Subjective Logic

Draft, 22 July 2010

Audun Jøsang

University of Oslo

Web: http://persons.unik.no/josang/
Email: josang@unik.no

Preface

Subjective logic is a type of probabilistic logic that allows probability values to be expressed with degrees of uncertainty. Probabilistic logic combines the strengths of logic and probability calculus, meaning that it has binary logic's capacity to express structured argument models, and it has the power of probabilities to express degrees of truth of those arguments. Subjective logic makes it possible to express uncertainty about the probability values themselves, meaning that it is possible to reason with argument models in presence of uncertain or partially incomplete evidence. This manuscript describes the central elements of subjective logic. More specifically, it first describes a set of equivalent representations of uncertain probabilities with their interpretations. It then describes the most important subjective logic operators. Finally, it describes how subjective logic can be applied in trust modelling and for analysing Bayesian networks. Subjective logic is directly compatible with binary logic, probability calculus and classical probabilistic logic. The advantage of using subjective logic is that real world situations can be more realistically modelled, and that conclusions more correctly reflect the ignorance and uncertainties that accompany the input arguments.

Contents

Chapter 1

Introduction

In standard logic, propositions are considered to be either true or false. However, a fundamental aspect of the human condition is that nobody can ever determine with absolute certainty whether a proposition about the world is true or false. In addition, whenever the truth of a proposition is assessed, it is always done by an individual, and it can never be considered to represent a general and objective belief. This indicates that important aspects are missing in the way standard logic captures our perception of reality, and that it is more designed for an idealised world than for the subjective world in which we are all living.

Probabilistic logic was first defined by Nilsson [13] with the aim of combining the capability of deductive logic to exploit the structure and relationship of arguments and events, with the capacity of probability theory to express degrees of truth about those arguments and events. This results in more realistic models of real world situations than is possible with binary logic.

The additivity principle of classical probability requires that the probabilities of mutually disjoint elements in a state space add up to 1. This requirement makes it nessesary to estimate a probability value for every state, even though there might not be a basis for it. On other words, it prevents us from explicitly expressing ignorance about the possible states, outcomes or statements. If somebody wants to express ignorance about the state x as *"I don't know"* this would be impossible with a simple scalar probability value. A probability $P(x) = 0.5$ would for example mean that x and $\overline{x}$ are equally likely, which in fact is quite informative, and very different from ignorance. Alternatively, a uniform probability density function would more closely express the situation of ignorance about the outcome of the outcome.

Arguments in subjecetive logic are called *"subjective opinions"* or *"opinions"* for short. An opinion can contain degrees of uncertainty in the sense of *"uncertainty about probability estimates"*. The uncertainty of an opinion can be interpreted as ignorance about the truth of the relevant states, or as second order probaility about the first order probabilities.

Subjective opinions are related to belief functions. Belief theory has its origin in a model for upper and lower probabilities proposed by Dempster in 1960. Shafer later proposed a model for expressing belief functions [16]. The main idea behind belief theory is to abandon the additivity principle of probability theory, i.e. that the sum of probabilities on all pairwise disjoint states must add up to one. Instead belief theory gives observers the ability to assign so-called belief mass to the powerset of the state space. The main advantage of this approach is that ignorance, i.e. the lack of evidence about the truth of the states, can be explicitly expressed e.g. by assigning belief mass to the whole state space. Shafer's book [16] describes various aspects of belief theory, but the two main elements are 1) a flexible way of expressing beliefs, and 2) a conjunctive method for fusing belief functions, commonly known as Dempster's Rule. We will not be concerned with Dempster's rule here.

From one perspective, subjecetive opinions are more restrictive than belief functions because opinions do not apply to the powerset of states whereas classical belief functions do. However, subjective opinions include base rates, and in that sense have a richer expressiveness than belief functions.

