From: AAAI-92 Proceedings. Copyright ©1992, AAAI (www.aaai.org). All rights reserved.

A Belief-Function Logic

Alessandro Saffiotti*

IRIDIA - Université Libre de Bruxelles
50 av. F. Roosevelt - CP 194/6 - 1050 Bruxelles - Belgium
asaffio@ulb.ac.be

Abstract

We present *BFL*, a hybrid logic for representing uncertain knowledge. *BFL* attaches a quantified notion of belief — based on Dempster-Shafer's theory of belief functions — to classical first-order logic. The language of *BFL* is composed of objects of the form F:[a,b], where F is a first-order sentence, and a and b are numbers in the [0,1] interval (with $a \leq b$). Intuitively, a measures the strength of our belief in the truth of F, and (1–b) that in its falseness. A number of properties of first-order logic nicely generalize to *BFL*; in return, *BFL* gives us a new perspective on some important points of Dempster-Shafer theory (e.g., the role of Dempster's combination rule).

Introduction

Logic plays a central role in the task of representing knowledge in artificial intelligence. Much of logical tradition is concerned with two-valued logics, i.e. logics in which we can only talk about propositions being completely true or completely false (possibly, according to a given believer). This contrasts with the widely recognized fact that real world knowledge is almost invariably affected by uncertainty, and judgments about the truth of propositions are rarely categorical. Techniques for handling uncertainty have long been studied within artificial intelligence, and a number of formalisms developed — ranging from those based on probability theory (e.g. Pearl, 1988), to possibility theory (Zadeh, 1978; Dubois & Prade, 1988), to Dempster-Shafer's (D-S) theory of belief functions (Dempster, 1967, Shafer, 1976, Smets, 1988). However, these formalisms are normally grounded on some mathematical, rather than logical, model. Recently, a few attempts at merging these formalisms with the logical tradition have been proposed in the AI literature (e.g., Nilsson, 1986; Ruspini, 1986; Bacchus, 1988; Dubois, Lang & Prade, 1989; Fagin & Halpern, 1989; Provan, 1990). While most of these approaches are based on the idea of defining some uncertainty measure over a set of possible worlds, the target is often different. Nilsson aims at (probabilistically) expressing uncertainty about truth of sentences: hence, he extends logic to have probability values as truth values (i.e., probabilities appear at the meta-level). From a different position, Bacchus focuses on the representation of probabilistic (but known with certainty) knowledge; accordingly, he puts probability values, and statements about them, inside the language of his logic (i.e., at the object level). Both Ruspini and Fagin & Halpern are more interested in investigating the foundations of uncertain reasoning: they conduct insightful possible-world analyses that, though grounded in probability theory, reach and scrutinize the theory of belief functions, and its representation and inference mechanisms. From a seemingly similar position, Provan analyzes D-S theory following a proof-theoretic approach.

In this paper, we take yet another position. We are interested in building a logic where partial belief can be represented and reasoned with. We follow what could be a usual schema for defining a first-order logic — going from language to semantics and to deduction procedures — but add a quantified notion of belief at each stage. The outcome is a "belief-function logic" (*BFL*, for short). *BFL* is similar in spirit to the logic proposed by Dubois, Lang and Prade (1989; also, Lang, 1991), but is based on belief functions rather than on possibility measures. The language of *BFL* is composed of objects of the form F:[a,b], where F is a first-order sentence, and a and b are numbers in the [0,1] interval (with $a \leq b$). Intuitively, a measures the strength of our belief in the truth of F, and (1-b) that in its falseness. *BFL* is aimed at modelling partial and incomplete belief: belief is "partial" in that we can partly believe in the truth of a proposition (e.g., a can be strictly smaller than 1); it is "incomplete" in that we can remain completely non-committal about the truth status of some propositions (i.e., both a=0 and b=1). We give semantics to *BFL* in a way that makes it a "coherent" hybrid of first-order logic and standard D-S theory. Many formal properties of first-order logic generalize (in a "graded" form) to *BFL* — including properties of (partial) inconsistency. Also, *BFL* gives us a new perspective on some important points of D-S theory (e.g., the role of Dempster's combination rule). More concretely, *BFL* is a hybrid knowledge representation language that combines the power of first-order logic for representing knowledge with that of D-S theory for representing uncertainty about this knowledge. The desirability of such a tool has been advocated in (Saffiotti, 1990). Moreover, and differently from most of the above logics, *BFL* is equipped with a (non-standard) deduction procedure. Finally, though *BFL* is based on first-order logic and belief functions, it can be extended to other languages and/or ucertainty formalisms.

In the rest of this paper, we describe the syntax and semantics of *BFL*, and discuss some of its properties. We also analyze a particular class of models for *BFL*, called *D-models*, based on Dempster's combination rule, and outline a deduction procedure for *BFL*. A full treatment of *BFL*, and the proofs of the theorems, can be found in (Saffiotti, 1991a).

* Currently at SRI International, AI Center, 333 Ravenswood Ave, Menlo Park CA 94025. E-mail: saffiotti@ai.sri.com.

Language

We consider a standard first-order (f.o.) language, with the usual operators ~, $\wedge$ and $\forall$, plus the abbreviations $F \vee G$ for $\sim(\sim F \wedge \sim G)$; $F \supset G$ for $\sim F \vee G$; and $\exists$x.P(x) for ~$\forall$x.~P(x). We define our language as:

$\mathscr{L} = \{F{:}[a,b] \mid F$ *first-order sentence, and* $0 \le a \le b \le 1\}$.

