's case. Therefore we formulate our knowledge about the two doctors as two belief sets $\mathbb{B}_1$ and $\mathbb{B}_2$, with each $\mathbb{B}_i$ containing our (intuitive) beliefs about a doctor:

$\mathbb{B}_1$ = {(AuthorityB ∧ BusyB) (1);
(BusyB → LessCaseStudyB) (.8) ∨
(BusyB → ¬LessCaseStudyB) (.2);
AuthorityB ∧ LessCaseStudyB → ReliableB) (.7) ∨
AuthorityB ∧ LessCaseStudyB → ¬ReliableB) (.3);
(AuthorityB ∧ ¬LessCaseStudyB → ReliableB) (.95) ∨
(AuthorityB ∧ ¬LessCaseStudyB → ¬ReliableB) (.05)}

$\mathbb{B}_2$ = {(SemiAuthorityC ∧ ReputationC) (1);
(SemiAuthorityC ∧ ReputationC → ReliableC) (.8) ∨
(SemiAuthorityC ∧ ReputationC → ¬ReliableC) (.2)}

After diagnosis, doctor B determines with full confidence that the patient has either illness X or illness Y. Also after diagnosis, doctor C determines with full confidence that the patient has either illness Y or illness Z. It is a medical fact that a person can not have any two of these three illnesses at the same time. Therefore we let the frame of (the main variable) Illness to be {X, Y, Z, OTHER}, and we specify three more categorical beliefs:

$\mathbb{B}_3$ = {(BSaysXY ∧ CSaysYZ) (1);
(BSaysXY ∧ ReliableB → (Illness = Y) ∨ (Illness = Z)) (1);
(CSaysYZ ∧ ReliableC → (Illness = Y) ∨ (Illness = Z)) (1)}.

With $\mathbb{B} = \mathbb{B}_1 \cup \mathbb{B}_2 \cup \mathbb{B}_3$, we get: BEL($\mathbb{B}$, (Illness = X)) = 0; BEL($\mathbb{B}$, (Illness = Z)) = 0; BEL($\mathbb{B}$, (Illness = Y)) = .6; BEL($\mathbb{B}$, (Illness = X) ∨ (illness = Y)) = .75; BEL($\mathbb{B}$, (Illness = Y) ∨ (Illness = Z)) = .8; BEL($\mathbb{B}$, (Illness = X) ∨ (Illness = Y) ∨ (Illness = Z)) = .95, etc.

4. Discussion

Appropriateness of the notation. The belief calculus serves as a (notational) alternative for the the use of the D-S formalism for uncertain reasoning. As such, the appropriateness of the belief calculus (as a notation) will have to depend on the application domain, and there may well be situations in which set-theoretic notations are more appropriate.

---

[9] The author thanks Nic Wilson and Philippe Smets for arriving at this characterization of distinctness.

[10] A source is reliable (with respect to the information it provides) if the information it provides is indeed true.

However, it may be worthwhile to point out a superficial but nevertheless important "difference" between the (uses of the) two notations: with the belief calculus, we can sometimes explicitly specify what our evidence is and how this evidence induces beliefs; whereas with the usual set-theoretic notations, the evidence is generally regarded as "outside of" our specifications of beliefs.

Tractability. The computational complexity of the belief calculus is exponential with respect to the number of variables in a belief set, and there are ways for improving the speed of this computation (e.g., [Kennes and Smets 90; Shafer et al. 87; Wilson 89]).

We might also look at this complexity problem from a different perspective: if we treat propositional provability (which is well known for its NP-completeness!) as the basic operator, then the complexity of the belief calculus is exponential with respect to the number of beliefs in a belief set. Thus, if we have many variables but only a few beliefs in a belief set, then a deduction-based approach such as ATMS + D-S (see [Provan 89b] for a complexity analysis of ATMS + D-S) may turn out to be a more attractive way for computing BEL.

5. Conclusion

We formulated the D-S formalism along the lines of natural deduction systems. This formulation (called the belief calculus) allows us to infer beliefs from beliefs without ever appealing to the use of set-theoretic notations.

To show how the belief calculus may be used for uncertain reasoning, we gave three examples. These examples suggested different ways for modelling real world situations.

