Artificial Reasoning with Subjective Logic*

Audun Jøsang **

Department of Telematics, NTNU, 7034 Trondheim, Norway email:
ajos@item.ntnu.no

Abstract. This paper defines a framework for artificial reasoning called *Subjective Logic*, which consists of a belief model called *opinion* and set of operations for combining opinions. Subjective Logic is an extension of standard logic that uses continuous uncertainty and belief parameters instead of only discrete truth values. It can also be seen as an extension of classical probability calculus by using a second order probability representation instead of the standard first order representation. In addition to the standard logical operations, Subjective Logic contains some operations specific for belief theory such as consensus and recommendation. In particular, we show that Dempster's consensus rule is inconsistent with Bayes' rule and therefore is wrong, and provide an alternative rule with a solid mathematical basis. Subjective Logic is directly compatible with traditional mathematical frameworks, but is also suitable for handling ignorance and uncertainty which is required in artificial intelligence.

1 Introduction

In standard logic, propositions are considered either true or false. However, a fundamental aspect of the human condition is that nobody can ever determine with absolute certainty whether a statement about the world is true or false. In addition, whenever the truth of a statement is assessed, it is always done by an individual, and it can never be considered to represent a general and objective opinion. This indicates that an important aspect is missing in the way standard logics capture our perception of reality, and that they are designed for an idealised world that we do not have access to.

Several alternative calculi and logics which take uncertainty into consideration have been proposed and quite successfully applied to practical problems where conclusions have to be made based on insufficient evidence (see for example [BK86] or [HH88] for an analysis of some uncertainty logics and calculi).

This paper defines *subjective logic* as a logic which operates on our subjective beliefs about the world. We use the term *opinion* to denote the representation of a belief. An opinion includes the concepts of disbelief and ignorance in addition to belief itself. Subjective logic contains the standard set of logic operations in addition to some non-standard operations which specifically use belief ownership as a parameter.

* Appears in the Proceedings of the 2nd Australian Workshop on Commonsense Reasoning. A. Nayak and M. Pagnucco (eds.), Perth, December 1997, Australian Computer Society.

** This research was supported by Norwegian Research Council Grant No.116417/410.

2 The Opinion Model

We assume the world to be in a particular state at any given time. Our knowledge about the world is never perfect so we can never determine the state exactly. For the purpose of believing a proposition about an aspect of the world, we assume that the proposition will either be true or false, and not something in between. However, because of our imperfect knowledge, it is impossible to know with certainty whether it is true or false, so that we can only have an *opinion* about it, which translates into degrees of belief or disbelief. In addition it is necessary to take into consideration degrees of ignorance, which can be described as a vacuous belief which fills the void in the absence of both belief and disbelief. For a single opinion about a proposition, we assume that

$$b + d + i = 1, \quad \{b, d, i\} \in [0, 1]^3 \tag{1}$$

where b, d and i designate belief, disbelief and ignorance respectively. Eq.(1) defines a triangle as illustrated in Fig.1, and an opinion can be uniquely described as a point $\{b, d, i\}$ in the triangle.

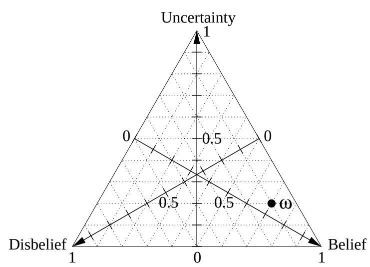

Fig. 1. Opinion Triangle

Definition 1. *Let $\pi = \{b, d, i\}$ be a triplet which satisfies Eq.(1), where the first, second and third element correspond to belief, disbelief and ignorance respectively. Then π is called an opinion. We will denote by Π the set of opinions.*

The bottom line between belief and disbelief in Fig.1 represents situations with zero ignorance and is equivalent to a traditional probability model. The degree of ignorance can be interpreted as the lack of evidence to support either belief or disbelief.

The difference between $\{0.5, 0.5, 0.0\}$ and $\{0, 0, 1\}$ is that in the first case there are equally strong reasons to believe that the proposition is true as false, whereas in the second case there is no reason to believe either. This way of modelling uncertain beliefs corresponds to the Dempster-Shafer Theory of Evidential Reasoning[Sha76] which will be briefly described.

The first step in applying evidential reasoning is to define a set of possible situations which is called the *frame of discernment*. A frame of discernment delimits a set of possible states of the world, exactly one of which is assumed to be true at any one time.

Definition 2. *Let Θ be a frame of discernment. If with each subset $x \subseteq \Theta$ a number $m_\Theta(x)$ is associated such that:*

$$\begin{array}{ll} \mathbf{1}. & m_\Theta(x) \geq 0 \\ \mathbf{2}. & m_\Theta(\emptyset) = 0 \\ \mathbf{3}. & \sum_{x \subseteq \Theta} m_\Theta(x) = 1 \end{array}$$

then m_Θ is called a belief mass distribution[1] *on Θ. For each subset $x \subseteq \Theta$, the number $m_\Theta(x)$ is called the* belief mass[2] *of x.*

A belief mass $m_\Theta(x)$ expresses the belief assigned to precisely the set x. It does not express any belief in subsets of x. Total ignorance can be expressed by not assigning any belief mass to any of the proper subsets of a frame of discernment.

The Dempster-Shafer theory provides a function for computing from two belief mass distributions a new belief mass distribution reflecting their combined influence. This function is known as *Dempster's consensus rule*. The definition is as follows.