We call a formula of $\mathscr{L}$ a *bf-formula* (for "belief-function formula"). A bf-formula $F{:}[a,b]$ may be thought of as expressing how a unitary amount of belief is distributed among the propositions "F is true" (a), and "F is false" ($1\text{-}b$). Roughly, a and b play the role of the *Bel* and *Pl* measures in Dempster-Shafer theory. The quantity $b\text{-}a$ may be read as the amount of belief we leave uncommitted (or our "ignorance" about the truth of F). In particular, $F{:}[1,1]$ represents complete confidence in F's being true; $F{:}[0,0]$ represents complete confidence in its being false; and $F{:}[0,1]$ represents complete ignorance about its truth state. Notice that we impose (internal) consistency of single items of knowledge by requiring that $a \le b$.

Example 1. "Italian(alex):[0.6, 0.8]" is a bf-formula, with intended meaning "We believe to the extent 0.6 that Alex is Italian; also, we believe to the extent 0.2 that he is not". "~Italian(alex):[0.2, 0.4]" represents the same information.

Example 2. "$\forall$x.drinker(x)$\supset$smoker(x):[0.7,1]" is a bf-formula, with intended meaning: "we believe with strength 0.7 that all drinkers are also smokers".

It is important to notice that we are concerned here with what we call "epistemic" uncertainty: what the 0.7 above measures is our *partial belief about the (complete) truth* of the given formula. Other readings are in principle possible: e.g., "we believe that the fact that all drinker are also smoker is partially (0.7) true" (*compete belief about partial truth*); or "we believe that 70% of drinkers are also smokers" (*complete belief about a statistical fact*). *BFL* is meant to model the epistemic uncertainty reading.

Example 3. "$\forall$x. ($\exists$y. alarm-report(y,x) $\wedge$ reliable(y)) $\supset$ alarm(x) : [0.95, 1]" is a bf-formula with intended meaning "We strongly believe (0.95) that, whoever is X, if some reliable person reports that X's alarm is ringing, then X's alarm is indeed ringing".

Notice that the item of knowledge in the last example (adapted from Pearl, 1988) could not be expressed in a standard Dempster-Shafer formulation, if not by enumerating all the possible Xs who have an alarm, and, for each one of them, all the possible Ys who could report about her alarm ringing. (similarly, Pearl must redefine his network every time a new element is added to his burglary example).

Given a set Φ of bf-formulae, we let $\widehat{\Phi} = \{F \mid F{:}[a,b] \in \Phi\}$ be the set of f.o. sentences obtained by dropping the uncertainty components from Φ, and we call it the "first-order component" of Φ. In particular, $\widehat{\mathscr{L}}$ denotes the f.o. language from which $\mathscr{L}$ has been built. We use $F, F_1, F_2, G,\ldots$ as metalinguistic variables ranging over f.o. formulae; ψ, ψ_1, ψ_2, φ,... for bf-formulae; Φ, Ψ,... for sets of bf-formulae.

Semantics

We give our language a semantics in the model-theoretic style. As it can be expected, this semantics makes use of concepts borrowed from both logical tradition, and Dempster-Shafer theory. A similar construction, however, could be used to generate logics for partial belief based on different languages and/or different uncertainty formalisms.

The Interpretation Structures

To start with, we need to find a suitable class of mathematical structures to act as models of our logic. Given our language $\mathscr{L}$, we focus on its first-order component $\widehat{\mathscr{L}}$, and the set $\mathscr{I}$ of standard f.o. interpretations for it;[1] we denote by $\models_{\text{fol}}$ the standard f.o. truth relation. Each element of $\mathscr{I}$ can be thought of as encoding a complete description of the state of the world. Given a f.o. formula F, $[\![F]\!]$ denotes the set of interpretations where F is true: $[\![F]\!] = \{I \in \mathscr{I} \mid I \models_{\text{fol}} F\}$.

In order to model incompleteness of belief, we consider (non empty) sets of interpretations, or "hyper-interpretations". We can think of an hyper-interpretation $s \subseteq \mathscr{I}$ as saying that the real "state of affairs" is one of those in s (but we do not know which one). The entailment relation can be extended to work on hyper-interpretations by:

$$s \models_{\text{fol}} F \text{ iff for each } I \in s,\ I \models_{\text{fol}} F$$

We draw a possible scenario in Fig.1: there, $s \models_{\text{fol}} F$, $s \not\models_{\text{fol}} \sim F$; $s' \not\models_{\text{fol}} F$, $s' \not\models_{\text{fol}} \sim F$; $s'' \not\models_{\text{fol}} F$, $s'' \models_{\text{fol}} \sim F$.

In order to enter partiality of belief (or "uncertainty") into the picture, we consider functions Cr from $\wp(\mathscr{I})$ (the power set of $\mathscr{I}$) to the unit interval [0,1]. Given a subset s of $\mathscr{I}$, we read $Cr(s)$ as the extent to which we believe that the real state of the world is one of the elements in s. Correspondingly, we read $Cr([\![F]\!])$ as the extent to which we believe that the real state of the world is one where F is true (i.e., our *confidence* in F). We may legitimately ask which class — if any — of the ($\wp(\mathscr{I}) \to [0,1]$) functions constitutes a reasonable representation of (partial and incomplete) belief. The answer, of course, depends from our notion of what a "reasonable" representation of belief is. Though we do not mean to enter here the philosophical debate on this

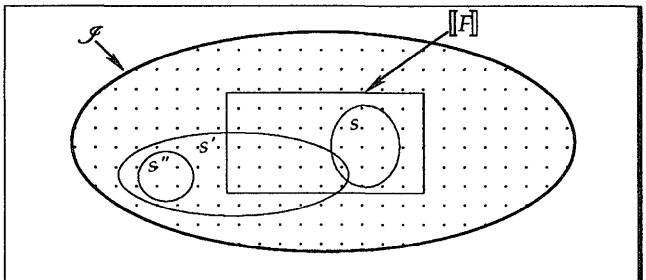


Figure 1. $\mathscr{I}$, $[\![F]\!]$, *and three hyper-interpretations*

[1] To avoid unnecessary complications, we pass over the issue of the cardinality of $\mathscr{I}$. However, we do assume to have countable domains.

issue, we suggest three possible requirements R1–3 for a measure of belief Cr that agrees with fol entailment.

