



Legal reasoning with subjective logic

AUDUN JØSANG*

Distributed Systems Technology Centre, Brisbane, Australia

E-mail: ajosang@dstc.edu.au

VIGGO A. BONDI

Norwegian Shipowners' Association, Oslo, Norway

E-mail: viggo.bondi@rederi.no

Abstract. Judges and jurors must make decisions in an environment of ignorance and uncertainty for example by hearing statements of possibly unreliable or dishonest witnesses, assessing possibly doubtful or irrelevant evidence, and enduring attempts by the opponents to manipulate the judge's and the jurors' perceptions and feelings. Three important aspects of decision making in this environment are the quantification of sufficient proof, the weighing of pieces of evidence, and the relevancy of evidence. This paper proposes a mathematical framework for dealing with the two first aspects, namely the quantification of proof and weighing of evidence. Our approach is based on subjective logic, which is an extension of standard logic and probability theory, in which the notion of probability is extended by including degrees of uncertainty. Subjective Logic is a framework for modelling human reasoning and we show how it can be applied to legal reasoning.

Key words: artificial, belief, court, law, probability, reasoning, uncertainty

1. Introduction

Every judicial decision is guided by a standard of proof. In civil cases the decision must be based on *a balance of probabilities*, while criminal cases require the tougher standard of *proof beyond reasonable doubt*. Despite the probabilistic language, these phrases have seldom been made more specific. Eggleston (1978, p. 102) comments that judges appear to avoid explaining or elaborating on these traditional phrases for fear that their attempts to improve them will be found wanting by their peers. Quantifying subjective probabilities related to such statements seems indeed problematic. Simon and Mahan (1971) asked mock jurors to provide numerical probability levels that in their opinion corresponded with the phrase *beyond reasonable doubt* and obtained levels ranging from 0.7 to 0.9. Professional judges required stricter probability levels when asked the same question.

* The work reported in this paper has been funded in part by the Co-operative Research Centre for Enterprise Distributed Systems Technology (DSTC) through the Australian Federal Government's CRC Programme (Department of Industry, Science & Resources)

In the same way as for standards of proof, the legal profession has been suspicious of formalising or quantification of the 'weighing of evidence', Eggleston (1978, pp. 2–3), although many attempts have been made to apply probability theory to legal problems. In fingerprinting for example, experts in various Western countries require from 8 to more than 16 matching characteristics and no unexplained points of difference to risk the claim that two prints have been made by the same person. More often, rules of evidence are qualitative and broad, and attempts to provide a 'scientific' basis for weighing evidence has been met with widespread scepticism, such as for example in connection with the use of the polygraph in the US, in the UK and in other countries in order to decide whether perjury is being committed.

Unlike the criteria for proof and evidence, the legal systems in most Western countries are quite specific about the limits imposed on what kind of evidence may be admitted or heard, and for how long. These limits amount to (ir)relevancy rules which distinguishes the matters that must be attended to by the judge and the jury, from those that may or should be ignored. A piece of evidence is considered relevant primarily if it may be used to prove statements concerning the case. This can for example happen if the evidence increases the credibility of acts the court wishes to (dis)confirm, or if the evidence is inconsistent with another relevant fact. Probativity is the dominant criterion for relevance insofar as it determines whether a piece of evidence has any bearing on the case, but veracity, credibility, adherence to proper procedures for obtaining evidence, and predujciality all define 'taboos' by which material may be banned from the court as inadmissible. For example, a defendant's prior criminal record, although verifiable, and perhaps relevant, may be ruled inadmissible because it would prejudice the jury against the defendant.

These aspects of judicial reasoning are well described by Smithson (1988), and his analysis shows that the legal profession deals with quantification of sufficient proof, weighing of evidence and determining evidence admissibility quite differently. As a general rule, there are concise and explicit relevancy criteria for determining whether a piece of evidence is admissible in court, whereas the methods for weighing of evidence and determining sufficient level of proof are subjective and rather inexplicable.

In the following sections we describe a mathematical framework called Subjective Logic that we will apply for weighing evidence and for making legal decisions in the presence of uncertainty. We do of course not propose to let legal decisions be based on mathematical analysis alone, but rather to let judges and jurors use it as a reference to see in which degree their intuitive decision can be formally and logically supported.

2. Probabilities in Court

An infamous example of the use of probabilities in court was the US 1967 trial of *The People vs. Collins* (1968), in which a number of erroneous appeals to probabil-

ity theory were utilised by the prosecution. Witnesses to a robbery indicated a male Negro with a beard and a female blonde Caucasian with hair worn in a pony tail, escaping in a yellow automobile. The police arrested the Collinses, a couple fitting the description. The prosecution pulled probability estimates of several of the characteristics out of thin air, multiplied them together under the assumption that they were all independent, and obtained a probability of 1 in 12,000,000 of finding one such couple possessing all the characteristics. The conviction was eventually overturned, but this case has often been cited as a warning to members of the legal profession concerning the use of probability theory.

The legal profession itself has criticised probability theory of being inapplicable in legal cases, by objecting to the requirement of quantifying uncertainty into a single numerical value, because for most judges, lawyers and jurors, uncertainty is multifaceted, as expressed for example in Cohen (1977) and James (1941). Eggleston comments that “the legal profession as a whole has been notably suspicious of the learning of mathematicians and actuaries. . .” Eggleston (1978, pp. 2–3), and he claims that lawyers and judges both find the ambiguities and vagueness of judicial standards for evidence and proof strategically useful.

The introduction of evidence from DNA analysis in court has been challenged by the jury’s difficulty of understanding and weighing this type statistical evidence. The correct model for using matching DNA as evidence against an accused is to use Bayes theorem where the probability of a DNA match given innocence (usually in the order of 1 in 100.000.000 or more) is balanced against the prior probability of innocence (i.e., without the DNA evidence) to produce the statistically correct probability of innocence given the DNA match. The human brain has a tendency of succumbing to the *prosecutor’s fallacy* (Thompson and Schumann 1987) in which the prior probability of innocence is ignored and the probability of innocence given a DNA match is assumed equal to the probability of a DNA match given innocence. Despite the fact that the prosecutor’s fallacy can be avoided by the correct application of Bayes theorem the relative complexity of the latter has led courts to dismiss it. It is for example difficult to assess the prior probability of innocence in most cases. In the English case R vs. Adams (1996) the prosecution rested entirely on expert evidence describing matching DNA profiles. Bayes theorem was explained to the jury but the conviction was quashed being regarded as unsafe, and the court took the view that the use of statistics trespassed on an area peculiarly and exclusively within the province of the jury, namely the way they evaluate the relationship between a piece of evidence and another. In the subsequent ruling R vs. Doheny and Adams (1997) the court set out the procedure for the introduction of DNA evidence in criminal trials. The court recognised that expert witnesses should give evidence in terms of the probability of a DNA match given innocence and not the opposite, but did not consider how the distinction between the two probabilities might best be brought to the attention of the jury.

Several other types of calculi and logics which take uncertainty and ignorance 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 Hunter (1996) or Motro and Smets (1997) for an analysis of some uncertainty logics and calculi).

Here we describe *Subjective Logic* (see Jøsang (2001) for an more detailed description) which is a logic of uncertain probabilities, and we show how this can be applied to judicial reasoning. This work builds on Dempster-Shafer belief theory (Shafer 1976) which will be briefly introduced next. Subjective Logic must not be confused with Fuzzy Logic. The latter operates on crisp and certain measures about linguistically vague and fuzzy propositions whereas Subjective Logic operates on uncertain measures about crisp propositions.

3. Uncertain Probabilities

3.1. THE BELIEF MODEL

The first step in applying the Dempster-Shafer belief model (Shafer 1976) is to define a set of possible situations, called the *frame of discernment*, which delimits a set of possible states of a given system. In the following, standard set theory will be used to describe frames of discernment, but the term “*state*” will be used instead of “*set*” because the former is more relevant to the field of application. It is assumed that only one state can be true at any one time. If a state is assumed to be true, then all superstates are considered true as well.

The elementary states in the frame of discernment Θ will be called atomic states because they do not contain substates. The powerset of Θ , denoted by 2^Θ , contains the atomic states and all possible unions of the atomic states, including Θ . A frame of discernment can be finite or infinite, in which cases the corresponding powerset is also finite or infinite respectively.

An observer who believes that some states in the powerset of Θ might be true can assign belief mass to these states. Belief mass on an atomic state $x \in 2^\Theta$ is interpreted as the belief that the state is true. Belief mass on a non-atomic state $x \in 2^\Theta$ is interpreted as the belief that one of the atomic states it contains is true, but that the observer is uncertain about which of them is true. The following definition is central in the Dempster-Shafer theory.