Definition 3. *Let Θ be a frame of discernment, and let m^1_Θ and m^2_Θ be belief mass distributions on Θ. Then $m^1_\Theta \oplus m^2_\Theta$ is a function $m^1_\Theta \oplus m^2_\Theta : 2^\Theta \longmapsto [0, 1]$ such that*

$$\begin{array}{ll} \mathbf{1}. & m^1_\Theta \oplus m^2_\Theta(\emptyset) = 0, \quad \textit{and} \\ \mathbf{2}. & m^1_\Theta \oplus m^2_\Theta(x) = \frac{\sum_{y \cap z = x} m^1_\Theta(y) \cdot m^2_\Theta(z)}{1 - \kappa} \quad \textit{for all } x \neq \emptyset \end{array}$$

where $\kappa = \sum_{y \cap z = \emptyset} m^1_\Theta(y) \cdot m^2_\Theta(z)$

Dempster's rule of combination has been criticised(see e.g. [Zad84] or [Coh86]), mainly because highly conflicting belief mass distributions produce counterintuitive results. The criticism seems justified, and in Secs.3 and 4.4 we will describe a combination rule based on Bayesian calculus and which therefore has a solid mathematical basis.

The opinion model is in reality a trivial instance of applying the Dempster-Shafer theory of evidence. In particular we see that the frame of discernment Θ is the trivial universe {true, false}, and that a belief mass distribution m_Θ according to Def.2 corresponds to an opinion π according to Def.1 in our model.

The opinion model is also equivalent to Baldwin's belief model for *Support Logic Programming* [Bal86]. In this model, a lower and upper support for a

[1] called *basic probability assignment* in [Sha76]
[2] called *basic probability number* in [Sha76]

proposition p is expressed as $p : [Sl, Su]$. The lower support Sl corresponds to the belief b in our model, and the upper support Su corresponds to belief plus ignorance expressed by $b + i$ in our model. Some of the operations presented in the next sections are therefore equivalent to operations defined in [Bal86] but we feel that our notation is much simpler and lends itself more easily to an intuitive interpretation.

3 Second Order Probability Representation

In this section we show that opinions can be represented as 2-order probability estimates, and define a bijective mapping between the two representations. By describing the *Bayesian consensus rule*, we then have the necessary formal basis for defining the equivalent operation for opinions in Sec.4.

3.1 Representation of 2-Order Probability Estimates

Let us consider the probability of throwing for example a "five" with a fair dice. This, most people would agree to be 1/6. Imagine now a dice which has been loaded, so that the probability of throwing a particular side, without knowing which, is 1/2, and the probability of throwing any of the other sides is 1/10. Let us again consider the probability of throwing a "five" with this loaded dice. A 1-order probability analysis would dictate this to be $1/2 \cdot 1/6 + 1/10 \cdot 5/6 = 1/6$, again the same result as for the fair dice. But with the available information, any observer would know better; it can impossibly be 1/6. As a matter of fact, it is either 1/2 (with probability 1/6) or 1/10 (with probability 5/6).

A 2-order probability is simply the probability of a 1-order probability. The above example illustrates a simple case of 2-order probability representation, i.e probabilities of probabilities, in which the 1-order probability could only have the two possible discrete values 1/2 and 1/10. In general, 2-order probabilities can be represented as a probability density function over a 1-order probability variable.

Definition 4. *Let θ be the 1-order probability variable of some binary event, then θ is defined in the interval $[0, 1]$. A 2-order probability density function on θ is a function $\psi(\theta)$ so that:*

$$\begin{array}{ll} \mathbf{a}.\ \psi(\theta) \geq 0 & \textit{for all}\ \ \theta \in [0,1], \\ \mathbf{b}.\ \int_0^1 \psi(\theta)\mathrm{d}\theta = 1 & \end{array}$$

We will designate by Ψ the class of 2-order probability density functions, or 2-order pdfs for short.

It has already been suggested in earlier work(see e.g. [GS82, Pea88, Chá96, Law97]) that this way of representing probabilities better reflect our perception of uncertain future events than standard probability does.

It is natural to choose the beta function as a mathematical approximation of a 2-order pdf, since it is one of the few well known distributions that give the cumulative probability 1 to the finite interval [0,1].

The beta-family of distributions is a continuous family of functions indexed by the two parameters α and β. The beta(α, β) pdf is:

$$f(\theta \,|\, \alpha, \beta) = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)}\theta^{\alpha-1}(1-\theta)^{\beta-1} \;, \quad 0 \leq \theta \leq 1, \;\alpha > 0, \;\beta > 0 \tag{2}$$

As the parameters α and β vary, the beta distribution takes on many shapes. The pdf can be strictly increasing ($\alpha > 1, \beta = 1$), strictly decreasing ($\alpha = 1, \beta > 1$), U-shaped ($\alpha < 1, \beta < 1$), or unimodal ($\alpha > 1, \beta > 1$). The case $\alpha = \beta$ yields a pdf symmetric about 1/2, and when $\alpha = \beta = 1$, the beta distribution reduces to the uniform distribution. For reasons which will become clear with Th.7 and also in Sec.3.2, we will only consider the subclass of beta distributions called *2-order B-pdfs.*

Definition 5. *Let Φ be the class of beta distributions for which $\alpha \geq 1$ and $\beta \geq 1$. Let φ be a beta distribution in Φ. Then φ is called a 2-order Bayesian probability density function, or 2-order B-pdf for short.*

In our notation, 2-order B-pdfs will be characterised by the parameters $\{r, s\}$ instead of $\{\alpha, \beta\}$ through the following correspondence:

$$\begin{array}{ll} \alpha = r+1, & r \geq 0 \quad \text{and} \\ \beta = s+1, & s \geq 0. \end{array}$$

Let φ be a 2-order B-pdf over the 1-order probability variable θ. In our notation φ will then be characterised by r and s according to:

$$\varphi(\theta \,|\, r, s) = \frac{\Gamma(r+s+2)}{\Gamma(r+1)\Gamma(s+1)}\theta^{r}(1-\theta)^{s} \;, \quad 0 \leq \theta \leq 1, \;\; r \geq 0, \;\; s \geq 0 \tag{3}$$

The choice of only considering beta distributions from Φ, which excludes all U-shaped distributions, will not cause any restrictions on the generality of our results. As a matter of fact, it can be shown that any 2-order probability distributions can be expressed as a linear combination of 2-order B-pdfs.