R1. (*monotonicity*) $F \models_{fol} G$ implies $Cr([\![F]\!]) \leq Cr([\![G]\!])$.

R2. (*deductive closure*)

If $Cr([\![F]\!]) > 0$ and $Cr([\![F \supset G]\!]) > 0$, then $Cr([\![G]\!]) > 0$.

R3. (*consistency*) $Cr([\![F]\!]) + Cr([\![\sim F]\!]) \leq 1$.

We use belief functions as Cr functions: given any function f: $\wp(\mathcal{I}) \to [0,1]$, and any $x \subseteq \mathcal{I}$, we let $Bel_f(x) = \Sigma_{y \subseteq x} f(y)$. Because $x \subseteq y$ implies $Bel_f(x) \leq Bel_f(y)$, the Bel_f functions satisfy R1. However, they do not satisfy R2 in general: $Bel_f([\![F]\!]) = a$ and $Bel_f([\![F \supset G]\!]) = b$ only imply $Bel_f([\![G]\!]) \geq \max\{0, a+b-1\}$. We might drop R2 on the ground that it is too strong a requirement, and model a weaker notion of belief where we may have, e.g., P:[.4,1] and Q:[.3,1], and yet P∧Q:[0,1].[2] However, we opt here for a stronger notion, where R2 always holds. In particular, we will require that $Bel_f(x \cap y) \geq Bel_f(x) \cdot Bel_f(y)$.[3] Finally, Bel_f satisfy R3 when $f(\emptyset) = 0$: we do not enforce this condition, as we want to deal with partial inconsistency. We are now ready to define the interpretation structures for *BFL*.

Definition. *A* *bf-interpretation for $\mathcal{L}$ is a function* $\mathcal{M}$: $\wp(\mathcal{I}) \to [0,1]$ *such that:*

(i) $Bel_{\mathcal{M}}(\mathcal{I}) = 1$

(ii) *If* $x \cup y \neq \mathcal{I}$, *then* $Bel_{\mathcal{M}}(x \cap y) \geq Bel_{\mathcal{M}}(x) \cdot Bel_{\mathcal{M}}(y)$.

A bf-interpretation $\mathcal{M}$ *is normal iff* $\mathcal{M}(\emptyset) = 0$.

Normal bf-interpretations correspond to a particular class of basic probability assignments in usual D-S theory: those satisfying condition (ii) above. Given a bf-interpretation $\mathcal{M}$, $Bel_{\mathcal{M}}([\![F]\!])$ measures the total amount of belief committed by $\mathcal{M}$ to the proposition *"F is true"*.

Entailment

We define satisfaction, validity and entailment for *BFL*.

Definition. *Let* $\mathcal{M}$ *be a bf-interpretation. Then*

(i) $\mathcal{M}$ *is a bf-model of* F:[a,b] (*written* $\mathcal{M} \models F$:[a,b]) *iff both* $Bel_{\mathcal{M}}([\![F]\!]) \geq a$ *and* $Bel_{\mathcal{M}}([\![\sim F]\!]) \geq 1-b$.

(ii) $\mathcal{M}$ *is a bf-model of* Φ ($\mathcal{M} \models \Phi$) *iff* $\mathcal{M} \models \varphi$ *for all* $\varphi \in \Phi$.

Example 4..Referring to Fig. 1, let $\mathcal{M}$ be s.t. $\mathcal{M}(s) = 0.6$, $\mathcal{M}(s') = 0.2$, $\mathcal{M}(s'') = 0.1$, and $\mathcal{M}(\mathcal{I}) = 0.1$. It is easy to check that $\mathcal{M}$ is a bf-interpretation, and that: $Bel_{\mathcal{M}}([\![F]\!]) = 0.6$ and $Bel_{\mathcal{M}}([\![\sim F]\!]) = 0.1$; i.e., $\mathcal{M} \models F$:[.6,.9].

Example 5. For any subset s of $\mathcal{I}$, the bf-interpretation $\mathbb{1}_s$ given by: $\mathbb{1}_s(s)=1$ and $\mathbb{1}_s(x)=0$ otherwise, encodes a categorical but incomplete state of belief. In particular, $\mathbb{1}_{\mathcal{I}}$ encodes the empty state of belief, and $\mathbb{1}_{\emptyset}$ the completely inconsistent one. If I is a fol interpretation, $\mathbb{1}_{\{I\}}$ encodes a categorical and complete state of belief.

We say that $\mathcal{M}$ *satisfies* ψ if $\mathcal{M} \models \psi$. Notice that $\mathcal{M} \models F$:[a,b] implies $\mathcal{M} \models F$:[c,d] for any interval [c,d] that contains [a,b]. In example 3, $\mathcal{M} \models F$:[a,b] for any [a,b]⊇[.6,.9]. A bf-formula ψ is said *bf-valid* (written $\models \psi$) iff $\mathcal{M} \models \psi$ for all $\mathcal{M}$. It is easy to see that all the bf-valid formulae take one of the forms: F:[0,1], for any F; F:[a,1], with F (fol) valid; or F:[0,b], with F unsatisfiable.

Definition. Φ *bf-entails* Ψ (*written* $\Phi \models \Psi$) *iff every bf-model of* Φ *is a bf-model of* Ψ.