DEFINITION 1 (Belief Mass Assignment). Let Θ be a frame of discernment. If with each substate $x \in 2^\Theta$ a number $m_\Theta(x)$ is associated such that:

1. $m_\Theta(x) \geq 0$
2. $m_\Theta(\emptyset) = 0$
3. $\sum_{x \in 2^\Theta} m_\Theta(x) = 1$

then m_Θ is called a *belief mass assignment*¹ on Θ , or BMA for short. For each substate $x \in 2^\Theta$, the number $m_\Theta(x)$ is called the *belief mass*² of x . \square

¹ called *basic probability assignment* in Shafer (1976).

² called *basic probability number* in Shafer (1976).

A belief mass $m_\Theta(x)$ expresses the belief assigned to the state x and does not express any belief in substates of x in particular.

In contrast to belief mass, the *belief* in a state must be interpreted as an observer's total belief that a particular state is true. The next definition from the Dempster-Shafer theory will make it clear that belief in x not only depends on $m_\Theta(x)$, but also on belief mass assigned to substates of x .

DEFINITION 2 (Belief Function). Let Θ be a frame of discernment, and let m_Θ be a BMA on Θ . Then the *belief function* corresponding with m_Θ is the function $b : 2^\Theta \mapsto [0, 1]$ defined by:

$$b(x) = \sum_{y \subseteq x} m_\Theta(y), \quad x, y \in 2^\Theta$$

□

Similarly to belief, an observer's *disbelief* must be interpreted as the total belief that a state is **not** true. The following definition is ours.

DEFINITION 3 (Disbelief Function). Let Θ be a frame of discernment, and let m_Θ be a BMA on Θ . Then the *disbelief function* corresponding with m_Θ is the function $d : 2^\Theta \mapsto [0, 1]$ defined by:

$$d(x) = \sum_{y \cap x = \emptyset} m_\Theta(y), \quad x, y \in 2^\Theta$$

□

The disbelief of x corresponds to the *doubt* of x in Shafer's book. However, we choose to use the term "disbelief" because we feel that for example the case when it is certain that a state is false can better be described by "absolute disbelief" than by "absolute doubt". Our next definition expresses uncertainty regarding a given state as the sum of belief masses on superstates or on partly overlapping states.

DEFINITION 4 (Uncertainty Function). Let Θ be a frame of discernment, and let m_Θ be a BMA on Θ . Then the *uncertainty function* corresponding with m_Θ is the function $u : 2^\Theta \mapsto [0, 1]$ defined by:

$$u(x) = \sum_{\substack{y \cap x \neq \emptyset \\ y \not\subseteq x}} m_\Theta(y), \quad x, y \in 2^\Theta$$

□

A BMA with zero belief mass assigned to Θ is called a dogmatic BMA. In later sections it is argued that dogmatic BMAs are unnatural in practical situations and strictly speaking can only be defended in idealised hypothetical situations. With the concepts defined so far a simple theorem can be stated. Proofs of theorems can be found in the appendix.

THEOREM 1 (Belief Function Additivity).

$$b(x) + d(x) + u(x) = 1 \quad (1)$$

□

For the purpose of expressing uncertain beliefs about particular states we will show that the relative number of atomic states is also needed in addition to belief functions. For any particular state x the *atomicity* of x is the number of states it contains, denoted by $|x|$. If Θ is a frame of discernment, the atomicity of Θ is equal to the total number of atomic states it contains.

Similarly, if $x, y \in 2^\Theta$ then the overlap between x and y relative to y can be expressed in terms of atomic states. Our next definition captures this idea of relative atomicity:

DEFINITION 5 (Relative Atomicity). Let Θ be a frame of discernment and let $x, y \in 2^\Theta$. Then the *relative atomicity* of x to y is the function $a : 2^\Theta \mapsto [0, 1]$ defined by:

$$a(x/y) = \frac{|x \cap y|}{|y|}, \quad x, y \in 2^\Theta$$

□

It can be observed that $x \cap y = \emptyset$ implies $a(x/y) = 0$, and that $y \subseteq x$ implies $a(x/y) = 1$. In all other cases the relative atomicity will be a value between 0 and 1.

The relative atomicity of an atomic state to its frame of discernment, denoted by $a(x/\Theta)$, can simply be written as $a(x)$. If nothing else is specified, the relative atomicity of a state then refers to the frame of discernment.

A frame of discernment with a corresponding BMA can be used to determine a probability expectation value for any given state. Uncertainty contributes to the probability expectation but will have different weight depending on the relative atomicities. The idea is that when the atomicity of a state is greater, its belief mass is spread out on more substates so that the contribution to each substate is reduced.

DEFINITION 6 (Probability Expectation). Let Θ be a frame of discernment with BMA m_Θ , then the *probability expectation function* corresponding with m_Θ is the function $E : 2^\Theta \mapsto [0, 1]$ defined by:

$$E(x) = \sum_y m_\Theta(y) a(x/y), \quad x, y \in 2^\Theta \quad (2)$$

□

Definition 6 is equivalent with the pignistic probability described in e.g., Smets and Kennes (1994), and is based on the principle of insufficient reason: a belief mass assigned to the union of n atomic states is split equally among these n states.

3.2. THE FOCUSED FRAME OF DISCERNMENT

The focused frame of discernment and the corresponding BMA will for a given state produce the same belief, disbelief and uncertainty functions as the original frame of discernment and BMA.

DEFINITION 7 (Focused Frame of Discernment). Let Θ be a frame of discernment and let $x \in 2^\Theta$. The frame of discernment denoted by $\tilde{\Theta}^x$ containing only x and $\neg x$, where $\neg x$ is the complement of x in Θ is then called a focused frame of discernment with focus on x . \square

When the original frame of discernment Θ contains more than 2 atomic states, the relative atomicity of x in the focused frame of discernment $\tilde{\Theta}^x$ will in general be different from $\frac{1}{2}$ although $\tilde{\Theta}^x$ per definition contains exactly two states.

DEFINITION 8 (Focused Belief Mass Assignment). Let Θ be a frame of discernment with BMA m_Θ and let $b(x)$, $d(x)$ and $u(x)$ be the belief, disbelief and uncertainty functions of x in 2^Θ . Let $\tilde{\Theta}^x$ be the the focused frame of discernment with focus on x . The focused BMA $m_{\tilde{\Theta}^x}$ on $\tilde{\Theta}^x$ is defined according to:

$$\begin{aligned} m_{\tilde{\Theta}^x}(x) &= b(x) \\ m_{\tilde{\Theta}^x}(\neg x) &= d(x) \\ m_{\tilde{\Theta}^x}(\tilde{\Theta}^x) &= u(x) \end{aligned} \quad (3)$$

The focused relative atomicity of x is defined by the following equation:

$$a_{\tilde{\Theta}^x}(x) = [E(x) - b(x)]/u(x) \quad (4)$$

\square

It can be seen that the belief, disbelief and uncertainty functions of x are identical in 2^Θ and $2^{\tilde{\Theta}^x}$. The focused relative atomicity is defined so that the probability expectation value of the state x is equal in Θ and $\tilde{\Theta}^x$, and the expression for $a_{\tilde{\Theta}^x}(x)$ in Definition 8 can be determined by using Definition 6.

The focused relative atomicity is a constructed value which does not correspond to real atomicities. It represents in fact the weighted average of relative atomicities of x to all other states in function of their uncertainty mass. Working with a focused frame of discernment makes it possible to represent the belief functions relative to x using a binary frame of discernment. This is a great advantage when operators on belief functions are introduced in Section 4.

3.3. THE OPINION SPACE

For purpose of having a simple and intuitive representation of uncertain beliefs we will define a 3-dimensional metric called *opinion* but which will contain a 4th redundant parameter in order to allow a compact operator definition.

DEFINITION 9 (Opinion). Let Θ be a binary frame of discernment with 2 atomic states x and $\neg x$, and let m_Θ be a BMA on Θ where $b(x)$, $d(x)$, $u(x)$, and $a(x)$ represent the belief, disbelief, uncertainty and relative atomicity functions on x in Θ respectively. Then the *opinion* about x , denoted by ω_x , is the quadruple defined by:

$$\omega_x \equiv (b(x), d(x), u(x), a(x)) \tag{5}$$

□

For compactness and simplicity of notation we will in the following denote the belief, disbelief, uncertainty and relative atomicity functions as b_x , d_x , u_x and a_x respectively. Also, the probability expectation of opinions can be denoted by $E(\omega_x)$ instead of $E(x)$. By using Definition 6 we can write:

$$E(\omega_x) = b_x + u_x a_x \tag{6}$$

The three coordinates (b_x, d_x, u_x) are dependent through Equation (1) so that one is redundant. As such they represent nothing more than the traditional (*Belief, Plausibility*) pair of Shaferian belief theory. However, it is useful to keep all three coordinates in order to obtain simple expressions when introducing operators on opinions in Section 4.

