

COMP 8920 Section 02

Computational Reasoning in AI

Midterm

Vijay Rajasekar Thirulokachander
105092311
thirulo@uwindsor.ca

April 8, 2019

1 Binary Division Constraint Derivation

1.1 Deriving $d_x \geq d_y$

Given,

$$d_{x\bar{\wedge}y} = \frac{d_x - d_y}{1 - d_y}$$

If $d_{x\bar{\wedge}y} = 0$

$$\begin{aligned} 0 &= \frac{d_x - d_y}{1 - d_y} \\ \Rightarrow \frac{d_y}{1 - d_y} &= \frac{d_x}{1 - d_y} \\ \Rightarrow d_x &= d_y x \end{aligned}$$

Since $0 \leq d_{x\bar{\wedge}y} \leq 1$, and $d_x = d_y$ for the lowest possible value of $0 \leq d_{x\bar{\wedge}y}$, therefore

$$d_x \geq d_y$$

1.2 Deriving $a_x < a_y$

Given,

$$b_{x\bar{\wedge}y} = \frac{a_y(b_x + a_x u_x)}{(a_y - a_x)(b_y + a_y u_y)} - \frac{a_x(1 - d_x)}{(a_y - a_x)(1 - d_y)}$$

If $a_x \geq a_y$

$$b_{x\bar{\wedge}y} = \frac{a_y(b_x + a_x u_x)}{\infty(b_y + a_y u_y)} - \frac{a_x(1 - d_x)}{\infty(1 - d_y)}$$

which is clearly undefined. Therefore, $a_x < a_y$

2 Defintions of concepts and their applications in reasoning

2.1 Probability Measure

Out of the many approaches to represent belief and uncertainty, probability has the advantage of being well understood. It is a powerful tool; many technical results have been proved that facilitate its use, and a number of arguments suggest that, under certain assumptions (whose reasonableness can be debated) probability is the only “rational” way to represent uncertainty. [Lecture Slides 1B]

Consider now a completely arbitrary nonempty set of possible outcomes to be Ω . A class \mathcal{F} of subsets of Ω is called a field if it contains Ω itself and is closed under the formation of complements and finite unions:

- i $\Omega \in \mathcal{F}$
- ii $A \in \mathcal{F}$ implies $\neg A \in \mathcal{F}$
- iii $A, B \in \mathcal{F}$ implies $A \cup B \in \mathcal{F}$

A set function is a real-valued function defined on some class of subsets of Ω . A set function P on a field \mathcal{F} is a *probability measure* if it satisfies these conditions:

- i $0 < P(A) < 1$ for $A \in \mathcal{F}$
- ii $P(\emptyset) = 0, P(\Omega) = 1$
- iii if A_1, A_2, \dots is a disjoint sequence of \mathcal{F} sets and if $\cup_{k=1}^{\infty} A_k \in \mathcal{F}$ i.e the union of all sequences of \mathcal{F} sets belongs to \mathcal{F} then

$$P\left(\bigcup_{k=1}^{\infty} A_k\right) = \sum_{k=1}^{\infty} P(A_k)$$

i.e the probability of the union of all disjoint sequences is equal to the sum of probability of each disjoint sequence. This property is referred to as *finite additivity*.

The above definitions were obtained from the book ***Probability and Measure, Third Edition, Patrick Billingsley, Pg. 19 - 23***

[<https://www.colorado.edu/amath/sites/default/files/attached-files/billingsley.pdf>]

Sets of probability measures have many of the advantages of probability, and may be more appropriate in a setting where there is uncertainty about the likelihood. Considering sets of weighted probability measures allows for more flexible modeling, but has the disadvantage of needing yet more numbers to characterize a situation.

2.2 Belief Function

Belief in a set of elements, say A , of a frame θ , represents the total belief that one has based on the evidence obtained. It is the sum of all the belief masses assigned to elements that are contained in the set A and the belief mass assigned to the set A . Mathematically, one can express the total belief in the set A as

$$Bel(A) = \sum_{B \subseteq A} m(B)$$

Unlike probability theory, $\text{Bel}(A) = 0$ represents lack of evidence about A , while $P(A) = 0$ represents the impossibility of A . However, $\text{Bel}(A) = 1$ represents certainty, that is A is certain to occur, similar to $P(A) = 1$, which also represents the certainty that A is true. [Srivastava, Rajendra P.. “The Dempster-Shafer Theory of Belief Functions for Managing Uncertainties : An Introduction and Fraud Risk Assessment Illustration 1.” (2012).]

Prof. Judea Pearl, in his work *Reasoning with Belief Functions: An Analysis of Compatibility*, provides three domains of reasoning in which belief functions are applied:

- i *Representation of incomplete knowledge*: In cases where fully specified probabilistic knowledge is not available, belief functions are often used to represent states of partial knowledge, with $\text{Bel}(A)$ interpreted as a strength of arguments in favor of A .
- ii *Belief updating*: Belief functions offer a method of assimilating the impact of new evidence into a state of partial knowledge or partial belief.
- iii *Pooling of evidence*: When multiple pieces of evidence are available, the BF combination method provides a faithful summary of the information carried by all the individual pieces. The method consists of encoding each piece of evidence as a belief function and combining these functions by Dempster’s rule of orthogonal sum.

2.3 Conditional Probability

Before defining conditional probability, we need to define what *unconditional probability* is. The unconditional or prior probability associated with a proposition A is the degree of belief accorded to it in the absence of any information. It is written as $P(A)$. However, once the agent has obtained some evidence concerning a previously unknown variable, prior probability no longer applies. Instead we use conditional probability. The notation $P(A|B)$ denotes “probability of A , given all that we know is B ”. [Lecture Slides 1B]

Conditional inference plays a central role in logical and Bayesian reasoning, and is used in a wide range of applications. It basically consists of expressing conditional relationship between parent and child propositions, and then to combine those conditionals with evidence about the parent propositions in order to infer conclusions about the child propositions. Probabilistic conditional reasoning is used extensively in areas where conclusions need to be derived from probabilistic input evidence, such as for making diagnoses from medical tests. [Conditional Reasoning with Subjective Logic, Audun Jøsang]

2.4 Probabilistic Reasoning

Probability theory has existed in one form or another for several hundred years. During this time various alternative formulations have been introduced, and it is now difficult to say where the definitive account may be found. Probability theory is built on three axioms or laws that define the behaviour of a probability measure, which may be used as an estimate of the degree to which an uncertain event is likely to occur. They are:

- i Convexity Law: $0 \leq Pr(A|H) \leq 1$
- ii Addition Law: $Pr(A \cup B|H) = Pr(A|H) + Pr(B|H)$
- iii Multiplication Law: $Pr(A \cap B|H) = Pr(A|H).Pr(B|A \cap H)$

2.5 Aleatory and Epistemic Opinions

[Lecture Slides 4]

Aleatory Uncertainty, which is the same as statistical uncertainty, expresses that we do not know the outcome each time we run the same experiment. we only know the long-term relative frequency of outcomes. An *aleatory opinion* applies to a variable governed by a frequentist process, and that represents the (uncertain) likelihood of values of the variable in any unknown past or future instance of the process.

Epistemic Uncertainty, aka systematic uncertainty, expresses that we could in principle know the outcome of a specific or future or past event, but that we do not have enough evidence to know it exactly. An *epistemic opinion* applies to a variable that is assumed to be non-frequentist, and that represents the (uncertain) likelihood of values of the variable in a specific unknown past or future instance.