Example 6. Let Φ = {∀x.drinker(x)⊃smoker(x) : [.7,1], drinker(peter) : [.8,.9]}; we want to know whether or not $\Phi \models$ smoker(peter):[.5,1]. In Fig. 2, we draw the set $\mathcal{I}$ partitioned among the interpretations where "drinker(peter)" holds (upper part) and those where "smoker(peter)" holds (right part): [\![drinker(peter)]\!] = $A \cup B$, and [\![smoker(peter)]\!] = $B \cup D$. As ∀x.drinker(x)⊃ smoker(x) $\models_{fol}$ drinker(peter)⊃ smoker(peter), [\![∀x.drinker(x)⊃smoker(x)]\!] is a subset of [\![drinker(peter)⊃smoker(peter)]\!] = $B \cup C \cup D$. Consider now any bf-model $\mathcal{M}$ of Φ. By definition of bf-model, $\mathcal{M}$ must be such that $Bel_{\mathcal{M}}(A \cup B) \geq 0.8$, $Bel_{\mathcal{M}}(C \cup D) \geq 0.1$, and $Bel_{\mathcal{M}}(B \cup C \cup D) \geq 0.7$. Moreover, by definition of bf-interpretations, we must also have $Bel_{\mathcal{M}}(B) \geq 0.8 \cdot 0.7 = 0.56$ (and $Bel_{\mathcal{M}}(C \cup D) \geq 0.07$). Hence, $Bel_{\mathcal{M}}$([\![smoker(peter)]\!])≥ 0.56. As this is true for any $\mathcal{M}$, $\Phi \models$ smoker(peter):[.5, 1].

We list some interesting properties of bf-entailment.

Theorem 1. *Let* Φ *be a set of bf-formulae*, F, G *fol formulae*, **t** *a ground term in* $\mathcal{L}$, *and a, b, c* ∈ [0,1].

(a) $\Phi \models F$:[a,b] *iff* $\Phi \models \sim F$:[1-b, 1-a]

(b) $\Phi \models F$:[a,b] *iff both* $\Phi \models F$:[a,1] *and* $\Phi \models F$:[0,b]

(c) $F \models_{fol} G$ *if and only if, for all a*, F:[a,1]$\models G$:[a,1]

(d) $\models_{fol} F$ *if and only if* $\models F$:[1,1]

(e) *If* $\Phi \models F \supset G$:[a,1] *and* $\Phi \models F$:[c,1] *then* $\Phi \models G$:[ac,1]

(f) *If* $\Phi \models F \supset G$:[a,1] *and* $\Phi \models G$:[0,b] *then* $\Phi \models$ F:[0,ab]

(g) *If* $\Phi \models$ ∀x.F(x):[a,1] *then* $\Phi \models F$(**t**):[a,1]

(h) *If* $\Phi \models F$(**t**):[0,b] *then* $\Phi \models$ ∀x.F(x):[0,b].

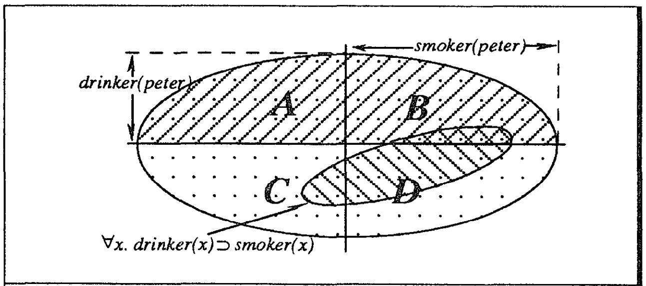


Figure 2. A partition of $\mathcal{I}$ *for example 6.*

---

[2] A possible reading of F:[a,b] under this notion would be "F is true at least 100a% of times, and false at least 100(1-b)%".

[3] This choice is clearly instrumental to have *BFL* behave according to Dempster-Shafer theory. Still other notions of belief can be captured by imposing different constraints. E.g., requiring $Bel_f(x \cap y) \geq min(Bel_f(x), Bel_f(y))$ (and replacing Σ by *sup* in the definition of Bel_f) would force belief to obey possibility theory. The situation is reminiscent of the one in Kripke-style semantics, where properties of the modal operators correspond to constraints on the accessibility relation in the models.

Point (a) shows that *BFL* treats negation according to D-S theory; (b) allows us to consider the *a* and *b* values separately wrt bf-entailment. (c) and (d) show that *BFL* is a conservative extension of standard fol. Also, they show that agents modelled by *BFL* are (partially) logically omniscient: they completely believe all the fol tautologies, and whenever a formula is believed to some extent, all its logical consequences must be believed to at least the same extent. (e) and (f) show that a graded *modus ponens* (and *modus tollens*) can be soundly used for performing deductions in *BFL*. (g) and (h) establish the relation between universal properties and single instances, and confirm that *BFL* models epistemic uncertainty. Believing, to some extent, a universal $\forall x.F(x)$ means to partially believe that F is true for *each* individual; dually, the existence of one single counter-example $\sim F(\mathbf{t})$ suffices for partially negating the validity of that universal.[4]

Example 7. Let $\Phi = \{\ \forall x.\mathrm{drinker}(x) \supset \mathrm{smoker}(x){:}[0.7, 0.9], \mathrm{drinker(peter)}{:}[0.8, 1], \mathrm{drinker(mary)}{:}[0.6, 0.7]\ \}$. The reader can use theorem 1 to check that the following are true:

$\Phi \models \sim\mathrm{smoker(peter)} : [0, 0.44]$

$\Phi \models \mathrm{smoker(peter)} \vee \sim\mathrm{smoker(peter)} : [1, 1]$

$\Phi \models \mathrm{smoker(mary)} : [0.42, 1]$

$\Phi \models \exists x.\ \mathrm{smoker}(x) : [0.56, 1]$

$\Phi \models \exists x.\ \sim\mathrm{smoker}(x) : [0.1, 1]$

(Hint: use (e) and (a) for the first bf-formula; the last bf-formula is entailed by the (negation of the) universal in Φ).