Equation (1) defines a triangle that can be used to graphically illustrate opinions as shown in Figure 1. As an example the position of the opinion $\omega_x = (0.40, 0.10, 0.50, 0.60)$ is indicated as a point in the triangle. Also shown are the probability expectation value and the relative atomicity.

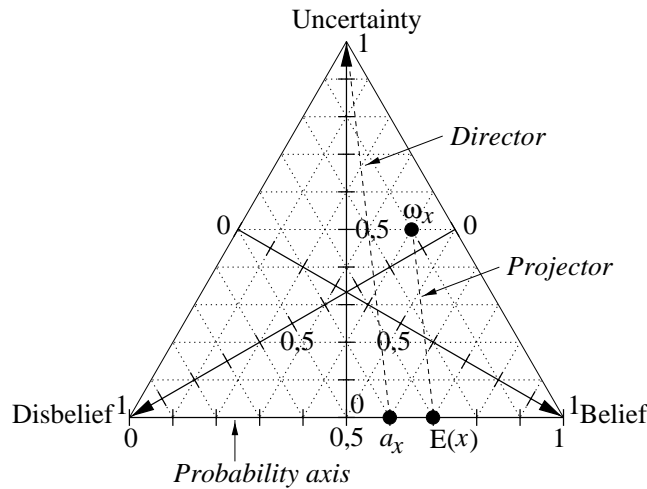


Figure 1. Opinion triangle with ω_x as example

The horizontal bottom line between the belief and disbelief corners in Figure 1 is called the *probability axis*. The relative atomicity can be graphically represented as a point on the probability axis. The line joining the top corner of the triangle

and the relative atomicity point becomes the *director*. In Figure 1 $a_x = 0.60$ is represented as a point, and the dashed line pointing at it represents the director.

The *projector* is parallel to the director and passes through the opinion point. Its intersection with the probability axis defines the probability expectation value which otherwise can be computed by the formula of Definition 6. The position of the probability expectation $E(x) = 0.70$ is shown. A hypothetical frame of discernment with infinite atomicity will make the relative atomicity most opinions equal to 0 or 1, producing a projector parallel to either the left or the right edge of the triangle.

Opinions situated on the probability axis are called *dogmatic opinions*. They represent situations without uncertainty and correspond to traditional probabilities. The distance between an opinion point and the probability axis can be interpreted as the degree of uncertainty.

Opinions situated in the left or right corner, i.e. with either $b = 1$ or $d = 1$ are called *absolute opinions*. They represent situations where it is absolutely certain that a state is either true or false, and correspond to TRUE or FALSE proposition in binary logic.

With the definitions established so far we are able to derive the fundamental Kolmogorov axioms of traditional probability theory as a theorem.

THEOREM 2 (Kolmogorov Axioms). Given a frame of discernment Θ with a BMA m_Θ , the probability expectation function E with domain 2^Θ satisfies:

1. $E(x) \geq 0$ for all $x \in 2^\Theta$
2. $E(\Theta) = 1$
3. If $x_1, x_2, \dots \in 2^\Theta$ are pairwise disjoint
then $E(\cup_{i=1}^{|2^\Theta|} x_i) = \sum_{i=1}^{|2^\Theta|} E(x_i)$

□

Opinions can be ordered according to probability expectation value, but additional criteria are needed in case of equal probability expectation values. The following definition determines the order of opinions:

DEFINITION 10 (Ordering of Opinions). Let ω_x and ω_y be two opinions. They can be ordered according to the following criteria by priority:

1. The greatest probability expectation gives the greatest opinion.
2. The least uncertainty gives the greatest opinion.
3. The least relative atomicity gives the greatest opinion.

□

The first criterion is self evident, and the second less so, but it is supported by experimental findings described by Ellsberg (1961). The third criterion is more an intuitive guess and so is the priority between the second and third criteria, and

before these assumptions can be supported by evidence from practical experiments we invite the readers to judge whether they agree.

4. Logical Operators

So far we have described the elements of a frame of discernment as states. In practice states will verbally be described as propositions; if for example Θ consists of possible colours of a ball when drawn from an urn with red and black balls, and x designates the state when the colour drawn from the urn is red then it can be interpreted as the verbal proposition x : ‘I will draw a red ball’.

Standard binary logic operates on binary propositions that can take the values TRUE or FALSE. *Subjective Logic* operates on opinions about binary propositions, i.e. opinions about propositions that are assumed to be either TRUE or FALSE. In this section we describe the traditional logical operators ‘AND’, ‘OR’ and ‘NOT’ applied to opinions, and it will become evident that binary logic is a special case of Subjective Logic for these operators.

Opinions are considered individual, and will therefore have an ownership assigned whenever relevant. In our notation, superscripts indicate ownership, and subscripts indicate the proposition to which the opinion applies. For example ω_x^A is an opinion held by agent A about the truth of proposition x .

4.1. PROPOSITIONAL CONJUNCTION AND DISJUNCTION

Forming an opinion about the conjunction of two propositions from different frames of discernment consists of determining from the opinions about each proposition a new opinion reflecting the truth of both propositions simultaneously. This corresponds to ‘AND’ in binary logic.

THEOREM 3 (Propositional Conjunction). Let Θ_X and Θ_Y be two distinct binary frames of discernment and let x and y be propositions about states in Θ_X and Θ_Y respectively. Let $\omega_x = (b_x, d_x, u_x, a_x)$ and $\omega_y = (b_y, d_y, u_y, a_y)$ be an agent’s opinions about x and y respectively. Let $\omega_{x \wedge y} = (b_{x \wedge y}, d_{x \wedge y}, u_{x \wedge y}, a_{x \wedge y})$ be the opinion such that

1. $b_{x \wedge y} = b_x b_y$
2. $d_{x \wedge y} = d_x + d_y - d_x d_y$
3. $u_{x \wedge y} = b_x u_y + u_x b_y + u_x u_y$
4. $a_{x \wedge y} = \frac{b_x u_y a_y + u_x a_x b_y + u_x a_x u_y a_y}{b_x u_y + u_x b_y + u_x u_y}$

Then $\omega_{x \wedge y}$ is called the propositional conjunction of ω_x and ω_y , representing the agent’s opinion about both x and y being true. The symbol ‘ \wedge ’ can designate this operator by defining $\omega_{x \wedge y} \equiv \omega_x \wedge \omega_y$. \square

Forming an opinion about the disjunction of two propositions from different frames of discernment consists of determining from the opinions about each proposition a new opinion reflecting the truth of one or the other or both propositions. This corresponds to ‘OR’ in binary logic.

THEOREM 4 (Propositional Disjunction). Let Θ_X and Θ_Y be two distinct binary frames of discernment and let x and y be propositions about states in Θ_X and Θ_Y respectively. Let $\omega_x = (b_x, d_x, u_x, a_x)$ and $\omega_y = (b_y, d_y, u_y, a_y)$ be an agent’s opinions about x and y respectively. Let $\omega_{x \vee y} = (b_{x \vee y}, d_{x \vee y}, u_{x \vee y}, a_{x \vee y})$ be the opinion such that:

1. $b_{x \vee y} = b_x + b_y - b_x b_y$
2. $d_{x \vee y} = d_x d_y$
3. $u_{x \vee y} = d_x u_y + u_x d_y + u_x u_y$
4. $a_{x \vee y} = \frac{u_x a_x + u_y a_y - b_x u_y a_y - u_x a_x b_y - u_x a_x u_y a_y}{u_x + u_y - b_x u_y - u_x b_y - u_x u_y}$

Then $\omega_{x \vee y}$ is called the propositional disjunction of ω_x and ω_y , representing the agent’s opinion about x or y or both being true. The symbol “ \vee ” can designate this operator by defining $\omega_{x \vee y} \equiv \omega_x \vee \omega_y$. \square

As would be expected, propositional conjunction and disjunction of opinions are both commutative and associative. Idempotence is not defined because that would assume that the arguments are identical and therefore belong to the same frame of discernment. It must be assumed that the opinions are independent and that the propositions to which they apply belong to distinct frames of discernment.

Propositional conjunction and disjunction are equivalent to the ‘AND’ and ‘OR’ operators of Baldwin’s support logic (Baldwin 1986) except for the relative atomicity parameter which is absent in Baldwin’s logic. When applied to absolute opinions, i.e., with either $b = 1$ or $d = 1$, propositional conjunction and disjunction are equivalent to ‘AND’ and ‘OR’ of binary logic, that is; they produce the truth tables of logical ‘AND’ and ‘OR’ respectively. When applied to dogmatic opinions, i.e opinions with zero uncertainty, they produce the same results as the product and co-product of probabilities respectively.