In traditional Bayesian theory, the difference between aleatory and epistemic uncertainty is mostly philosophical and has no practical implication for modelling and analysing real situations, because the two types of opinions are formally and syntactically indistinguishable. In subjective logic however, the difference between aleatory and epistemic opinions is syntactically explicit, in the sense that epistemic opinions are uncertainty-maximised, whereas aleatory opinions can have arbitrary relative levels of uncertainty.

3 Beta and Dirichlet Probability Density Functions

The below definitions were obtained from *Subjective Logic: A Formalism for Reasoning under Uncertainty*, Audun Jøsang, Chapter 3

3.1 Beta PDFs

The Beta PDF is a probability density function denoted $Beta(p(x), \alpha, \beta)$ with variable $p(x)$, and two ‘strength parameters’ α and β . The Beta PDF is defined below:

Assume a binary domain $\mathbb{X} = \{x, \bar{x}\}$ and a random variable X in \mathbb{X} . Let p denote the continuous probability function $p : X \rightarrow [0, 1]$ where $p(x) + p(\bar{x}) = 1$.

The parameter α represents evidence/observations of $X = x$ and the parameter β represents evidence/observations of $X = \bar{x}$. With p_x as variable, the Beta probability density function $Beta(p_x, \alpha, \beta)$ is the function expressed as:

$$Beta(p_x, \alpha, \beta) : [0, 1] \rightarrow \mathbb{R}_{\geq 0}$$
$$Beta(p_x, \alpha, \beta) = \frac{\Gamma(\alpha + \beta)}{\Gamma(\alpha)\Gamma(\beta)} (p_x)^{\alpha-1} (1 - p_x)^{\beta-1}, \alpha > 0, \beta > 0$$

The PDF is normalized under the integration,

$$\int_0^1 Beta(p, \alpha, \beta) dp = 1$$

Assume that x represents a frequentist event, such as getting heads when flipping a coin. Let r_x denote the number of observations

$$\alpha = r_x + a_x W$$

$$\beta = s_x + (1 - a_x)W$$

Using the α and β parameters, we can express Beta PDF as an evidence notation,

$$Beta^e(p_x, r_x, s_x, a_x) = \frac{\Gamma(r_x + s_x + W)}{\Gamma(r_x + a_x W) \Gamma(s_x + (1 - a_x)W)} (p_x)^{r_x} (1 - p_x)^{s_x + (1 - a_x)W - 1}$$

The expected probability $E(x)$ as a function of Beta PDF parameters is

$$E(x) = \frac{\alpha}{\alpha + \beta} = \frac{r_x + a_x W}{r_x + s_x + W}$$

The traditional interpretation of probability as likelihood of events gets a richer expression through probability density, which can be interpreted as second-order probability. This interpretation is the basis for mapping (high) probability density in Beta PDFs into (high) belief mass in opinions, through bijective mapping. As a consequence, flat probability density in Beta PDFs is represented as uncertainty in opinions.

The Beta PDF is important in subjective logic, because a bijective mapping can be defined between the projected probability of a binomial opinion and the expected probability of a Beta PDF.

The bijective mapping between a binomial opinion and a Beta PDF emerges from the intuitive requirement that $P(x) = E(x)$, i.e. that the projected probability of a binomial opinion must be equal to the expected probability of a Beta PDF.

Let $\omega_x = (b_x, d_x, u_x, a_x)$ be a binomial opinion and let $p(x)$ be a probability distribution, both over the same binomial random variable X . Let $Beta^e(p_x, r_x, s_x, a_x)$ be a Beta PDF. The opinion ω_x and the Beta PDF $Beta^e(p_x, r_x, s_x, a_x)$ are equivalent as follows:

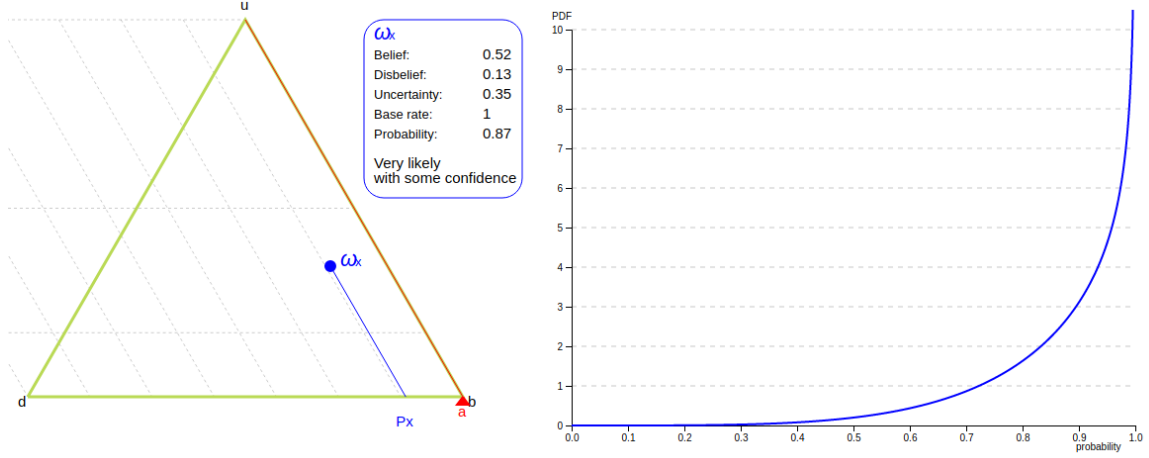
For $u_x \neq 0$:

$$\begin{cases} b_x = \frac{r_x}{W + r_x + s_x} \\ d_x = \frac{s_x}{W + r_x + s_x} \\ u_x = \frac{W}{W + r_x + s_x} \end{cases} \Leftrightarrow \begin{cases} r_x = \frac{b_x W}{u_x} \\ s_x = \frac{d_x W}{u_x} \\ 1 = b_x + d_x + u_x \end{cases}$$

For $u_x = 0$:

$$\begin{cases} b_x = \frac{r_x}{W + r_x + s_x} \\ d_x = \frac{s_x}{W + r_x + s_x} \\ u_x = \frac{W}{W + r_x + s_x} \end{cases} \Leftrightarrow \begin{cases} r_x = b_x \cdot \infty \\ s_x = d_x \cdot \infty \\ 1 = b_x + d_x \end{cases}$$

The below figure illustrates a binomial opinion $\omega_x = (0.52, 0.13, 0.35, 1)$ mapped into a Beta PDF graph.



The equivalence between binomial opinions and Beta PDFs is very powerful, because subjective-logic operators (SL operators) can then be applied to Beta PDFs, and statistics operations for Beta PDFs can be applied to opinions. In addition, it makes it possible to determine binomial opinions from statistical observations.

3.2 Dirichlet PDFs

Let \mathbb{X} be a domain consisting of k mutually disjoint values. Let α_x represent the strength vector over the values of \mathbb{X} and let p_x denote the probability distribution over \mathbb{X} . With p_x as a k -dimensional variable the Dirichlet PDF denoted $Dir(p_x, \alpha_x)$ is expressed as:

$$Dir(p_x, \alpha_x) = \frac{\Gamma(\sum_{x \in \mathbb{X}} \alpha_X(x))}{\prod_{x \in \mathbb{X}} (\alpha_X(x))} \prod_{x \in \mathbb{X}} p_x(x)^{\alpha_X(x)-1}, \alpha_X(x) \geq 0$$

The strength vector α_X represents the prior, as well as the observation evidence. The non-informative prior weight is expressed as a constant W , and this weight is distributed over all the outcomes as a function of base rate. As already mentioned, it is normally assumed that $W = 2$.