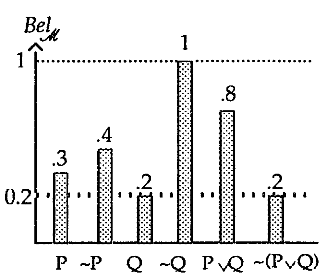
Consistency and Partial Inconsistency

Unlike most approaches to belief functions, we allow a basic probability assignment to assign a non-zero value to the empty set, and have $Bel_{\mathcal{M}}$ "count" this value. Thanks to these peculiarities, *BFL* preserves (in a "graded" form) many of the properties that accompany inconsistency in standard fol.

Definition. *Φ is <u>bf-consistent</u> if Φ has a normal bf-model. It is <u>α-consistent</u> ($0<\alpha\leq 1$), if it has a bf-model $\mathcal{M}$ such that $\mathcal{M}(\emptyset) < \alpha$.*

We say that Φ is *α-inconsistent* if Φ is not α-consistent, and that Φ is *bf-inconsistent* if it is 1-inconsistent.

Example 8. The set $\Phi = \{P\vee Q{:}[.8,1], P{:}[.3,.6], \sim Q{:}[1,1]\}$ is 0.2-inconsistent. In fact, any bf-model $\mathcal{M}$ of Φ must be such that $Bel_{\mathcal{M}}([\![P\vee Q]\!])\geq 0.8$, $Bel_{\mathcal{M}}([\![P]\!])\geq 0.3$, $Bel_{\mathcal{M}}([\![\sim P]\!])\geq 0.4$, and $Bel_{\mathcal{M}}([\![\sim Q]\!])\geq 1$. The last two conditions imply, by definition of bf-interpretation, $Bel_{\mathcal{M}}([\![\sim P \wedge \sim Q]\!]) \geq 0.4$. As $[\![P\vee Q]\!]$ and $[\![\sim P\wedge\sim Q]\!]$ are disjoint, the only way $\mathcal{M}$ can satisfy all the above constraints is by having $\mathcal{M}(\emptyset) \geq 0.2$ (remind that $Bel_{\mathcal{M}}(\mathcal{S})=1$). The lower bounds for the $Bel_{\mathcal{M}}$ values of for any bf-model of Φ are shown on the bottom left corner. Notice that, for any F, $Bel_{\mathcal{M}}([\![F]\!]) \geq 0.2$, and hence $\Phi\models F{:}[.2,1]$ (in particular, $\Phi\models Q{:}[.2,1]$).



α-inconsistency can be thought of as a "noise" that covers all our beliefs below the threshold α: if Φ is α-inconsistent, then $\Phi\models F{:}[\alpha,1]$ (and $\Phi\models F{:}[0,1-\alpha]$) for any F (graded *ex falso quodlibet*). *BFL* can live with this "noise", and still express meaningful information, without degenerating into "logical chaos" (Rescher & Brandom, 1979; Lang, 1991):

Theorem 2. *Let Φ be α-consistent. Then there is a fol formula G such that $\Phi \not\models G{:}[\alpha,1]$.*

Finally, α-inconsistency can be used in proofs by refutation (recall that, in fol, $\Gamma\models_{\mathrm{fol}} F$ iff $\Gamma\cup\{\sim F\}$ is inconsistent):

Theorem 3. *For any Φ, F and a, $\Phi\models F{:}[a,1]$ if and only if $\Phi\cup\{\sim F{:}[1,1]\}$ is a-inconsistent.*

D-models and D-consistency

Bf-formulae represent uncertain information; bf-interpretations encode states of partial belief. We may wonder whether there is, for any given set Φ of bf-formulae, a bf-model of Φ that encompasses all and only the information encoded by Φ. I.e., we may want to seek a least informative bf-model of Φ. We show that, when Φ satisfies a condition called D-consistency, such a bf-model — called a D-model — exists, and can be constructively defined. Interestingly enough, D-models and D-consistency are both based on Dempster's rule of combination (hence the "D"), the pivot mechanism for aggregating knowledge in Dempster-Shafer theory. We start by making the notion of "least informative" precise.

Definition. *Let $\mathcal{M}$ and $\mathcal{M}'$ be two bf-interpretations.*
(i) $\mathcal{M}$ is <u>less informative than</u> $\mathcal{M}'$ (written $\mathcal{M}\sqsubseteq\mathcal{M}'$) if, for any ψ, $\mathcal{M}\models\psi$ implies $\mathcal{M}'\models\psi$.
(ii) For a given Φ, $\mathcal{M}$ is a <u>least informative bf-model</u> of Φ if $\mathcal{M}\models\Phi$, and for any bf-model $\mathcal{M}'$ of Φ, $\mathcal{M}\sqsubseteq\mathcal{M}'$.