Acknowledgements. The author thanks Mike Clarke, Robert Kennes, Alessandro Saffiotti, Philippe Smets, and Nic Wilson for stimulating discussions, and two anonymous referees for very helpful comments.

References

D'Ambrosio, B. (1987). Truth maintenance with numeric certainty estimates. *Proceedings of the Third IEEE Conference on AI Applications*, Orlando, Florida, 244-249.

D'Ambrosio, B. (1988). A hybrid approach to reasoning under uncertainty. *International Journal of Approximate Reasoning*, **2**, 1, 29-45.

de Kleer, J. (1986a). An assumption-based truth maintenance system. *Artificial Intelligence*, **28**, 2, 127-162.

de Kleer, J. (1986b). Extending the ATMS. *Artificial Intelligence*, **28**, 2, 163-196.

Dubois, D. and Prade, H. (1986). The principle of minimum specificity as a basis for evidential reasoning. In *Uncertainty in Knowledge-Based Systems* (Bouchon and Yager eds.), Springer-Verlag, Berlin, 75-84.

Fagin, R. and Halpern, J. Y. (1989). Uncertainty, belief, and probability. *Proceedings of the Eleventh International Joint Conference on Artificial Intelligence*, Detroit, Michigan, Aug. 20-25, 1161-1167.

Gordon, J. and Shortliffe, E. H. (1984). The Dempster-Shafer theory of evidence. In *Rule-Based Expert Systems: The MYCIN Experiments of the Stanford Heuristic Programming Project* (Buchanan and Shortliffe eds.), Addison-Wesley, Reading, Mass., 272-292.

Kennes, R. and Smets, P. (1990). Fast algorithms for Dempster-Shafer theory. (submitted for publication).

Kong, A. (1986). Multivariate belief functions and graphical models. Doctoral dissertation, Department of Statistics, Harvard University.

Laskey, K. B. and Lehner, P. E. (1988). Belief maintenance: An integrated approach to uncertainty management. *Proceedings of the Seventh AAAI Conference*, Saint Paul, Minnesota, Aug. 21-26, 210-214.

Laskey, K. B. and Lehner, P. E. (1989). Assumptions, beliefs and probabilities. *Artificial Intelligence*, **41**, 1, 65-77.

Pearl, J. (1986). Fusion, propagation, and structuring in belief networks. *Artificial Intelligence*, **29**, 241-288.

Provan, G. (1989a). An analysis of ATMS-based techniques for computing Dempster-Shafer belief functions. *Proceedings of the Eleventh International Joint Conference on Artificial Intelligence*, Detroit, Michigan, Aug. 20-25, 1115-1120.

Provan, G. (1989b). A logic-based analysis of Dempster Shafer theory. Technical Report 89-08, Computer Science Dept., University of British Columbia.

Ruspini, E. H. (1987). Epistemic logics, probability, and the calculus of evidence. *Proceedings of the Tenth International Joint Conference on Artificial Intelligence*.

Shafer, G. (1976). *A Mathematical Theory of Evidence*. Princeton University Press.

Shafer, G., Shenoy, P. P. and Mellouli, K. (1987). Propagating belief functions in qualitative Markov trees. *International Journal of Approximate Reasoning*, **1**, 349-400.

Smets, P. (1978). Un modèle mathématico-statistique simulant le processus du diagnostic médical. Doctoral dissertation, Université Libre de Bruxelles, Bruxelles.

Smets, P. (1988). Belief functions. In *Non-Standard Logics for Automated Reasoning* (P. Smets, E. H. Mamdani, D. Dubois and H. Prade eds.). Academic Press, London.

Wilson, N. (1989). Justification, computational efficiency and generalisation of the Dempster-Shafer theory.

Yen, J. (1989). Gertis: A Dempster-Shafer approach to diagnosing hierarchical hypotheses. *Communications of the ACM*, **32**, 5, 573-585.

Zarley, D., Hsia, Y.-T., and Shafer, G. (1988). Evidential reasoning using DELIEF. *Proceedings of the Seventh AAAI Conference*, St. Paul, Minnesota, **1**, 205-209.