Evidence representation in Dirichlet PDF can be done using the evidence vector r_x , base rate distribution a_x and non-informative prior W can be done as below:

$$Dir_x^e(p_x, r_x, a_x) = \frac{\Gamma(\sum_{x \in \mathbb{X}} (r_X(x) + a_X(x)W))}{\prod_{x \in \mathbb{X}} (r_X(x) + a_X(x)W)} \prod_{x \in \mathbb{X}} p_x(x)^{(r_X(x) + a_X(x)W)-1}, \alpha_X(x) \geq 0$$

This notation is useful because it allows the determination of the probability densities over variables, where each value can have an arbitrary base rate.

The bijective mapping between ω_x and $Dir_x^e(p_x, r_x, a_x)$ is based on the requirement for equality between the projected probability distribution P_x derived from ω_x and the expected probability distribution E_x derived from $Dir_x^e(p_x, r_x, a_x)$. This requirement is expressed as:

$$P_x = E_x \Leftrightarrow b_X(x) + a_X(x)u_X = \frac{r_X(x) + W a_X(x)}{W + \sum_{x_j \in \mathbb{X}} r_X(x_j)}$$

Therefore the bijective mapping can be described as below:

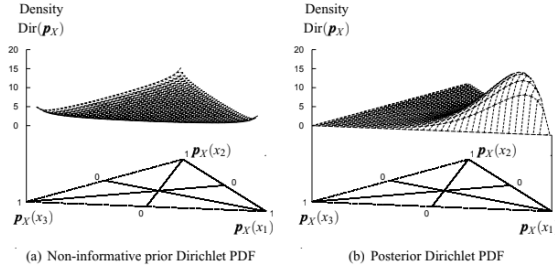
For $u_x \neq 0$:

$$\begin{cases} b_X(x) = \frac{r_X(x)}{W + \sum_{x_j \in \mathbb{X}} r_X(x_j)} \\ u_X = \frac{r_X(x)}{W + \sum_{x_j \in \mathbb{X}} r_X(x_j)} \end{cases} \Leftrightarrow \begin{cases} r_X(x) = \frac{W b_X(x)}{u_x} \\ 1 = u_x + \sum_{x_j \in \mathbb{X}} b_X(x_j) \end{cases}$$

For $u_x = 0$:

$$\begin{cases} b_X(x) = \frac{r_X(x)}{W + \sum_{x_j \in \mathbb{X}} r_X(x_j)} \\ u_X = \frac{r_X(x)}{W + \sum_{x_j \in \mathbb{X}} r_X(x_j)} \end{cases} \Leftrightarrow \begin{cases} r_X(x) = b_X(x) \cdot \infty \\ 1 = \sum_{x_j \in \mathbb{X}} b_X(x_j) \end{cases}$$

The below figure illustrates Dirichlet PDF over $(k - 1)$ dimensions



4 Subjective Logic Operators

All the below definitions were obtained from *Subjective Logic: A Formalism for Reasoning under Uncertainty*, Audun Jøsang, Chapters 6 and 7

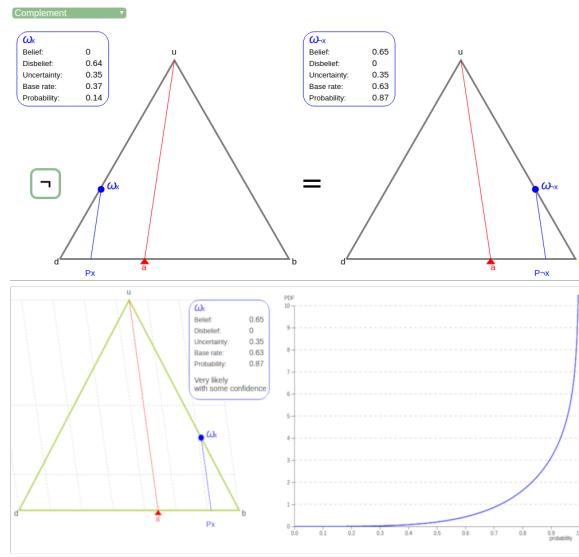
4.1 Complement Operator

Assume the domain $X = \{x, \bar{x}\}$ which can be assumed to be binary or to be a binary partition of two subsets, where $\omega_x = (b_x, d_x, u_x, a_x)$ is a binomial opinion about x . Its complement is the binomial opinion $\omega_{\bar{x}}$ expressed as

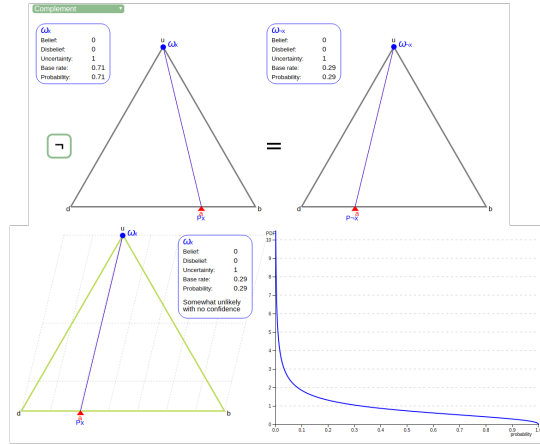
$$\omega_{\bar{x}} : \begin{cases} b_{\bar{x}} = d_x \\ d_{\bar{x}} = b_x \\ u_{\bar{x}} = u_x \\ a_{\bar{x}} = 1 - a_x \end{cases}$$

The complement operator corresponds to binary logic NOT, and to complement of probabilities.

Complement Example 1: Input data: $\omega_x = (0, 0.64, 0.35, 0.37)$



Complement Example 2: Input data: $\omega_x = (0, 0, 1, 0.29)$ (Vacuous Opinion)



4.2 Addition

Addition of opinions in subjective logic is a binary operator that takes opinions about two mutually exclusive values (i.e. two disjoint subsets of the same domain) as arguments, and outputs an opinion about the union of the values.

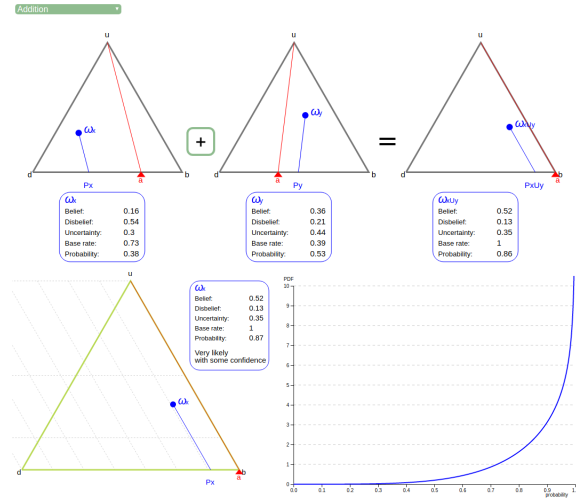
Let $\omega_{x_1} = (b_{x_1}, d_{x_1}, u_{x_1}, a_{x_1})$ and $\omega_{x_2} = (b_{x_2}, d_{x_2}, u_{x_2}, a_{x_2})$ be two binomial opinions. The sum of

these two opinions can be expressed as:

$$\omega_{(x_1 \cup x_2)} : \begin{cases} b_{(x_1 \cup x_2)} = b_{x_1} + b_{x_2} \\ d_{(x_1 \cup x_2)} = \frac{a_{x_1}(d_{x_1} - b_{x_2}) + a_{x_2}(d_{x_2} - b_{x_1})}{a_{x_1} + a_{x_2}} \\ u_{(x_1 \cup x_2)} = u_x \\ a_{(x_1 \cup x_2)} = 1 - a_x \end{cases}$$