It can be observed that for dogmatic opinions the denominator becomes zero in the expressions for the relative atomicity in Theorems 3 and 4. However, the limits do exist and can be computed in such cases. See also comment about dogmatic opinions in Section 4.4 below.

Propositional conjunction and disjunction of opinions are not distributive on each other. If for example ω_x , ω_y and ω_z are independent opinions we have:

$$\omega_x \wedge (\omega_y \vee \omega_z) \neq (\omega_x \wedge \omega_y) \vee (\omega_x \wedge \omega_z) \quad (7)$$

This result which may seem surprising is due to the fact that that ω_x appears twice in the expression on the right side so that it in fact represents the propositional

disjunction of partially dependent arguments. Only the expression on the left side is thus correct.

Propositional conjunction decreases the relative atomicity whereas propositional disjunction increases it. What really happens is that the product of the two frames of discernment produces a new frame of discernment with atomicity equal to the product of the respective atomicities. However, as opinions only apply to binary frames of discernment, a new frame of discernment with corresponding relative atomicity must be derived both for propositional conjunction and propositional disjunction. The expressions for relative atomicity in Theorems 3 and 4 are in fact obtained by forming the product of the two frames of discernment and applying Equation (3) and Definition 6 .

In order to show that Subjective Logic is compatible with probability calculus regarding multiplication and co-multiplication of probabilities we state the following theorem.

THEOREM 5 (Product and Co-product). Let Θ_X and Θ_Y be two distinct binary frames of discernment and let x and y be propositions about states in Θ_X and Θ_Y respectively. Let $\omega_x = (b_x, d_x, u_x, a_x)$ and $\omega_y = (b_y, d_y, u_y, a_y)$ be an agent's respective opinions about the propositions x and y , and let $\omega_{x \wedge y} = (b_{x \wedge y}, d_{x \wedge y}, u_{x \wedge y}, a_{x \wedge y})$ and $\omega_{x \vee y} = (b_{x \vee y}, d_{x \vee y}, u_{x \vee y}, a_{x \vee y})$ be the propositional conjunction and disjunction of ω_x and ω_y respectively. The probability expectation function E satisfies:

1. $E(\omega_{x \wedge y}) = E(\omega_x)E(\omega_y)$
2. $E(\omega_{x \vee y}) = E(\omega_x) + E(\omega_y) - E(\omega_x)E(\omega_y)$

□

4.2. NEGATION

The negation of an opinion about proposition x represents the agent's opinion about x being false. This corresponds to 'NOT' in binary logic.

THEOREM 6 (Negation). Let $\omega_x = (b_x, d_x, u_x, a_x)$ be an opinion about the proposition x . Then $\omega_{\neg x} = (b_{\neg x}, d_{\neg x}, u_{\neg x}, a_{\neg x})$ is the negation of ω_x where:

1. $b_{\neg x} = d_x$
2. $d_{\neg x} = b_x$
3. $u_{\neg x} = u_x$
4. $a_{\neg x} = 1 - a_x$

The symbol ' \neg ' can designate this operator by defining $\neg\omega_x \equiv \omega_{\neg x}$. □

Negation can be applied to expressions containing propositional conjunction and disjunction, and it can be shown that De Morgans laws are valid.

4.3. DISCOUNTING

Assume two agents A and B where A has an opinion about B in the form of the proposition: ' B is knowledgeable and will tell the truth'. In addition B has an opinion about a proposition x . Agent A can then form an opinion about x by discounting B 's opinion about x with A 's opinion about B . There is no such thing as physical belief discounting, and discounting 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. This we will interpret as if A thinks that B is uncertain about the truth value of x so that A also is uncertain about the truth value of x no matter what B 's actual advice is. Our next definition captures this idea.

DEFINITION 11 (Discounting). Let A and B be two agents where $\omega_B^A = (b_B^A, d_B^A, u_B^A, a_B^A)$ is A 's opinion about B 's advice, and let x be a proposition where $\omega_x^B = (b_x^B, d_x^B, u_x^B, a_x^B)$ is B 's opinion about x expressed in an advice to A . Let $\omega_x^{AB} = (b_x^{AB}, d_x^{AB}, u_x^{AB}, a_x^{AB})$ be the opinion such that:

1. $b_x^{AB} = b_B^A b_x^B$,
2. $d_x^{AB} = b_B^A d_x^B$
3. $u_x^{AB} = d_B^A + u_B^A + b_B^A u_x^B$
3. $a_x^{AB} = a_x^B$

then ω_x^{AB} is called the discounting of ω_x^B by ω_B^A expressing A 's opinion about x as a result of B 's recommendation about x to A . By using the symbol ' \otimes ' to designate this operator, we define $\omega_x^{AB} \equiv \omega_B^A \otimes \omega_x^B$. \square

It is easy to prove that \otimes is associative but not commutative. This means that in case of chains of two or more recommenders the computation can start in either end of the chain, but that the order of opinions is significant. Opinion independence must be assumed, which for example translates into not allowing the same entity to appear more than once in a chain.

4.4. INDEPENDENT CONSENSUS

The consensus opinion of two opinions is an opinion that reflects both opinions in a fair and equal way. For example if two agents have observed an unreliable machine over two different time intervals they might have different opinions about its reliability depending on the behaviour of the machine in the respective periods. The consensus opinion must then be the opinion that a single agent would have after having observed the machine during both periods.

DEFINITION 12 (Independent Consensus). Let $\omega_x^A = (b_x^A, d_x^A, u_x^A, a_x^A)$ and $\omega_x^B = (b_x^B, d_x^B, u_x^B, a_x^B)$ be opinions respectively held by agents A and B about

the same proposition x . Let $\omega_x^{A,B} = (b_x^{A,B}, d_x^{A,B}, u_x^{A,B}, a_x^{A,B})$ be the opinion such that:

1. $b_x^{A,B} = (b_x^A u_x^B + b_x^B u_x^A) / \kappa$
2. $d_x^{A,B} = (d_x^A u_x^B + d_x^B u_x^A) / \kappa$
3. $u_x^{A,B} = (u_x^A u_x^B) / \kappa$
4. $a_x^{A,B} = \frac{a_x^B u_x^A + a_x^A u_x^B - (a_x^A + a_x^B) u_x^A u_x^B}{u_x^A + u_x^B - 2u_x^A u_x^B}$

where $\kappa = u_x^A + u_x^B - u_x^A u_x^B$ such that $\kappa \neq 0$, and $a_x^{A,B} = (a_x^A + a_x^B) / 2$ when $u_x^A, u_x^B = 1$. Then $\omega_x^{A,B}$ is called the consensus between ω_x^A and ω_x^B , representing an imaginary agent $[A, B]$'s opinion about x , as if she represented both A and B . By using the symbol ' \oplus ' to designate this operator, we define $\omega_x^{A,B} \equiv \omega_x^A \oplus \omega_x^B$. \square

Justification for the consensus operator can be found in (Jøsang 2001). It is easy to show that \oplus is both commutative and associative which means that the order in which opinions are combined has no importance. Opinion independence must be assumed, which obviously translates into not allowing an agent's opinion to be counted more than once

The effect of the consensus operator is to reduce the uncertainty. For example the case where several witnesses give consistent testimony should amplify the judge's opinion, and that is exactly what the operator does. Consensus between an infinite number of independent non-dogmatic opinions would necessarily produce a consensus opinion with zero uncertainty.

Two dogmatic opinions can not be combined according to Definition 12. This can be explained by interpreting uncertainty as *room for influence*, meaning that it is only possible to reach consensus with somebody who maintains some uncertainty. A situation with conflicting dogmatic opinions is philosophically counterintuitive, primarily because opinions about real situations can never be certain, and secondly, because if they were they would necessarily be equal. The consensus of two absolutely uncertain opinions results in a new absolutely uncertain opinion, although the relative atomicity is not well defined. The limit of the relative atomicity when both $u_x^A, u_x^B \rightarrow 1$ is $(a_x^A + a_x^B) / 2$, i.e. the average of the two relative atomicities, which intuitively makes sense.

The consensus operator will normally be used in combination with the discounting operator, so that if dogmatic opinions are recommended, the recipient should not have absolute trust in the recommending party and thereby introduce uncertainty before combining the advice by the consensus operator. This is illustrated by Example A in Section 6.1 below.