It is easy to verify that $\sqsubseteq$ is a partial order, and that $\mathbb{1}_{\mathcal{S}}\sqsubseteq\mathcal{M}\sqsubseteq\mathbb{1}_{\emptyset}$ for any $\mathcal{M}$.[5] It follows from the above definition that least informative bf-models completely characterize bf-entailment: if $\mathcal{M}$ is the least informative bf-model of Φ, then, for any bf-formula ψ, $\Phi\models\psi$ iff $\mathcal{M}\models\psi$

We now define D-models. We consider the single item of evidence represented by a bf-formula $\psi = F{:}[a,b]$, and define the *ψ-induced evidence* to be the function $\mathcal{E}_\psi$ given by: $\mathcal{E}_\psi([\![F]\!]) = a$; $\mathcal{E}_\psi([\![\sim F]\!]) = 1-b$; $\mathcal{E}_\psi(\mathcal{S}) = b-a$; and $\mathcal{E}_\psi(x) = 0$ otherwise.[6] Intuitively, $\mathcal{E}_\psi$ says that we believe to the extent a that the "true" state of the world is one where F holds, and to the extent $(1\text{-}b)$ that it is one where $\sim F$ does. We extend the idea of "induced evidence" to sets of bf-formulae as follows.

[4] Hence, *BFL* cannot handle (nor is it meant to!) non-monotonicity.

[5] Notice that $\mathcal{M}\sqsubseteq\mathcal{M}'$ iff $Bel_{\mathcal{M}}(x) \leq Bel_{\mathcal{M}'}(x)$ for all x; as a consequence, least informative bf-models are unique. Shafer, Dubois & Prade, and Smets independently defined equivalent orders.

[6] $\mathcal{E}_\psi$ slightly generalize Shafer's (1976) *simple support functions*.

Definition. *Let Φ be a set of bf-formulae. The* <u>*D-model*</u> *of Φ is the function $\mathcal{M}_D(\Phi)$ given by:*

$$\mathcal{M}_{\mathcal{D}}(\Phi) = \oplus\{\, \mathcal{E}_\psi \mid \psi \in \Phi \},$$

where : $(f_1 \oplus f_2)(x) = \sum_{y \cap z = x} f_1(y) f_2(z)$

The $\oplus$ is the usual (but un-normalized) Dempster's rule of combination. Intuitively, $\mathcal{M}_{\mathcal{D}}(\Phi)$ distributes our credibility over all the possible interpretations in such a way that all the information contained in Φ is considered — through its induced evidence. $\mathcal{M}_{\mathcal{D}}(\Phi)$ is indeed a bf-model of Φ.

Theorem 4. *For any* Φ, $\mathcal{M}_{\mathcal{D}}(\Phi) \models \Phi$.

Example 9. Consider again the Φ in Example 6, and let $\boldsymbol{B'} \cup \boldsymbol{C'} \cup \boldsymbol{D'} = [\![\forall x.\text{drinker}(x) \supset \text{smoker}(x)]\!]$. $\mathcal{E}_{\text{drinker(peter)}:[.8,.9]}$ assigns 0.8 to the set $\boldsymbol{A} \cup \boldsymbol{B}$, 0.1 to $\boldsymbol{C} \cup \boldsymbol{D}$, and 0.1 to $\mathcal{T}$. $\mathcal{E}_{\forall x.\text{drinker}(x) \supset \text{smoker}(x):[.7,1]}$ assigns 0.7 to $\boldsymbol{B'} \cup \boldsymbol{C'} \cup \boldsymbol{D'}$, and 0.3 to $\mathcal{T}$. By combining these two functions through $\oplus$ we get:

	B'	$C' \cup D'$	$B' \cup C' \cup D'$	$A \cup B$	$C \cup D$	$\mathcal{T}$	$\emptyset$
$\mathcal{M}_{\mathcal{D}}(\Phi)$	0.24	0.03	0.07	0.56	0.07	0.03	0

The reader can verify that $\mathcal{M}_{\mathcal{D}}(\Phi)$ is a bf-model of Φ by computing the values of $Bel_{\mathcal{M}_{\mathcal{D}}(\Phi)}$.

In the above example, $\mathcal{M}_{\mathcal{D}}(\Phi)(\emptyset)$. However, because our $\oplus$ is not normalized, we have no guarantee in general that $\mathcal{M}_{\mathcal{D}}(\Phi)$ is normal. We give the following:

Definition. Φ *is* D-consistent *iff* $\mathcal{M}_{\mathcal{D}}(\Phi)(\emptyset) = 0$.

When Φ is D-consistent, its D-model has some very interesting properties.

Theorem 5. *If Φ is D-consistent, then:*
(i) $\mathcal{M}_{\mathcal{D}}(\Phi)$ *is the least informative bf-model of* Φ.
(ii) $\Phi \models F{:}[a,b]$ *if and only if* $\mathcal{M}_{\mathcal{D}}(\Phi) \models F{:}[a,b]$
(iii) $\Phi \cup F{:}[1,1]$ *α-inconsistent iff* $\mathcal{M}_{\mathcal{D}}(\Phi \cup F{:}[1,1])(\emptyset) \geq \alpha$.

From a D-S viewpoint, (i) says that the result of combining items of information by Dempster's rule is the least informative basic probability assignment that subsumes the given information, *provided that* it is D-consistent. From the viewpoint of *BFL*, (ii) guarantees that we can build a bf-model of any D-consistent set Φ that fully characterize its entailment set. We will see in the next section the relevance of (iii) to automatic deduction.

Example 10. The willing reader who has computed the $Bel_{\mathcal{M}_{\mathcal{D}}(\Phi)}$ values in example 9 can now verify that the corresponding $\mathcal{M}_{\mathcal{D}}(\Phi)$ is least informative, by comparing these values with the lower bounds for the $Bel_{\mathcal{M}}$ values of any bf-model of Φ computed in example 6.

In order to conveniently use the results of Theorem 5, we need to find a more syntactical characterization of D-consistency. Notice that a bf-formula $F{:}[a,b]$ "says something" about (the truth of) both F and $\sim F$. We write $\pm F$ to denote either F or $\sim F$. Given a set Φ of bf-formulae, we focus on the set of all the sets of sentences about which Φ says something: we let $\Phi^* = \{\{\pm F_1, \ldots, \pm F_n\} \mid F_i \in \widehat{\Phi},\ i=1,\ldots,n\}$.