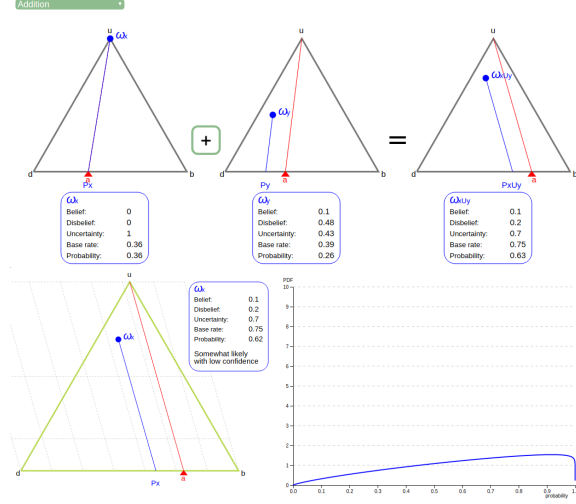
Addition Example 1:

Input data: $\omega_x = (0.16, 0.54, 0.3, 0.73)$, $\omega_y = (0.36, 0.21, 0.44, 0.39)$



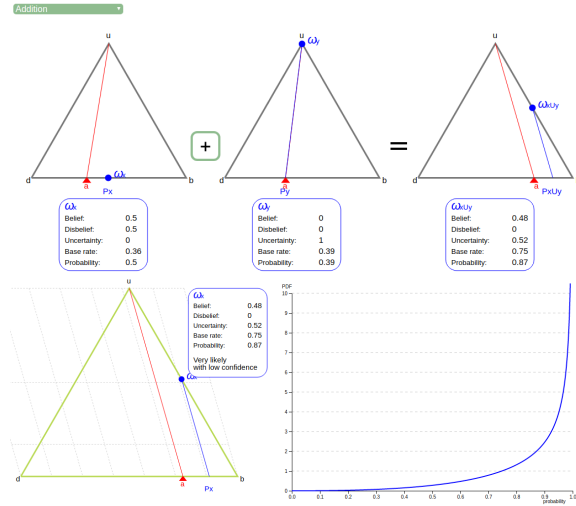
Addition Example 2:

Input data: $\omega_x = (0, 0, 1, 0.36)$, $\omega_y = (0.1, 0.48, 0.43, 0.39)$



Addition Example 3:

Input data: $\omega_x = (0.5, 0.5, 0, 0.36)$ (Dogmatic), $\omega_y = (0, 0, 1, 0.39)$ (Vacuous)



4.3 Subtraction

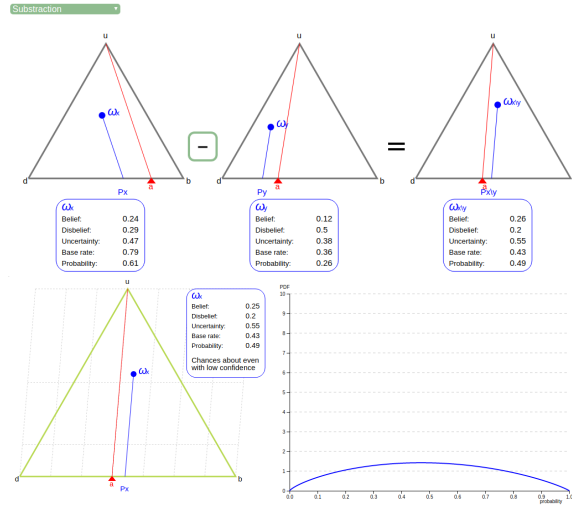
The inverse operation to opinion addition is opinion subtraction. Since addition of opinions yields the opinion about $x_1 \cup x_2$ from the opinions about disjoint subsets of the domain, then the difference between the opinions about $x_1 \cup x_2$ and x_2 (i.e. the opinion about $((x_1 \cup x_2)/x_2)$) can only be defined

if $x_2 \subseteq (x_1 \cup x_2)$ where x_2 and $(x_1 \cup x_2)$ are subsets of the domain X , i.e. the system must be in the state $(x_1 \cup x_2)$ whenever it is in the state x_2 . The sum of these two opinions can be expressed as:

$$\omega_{((x_1 \cup x_2)/x_2)} : \begin{cases} b_{((x_1 \cup x_2)/x_2)} = b_{(x_1 \cup x_2)} - b_{x_2} \\ d_{((x_1 \cup x_2)/x_2)} = \frac{a_{(x_1 \cup x_2)}(d_{(x_1 \cup x_2)} + b_{x_2}) - a_{x_2}(1 + b_{x_2} - b_{(x_1 \cup x_2)} - u_{x_2})}{a_{x_1} + a_{x_2}} \\ u_{((x_1 \cup x_2)/x_2)} = \frac{a_{(x_1 \cup x_2)}u_{(x_1 \cup x_2)} - a_{x_2}u_{x_2}}{a_{(x_1 \cup x_2)} - a_{x_2}} \\ a_{((x_1 \cup x_2)/x_2)} = a_{(x_1 \cup x_2)} - a_{x_2} \end{cases}$$

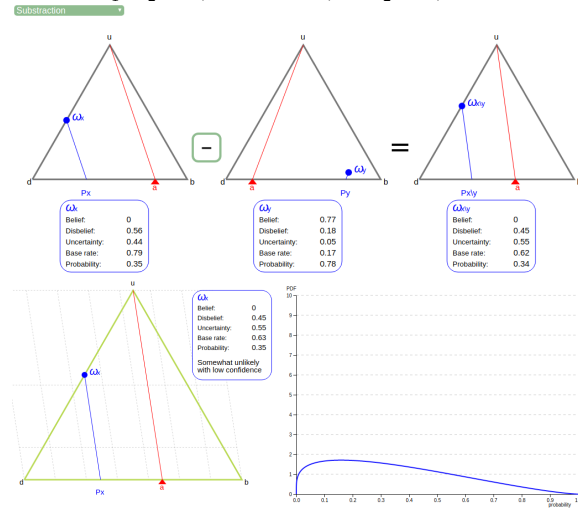
Subtraction Example 1:

Input data: $\omega_{x_1 \cup x_2} = (0.16, 0.54, 0.3, 0.73)$, $\omega_{x_2} = (0.36, 0.21, 0.44, 0.39)$