Definition 6. *As before Φ is the set of 2-order B-pdfs. Let ${}_{\epsilon}\varphi$ be a 2-order pdf defined according to*

$${}_{\epsilon}\varphi = \sum_{i=1}^{n} \varphi_i \delta_i, \quad n \textit{ is positive integer}, \;\; \delta_i \in (0,1], \;\; \sum_{i=1}^{n} \delta_i = 1 \tag{4}$$

Then ${}_{\epsilon}\varphi$ is called a linear combination of 2-order B-pdfs. We will let ${}_{\epsilon}\Phi$ designate the set of 2-order pdfs so defined.

It is obvious that ${}_\epsilon\Phi \subseteq \Psi$. Whether ${}_\epsilon\Phi \subset \Psi$ or ${}_\epsilon\Phi = \Psi$ remains unclear but will not have any practical consequences on our results.

Theorem 7. *As before, Ψ is the set of 2-order pdfs. For any element $\psi \in \Psi$, ψ can be approximated with arbitrary accuracy as $\psi \approx {}_\epsilon\varphi$, where ${}_\epsilon\varphi \in {}_\epsilon\Phi$.*

The example with the loaded dice can in fact be approximated and represented as a linear combination of two 2-order B-pdfs in the form of peak-functions, which is graphically illustrated in Fig.2.

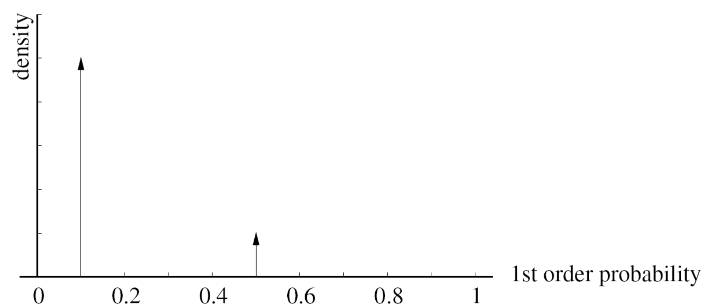

Fig. 2. 2-order probability estimate of throwing a "five" with loaded dice

3.2 Determining the Posteriori 2-Order Probability Estimate

Assume an entity which produce events where the outcome can either be positive or negative. Let the entity produce the event a certain number of times, and let p and n designate the number of positive and negative results respectively.

We assume that the priori probability estimate can be described as the uniform distribution $\varphi(\theta \,|\, 0, 0)$. It can then be shown that the posteriori probability estimate is a 2-order B-pdf in the form of (5) with $\{r, s\}$ parameters equal to $\{p, n\}$.

$$\varphi(\theta \,|\, p, n) = \frac{\Gamma(p+n+2)}{\Gamma(p+1)\Gamma(n+1)} \theta^p (1-\theta)^n \;, \quad 0 \leq \theta \leq 1, \;\; p \geq 0, \;\; n \geq 0 \tag{5}$$

The mean value of a 2-order probability estimate represented in the general form of (3) is given by:

$$\mu(\theta) = \frac{(r+1)}{(r+s+2)} \tag{6}$$

3.3 Consensus Between Independent 2-Order Probability Estimates

Assume two agents A and B having observed an entity produce a binary event over two different periods respectively. According to (5), their respective 2-order B-pdfs are then $\varphi(p^A, n^A)$ and $\varphi(p^B, n^B)$. Imagine now that they combine their observations to form a better estimate of the event's probability. This is equivalent to an imaginary agent $[A, B]$ having made all the observations and who therefore can form the 2-order B-pdf defined by $\varphi(p^A + p^B, n^A + p^B)$. This result can be generalised to cover real pdf parameters.

Definition 8. *Let $\varphi(r_p^A, s_p^A)$ and $\varphi(r_p^B, s_p^B)$ be two 2-order probability estimates respectively held by the agents A and B regarding the truth of a proposition p. The 2-order probability estimate $\varphi(r_p^{A,B}, s_p^{A,B})$ defined by*

$$\begin{array}{l} \mathbf{1}.\ r_p^{A,B} = r_p^A + r_p^B \\ \mathbf{2}.\ s_p^{A,B} = s_p^A + s_p^B \end{array}$$

is then called the Bayesian consensus rule for combining A's and B's estimates, as if it was an estimate held by an imaginary agent $[A, B]$. By using the symbol $\oplus$ to designate this operation, we get $\varphi(r_p^{A,B}, s_p^{A,B}) = \varphi(r_p^A, s_p^A) \oplus \varphi(r_p^B, s_p^B)$.

It is easy to prove that $\oplus$ is both commutative and associative which means that the order in which probability estimates are combined has no importance. Probability estimate independence is explicitly assumed, which obviously translates into not allowing an entity's probability estimate to be counted more than once.