We say that Φ is *coherent* if all the sets in Φ^* are (fol) consistent. Two sets of bf-formulae Φ and Ψ are *distinct* if $\Phi \cup \Psi$ is coherent. Intuitively, Φ and Ψ are distinct if Φ does not say anything about any formula that is entailed by some set in Ψ^* (recall that a set Γ of (fol) formulae entails a formula F if and only if $\Gamma \cup \{\sim F\}$ is inconsistent), and vice-versa. Φ and Ψ "speak about" different formulae. Notice that Φ is coherent iff all its subsets are mutually distinct.

Theorem 6. *If Φ is coherent, then it is D-consistent.*

Coherence may seem too strong a condition to be useful: e.g., even the propositional set $\{P, P \supset Q\}$ is not coherent ($\{\sim P, \sim(P \supset Q)\}$ is inconsistent). The situation is however less extreme when we consider quantified knowledge: $\{\forall x. P(x) \supset Q(x), P(\mathbf{a})\}$, is coherent (intuitively, $\forall x.P(x) \supset Q(x)$ and $P(\mathbf{a})$ do speak about different things). Similarly, the Φ's in examples 6 and 7 are coherent, and the one in ex. 8 is not. Our definition of coherence and distinctness play the role of the intuitive "distinctness" condition required by Shafer for applying Dempster's combination $\oplus$ (also, cf. the notion of evidential independence in Shafer, 1976, §7.4): if Φ has been built up from *distinct* items of information (hence, it is *coherent*), then our $\oplus$-based D-models "behave well" as representations of Φ. Otherwise, D-models are not least informative (but they are still bf-models). Said differently, D-consistency formalizes a well known (but otherwise poorly understood) precondition in Dempster-Shafer theory.

Automatic Deduction

We outline a technique for performing automatic deduction in *BFL* (a full account, and an algorithm, are given in Saffiotti, 1991a). By automated deduction we mean here the ability to decide whether $\Phi \models F{:}[a,b]$ holds. The approach taken here is peculiar in that it relies on the construction of an uncertainty network that corresponds — in a precise sense — to the given entailment problem. The technique is sound and complete under the hypothesis of D-consistency of Φ. We answer the question $\Phi \models F{:}[a,b]$ by answering the two questions $\Phi \models F{:}[a,1]$ and $\Phi \models F{:}[0,b]$ separately (Theorem 1). We test if $\Phi \models F{:}[a,1]$ by testing if $\Phi \cup \{\sim F{:}[1,1]\}$ is a-inconsistent (Th. 3), and we test the latter by testing if $\mathcal{M}_{\mathcal{D}}(\Phi \cup \{\sim F{:}[1,1]\})(\emptyset) \geq a$ (Th. 5). (similarly for $F{:}[0,b]$).

Uncertainty networks enter into play at this point: using Shafer and Shenoy's valuation system formalism[7], we construct a valuation system $VS_{\mathcal{D}}(\Phi \cup \{\sim F{:}[1,1]\})$ in such a way that evaluating it produces (for a certain variable) exactly the value of $\mathcal{M}_{\mathcal{D}}(\Phi \cup \{\sim F{:}[1,1]\})(\emptyset)$. Rather than exposing the construction technique in general, we illustrate it by an example. We let $\Phi = \{\forall x.\text{P}(x) \supset \text{Q}(x){:}[.7,.9],\ \text{P}(\mathbf{a}){:}[.8,0],\ \text{P}(\mathbf{b}){:}[.6,.7]\}$ (cf. ex. 7), and wonder if $\Phi \models \exists x.\text{Q}(x){:}[.6,\ 1]$. We need to check if $\mathcal{M}_{\mathcal{D}}(\Phi \cup \{\forall x.\sim\text{Q}(x){:}[1,1]\})(\emptyset) \geq 0.6$. We take each set of (fol) formulae in Φ^*, and translate it into Skolem clausal form; let Γ be the set of resulting sets of

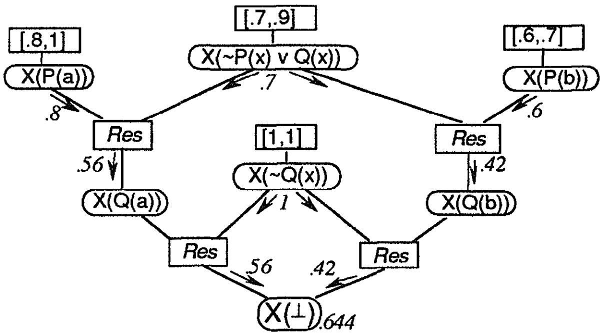
---

7 Space precludes us from giving even a short introduction to valuation systems (e.g., Shenoy and Shafer, 1988). We want to emphasize that valuations systems are not limited belief functions, but are intended as a general framework for local computation.

clauses. We also translate $\forall x.{\sim}Q(x)$ into clausal form, and look for resolution derivations of the empty clause $\perp$ from it and any of the sets in Γ. Two derivations are found:

$$\dfrac{{\sim}Q(x) \qquad \dfrac{P(a) \quad {\sim}P(x) \vee Q(x)}{Q(a)}}{\perp} \qquad\qquad \dfrac{{\sim}Q(x) \qquad \dfrac{P(b) \quad {\sim}P(x) \vee Q(x)}{Q(b)}}{\perp}$$