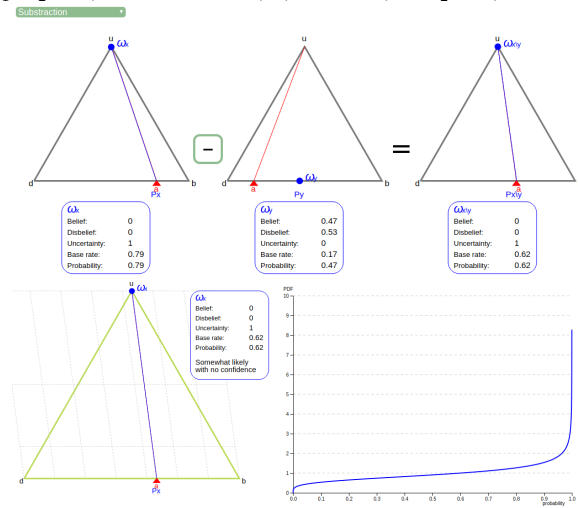
Subtraction Example 2:

Input data: $\omega_{x_1 \cup x_2} = (0, 0, 1, 0.36)$, $\omega_{x_2} = (0.1, 0.48, 0.43, 0.39)$



Subtraction Example 3:

Input data: $\omega_{x_1 \cup x_2} = (0.5, 0.5, 0, 0.36)$ (Vacuous), $\omega_{x_2} = (0, 0, 1, 0.39)$ (Dogmatic)



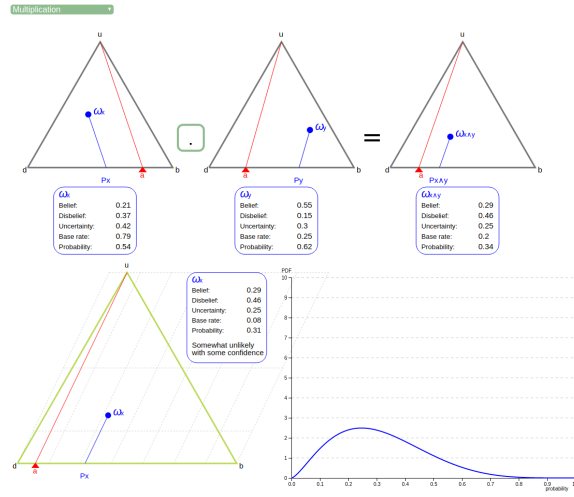
4.4 Multiplication

Let $\omega_x = (b_x, d_x, u_x, a_x)$ and $\omega_y = (b_y, d_y, u_y, a_y)$ be two independent binomial opinions on x and y respectively. The binomial opinion $\omega_{x \wedge y}$ on the conjunction $(x \wedge y)$ is:

$$\omega_{x \wedge y} \begin{cases} b_{x \wedge y} = b_x b_y + \frac{(1-a_x)a_y b_x u_y + a_x(1-a_y)u_x b_y}{1-a_x a_y} \\ d_{x \wedge y} = d_x + d_y - d_x d_y \\ u_{x \wedge y} = u_x u_y + \frac{(1-a_x)a_y b_x u_y + a_x(1-a_y)u_x b_y}{1-a_x a_y} \\ a_{x \wedge y} = a_x a_y \end{cases}$$

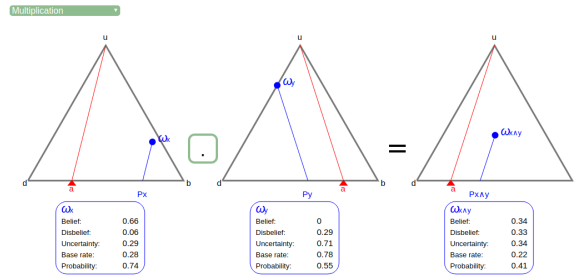
Multiplication Example 1:

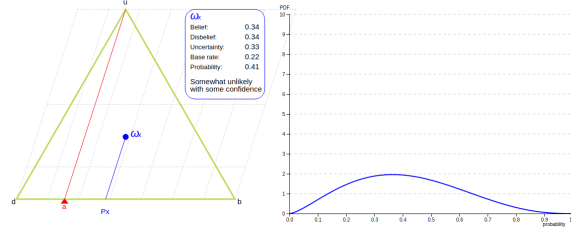
Input data: $\omega_{x_1} = (0.21, 0.37, 0.42, 0.54)$, $\omega_{x_2} = (0.55, 0.15, 0.3, 0.25)$



Multiplication Example 2:

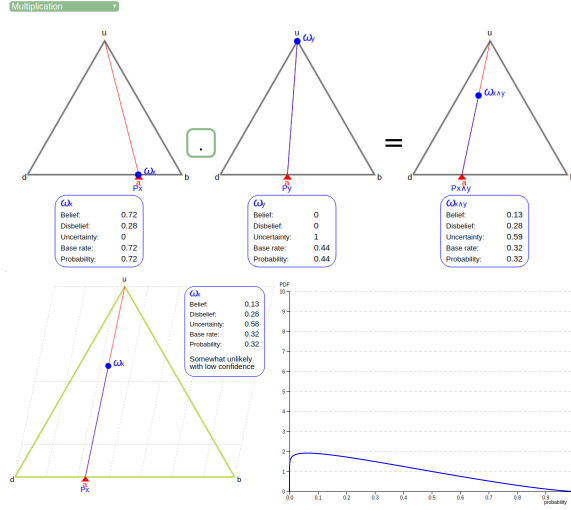
Input data: $\omega_{x_1} = (0.66, 0.06, 0.29, 0.28)$, $\omega_{x_2} = (0, 0.29, 0.71, 0.55)$





Multiplication Example 3:

Input data: $\omega_{x_1} = (0.72, 0.28, 0, 0.72)$ (Dogmatic), $\omega_{x_2} = (0, 0, 1, 0.44)$ (Vacuous)



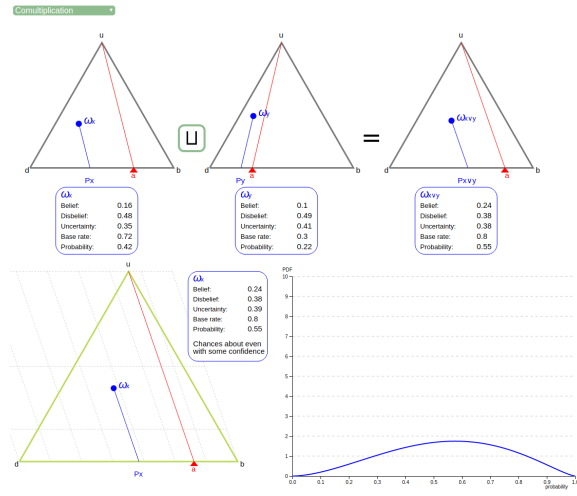
4.5 Co-Multiplication

Let $\omega_x = (b_x, d_x, u_x, a_x)$ and $\omega_y = (b_y, d_y, u_y, a_y)$ be two independent binomial opinions on x and y respectively. The binomial opinion $\omega_{x \vee y}$ on the conjunction $(x \vee y)$ is:

$$\omega_{x \vee y} \begin{cases} b_{x \vee y} = d_x + d_y - d_x d_y \\ d_{x \vee y} = b_x b_y + \frac{a_x(1-a_y)d_x u_y + (1-a_x)a_y u_x d_y}{a_x + a_y - a_x a_y} \\ u_{x \vee y} = u_x u_y + \frac{a_y d_x u_y + a_x u_x d_y}{a_x + a_y - a_x a_y} \\ a_{x \vee y} = a_x + a_y - a_x a_y \end{cases}$$