3.4 Equivalence between Opinions and 2-Order B-Pdfs

We have defined Φ to be the class of 2-order B-pdfs, and Π to be the class of opinions. Let $\pi_p = \{b_p, d_p, i_p\}$ be an agent's opinion about proposition p, and let $\varphi(r_p, s_p)$ be the same agent's estimate of p being true expressed as a 2-order B-pdf. Let $\varphi(r_p, s_p)$ be defined as a function of π_p according to:

$$\begin{cases} r_p = \frac{b_p}{i_p} \\ s_p = \frac{d_p}{i_p} \end{cases} \tag{7}$$

We see for example that $\pi = \{0, 0, 1\}$ which expresses total ignorance corresponds to the uniform $\varphi(0, 0)$, that $\pi = \{1, 0, 0\}$ which expresses absolute belief corresponds to $\varphi(\infty, 0)$ or the absolute probability, and that $\pi = \{0, 1, 0\}$ which expresses absolute disbelief corresponds to $\varphi(0, \infty)$ or the zero probability. By defining φ as a function of π according to (7), the interpretation of φ corresponds exactly to the interpretation of π. The correspondence in the opposite direction is established by including Eq.(1) in (7) and solving the system for $\{b, d, i\}$ to obtain:

$$\begin{cases} b_p = \frac{r_p}{r_p+s_p+1} \\ d_p = \frac{s_p}{r_p+s_p+1} \\ i_p = \frac{1}{r_p+s_p+1} \end{cases} \tag{8}$$

Eq.(8) defines a bijective mapping between Φ and Π so that any opinion has an equivalent mathematical and interpretative representation as a 2-order B-pdf and vice versa. We will call this mapping the *probability-opinion mapping*.

Definition 9. *As before,* Π *is the class of opinions and* Φ *is the class of 2-order B-pdfs. The function* $\widehat{\Phi\Pi} : \Phi \longmapsto \Pi$ *defined by* $\widehat{\Phi\Pi}(\varphi(r_p, s_p)) = \pi_p$*, where the components of* π_p *are defined according to (8), is called the probability-opinion mapping between* Φ *and* Π.

The $\widehat{\Phi\Pi}$-mapping, which also has an inverse, will make it possible to define the equivalent of consensus of opinions which, in contrast to Dempster's rule, will have a sound mathematical basis.

3.5 Eccentric Opinions and 2-Order Pdfs

Having established the mathematical and interpretative equivalence between elements from Φ and Π, this section will describe the correspondence between elements from ${}_\epsilon\Phi$ and what we will call the class of *eccentric opinions*.

Definition 10. *As before,* Π *is the class of opinions. Let* ${}_\epsilon\pi$ *be a linear combination of elements in* Π *such that*

$${}_\epsilon\pi = \sum_{i=1}^{n} \pi_i\delta_i, \;\; n \textit{ is positive integer}, \;\; \delta_i \in (0,1], \; \sum_{i=1}^{n} \delta_i = 1$$

then ${}_\epsilon\pi$ *is called an eccentric opinion. We will denote by* ${}_\epsilon\Pi$ *the class of eccentric opinions.*

We have not been able to conceive a graphical illustration of eccentric opinions. However they correspond to linear combinations of 2-order B-pdfs, and as such, an opinion about the proposition p : "*I will throw a "five" with the loaded dice*" is an eccentric opinion, because it can not be expressed as a simple opinion. Numerically, the eccentric opinion, which corresponds to Fig.2, can be expressed as:

$${}_\epsilon\pi_p = \{\frac{1}{2}, \frac{1}{2}, 0\} \cdot \frac{1}{6} + \{\frac{1}{10}, \frac{9}{10}, 0\} \cdot \frac{5}{6}$$

The $\widehat{\Phi\Pi}$-mapping defined above can naturally be extended to map elements from ${}_\epsilon\Phi$ to elements in ${}_\epsilon\Pi$.

Definition 11. *Let* ${}_\epsilon\varphi$ *and* ${}_\epsilon\pi$ *be elements from* ${}_\epsilon\Phi$ *and* ${}_\epsilon\Pi$ *respectively. Let* ${}_\epsilon\varphi$ *and* ${}_\epsilon\pi$ *be defined by:*

$$\begin{array}{ll} {}_\epsilon\varphi = \sum_{i=1}^{n} \varphi_i\delta_i \;, & \varphi_i \in \Phi \\ {}_\epsilon\pi = \sum_{i=1}^{n} \pi_i\delta_i \;, & \pi_i \in \Pi \end{array}$$

where n is positive integer, $\delta_i \in (0,1]$, and $\sum_{i=1}^{n} \delta_i = 1$. The function $\widehat{{}_{\epsilon}\Phi_{\epsilon}\Pi} : {}_{\epsilon}\Phi \longmapsto {}_{\epsilon}\Pi$ defined by $\widehat{{}_{\epsilon}\Phi_{\epsilon}\Pi}({}_{\epsilon}\varphi) = {}_{\epsilon}\pi$, where each π_i is defined in function of φ_i according to the $\widehat{\Phi\Pi}$-mapping, is then called the probability-opinion mapping between ${}_{\epsilon}\Phi$ and ${}_{\epsilon}\Pi$.

Since $\widehat{\Phi\Pi}$ is bijective, so is $\widehat{{}_{\epsilon}\Phi_{\epsilon}\Pi}$.

4 Subjective Logic

Standard propositional logic operates on propositions about a world that we do not have direct access to, and the logic's variables can take the discrete values of *true* or *false*. Subjective logic which we will present in this section operates on our subjective perception about the world. The logic uses our individual opinions about the truth of propositions as variables.

Opinions, as defined in Sec.2, are considered subjective, and will therefore have an ownership assigned whenever relevant. In our notation, superscripts indicate ownership, and subscripts indicate the proposition to which the opinion apply. For example

$$\pi_p^A$$

is an opinion held by agent A about the truth of proposition p. Presently, subjective logic contains about 10 different operations, where the most important will be described below.