Defining logic operators on subjective opinions is normally quite simple, and a relatively large set of practical logic operators exists. This provides the necessary framework for reasoning in a large variety of situations where input arguments can be incomplete or affected by uncertainty. Subjective opinions are equiavalent to Dirichlet and Beta probability density functions. Through this equivalence subjective logic provides a calculus for reasoning with probability density functions.

In this manuscript we describe the general principles of subjective logic. Four different but equivalent representations of subjective opinions are presented together with their interpretation. This allows uncertain probabilities to be seen from different angles, and allows an analyst to define models according to the formalisms that they are most familiar with, and that most naturally represents a specific real world situation. Subjective logic contains the same set of basic operators known from binary logic and classical probability calculus, but also contains some non-traditional operators which are specific to subjective logic.

The advantage of subjective logic over traditional probability calculus and probabilistic logic is that real world situations can be modeled and analysed more realistically. The analyst's partial ignorance and lack of information can be taken explicitly into account during the analysis, and explicitly expressed in the conclusion. When used for decision support, subjective logic allows decision makers to be better informed about uncertainties affecting the assessment of specific situations and future outcomes.

Chapter 2

Belief Representations in Subjective Logic

A fundamental aspect of the human condition is that nobody can ever determine with absolute certainty whether a proposition about the world is true or false. In addition, whenever the truth of a proposition is expressed, it is always done by an individual, and it can never be considered to represent a general and objective belief. These philosophical ideas are directly reflected in the mathematical formalism and belief representation of subjective logic.

Explicit expression of uncertainty is one of the main characteristics of subjective logic. Uncertainty comes in many flavours, and a good taxonomy is described in in [17]. In subjective logic, the uncertainty relates to probability values. For example, let the probability estimate of a future event x be expressed as $P(x) = 0.5$, e.g. for obtaining heads when flipping a coin. In subjective logic, the probability P as it stands without any additional parameters would be interpreted as totally certain because it expresses a crisp value, even though the outcome of the event is totally *uncertain*. The probability of an event is thus separated from the certainty/uncertainty of its probability. With this separation subjective logic can be applied in case of events with highly certain outcomes but where the probability can still be totally uncertain. This is possible by including the base rate of an event in the belief representation. The extreme case of a totally certain event combined with a totally uncertain probability is theoretically possible but is at the same time a singularity in subjective logic.

This chapter describes four different syntactic represention of beliefs that can be applied in subjective logic. Although quite different in notation, these representations are mathematically and semantically equivalent. The subjective opinion notaion is the classical and original representation used in subjecetive logic. Subjective opinions can be visualised in the form of opinion triangels and opinion simplexes which can aid human interpretation. The subjective opinion representation forms the basis for the subjective logic operators, and the other representations are useful to better understand the correspondence between subjective logic and other mathematical formalisms, for sollicitation of beliefs. The evidence representation, which is the second type, provides a classical mathematical representation often used in statistics which can also give useful and intuitive visualisations in the form of probability density functions. The evidence representation also provides the most intuitive way of including new evidence an observations into opinions. The probabilistic representation, which is the third type, might seem simple because it explicitly contains the probability expectation value. This representiation provides the most direct correspondence with probability calculus, but it does not seem to facilitate any particularly intuitive visualisations of uncertain probabilities. The fuzzy category representation is the fourth type and provides a way of expressing opinions in terms of common verbal expressions such as *"unlikely"* or *"very likely"*. This representation is mos useful for sollicitating input from humans, and for easy interpration of results by analysts. This variety of alternative but equivalent representations provides a high level of flexibility for collecting input arguments from various sources and for facilitating optimal human interpretation in various applications.