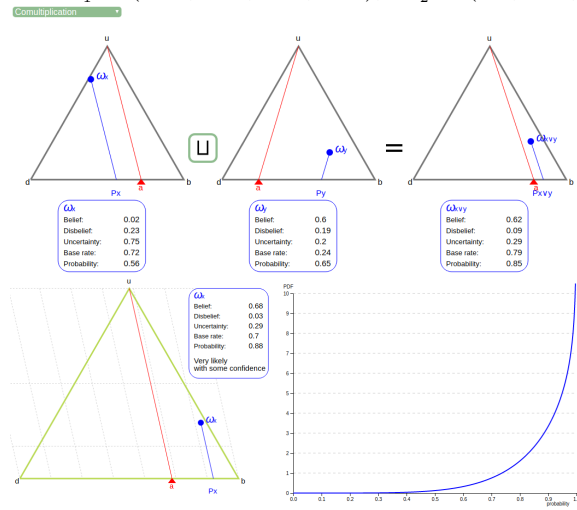
Co-Multiplication Example 1:

Input data: $\omega_{x_1} = (0.16, 0.48, 0.35, 0.72)$, $\omega_{x_2} = (0.1, 0.49, 0.41, 0.3)$



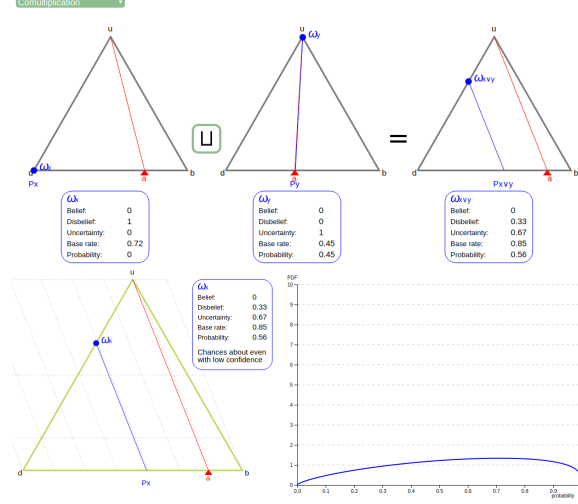
Co-Multiplication Example 2:

Input data: $\omega_{x_1} = (0.02, 0.23, 0.75, 0.72)$, $\omega_{x_2} = (0.6, 0.19, 0.2, 0.24)$



Co-Multiplication Example 3:

Input data: $\omega_{x_1} = (0, 1, 0, 0.72)$ (Dogmatic), $\omega_{x_2} = (0, 0, 1, 0.45)$ (Vacuous)



4.6 Division

The inverse operation to binomial multiplication is binomial division. The quotient of opinions about propositions x and y represents the opinion about a proposition z which is independent of y such that $\omega_x = \omega_{y \wedge z}$.

Let $\omega_x = (b_x, d_x, u_x, a_x)$ and $\omega_y = (b_y, d_y, u_y, a_y)$ be two independent binomial opinions on x and y respectively. The division of ω_x by ω_y produces the quotient opinion: $\omega_{x \bar{\wedge} y} = (b_{x \bar{\wedge} y}, d_{x \bar{\wedge} y}, u_{x \bar{\wedge} y}, a_{x \bar{\wedge} y})$

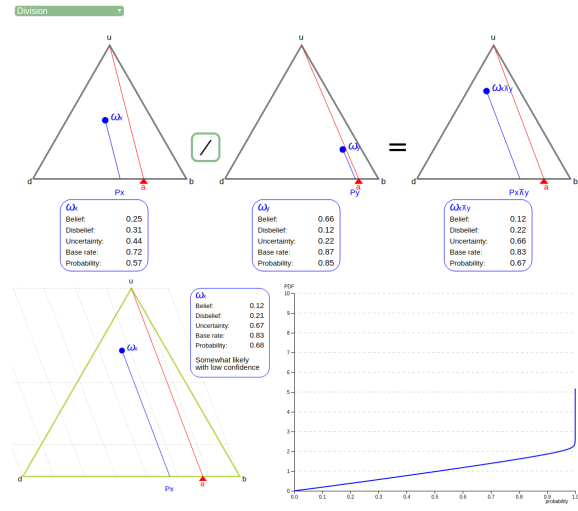
$$\omega_{x \bar{\wedge} y} \begin{cases} b_{x \bar{\wedge} y} = \frac{a_y(b_x + a_x u_x)}{(a_y - a_x)(b_y + a_y u_y)} - \frac{a_x(1 - d_x)}{(a_y - a_x)(1 - d_y)} \\ d_{x \bar{\wedge} y} = \frac{d_x - d_y}{1 - d_y} \\ u_{x \bar{\wedge} y} = \frac{a_y(1 - d_x)}{(a_y - a_x)(1 - d_y)} - \frac{a_y(b_x + a_x u_x)}{(a_y - a_x)(b_y + a_y u_y)} \\ a_{x \bar{\wedge} y} = \frac{a_x}{a_y} \end{cases}$$

where the following constraints are satisfied:

$$\text{Constraints} \begin{cases} b_x \geq \frac{a_x(1 - a_y)(1 - d_x)b_y}{(1 - a_x)a_y(1 - d_y)} \\ d_x \geq d_y \\ u_x \geq \frac{u_x(1 - a_y)(1 - d_x)b_y}{(1 - a_x)(1 - d_y)} \\ a_x < a_y \end{cases}$$

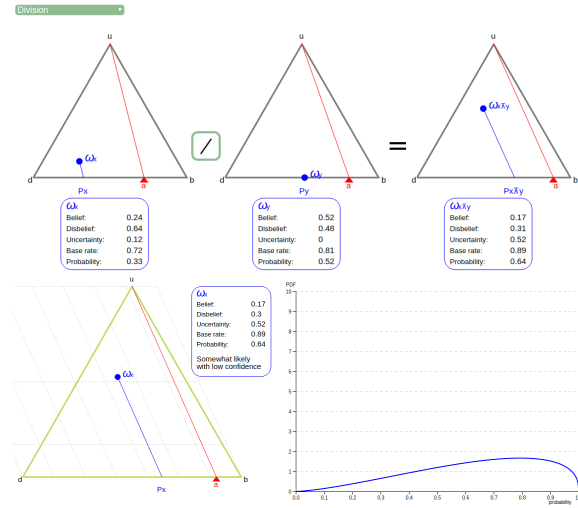
Division Example 1:

Input data: $\omega_{x_1} = (0.25, 0.31, 0.44, 0.72)$, $\omega_{x_2} = (0.66, 0.12, 0.22, 0.87)$



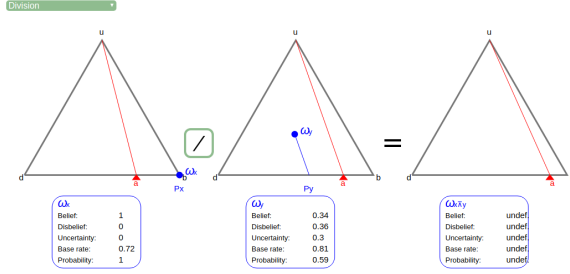
Division Example 2:

Input data: $\omega_{x_1} = (0.24, 0.64, 0.12, 0.72)$, $\omega_{x_2} = (0.52, 0.48, 0, 0.81)$ (Dogmatic)



Division Example 3:

Input data: $\omega_{x_1} = (1, 0, 0, 0.72)$ (Dogmatic), $\omega_{x_2} = (0.34, 0.36, 0.3, 0.81)$



It was not possible to obtain a visualization for this sample case.