1. conjunction
2. disjunction
3. negation
4. consensus
5. recommendation
6. ordering

Operations 1), 2) and 3) are equivalent to operations defined in [Bal86]. The definitions of the operations 4), 5), and 6) have as far as we know not been proposed before.

4.1 Conjunction

A conjunction of two opinions about propositions consists of determining from the two opinions a new opinion reflecting the conjunctive truth of both propositions. This corresponds to the logical binary “AND” operation in standard logic.

Definition 12. *Let $\pi_p = \{b_p, d_p, i_p\}$ and $\pi_q = \{b_q, d_q, i_q\}$ be an agent's opinions about two distinct propositions p and q. Let $\pi_{p\wedge q} = \{b_{p\wedge q}, d_{p\wedge q}, i_{p\wedge q}\}$ be the opinion such that*

1. $b_{p\wedge q} = b_p b_q$
2. $d_{p\wedge q} = d_p + d_q - d_p d_q$
3. $i_{p\wedge q} = b_p i_q + i_p b_q + i_p i_q$

Then $\pi_{p\wedge q}$ is called the conjunction of π_p and π_q, representing the agents opinion about both p and q being true. By using the symbol "$\wedge$" to designate this operation, we get $\pi_{p\wedge q} = \pi_p \wedge \pi_q$.

As would be expected, conjunction of opinions is both commutative and associative. It must be assumed that the opinion arguments in a conjunction are independent. This means for example that the conjunction of an opinion with itself will be meaningless, because the conjunction rule will see them as if they were opinions about distinct propositions.

When applied to opinions with absolute belief or disbelief, the conjunction rule produces the same results as the "AND" operator in standard logic. In addition, when applied to opinions with zero ignorance, it is equivalent with the product of 1-order probabilities.

4.2 Disjunction

A disjunction of two opinions about propositions consists of determining from the two opinions a new opinion reflecting the disjunctive truth of both propositions. This corresponds to the logical binary "OR" operation in standard logic.

Definition 13. *Let $\pi_p = \{b_p, d_p, i_p\}$ and $\pi_q = \{b_q, d_q, i_q\}$ be an agent's opinions about two distinct propositions p and q. Let $\pi_{p\vee q} = \{b_{p\vee q}, d_{p\vee q}, i_{p\vee q}\}$ be the opinion such that*

1. $b_{p\vee q} = b_p + b_q - b_p b_q$
2. $d_{p\vee q} = d_p d_q$
3. $i_{p\vee q} = d_p i_q + i_p d_q + i_p i_q$

Then $\pi_{p\vee q}$ is called the disjunction of π_p and π_q, representing the agents opinion about either p, q or both p and q being true. By using the symbol "$\vee$" to designate this operation, we get $\pi_{p\vee q} = \pi_p \vee \pi_q$.

Disjunction of opinions is both commutative and associative. As for conjunction, it must be assumed that the opinion arguments in a disjunction are independent. This means for example that the disjunction of an opinion with itself will be meaningless, because the disjunction rule will see them as if they were opinions about distinct propositions.

4.3 Negation

A negation of an opinion about a proposition consists of inverting the belief and disbelief components while keeping the ignorance component unchanged. This corresponds to the logical unary "NOT" operation in standard logic.

Definition 14. *Let $\pi_p = \{b_p, d_p, i_p\}$ be an agent's opinion about the proposition p being true. Let $\pi_{\neg p} = \{b_{\neg p}, d_{\neg p}, i_{\neg p}\}$ be the opinion such that*

1. $b_{\neg p} = d_p$
2. $d_{\neg p} = b_p$
3. $i_{\neg p} = i_p$

Then $\pi_{\neg p}$ is called the negation of π_p, representing the agents opinion about p being false. By using the symbol "$\neg$" to designate this operation, we get $\pi_{\neg p} = \neg\pi_p$.

Negation is involutive so that $\neg(\neg\pi) = \pi$ for any opinion π.

4.4 Consensus between Independent Opinions

The consensus rule for combining independent opinions consists of combining two or more independent opinions about the same proposition into a single opinion. This is what Dempster's rule is supposed to do, but we will provide an alternative rule with sound basis in conditional probability calculus.

The Bayesian consensus rule for combining independent opinions is obtained by using Def.8 and the probability-opinion mapping (8) from Sec.3.4.

Definition 15. *Let $\pi_p^A = \{b_p^A, d_p^A, i_p^A\}$ and $\pi_p^B = \{b_p^B, d_p^B, i_p^B\}$ be opinions respectively held by agents A and B about the same proposition p. Let $\pi_p^{A,B} = \{b_p^{A,B}, d_p^{A,B}, i_p^{A,B}\}$ be the opinion such that*

$$\begin{array}{ll} \mathbf{1}. & b_p^{A,B} = (b_p^A i_p^B + b_p^B i_p^A)/\kappa \\ \mathbf{2}. & d_p^{A,B} = (d_p^A i_p^B + d_p^B i_p^A)/\kappa \\ \mathbf{3}. & i_p^{A,B} = (i_p^A i_p^B)/\kappa \end{array}$$

where $\kappa = i_p^A + i_p^B - i_p^A i_p^B$ such that $\kappa \neq 0$. Then $\pi_p^{A,B}$ is called the Bayesian consensus between π_p^A and π_p^B, representing an imaginary agent $[A, B]$'s opinion about p, as if she represented both A and B. By using the symbol $\oplus$ to designate this operation, we get $\pi_p^{A,B} = \pi_p^A \oplus \pi_p^B$.

It is easy to prove that $\oplus$ is both commutative and associative which means that the order in which opinions are combined has no importance. Opinion independence is specifically assumed, which obviously translates into not allowing an entity's opinion to be counted more than once

Two opinions which both contain zero ignorance can not be combined according to Def.15. This can be explained by interpreting ignorance as *room for influence*, meaning that it is only possible to influence an opinion which has not yet been committed to belief or disbelief.