The consensus operator has the same purpose as Dempster's rule (Shafer 1976), but is quite different from it. Dempster's rule has been criticised for producing counterintuitive results (see e.g., Zadeh 1984; Cohen 1986), and (Jøsang 2001) compares our consensus operator with Dempster's rule.

4.5. DEPENDENT CONSENSUS

Assume two agents A and B having observed the same evidence. Because their observations are identical, their respective opinions will necessarily be dependent, and a consensus according to Definition 12 would be meaningless.

If the two observers have made exactly the same observations and their opinions are equal, it would be sufficient to take only one of the opinions into account. However, although two observers witness the same phenomenon, it is possible that they record and interpret it differently. Some observers may temporarily lose concentration when a piece of evidence is presented and thereby miss or misinterpret it, resulting in different, but still dependent opinions. We will define a consensus operator for dependent opinions based on the average of positive and negative evidence supporting the opinions. By “evidence” we here mean the statistical evidence that would produce a given opinion as explained in Section 5.2 below.

DEFINITION 13 (Dependent Consensus). Let $\omega_x^{A_i} = (b_x^{A_i}, d_x^{A_i}, u_x^{A_i}, a_x^{A_i})$ where $i \in [1, n]$, be n dependent opinions respectively held by agents A_i about the same proposition x . The relative atomicities $a_x^{A_i}$ will normally be equal but for generality we provide an expression that equals the average value. Let $\overline{\omega_x^{A_1, \dots, A_n}} = (b_x^{\overline{A_1, \dots, A_n}}, d_x^{\overline{A_1, \dots, A_n}}, u_x^{\overline{A_1, \dots, A_n}}, a_x^{\overline{A_1, \dots, A_n}})$ be the opinion such that:

1. $b_x^{\overline{A_1, \dots, A_n}} = \frac{\sum_1^n (b_x^{A_i} / u_x^{A_i})}{\sum_1^n (b_x^{A_i} / u_x^{A_i}) + \sum_1^n (d_x^{A_i} / u_x^{A_i}) + n}$
2. $d_x^{\overline{A_1, \dots, A_n}} = \frac{\sum_1^n (d_x^{A_i} / u_x^{A_i})}{\sum_1^n (b_x^{A_i} / u_x^{A_i}) + \sum_1^n (d_x^{A_i} / u_x^{A_i}) + n}$
3. $u_x^{\overline{A_1, \dots, A_n}} = \frac{n}{\sum_1^n (b_x^{A_i} / u_x^{A_i}) + \sum_1^n (d_x^{A_i} / u_x^{A_i}) + n}$
4. $a_x^{\overline{A_1, \dots, A_n}} = \frac{\sum_1^n a_x^{A_i}}{n}$

where all the $u_x^{A_i}$ are different from zero. Then $\overline{\omega_x^{A_1, \dots, A_n}}$ is called the dependent consensus between all the $\omega_x^{A_i}$. By using the symbol $\overline{\oplus}$ to designate this operator, we define $\overline{\omega_x^{A_1, \dots, A_n}} \equiv \omega_x^{A_1} \overline{\oplus} \dots \overline{\oplus} \omega_x^{A_n}$. \square

It is easy to show that $\overline{\oplus}$ is both commutative and associative which means that the order in which opinions are combined has no importance.

The effect of the dependent consensus operator is to produce an opinion which is based on an average of observed positive and an average of observed negative evidence. Opinions based on little observed evidence will carry less weight than those with a large evidence basis. If the operator instead was defined to take the average of the (b, d, u) components, uncertain opinions would carry as much weight as certain opinions, but we feel that such an operator would be counterintuitive. Jurors will for example have dependent opinions, because they have received the

same evidence. However, jurors who for example have had problems concentrating during court proceedings will have more uncertain opinions regarding guilt than jurors who have followed the proceedings with interest and concentration. It is then natural that the more uncertain opinions carry less weight than the more certain opinions. A more detailed justification for the operator can be found in Jøsang and Knapskog (1998).

Opinions without uncertainty can not be combined according to Definition 13 because the (b, d, u, a) components would be undefined according to Th.13. A situation with conflicting opinions without uncertainty is intuitively meaningless, primarily because no such opinion really exists, and secondly, because if it did, then they would necessarily be equal. Hypothetically seen, if only one of the opinions is without uncertainty, then it would dictate the consensus.

4.6. MIXING DISCOUNTING AND CONSENSUS

It is possible that several chains of discounted advice produce opinions about the same proposition. Under the condition of opinion independence, these opinions can be combined with the consensus rule to produce a single opinion about the target proposition. An example of mixed consensus and discounting is illustrated in Figure 2.

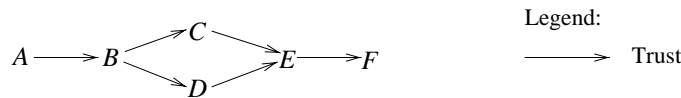


Figure 2. Mixing discounting and consensus

The discounting rule is not distributive relative to the consensus rule because it would violate the independence requirement. To see this let $\omega_B^A, \omega_C^B, \omega_D^B, \omega_E^C, \omega_E^D$ and ω_F^E represent the opinion relationships in Figure 2. We then have

$$\omega_B^A \otimes ((\omega_C^B \otimes \omega_E^C) \oplus (\omega_D^B \otimes \omega_E^D)) \otimes \omega_F^E \neq (\omega_B^A \otimes \omega_C^B \otimes \omega_E^C \otimes \omega_F^E) \oplus (\omega_B^A \otimes \omega_D^B \otimes \omega_E^D \otimes \omega_F^E) \tag{8}$$

which according to the notation in Definition 12 and Definition 11 can be written as

$$\omega_F^{AB(C,D)E} \neq \omega_F^{ABCE, ABDE} \tag{9}$$

The not-equal sign may seem surprising, but the right sides of (8) and (9) violate the requirement of independent opinions because both ω_B^A and ω_F^E appear twice. Only the left sides of (8) and (9) represent the graph of Figure 2 correctly.

There will always be cases which can not be analysed directly. Figure 3 illustrates a situation where agent A needs to determine her trust in agent F, of which she only has second-hand evidence trough a network of agents.

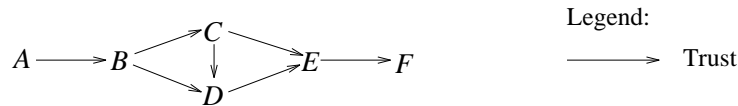


Figure 3. Network which can not be analysed

Whether some advice are ignored, and thereby leaving out some of the evidence, or all advice are included, and thereby violating the independence requirement, the result will never be as correct as one could wish. We will leave this problem open and simply mention that standard probability theory also does not give a clear answer in such situations.

5. How to Determine Opinions

The major difficulty with applying Subjective Logic is to find a way to consistently determine opinions to be used as input parameters. People may find the opinion model unfamiliar and difficult to relate to, and different individuals may produce conflicting opinions when faced with the same evidence.

5.1. DETERMINING OPINIONS USING STATISTICAL EVIDENCE

In Jøsang (2001) it is describes how opinions can be formed based on statistical evidence. Assume a process that can produce positive or negative outcomes. When the process previously has produced r positive and s negative outcomes the opinion that any random outcome is positive can be expressed as $\omega = (b, d, u, a)$ where:

$$\begin{aligned}
 b &= r/(r + s + 2) \\
 d &= s/(r + s + 2) \\
 u &= 2/(r + s + 2) \\
 a &= \text{relative atomicity of event}
 \end{aligned}
 \tag{10}$$

The numbers of positive and negative outcomes, r and s , represents the statistical evidence supporting the opinion. Unfortunately, this method can only be applied in idealised cases such as for example when picking red and black balls from an urn. In most real life situations, the evidence at hand can only be analysed intuitively, and this will be discussed next.

5.2. USING GUIDELINES FOR DETERMINING OPINIONS

In this section we will attempt to formulate a questionnaire for guiding people in expressing their beliefs as opinions. The idea behind the questionnaire is to make the subject consider each component of her opinion separately.

An opinion is always about a proposition, so the first task when trying to determine an opinion intuitively is to express the proposition clearly. The subject should

be informed that it is assumed that nobody can be absolutely sure about anything, so that opinions with absolute belief ($b = 1$) or disbelief ($d = 1$) should never be specified. The questionnaire below will help subjects isolate the components of their opinions in the form of belief, disbelief and uncertainty.