4.7 Co-Division

The inverse operation to binomial co-multiplication is binomial co-division. The co-quotient of opinions about propositions x and y represents the opinion about a proposition z which is independent of y such that $\omega_x = \omega_y \wedge z$.

Let $\omega_x = (b_x, d_x, u_x, a_x)$ and $\omega_y = (b_y, d_y, u_y, a_y)$ be two independent binomial opinions on x and y respectively. The division of ω_x by ω_y produces the coquotient opinion: $\omega_{x \nabla y} = (b_{x \nabla y}, d_{x \nabla y}, u_{x \nabla y}, a_{x \nabla y})$

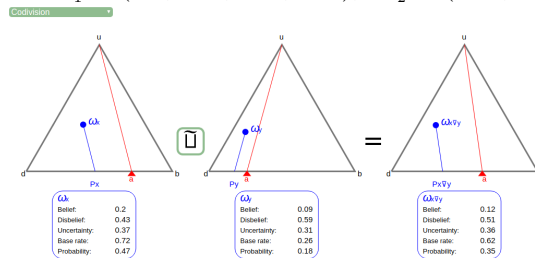
$$\omega_{x \nabla y} \begin{cases} b_{x \nabla y} = \frac{b_x - b_y}{1 - b_y} \\ u_{x \nabla y} = \frac{(1 - a_y)(1 - b_x)}{(a_x - a_y)(1 - b_y)} - \frac{(1 - a_y)(1 - b_x)}{(a_x - a_y)(1 - b_y)} \\ d_{x \nabla y} = \frac{(a_x - a_y)(1 - b_y)}{(1 - a_y)(1 - b_x)} - \frac{(1 - a_y)(d_x + (1 - a_x)u_x)}{(a_x - a_y)(1 - b_y)} \\ a_{x \nabla y} = \frac{a_x - a_y}{1 - a_y} \end{cases}$$

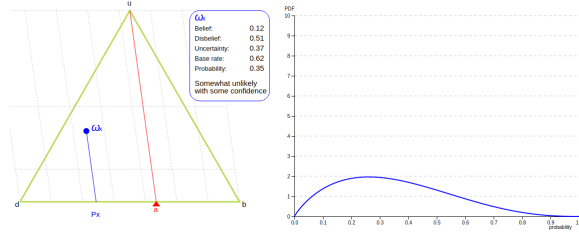
where the following constraints are satisfied:

$$Constraints \begin{cases} d_x \geq \frac{(1 - a_x)a_y(1 - b_x)d_y}{a_x(1 - a_y)(1 - b_y)} \\ u_x \geq \frac{a_y(1 - b_x)u_y}{a_x(1 - b_y)} \\ b_x \geq b_y \\ a_x > a_y \end{cases}$$

Co-Division Example 1:

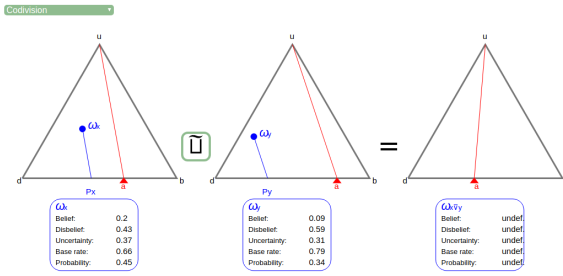
Input data: $\omega_{x_1} = (0.2, 0.43, 0.37, 0.72)$, $\omega_{x_2} = (0.09, 0.59, 0.31, 0.26)$





Co-Division Example 2:

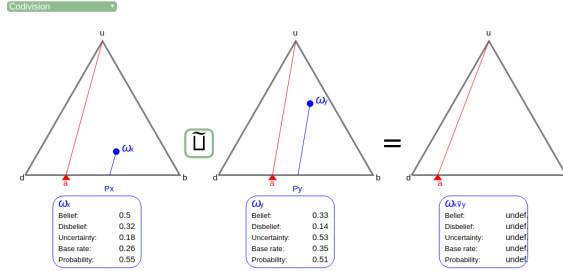
Input data: $\omega_{x_1} = (0.2, 0.43, 0.37, 0.66)$, $\omega_{x_2} = (0.09, 0.59, 0.31, 0.79)$



It was not possible to obtain a visualization for this sample case.

Co-Division Example 3:

Input data: $\omega_{x_1} = (0.5, 0.32, 0.18, 0.26)$, $\omega_{x_2} = (0.33, 0.14, 0.53, 0.35)$



It was not possible to obtain a visualization for this sample case.

5 Approaches to Reasoning

Various parts of the below essay have been adapted from *A Review of Uncertainty Handling Formalisms*, *Simon Parsons and Anthony Hunter*.

Over the decades, various approaches to reasoning have been developed, each of which provide a formalism to handle uncertainty and ignorance. Uncertainty is generally considered to be a subjective measure of the certainty of something and is thus modelled using a numerical value, typically between 0 and 1 with 0 denoting falsity and 1 denoting truth. Absence is the occurrence of missing facts, and is usually dealt with by essentially logical methods. The wide acceptance of the suggestion

that uncertainty and absence are essentially different, and must therefore be handled by different techniques has lead to a schism in approximate reasoning between the "symbolic camp" who use logical methods to deal with absence and the "numerical camp" who use quantitative measures to deal with uncertainty. Let us look at these approaches in detail and compare their advantages and disadvantages with Subjective Logic as proposed by Prof. Jøsang.

5.1 Numerical Approaches

Numerical Approaches towards handling uncertainty can be divided into three categories: Probability Theory, Possibility Theory and Evidence Theory. Even though these theories may appear significantly different, they are remarkably similar. The basic problem is how to weigh up the degree to which several uncertain events are believed to occur so that the most believed may be unambiguously identified. The basis on which the "belief" is assigned is a contentious issue, though all the theories that we shall consider assume allocation by an assignment function that distributes belief to possible events under consideration. Belief may be distributed on the basis of statistical information, physical possibility, or purely subjective assessment by an expert or otherwise. The belief assigned is a number between 0 and 1, with 0 being the belief assigned to a fact that is known to be false, and 1 the belief assigned to a fact known to be true. The infinite number of degrees of belief between the limits represent various shades of uncertainty.

5.1.1 Probability Theory

Probability theory has existed in one form or another for several hundred years. During this time various alternative formulations have been introduced, and it is now difficult to say where the definitive account may be found. Probability theory is built on three axioms or laws that define the behaviour of a probability measure, which may be used as an estimate of the degree to which an uncertain event is likely to occur. They are:

- i Convexity Law: $0 \leq Pr(A|H) \leq 1$
- ii Addition Law: $Pr(A \cup B|H) = Pr(A|H) + Pr(B|H)$
- iii Multiplication Law: $Pr(A \cap B|H) = Pr(A|H).Pr(B|A \cap H)$

The above laws form the basis for another set of important laws that define the use of probability theory in artificial intelligence:

- i Jeffrey's rule: $\sum_{i=1, \dots, n} Pr(A|B_i)Pr(B_i)$
- ii Bayes' Theorem: $Pr(A|B) = \frac{Pr(B|A).Pr(A)}{Pr(B)}$

5.1.2 Evidence Theory

Evidence Theory, otherwise known as Dempster Shafer Theory, was an attempt at remedying the limitations of probability theory. The theory deals with a frame of discernment: the set of elements $\theta = \theta_1, \theta_2$ which is assigned a *belief mass*, wherein:

$$m(\emptyset) = 0$$

$$\sum_{A \subseteq \theta} m(A) = 1$$

We can define the belief we have in a subset A as the sum of all the probability masses that support its constituents:

$$Bel(A) = \sum_{B \subseteq A} m(B)$$

$$Pl(A) = \sum_{B \cap A \neq \emptyset} m(B)$$

$$Pl(A) = 1 - Bel(\neg A)$$

The interval $[Bel(A), Pl(A)]$ can be considered to be a measure of our ignorance about A, and can vary from zero when we have the same degree of belief in A as would be generated by probability theory, to 1 when A has belief 0 and plausibility 1. This means that no mass is assigned to A or any of its subsets, but equally no mass is assigned to $\neg A$.