4.5 Recommendation

Assume two agents A and B where A has an opinion about B, and B has an opinion about a proposition p. A recommendation of these two opinions consists of combining A's opinion about B with B's opinion about p in order for A to get an opinion about p.

There is no such thing as physical recommendation, and recommendation of opinions therefore lends itself to different interpretations. The main difficulty lies with describing the effect of A disbelieving that B will give a good advice. What does this exactly mean? We will give three different interpretations.

1. *Ignorance Favouring.* A's disbelief in the recommending agent B means that A thinks that B ignores the truth value of p. As a result A also ignores the truth value of p.

2. *Disbelief Favouring.* A's disbelief in the recommending agent B means that A thinks that B consistently tries to misinform A through malicious intent. As a result A disbelieves in p.

3. *Belief and Disbelief Favouring.* A's disbelief in the recommending agent B means that A thinks that B consistently recommends the negation of his real opinion about the truth value of p. As a result, A not only disbelieves in p to the degree that B recommends belief, but she also believes in p to the degree that B recommends disbelief in p, because the combination of two disbeliefs results in belief in this case.

We will only define the recommendation rule for the first interpretations above, because we find the second and third interpretations less intuitive.

Definition 16. *Let A, B and be two agents where $\pi_B^A = \{b_B^A, d_B^A, i_B^A\}$ is A's opinion about B's recommendations, and let p be a proposition where $\pi_p^B = \{b_p^B, d_p^B, i_p^B\}$ is B's opinion about p expressed in a recommendation to A. Let $\pi_p^{AB} = \{b_p^{AB}, d_p^{AB}, i_p^{AB}\}$ be the opinion such that*

$$\begin{aligned}
&\mathbf{1}.\ b_p^{AB} = b_B^A b_p^B,\\
&\mathbf{2}.\ d_p^{AB} = b_B^A d_p^B\\
&\mathbf{3}.\ i_p^{AB} = d_B^A + i_B^A + b_B^A i_p^B
\end{aligned}$$

then π_p^{AB} is called the recommendation rule for combining π_B^A and π_p^B expressing A's opinion about p as a result of the recommendation from B. By using the symbol $\otimes$ to designate this operation, we get $\pi_p^{AB} = \pi_B^A \otimes \pi_p^B$.

It is easy to prove that π_p^{AB} satisfies Eq.(1) and that $\otimes$ is associative but not commutative. This means that the combination of opinions can start in either end of the chain, and that the order in which opinions are combined is significant. In a chain with more than one recommending entity, opinion independence must be assumed, which for example translates into not allowing the same entity to appear more than once in a chain.

B's recommendation must be interpreted as what B actually recommends to A, and *not* necessarily as B's real opinion. It is obvious that these can be totally different if B for example defects.

It is important to notice that the recommendation rule can only be justified when it can be assumed that recommendation is transitive. More precisely it must be assumed that the agents in the chain do not change their behaviour (i.e. what they recommend) as a function of which entities they interact with. However, this can not always be assumed, because defection can be motivated for example by antagonism between certain agents. The recommendation rule must therefore be used with care, and can only be applied in environments where behaviour invariance can be assumed.

4.6 Ordering

Assume an agent having opinions about different propositions. The maximum of these opinions consists of selecting the opinion containing the "strongest" belief. In this way, a set of opinions can be ordered.

The definition of the Max operation is obtained by using the mean value for 2-order pdfs given by Eq.(6) and the probability-opinion mapping (8).

Definition 17. *Let* $\pi_p = \{b_p, d_p, i_p\}$ *and* $\pi_q = \{b_q, d_q, i_q\}$ *be an agent's opinions about two distinct propositions* p *and* q*. Let* $\pi_{p\uparrow q} = \{b_{p\uparrow q}, d_{p\uparrow q}, i_{p\uparrow q}\}$ *be the opinion which the following algorithm would return:*

```
Max(INPUT: π_p, π_q, OUTPUT: π_{p↑q})
BEGIN PROCEDURE
   IF π_p and π_q have different (b+i)/(b+d+2i) ratios
   THEN
     RETURN opinion with greatest (b+i)/(b+d+2i) ratio;
   ELSE
     RETURN opinion with the least i;
   ENDIF;
END PROCEDURE;
```

Then $\pi_{p\uparrow q}$ *is called the maximum of* π_p *and* π_q*, representing the opinion with the strongest belief. By using the symbol "*$\uparrow$*" to designate this operation, we get* $\pi_{p\uparrow q} = \pi_p \uparrow \pi_q$.

As an example, Fig.3 illustrates how π_p, π_q and π_r are ordered. The opinions π_p and π_q have the greatest $(b+i)/(b+d+2i)$ ratio, so π_r is the weakest opinion. π_p and π_q have in fact equal $(b+i)/(b+d+2i)$ ratios, but π_q has the least i so finally π_q is the strongest opinion.

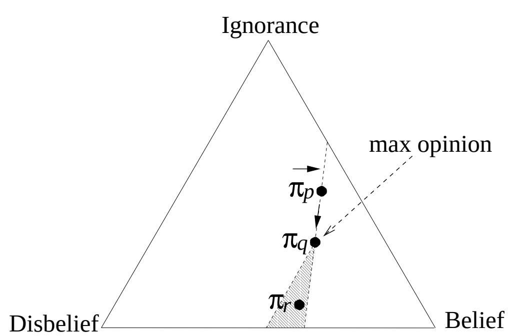

Fig. 3. Maximum of opinions

By using the symbol "<" for ordering opinions, we can write $\pi_r < \pi_p < \pi_q$.