5.2.1. Questionnaire for Determining Subjective Opinions

1. Is the proposition clearly expressed?
Yes: \rightarrow (2)
No: Do it, and start again. \rightarrow (1)
2. Is there any evidence, or do you have an intuitive feeling in favour of or against the proposition?
Yes: \rightarrow (3)
No: You are totally uncertain. $b := 0, d := 0, u := 1. \rightarrow$ (7)
3. How conclusive is this evidence or how strong is this feeling?
Give a value $0 \leq x \leq 1. \rightarrow$ (4)
4. How strong is the evidence or the intuitive feeling against the proposition?
Give a value $0 \leq y \leq 1. \rightarrow$ (5)
5. How strong is the evidence or the intuitive feeling in favour of the proposition?
Give a value $0 \leq z \leq 1. \rightarrow$ (6)
6. Normalisation of results:

$$\begin{aligned} b &:= \frac{z}{z+y+(1-x)} \\ d &:= \frac{y}{z+y+(1-x)} \\ u &:= \frac{1-x}{z+y+(1-x)} \quad \rightarrow (7) \end{aligned}$$

7. What is the total number of states in the event space (e.g. if you are picking coloured balls from an urn containing balls of n different colours then there are n states)?
 $n :=$ total number of states \rightarrow (8)
8. How many states does the proposition cover (e.g. of you are interested in picking any out of m different colours where $m \leq n$ then the proposition covers m states)?
 $m :=$ number of states covered by the propositions, \rightarrow (9)
9. $a := \frac{m}{n}, \rightarrow$ (10)
10. $\omega = (b, d, u, a)$ is the subjective opinion.

The relative atomicity of propositions about statements that are either true or false is always 0.5, and in other cases the relative atomicity can usually be specified when giving the proposition. Steps 7, 8 and 9 can therefore be skipped in most cases.

6. Examples of Applying Subjective Logic

Legal reasoning in criminal law cases consists of considering evidence in favour of or against claims expressed either by the defence or by the prosecution, on the background of the relevant legislation. The judge or the jury then has to decide whether the claims are true or false, and ultimately whether the accused is guilty or not according to the accusation and the relevant legislation.

The truth value of this type of claims, and therefore the frame of discernment to which the proposition belongs, can be characterised as binary, because it is assumed that the claims are either true or false, and not something in between. In the same manner, the accused is assumed to be innocent or guilty, and not for example half guilty or half innocent.

Although it is assumed that a statement made in court is true or false, it is impossible to be absolutely certain about its truth value, and the judge or the jury can only ever have an *opinion* about it. Subjective Logic seems suitable to this type of reasoning.

Many judicial systems allow the severity of guilt to be graded, for example by expressing whether the crime was premeditated or not, but our examples will only cover the purely binary situation.

6.1. EXAMPLE A: ASSESSMENT OF TESTIMONY FROM WITNESSES

The evidence can be presented in a multitude of forms, but usually the judge or the jury will only be presented with second-hand evidence. That is, they will rarely be able to assess the physical evidence themselves, but have to accept the statement about the first-hand evidence from witnesses. In this case, the judge or the jury members will have to determine the witnesses' credibility, and take the statements from the witnesses as recommendations about the first-hand evidence they have observed. The discounting operator can be used to model this process.

In this example we consider a court case where 3 witnesses W_1 , W_2 and W_3 are giving testimony to express their opinions about a verbal proposition x which has been made about the accused. We assume that the verbal proposition is either true or false, and let each witness express his or her opinion about the truth of the proposition as an opinion ω_x^W , to the courtroom. The judge J then has to determine her own opinion about x as a function of her trust ω_W^J : 'Witness W is reliable and will tell the truth' in each individual witness. This situation is illustrated in Figure 4 where the arrows denote trust or opinions about truth.

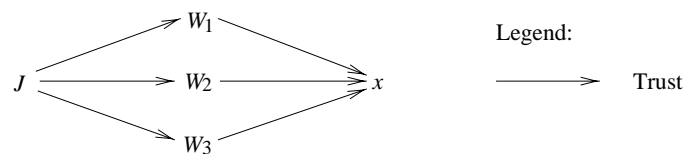


Figure 4. Trust in testimony from witnesses

The effect of each individual testimony on the judge can be computed using the discounting operator, so that for example W_1 's testimony causes the judge to have the opinion

$$\omega_x^{JW_1} = \omega_{W_1}^J \otimes \omega_x^{W_1}$$

about the truth of x . Assuming that the opinions resulting from each witness are independent, they can finally be combined using the consensus operator to produce the judge's own opinion about x :

$$\omega_x^{J(W_1, W_2, W_3)} = (\omega_{W_1}^J \otimes \omega_x^{W_1}) \oplus (\omega_{W_2}^J \otimes \omega_x^{W_2}) \oplus (\omega_{W_3}^J \otimes \omega_x^{W_3}) \quad (11)$$

As a numerical example, let J 's opinion about the witnesses, and the witnesses' opinions about the truth of proposition x be given by:

$$\omega_{W_1}^J = (0.90, 0.00, 0.10, 0.50) \quad \omega_x^{W_1} = (0.90, 0.00, 0.10, 0.50)$$

$$\omega_{W_2}^J = (0.00, 0.90, 0.10, 0.50) \quad \omega_x^{W_2} = (0.90, 0.00, 0.10, 0.50)$$

$$\omega_{W_3}^J = (0.10, 0.00, 0.90, 0.50) \quad \omega_x^{W_3} = (0.90, 0.00, 0.10, 0.50)$$

It can be seen that the judge has a high degree of trust in W_1 , that she distrusts W_2 , and that her opinion about W_3 is highly uncertain, meaning that she is very uncertain about whether W_3 should be trusted or not.

The judge's separate opinions about the proposition x as a function of the advice from each witness then become:

$$\omega_x^{JW_1} = (0.81, 0.00, 0.19, 0.50)$$

$$\omega_x^{JW_2} = (0.00, 0.00, 1.00, 0.50)$$

$$\omega_x^{JW_3} = (0.09, 0.00, 0.91, 0.50)$$

It can be seen that $\omega_x^{JW_2}$ is totally uncertain due to the fact that the judge distrusts W_2 , and that $\omega_x^{JW_3}$ is highly uncertain because the judge is very uncertain about testimonies from W_3 . Only $\omega_x^{JW_1}$ represents an opinion that can be useful for making a decision. By combining all three independent opinions into one the judge gets:

$$\omega_x^{J(W_1, W_2, W_3)} = (0.8135, 0.0000, 0.1865, 0.5000)$$

Obviously the combined opinion is mainly based on the advice from W_1 .

6.2. EXAMPLE B: QUANTIFICATION OF SUFFICIENT PROOF

In case a judge decides on the verdict alone, she alone must weigh the presented evidence and from this determine an opinion. In case a jury decides on the verdict,

the members must somehow reach a consensus. In some judicial systems, absolute consensus is required, whereas others only need a majority of votes to pronounce the accused as guilty. This example introduces a new principle, whereby an opinion about guilt is determined with the consensus operator. The consensus opinion is finally compared with a predetermined threshold value to decide whether the accused shall be pronounced guilty.

In this example we consider a jury which has to decide whether the accused in a criminal case is guilty or not. For simplicity, the jury consists of only the three jurors J_1, J_2 and J_3 . They each have an individual opinion $\omega_g^{J_i}$ about whether the accused is guilty according to the charge. For the jury to reach a consensus, the opinions can be combined using the consensus operator. The jury must be considered as an imaginary agent $[J_1, J_2, J_3]$ consisting of all the jury members combined. This is illustrated in Figure 5.

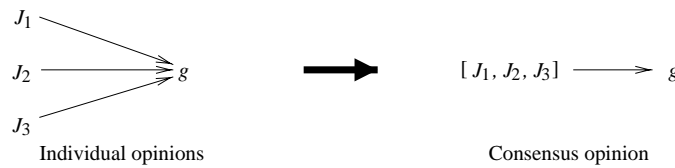


Figure 5. Consensus in a jury

The jurors’ opinions will necessarily be dependent, since they have been listening to the court proceedings together and therefore have received the same evidence. If the operator for independent consensus was used, a larger jury would produce a stronger opinion of guilt (or innocence), simply because the number of jury members is large. The operator for dependent consensus must therefore be used. This has the effect of producing an average opinion, where the most certain opinions carry the most weight. The opinion of the jury can then be expressed as:

$$\omega_g^{\overline{J_1, J_2, J_3}} = \omega_g^{J_1} \oplus \omega_g^{J_2} \oplus \omega_g^{J_3} \tag{12}$$

The final part of the analysis is to compare the jury’s opinion with threshold value in order to determine whether the opinion about guilt is sufficiently strong to pronounce the accused as guilty. This can be done by using the ordering operator, but first, a threshold value must be determined.

The topic of fixing a threshold value for guilt is likely to cause considerable controversy. Our proposal is to let the threshold value be determined by the acceptable rate of wrongful convictions. Some will say that wrongful convictions can never be accepted, but then we claim that convictions can never be made, except perhaps when the accused pleads guilty.