5.1.3 Possibility Theory

The economist G.L.S Shackle proposed a formalism wherein the calculus of *potential surprise* was defined as a subjective of the degree by which the observer would be surprised. If an event is entirely possible, then there would be no surprise attached to its occurrence.

A fuzzy set is a set whose membership is not absolute, but a matter of degree, such as the set of tall people. A fuzzy set F is characterised by the membership function μ_F which specifies the degree to which each object in a universe U is a member of F . The assignment of a value u to X has the form:

$$X = u : \mu_F(u)$$

A possibility distribution function π_x which is defined to be equal to the membership function of F:

$$\pi_x = \mu_F$$

Thus π_x , the possibility that $X = u$ is taken to be equal to $\mu_F(u)$.

5.1.4 Limitations of numerical techniques

Parsons and Hunter describe three limitations of numerical techniques and how various mathematicians have pointed them out:

- Lack of development of fuzzy sets and possibility theory
- Interpretation of results as they are in the form of numbers
- Computational expense

5.2 Symbolic Approaches

Nonmonotonic logics were designed with an intention to create sound formal mechanisms for reasoning in such a way that they overcome the limits of numerical approaches. There are many problems in AI that cannot be solved using classical logic, for example, problems that require "commonsense".

Nonmonotonic approaches can be divided into three categories: Default Logic, Argumentation and Truth Maintenance Systems

5.2.1 Default Logic

Default logic allows the modelling of prototypical reasoning by allowing special inference rules, known as default rules to be added to a standard first order logic.

Given a set of default rules D and a first order theory W , it is possible to define an extension of the default theory as the closure of W plus a maximal consistent set of consequences of D . An extension E of a default theory is a minimal set of beliefs that contain W are deductively closed and maximally consistent with the rules in D . Thus if S is a sentence then, $\Gamma(S)$ is a minimal set such that:

$$\begin{aligned} W &\subseteq \Gamma(S) \\ Th(\Gamma(S)) &= \Gamma(S) \end{aligned}$$

where $Th(T)$ is the deductive closure of T .

Despite the maturity of the theoretical work on default logic, there are as yet few applications because there has been less attention paid to providing prospective application builders with useful tools for using default logic.

5.2.2 Argumentation

An argument can be structured so that from facts a qualified claim can be argued only if:

- 1 there is some warrant that can be used with the facts to logically derive the claim
- 2 there is no other argument that would act as a rebuttal of the claim

An argument can be modelled by a pair (ϕ, α) where ϕ is a set of formulae, and α is a formula derived as a conclusion from the assumptions ϕ also known as grounds of the argument.

Argumentation can be used to handle uncertain information by extending the pair to a triple in which the third quantity is a measure of the degree to which α is believed to be true.

5.2.3 Truth Maintenance System

A Truth Maintenance System records information about each inference that is generated from a set of assumptions. There are two types of TMS:

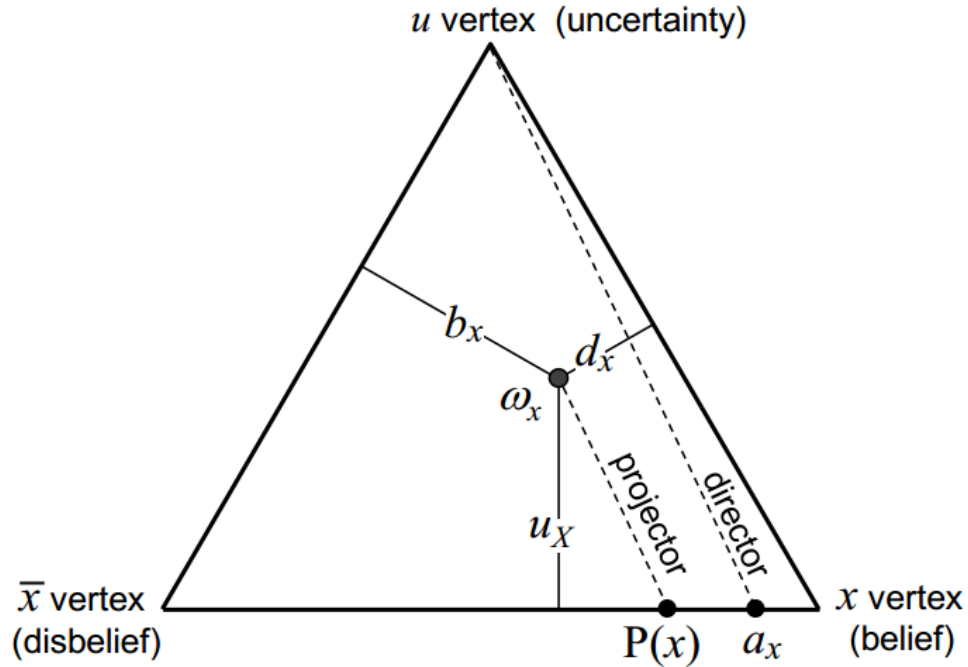
- Justification Based Truth Maintenance System: When an inconsistency in stored rules is detected, some external system is invoked to resolve the inconsistency.
- Assumption Based Truth Maintenance System: inconsistency is handled by creating new consistent subsets and identifying which inferences may be made from them.

5.3 Subjective Logic

The below definitions and diagram were obtained from *Subjective Logic: A Formalism for Handling Uncertainty*, Audun Jøsang, 2016

In binary logic a proposition about the state of the world must be either true or false, which fits well with an assumed objective world. Probability calculus takes argument probabilities in the range $[0,1]$, and hence to some extent reflects subjectivity by allowing propositions to be partially true. However, we are often unable to estimate probabilities with confidence because we lack the necessary evidence. A formalism for expressing degrees of uncertainty about beliefs is therefore needed in order to more faithfully reflect the perceived world in which we are all immersed. In addition, whenever a belief about a proposition is expressed, it is always done by an individual, and it can never be considered to represent a general and objective belief. It is therefore necessary that the formalism also includes belief ownership in order to reflect the fundamental subjectivity of all beliefs.

In subjective logic, we can represent an opinion as $\omega_x = (b_x, d_x, u_x, a_x)$, where b_x represents belief mass distribution, d_x represents disbelief, u_x represents uncertainty mass and a_x represents base rate.



The advantages of using subjective logic are that real-world situations can be realistically modelled with regard to how those situations are perceived, and that conclusions more correctly reflect the ignorance and uncertainties that necessarily result from partially uncertain input arguments. Subjective logic has the advantage of providing advanced operators for Dirichlet PDFs, for which no

practical analytical solutions previously existed. It should be mentioned that the simplicity of some SL operators comes at the cost of allowing those operators to be approximations of the analytically correct operators.