It is interesting to notice that π_r's belief component in fact is greater that the belief components of π_p and π_q. The fact that π_r nevertheless is the the weakest opinion can be explained by π_p's and π_q's greater ignorance which potentially can be transformed into greater belief. In other words, it is better to select an opinion with a somewhat lesser belief component if only the ignorance component is much greater. The shaded area in Fig.3 covers the set of opinions with a greater belief component than π_q's, but which nevertheless are weaker than π_q.

4.7 Generalisation of Results

The above described operations were defined on elements from Π. It is easy to prove that the operations also can be applied to eccentric opinions from ${}_\epsilon\Pi$.

Let ${}_\epsilon\pi_p$ and ${}_\epsilon\pi_q$ be two eccentric opinions defined by

$$\begin{array}{l} {}_\epsilon\pi_p = \sum_{i=1}^{n} \pi_{p,i}\delta_{p,i}, \quad n \text{ positive integer}, \;\; \pi_{p,i} \in \Pi, \;\; \delta_{p,i} \in (0,1], \; \sum_{i=1}^{n} \delta_{p,i} = 1 \\ {}_\epsilon\pi_q = \sum_{j=1}^{m} \pi_{q,j}\delta_{q,j}, \quad m \text{ positive integer}, \;\; \pi_{q,j} \in \Pi, \;\; \delta_{q,j} \in (0,1], \; \sum_{j=1}^{m} \delta_{q,j} = 1 \end{array}$$

We can then for example express the conjunction of ${}_\epsilon\pi_p$ and ${}_\epsilon\pi_q$ as:

$${}_\epsilon\pi_p \wedge {}_\epsilon\pi_q = \sum_{i=1}^{n}\sum_{j=1}^{m}(\pi_{p,i} \wedge \pi_{q,j}) \cdot \delta_{p,i}\,\delta_{q,j}$$

Disjunction, negation and recommendation can be applied to eccentric opinions in a similar way. For consensus and ordering, the definition given above are not directly applicable, but have to be slightly modified and extended. Unfortunately we have to omit the description here due to the limited space.

5 Subjective Algebra

5.1 Boolean Functions

We have already mentioned that conjunction and disjunction are both commutative and associative, and that negation is involutive. But not all the laws of Boolean algebra are valid with the logical operations defined above.

Conjunction and disjunction can not be combined using the distributive laws. This is due to the fact that opinions must be assumed to be independent, whereas distribution always introduces an element of dependence. Take for example

$$\pi_p \wedge (\pi_q \vee \pi_r) \neq (\pi_p \wedge \pi_q) \vee (\pi_p \wedge \pi_r)$$

The right side of the not-equal sign is a disjunction of two dependent opinions, because they both contain π_p. Thus only the left side represents a correct combination of conjunction and disjunction.

Conjunction and disjunction are not idempotent, because that would imply combining an opinion with itself, which would also violate the independence requirement. However, it is easy to prove that De Morgan's laws are applicable in subjective logic. Let π_p and π_q be two independent opinions. We then have:

$$\begin{array}{l} \mathbf{1}.\;\; \neg(\pi_p \wedge \pi_q) = (\neg\pi_p) \vee (\neg\pi_q) \\ \mathbf{2}.\;\; \neg(\pi_p \vee \pi_q) = (\neg\pi_p) \wedge (\neg\pi_q) \end{array}$$

5.2 Mixing Consensus and Recommendation

It can be imagined that several recommendation chains produce opinions about the same proposition. Under the condition of opinion independence, these opinions generated through recommendation can be combined with the consensus rule to produce a single opinion about the target proposition.

It can also be imagined that within a recommendation chain, it is possible to obtain several recommended opinions about some of the agents in the chain. Again by assuming opinion independence, opinions obtained through the consensus rule can be recommended. An example of mixed consensus and recommendation is illustrated in Fig.4.

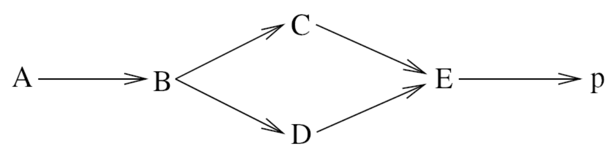

Fig. 4. Mixing consensus and recommendation

For the same reason as for conjunction and disjunction, the recommendation rule is not distributive relative to the consensus rule. Let π^A_B, π^B_C, π^B_D, π^C_E, π^D_E and π^E_p represent the opinion relationships in Fig.4. We then have

$$\pi^A_B \otimes ((\pi^B_C \otimes \pi^C_E) \oplus (\pi^B_D \otimes \pi^D_E)) \otimes \pi^E_p \neq (\pi^A_B \otimes \pi^B_C \otimes \pi^C_E \otimes \pi^E_p) \oplus (\pi^A_B \otimes \pi^B_D \otimes \pi^D_E \otimes \pi^E_p) \tag{9}$$

which according to the notation in Defs.15 and 16 can be written as

$$\pi_p^{A(BC,BD)E} \neq \pi_p^{ABCE,ABDE} \tag{10}$$

This result may seem counterintuitive at first, but the right side of (9) and (10) violates the requirement of independent opinions because they both contain π^A_B and π^E_p and thereby contain consensus combination of dependent opinions. Only the left sides of (9) and (10) thus represent correct mixing of consensus and recommendation.

6 Example: Taking Advice from Multiple Sources

A typical task in artificial reasoning systems is to combine several opinions into a single opinion, where each opinion is held by a different agent. In subjective logic, this can be done by letting one agent receive advice in the form of opinions from other agents and combine these into a single opinion. This would include consensus and recommendation. In addition, it is often necessary to determine an opinion about complex systems based on an observers opinions about the subsystems only.