We then build a valuation system that "mirrors" these derivations. For the present goals, we can think of a valuation system as a network $\langle \mathcal{X}, \mathcal{V} \rangle$, where $\mathcal{X}$ is a set of nodes, each associated with a set of possible values, and $\mathcal{V}$ a set of basic probability assignments (called here *valuations*) that express information about the values taken by some (subsets of) nodes. A propagation mechanism (*evaluation*) can be used to compute the effect of the given valuations on the value of some "unknown" nodes. We build a valuation system $VS_{\mathcal{L}}(\Phi \cup \{\forall x.{\sim}Q(x){:}[1,1]\}) = \langle \mathcal{X}^*, \mathcal{V}^* \rangle$ as follows. For each clause C appearing in the above deductions, we put a node $X(C)$ in $\mathcal{X}^*$, with possible values {in, out}. For each resolution step $C_1,C_2/R$, we put in $\mathcal{V}^*$ a valuation *Res* on $\{X(C_1), X(C_2), X(R)\}$ that encodes the relation "whenever both $X(C_1)$ and $X(C_2)$ are in, $X(R)$ must be in". Finally, we put in $\mathcal{V}^*$ valuations corresponding to the $[a,b]$ values in Φ. The following picture shows the $\langle \mathcal{X}^*, \mathcal{V}^* \rangle$ for our example.[8]



We are almost done. All we have to do now is to evaluate this valuation system to find a value for the event $X(\perp)$ = in: here, 0.644. This value is exactly the value given by the D-model of $\Phi \cup \{\forall x.{\sim}Q(x){:}[1,1]\}$ to the empty set. The full paper proves this result in general. (Saffiotti & Umkehrer, 1991) applies this technique to the generation of uncertainty networks from first-order clauses.

Conclusion

BFL extends first-order logic with a notion of quantified belief based on the belief function formalism. From a dual viewpoint, *BFL* extends Dempster-Shafer theory with the ability of expressing belief about first-order (rather than propositional) statements. A number of results define the behavior of *BFL* as an integrated logic; as a by-product, these results reveal a new perspective on Dempster-Shafer theory.

Though contructed from first-order logic and belief functions, *BFL* illustrates a general approach to coupling a logical language and an uncertainty formalism (discussed in Saffiotti, 1990). On the one hand, the properties of the measures of strength of belief can be changed, by imposing different constraints on the interpretation structures. On the other hand, any language can be used for representing the objects of belief, provided that a (recursively enumerable) entailment relation is defined for it. The network generation procedure described above can similarly be extended to other languages and/or uncertainty formalism (Saffiotti, 1991b).

Acknowledgements. Yen-Teh Hsia, Robert Kennes, Victor Poznanski, Enrique Ruspini, Philippe Smets, Elisabeth Umkehrer, and two anonymous referees provided very useful comments.

References

Bacchus, F. (1988) *Representing and Reasoning with Probabilistic Knowledge*. PhD Thesis (Univ. of Alberta).

Dempster, A.P. (1967). "Upper and Lower Probabilities Induced by a Multivalued Mapping". *Annals of Math. Statistic* **38**.

Dubois, D. and Prade, H. (1988) *Possibility Theory: An Approach to Computerized Processing of Uncertainty* (Plenum Press, NY).

Dubois, D., Lang J. and Prade, H. (1989) "Automated Reasoning Using Possibilistic Logic:". *Procs. of the 5th Workshop on Unc. in AI* (Winsdor, Ontario)

Fagin, R. and Halpern, J. Y. (1989) "Uncertainty, Belief and Probability". *Procs. of IJCAI 89*.

Lang, J. (1991) *Logique Possibiliste: aspect formels, déduction automatique et applications*. PhD Thesis (Toulouse, Fr).

Nilsson, N. J. (1986) "Probabilistic Logic". *Artif. Intell.* **28**.

Pearl, J. (1988). *Probabilistic Reasoning in Intelligent Systems* (Morgan Kaufman, CA).

Provan, G. M. (1990) "A Logic-Based Analysis of Dempster-Shafer Theory". *Int. J. of Approximate Reasoning* **4**.

Rescher, N. and Brandom, R. (1979) *The Logic of Inconsistency*. (Billings & Sons Ltd., GB)

Ruspini, E. H. (1986) "The Logical Foundations of Evidential Reasoning". *Technical Note 408*. SRI Int. (Menlo Park, CA).

Saffiotti, A. (1990). "A Hybrid Framework for Representing Uncertain Knowledge". *Procs. of AAAI-90* (Boston, MA).

Saffiotti, A. (1991a) "A Belief-Function Logic". *Technical Report TR/IRIDIA/91-25* (Université Libre de Bruxelles, Belgium).

Saffiotti, A. (1991b) "Dynamic Construction of Valuation Systems". *Tech. Rep. TR/IRIDIA/91-18* (Univ. L. Bruxelles).

Saffiotti, A. and Umkehrer, E. (1991) "Automatic Construction of Valuation Systems from General Clauses". *Procs of IPMU-92. Also Tech. Rep. TR/IRIDIA/91-24* (Univ. Libre de Bruxelles).

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

Shenoy, P.P. and Shafer, G.R. (1988). "An Axiomatic Framework for Bayesian and Belief-Function Propagation". *Procs. of the 4th Workshop on Unc. in AI* (Minneapolis, MN).

Smets, Ph. (1988). "Belief Functions". In: Smets, Mamdani, Dubois, and Prade (Eds.) *Non-Standard Logics for Automated Reasoning* (Academic Press, London).

Zadeh, L. A. (1978) "Fuzzy Sets as a Basis for a Theory of Possibility". *Fuzzy Sets and Systems* **1**.

[8] Rounds represent nodes, rectangles valuations. The real system is more complex: (Saffiotti, 1991a) reports the details. Arrows are suggestive of the propagated values: the actual propagation works differently.