The question of acceptable wrongful conviction within a judicial system could be partly politically and partly judicially determined. As an example we will let the threshold opinion about guilt be fixed as $\omega_{\text{guilty}}^T = (0.999, 0.001, 0.000, 0.500)$, meaning that it is acceptable that 1 in 1000 convictions is wrongful.

Assume further that the jurors have valued their individual opinions about guilt to be:

$$\begin{aligned}\omega_g^{J_1} &= (0.9993, 0.0005, 0.0002, 0.5) \\ \omega_g^{J_2} &= (0.9985, 0.0010, 0.0005, 0.5) \\ \omega_g^{J_3} &= (0.9990, 0.0005, 0.0005, 0.5)\end{aligned}$$

By using these values in (12), the jury's opinion about guilt can be computed as:

$$\omega_g^{\overline{J_1, J_2, J_3}} = (0.99906, 0.00061, 0.00033, 0.5)$$

It can be observed that $\omega_{\text{guilty}}^T < \omega_g^{\overline{J_1, J_2, J_3}}$, meaning that the jury's opinion about guilt is sufficiently strong to pronounce the accused as guilty.

6.3. EXAMPLE C: THE PEOPLE VS. COLLINS REVISITED

In this example we will apply Subjective Logic to the case of *The People vs. Collins* described in Section 2 and show that by taking a subjective approach, the results are much less conclusive. To repeat the case, witnesses to a robbery indicated a male Negro with a beard and a female blonde Caucasian with hair worn in a pony tail, escaping in a yellow automobile. The probability of finding a couple possessing these characteristics was estimated to 1/12,000,000. The Collins couple apparently matched the characteristics, and were therefore thought to have committed the robbery with overwhelming certainty.

Instead of accepting the testimonies as absolutely reliable (which they never are) we will assume that the witnesses are invited to numerically express opinions about each statement they make and that the judge expresses her trust in the witnesses in the same way. The statements made by the witnesses can be expressed as:

- x_1 : *The robbery was committed by a couple consisting of a male Negro and a female Caucasian*
- x_2 : *The male Negro had a beard*
- x_3 : *The female Caucasian wore her hair in a pony tail*
- x_4 : *The robbers escaped in a yellow automobile*

According to the argument put forward by the prosecution, the Collinses are guilty if they fit the description. This implicitly means that they can only be guilty according to the same argument if the description is correct. Logically, the description of the robbers is correct only if all the statements are true. To simplify our example, we will assume that there are only two witnesses W_1 and W_2 , and that their opinions about the respective statements are:

Assume further that the judge J 's trust in W_1 and W_2 is given by $\omega_{W_1}^J = (0.82, 0.03, 0.15, 0.5)$ and $\omega_{W_2}^J = (0.80, 0.05, 0.15, 0.5)$ respectively.

Table I. The witnesses' opinions about the statements

	W_1 's opinions:	W_2 's opinions:
x_1 :	(0.95, 0.00, 0.05, 0.50)	(0.97, 0.00, 0.03, 0.50)
x_2 :	(0.90, 0.00, 0.10, 0.50)	(0.92, 0.00, 0.08, 0.50)
x_3 :	(0.90, 0.00, 0.10, 0.50)	(0.92, 0.00, 0.08, 0.50)
x_4 :	(0.96, 0.00, 0.04, 0.50)	(0.98, 0.00, 0.02, 0.50)

For witness W_1 , the opinion about the correctness of the description can be expressed as:

$$\omega_{x_1 \wedge x_2 \wedge x_3 \wedge x_4}^{W_1} = \omega_{x_1}^{W_1} \wedge \omega_{x_2}^{W_1} \wedge \omega_{x_3}^{W_1} \wedge \omega_{x_4}^{W_1}$$

and similarly for witness W_2

The judge must discount the witnesses' advice by her opinions about the witnesses and use the consensus operator in order to determine her own opinion about the correctness of the robbers' description. This can be expressed and computed as:

$$\omega_{x_1 \wedge x_2 \wedge x_3 \wedge x_4}^{J(W_1, W_2)} = (0.77, 0.00, 0.23, 0.48)$$

If the Collinses really fit the description, they can only be guilty to the degree that the description is correct. The value above can be interpreted as the judge's opinion about the correctness of the description. The mean value of the judge's opinion which is equal to 0.88 gives a probability estimate of the correctness of the description. This means that there is about 12% chance that the description actually is wrong. Even if the Collinses fit the description the likelihood of them actually being guilty is then substantially reduced.

Although the values used here are purely hypothetical, the example illustrates the effect of interpreting statements from witnesses as advice, and how the judge's opinions about the witnesses are applied to discount the witnesses advice.

7. Conclusion

There seems to be a consensus between the judicial and statistical professions that probability theory is insufficient for modelling legal reasoning, mainly because probability is not able to express uncertainty. In the present paper we have described a calculus for uncertain probabilities called Subjective Logic, and explored how this calculus can be applied to legal reasoning.

The main difficulty with applying Subjective Logic is that there is no consistent way of determining opinions when the evidence at hand can not be analysed statistically. If for example two individuals value their opinions differently when confronted with the same evidence, and their opinions can change over time

without any clear reason, it becomes evident how subjective, and ultimately unmanageable a mathematical analysis can become.

Judicial decisions are influenced by an infinity of factors, of which only a fraction is contemplated in a conscious way by the judges and jurors. It can therefore not be expected that these complex mental mechanisms can be modelled by a dozen formulas. The present contribution is more an attempt to understand this process than it is an attempt to clone and eventually replace it. However, we believe it would be interesting to compare results produced by a formal model with actual legal decisions, to see whether the latter can be formally and logically supported.

Appendix

PROOF 1 (Belief Function Additivity). The sum of the belief, disbelief and uncertainty functions is equal to the sum of the belief masses in a BMA which according to Definition 1 sums up to 1. \square

PROOF 2 (Kolmogorov Axioms). Each property can be proved separately.

1. Immediate results of Defs.1 & 2 are that $b_x \geq 0$, that $u_x \geq 0$, and that $a_x \geq 0$ for all x . As a consequence any probability expectation according to Definition 6 will satisfy $0 \leq E(x) \leq 1$.
2. Immediate results of Defs.1 are that $b_\Theta = 1$ and that $u_\Theta = 0$, resulting in $E(\Theta) = 1$.
3. Let $x_1, x_2 \dots \in 2^\Theta$ be a set of disjoint states, i.e. so that $x_i \cap x_j = \emptyset$ for $i \neq j$. According to Definition 6 we can write:

$$E((x_i \cup x_j)) = \sum_y m_\Theta(y) a((x_i \cup x_j)/y), \quad y \in 2^\Theta \quad (13)$$

Because x_i and x_j are disjoint the following holds:

$$a((x_i \cup x_j)/y) = a(x_i/y) + a(x_j/y) \quad (14)$$

The sum in (13) can therefore be split in two so that $E((x_i \cup x_j))$ can be written as:

$$\begin{aligned} E((x_i \cup x_j)) &= \sum_y m_\Theta(y) a(x_i/y) + \sum_y m_\Theta(y) a(x_j/y), \quad y \in 2^\Theta \\ &= E(x_i) + E(x_j) \end{aligned} \quad (15)$$

This can be generalised to cover arbitrary sets of disjoint states. \square

PROOF 3 and 4. (Propositional Conjunction and Disjunction). Let Θ_X and Θ_Y be two binary frames of discernment, where $x, \neg x \in \Theta_X$ and $y, \neg y \in \Theta_Y$. The product

frame of discernment of Θ_X and Θ_Y , denoted by $\Theta_{X \times Y}$ is obtained by conjugating each element of 2^{Θ_X} with each element of 2^{Θ_Y} . This produces:

$$\begin{aligned}\Theta_{X \times Y} &= \{x, \neg x, \Theta_X\} \times \{y, \neg y, \Theta_Y\} \\ &= \{x \cap y, x \cap \neg y, x \cap \Theta_Y, \neg x \cap y, \neg x \cap \neg y, \\ &\quad \neg x \cap \Theta_Y, \Theta_X \cap y, \Theta_X \cap \neg y, \Theta_X \cap \Theta_Y\}\end{aligned}$$

Let m_{Θ_X} and m_{Θ_Y} be BMAs on Θ_X and Θ_Y respectively. Because Θ_X and Θ_Y are binary, the belief masses can be expressed according to Equation (3) as simple belief functions such that:

$$\begin{array}{ll} m_{\Theta_X}(x) &= b_x & m_{\Theta_Y}(y) &= b_y \\ m_{\Theta_X}(\neg x) &= d_x & m_{\Theta_Y}(\neg y) &= d_y \\ m_{\Theta_X}(\Theta_X) &= u_x & m_{\Theta_Y}(\Theta_Y) &= u_y \end{array}$$