We will analyse a situation where a newly designed industrial process c depends on two subprocesses a and b to produce correct result. This conjunctive situation is illustrated in Fig.5 below. Agent A needs to determine an opinion

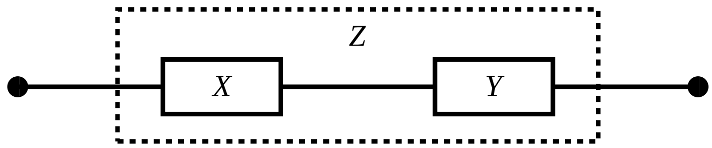

Fig. 5. Conjunctive system

about the proposition $p(c)$: "*Process c will produce correct result*". From earlier experience, A knows that process a produced correct result in 8 out of 10 trials. However, for process b, A has only second-hand evidence from the experts E_1 and E_2. First A must form an opinion about $p(a)$: "*Process a will produce correct result*" based on the statistical evidence. Using Eq.(8) and Def.8 she gets

$$\pi^{A}_{p(a)} = \{0.73,\ 0.18,\ 0.09\}$$

Now A must consider the advice from the experts E_1 and E_2 to form two recommended opinions about $p(b)$: "*Process b will produce correct result*" using the recommendation rule, and then combine these into a single opinion about $p(b)$ using the consensus rule. Her final opinion about $p(b)$ will then be:

$$\pi^{AE_1,AE_2}_{p(b)} = (\pi^{A}_{E_1} \otimes \pi^{E_1}_{p(b)}) \oplus (\pi^{A}_{E_2} \otimes \pi^{E_2}_{p(b)}) \tag{11}$$

It can be seen that this is a correct expression which does not violate the independence requirements, because no opinion appears more than once.

Let the experts opinions about $p(b)$ and A's opinion about the experts be defined by

$$\begin{array}{ll} \pi^{A}_{E_1} = \{0.90,\ 0.05,\ 0.05\} & \qquad \pi^{E_1}_{p(b)} = \{0.80,\ 0.00,\ 0.20\} \\ \pi^{A}_{E_2} = \{0.80,\ 0.10,\ 0.10\} & \qquad \pi^{E_2}_{p(b)} = \{0.70,\ 0.10,\ 0.20\} \end{array}$$

A's opinions about the experts can for example example be based on her statistical records of the experts' earlier performance. A's opinion about $p(b)$ as a function of her opinion about the experts combined with the experts' advice, can now be calculated using (11).

$$\pi^{AE_1,AE_2}_{p(b)} = \{0.77,\ 0.04,\ 0.19\}$$

Finally A's opinions about $p(a)$ and $p(b)$ can be combined using the conjunctive combination rule to produce her opinion about $p(c)$.

$$\begin{aligned}\pi^A_{p(c)} &= \pi^A_{p(a)} \wedge \pi^{AE_1,AE_2}_{p(b)} \\ &= \{0.56,\ 0.22,\ 0.22\}\end{aligned} \tag{12}$$

7 Conclusion

The subjectivity of the logic described in this paper is based on the fact that the opinion model includes ignorance as an essential part of subjective human beliefs and that ownership of opinions is assigned to individuals.

The simple opinion model can be seen as an instance of a Shaferian belief model. However, a Shaferian model is incapable of representing eccentric opinions, and therefore has limited capabilities for representing beliefs. We have also demonstrated that Dempster's rule is inconsistent with Bayesian probability calculus, and have proposed an alternative rule with a sound mathematical basis. We believe that the simplicity and conciseness of subjective logic makes it very attractive for applications not only in artificial intelligence, but also in other areas such as reliability analysis, information security, risk analysis, public opinion polls etc.

References

[Bal86] J.F. Baldwin. Support logic programming. In A.I Jones et al., editors, *Fuzzy Sets: Theory and Applications*. Reidel, 1986.

[BK86] Raj K. Bhatnagar and Laveen N. Kanal. Handling uncertain information: A review of numeric and non-numeric methods. In L.N. Kanal and J.F. Lemmer, editors, *Uncertainty in Artificial Intelligence 1*. North-Holland, 1986.

[Chá96] Tom Chávez. Modelling and measuring the effects of vagueness in decision models. *IEEE Transactions on Systems, Man and Cybernetics - Part A*, 26(3):311–323, 1996.

[Coh86] M.S. Cohen. An expert system framework for non-monotonic reasoning about probabilistic assumptions. In L.N. Kanal and J.F. Lemmer, editors, *Uncertainty in Artificial Intelligence 1*. North-Holland, 1986.

[GS82] P. Gärdenfors and N.-E. Sahlin. Unreliable probabilities, risk taking, and decision making. *Synthese*, 53:361–386, 1982.

[HH88] S.J. Henkind and M.C. Harrison. An analysis of four uncertainty calculi. *IEEE Transactions on Systems, Man and Cybernetics*, 18:700–714, Sept./Oct. 1988.

[Law97] J. Lawry. An inductive method for inexact reasonings. *International Journal of Approximate Reasoning*, 16:205–221, 1997.

[Pea88] J. Pearl. Do we need higher-order probabilities, and if so, what do they mean? In *Proc. of the 1988 Conf. on Uncertainty in Artificial Intelligence*, 1988.

[Sha76] G. Shafer. *A Mathematical Theory of Evidence*. Princeton University Press, 1976.

[Zad84] L.A. Zadeh. Review of Shafer's A Mathematical Theory of Evidence. *AI Magazine*, 5:81–83, 1984.

This article was processed using the LaTeX macro package with LLNCS style