The BMA on $\Theta_{X \times Y}$ is obtained by multiplying the respective belief masses on the elements of 2^{Θ_X} with the belief masses on the elements of 2^{Θ_Y} . This produces:

$$\begin{array}{ll} m_{\Theta_{X \times Y}}(x \cap y) &= b_x b_y & m_{\Theta_{X \times Y}}(\Theta_X \cap y) &= u_x b_y \\ m_{\Theta_{X \times Y}}(x \cap \neg y) &= b_x d_y & m_{\Theta_{X \times Y}}(\Theta_X \cap \neg y) &= u_x d_y \\ m_{\Theta_{X \times Y}}(x \cap \Theta_Y) &= b_x u_y & m_{\Theta_{X \times Y}}(\Theta_X \cap \Theta_Y) &= u_x u_y \\ m_{\Theta_{X \times Y}}(\neg x \cap y) &= d_x b_y & & \\ m_{\Theta_{X \times Y}}(\neg x \cap \neg y) &= d_x d_y & & \\ m_{\Theta_{X \times Y}}(\neg x \cap \Theta_Y) &= d_x u_y & & \end{array}$$

– Propositional Conjunction:

The conjunction between $x \in \Theta_X$ and $y \in \Theta_Y$ is simply $x \cap y \in \Theta_{X \times Y}$. The derived frame of discernment with focus on $x \cap y$ then becomes $\tilde{\Theta}_{X \times Y}^{x \cap y} = \{x \cap y, \neg\{x \cap y\}\}$, where $\neg\{x \cap y\} = \{x \cap \neg y, \neg x \cap y, \neg x \cap \neg y\}$. According to Definition 8 the BMA $m_{\tilde{\Theta}_{X \times Y}^{x \cap y}}$ is such that:

$$\begin{array}{ll} 1. & m_{\tilde{\Theta}_{X \times Y}^{x \cap y}}(x \cap y) = b_{x \wedge y} \\ 2. & m_{\tilde{\Theta}_{X \times Y}^{x \cap y}}(\neg\{x \cap y\}) = d_{x \wedge y} \\ 3. & m_{\tilde{\Theta}_{X \times Y}^{x \cap y}}(\tilde{\Theta}_{X \times Y}^{x \cap y}) = u_{x \wedge y} \end{array}$$

By using Equation (4) it can also be observed that the derived relative atomicity of $x \cap y$ is such that:

$$4. \quad a_{\tilde{\Theta}_{X \times Y}^{x \cap y}}(x \cap y) = a_{x \wedge y}$$

These four parameters define $\omega_{x \wedge y}$ as specified in Theorem 3.

– Propositional Disjunction:

Similarly to propositional conjunction, the propositional disjunction between $x \in \Theta_X$ and $y \in \Theta_Y$ is simply $x \cup y = \{x \cap y, x \cap \neg y, \neg x \cap y\}$, with $x \cup y \in \Theta_{X \times Y}$. The derived frame of discernment with focus on $x \cup y$ then

becomes $\tilde{\Theta}_{X \times Y}^{x \cup y} = \{x \cup y, \neg\{x \cup y\}\}$, where $\neg\{x \cup y\} = \{\neg x \cap \neg y\}$. According to Definition 8 the belief mass assignment $m_{\Theta_{x \cup y}}$ is such that:

1. $m_{\tilde{\Theta}_{X \times Y}^{x \cup y}}(x \cup y) = b_{x \vee y}$
2. $m_{\tilde{\Theta}_{X \times Y}^{x \cup y}}(\neg\{x \cup y\}) = d_{x \vee y}$
3. $m_{\tilde{\Theta}_{X \times Y}^{x \cup y}}(\tilde{\Theta}_{X \times Y}^{x \cup y}) = u_{x \vee y}$

By using Equation (4) it can also be observed that the derived relative atomicity of $x \cup y$ is such that:

4. $a_{\tilde{\Theta}_{X \times Y}^{x \cup y}}(x \cup y) = a_{x \vee y}$

These four parameters define $\omega_{x \vee y}$ as specified in Theorem 4. □

PROOF 5 (Product and Co-product). Each property can be proved separately.

1. Equation 1 corresponds to the product of probabilities. By using Definition 6 and Theorem 3 we get:

$$\begin{aligned}
 E(\omega_{x \wedge y}) &= b_{x \wedge y} + u_{x \wedge y} a_{x \wedge y} \\
 &= b_x b_y + b_x u_y a_y + u_x a_x b_y + u_x a_x u_y a_y \\
 &= (b_x + u_x a_x)(b_y + u_y a_y) \\
 &= E(\omega_x)E(\omega_y)
 \end{aligned} \tag{16}$$

2. Equation 2 corresponds to the co-product of probabilities. By using Definition 6, Theorem 4 and Equation (1) we get:

$$\begin{aligned}
 E(\omega_{x \vee y}) &= b_{x \vee y} + u_{x \vee y} a_{x \vee y} \\
 &= b_{x \vee y} + (d_x u_y + u_x d_y + u_x u_y) a_{x \vee y} \\
 &= b_{x \vee y} + (u_x + u_y - b_x u_y - u_x b_y - u_x u_y) a_{x \vee y} \\
 &= b_x + b_y - b_x b_y + u_x a_x + u_y a_y - b_x u_y a_y \\
 &\quad - u_x a_x b_y - u_x a_x u_y a_y \\
 &= b_x + u_x a_x + b_y + u_y a_y - (b_x + u_x a_x)(b_y + u_y a_y) \\
 &= E(\omega_x) + E(\omega_y) - E(\omega_x)E(\omega_y)
 \end{aligned} \tag{17}$$

□

PROOF 6 (Negation). The opinion about the negation of the proposition is the opinion about the complement state in the frame of discernment. An immediate result of Equation (3) is then that $b_{\neg x} = d_x$, $d_{\neg x} = b_x$ and $u_{\neg x} = u_x$. The probability expectation values of x and $\neg x$ satisfy $E(\omega_x) + E(\omega_{\neg x}) = 1$ which when used in Equation 3 results in $a_{\neg x} = 1 - a_x$. □

References

Baldwin, J. 1986. Support logic programming. In: A. Jones et al. (eds.), *Fuzzy Sets: Theory and Applications*. Reidel.

- Cohen, L. 1977. *The Probable and the Provable*. Clarendon Press. Oxford.
- Cohen, M. 1986. An expert system framework for non-monotonic reasoning about probabilistic assumptions. In: L. Kanal and J. Lemmer (eds.), *Uncertainty in Artificial Intelligence 1*. North-Holland.
- Eggleston, R. 1978. *Evidence, Proof, and Probability*. London: Weidenfeld and Nicholson.
- Ellsberg, D. 1961. Risk, Ambiguity, and the Savage Axioms. *Quarterly Journal of Economics* **75**, 643–669.
- Hunter, A. 1996. *Uncertainty in information systems*. London: McGraw-Hill.
- James, W. 1941. Relevance, probability and the law. *California Law Review* **29**, 689–699.
- Jøsang, A. 2001. A Logic for Uncertain Probabilities. *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems* **9.3**, June 2001.
- Jøsang, A. & Knapkog, S. 1998. A Metric for Trusted Systems (full paper). In: *Proceedings of the 21st National Information Systems Security Conference*. NSA.
- Motro, A. & Smets, P. 1997. *Uncertainty management in information systems: from needs to solutions*. Boston: Kluwer.
- PvC. 1968. People vs. Collins. 68 Cal 2d 319; 438 P2d 33.
- RvA. 1996. R vs. Adams. 2 Court of Appeal (Criminal Division), R.467.
- RvDA. 1997. R vs. Doheny and Adams. 1 Court of Appeal (Criminal Division), R.369.
- Shafer, G. 1976. *A Mathematical Theory of Evidence*. Princeton University Press.
- Simon, R. & Mahan, L. 1971. Quantifying burdens of proof. *Law and Society Review* **5**, 319–322.
- Smets, P. & Kennes, R. 1994. The transferable belief model. *Artificial Intelligence* **66**, 191–234.
- Smithson, M. 1988. *Ignorance and Uncertainty: Emerging Paradigms*. Springer.
- Thompson, W. & Schumann, E. 1987. Interpretation of statistical evidence in criminal trials: The prosecutor's fallacy and the defense attorney's fallacy. *Law and Human Behavior* **11**, 167–187.
- Zadeh, L. 1984. Review of Shafer's a mathematical theory of evidence. *AI Magazine* **5**, 81–83.