2.1 Subjective Opinion Representation

Subjective opinions express subjective beliefs about the truth of propositions with degrees of uncertainty, and can indicate subjective belief ownership whenever required. A distinction can be made between multinomial and binomial opinions. An multinomial opinion is denoted as ω_X^A where A is the belief owner, also called the subject, and X is the target frame, also called state space, to which the opinion applies. An alternative notation is $\omega(A : X)$. In case of binomial opinions, the notation is ω_x^A, or alternatively $\omega(A : x)$, where x is a single proposition that is assumed to belong to a frame e.g. denoted as X, but the frame is usually not included in the notation for binomial opinions. The propositions of a frame are normally assumed to be exhaustive and mutually disjoint, and belief owners are assumed to have a common semantic interpretation of propositions. The belief owner (subject) and the propositions (object) are attributes of an opinion. Indication of subjective belief ownership can be omitted whenever irrelevant.

2.1.1 Multinomial Opinions

A general multinomial opinion is a composite function consisting of a belief vector $\vec{b}$, an uncertainty mass u and a base rate vector $\vec{a}$. These components are defined next.

Definition 1 Belief Mass Vector
Let $X = \{x_i | i = 1, \dots k\}$ be a frame and let $\vec{b}$ be a vector function from X to $[0, 1]^k$ representing belief masses over X satisfying:

$$\vec{b}(\emptyset) = 0 \quad \textit{and} \quad \sum_{x \in X} \vec{b}(x) \leq 1 \, . \tag{2.1}$$

Then $\vec{b}$ is called a belief mass vector, or belief vector for short.

An element $\vec{b}(x_i)$ is interpreted as belief mass over x, i.e. the amount of positive belief that x is true. The belief vector can be interpreted as a sub-additive probability function because the sum can be less than one. Additivity is achieved by including the uncertainty mass defined below.

Definition 2 Uncertainty Mass
Let $X = \{x_i | i = 1, \dots k\}$ be a frame with a belief vector $\vec{b}$. Let u be a function from X to $[0,1]$ representing uncertainty over X satisfying:

$$u + \sum_{x \in X} \vec{b}(x) = 1\ . \tag{2.2}$$

The parameter u is then called an uncertainty mass.

The uncertainty mass can be interpreted as the lack of committed belief mass in the truth of any of the propositions of X. In other words, uncertainty mass reflects that the belief owner does not know which of the propositions of X in particular is true, only that one of them must be true.

In case the belief vector is subadditive, i.e. $\sum_{x \in X} \vec{b}(x) < 1$, the base rate vector will play a role in determining probability expecation values over X. The base rate vector is defined below.

Definition 3 Base Rate Vector
Let $X = \{x_i | i = 1, \dots k\}$ be a frame and let $\vec{a}$ be a vector function from X to $[0,1]^k$ representing non-informative a priori *probability over X before any evidence has been received, satisfying:*

$$\vec{a}(\emptyset) = 0 \quad \textit{and} \quad \sum_{x \in X} \vec{a}(x) = 1\ . \tag{2.3}$$

Then $\vec{a}$ is called a base rate vector.

Having defined the belief vector, the uncertainty mass and the base rate vector, the general opinion can be defined.

Definition 4 Subjective Opinion
Let $X = \{x_i | i = 1, \dots k\}$ be a frame, i.e. a set of exhaustive and mutually disjoint propositions x_i. Let $\vec{b}$ be a belief vector, let u be the corresponding uncertainty mass, and let $\vec{a}$ be a base rate vector over X, all seen from the viewpoint of a subject entity A. The composite function $\omega_X^A = (\vec{b}, u, \vec{a})$ is then A's subjective opinion over X. This represents the traditional belief notation of opinions.

We use the convention that the subscript on the multinomial opinion symbol indicates the frame on which the opinion applies, and that a superscript indicates the subject owner of the opinion. Subscripts can be omitted when it is clear and implicitly assumed to which frame an opinion applies, and superscripts can be omitted when it is irrelevant who the owner is.

Assuming that the frame X has cardinality k, the belief vector $\vec{b}$ and the base rate vector $\vec{a}$ will have k parameters each. The uncertainty parameter u is a simple scalar. A multinomial opinion in belief notation over a frame of cardinality k will thus contain

$2k+1$ parameters. However, given the constraints of Eq.(2.2) and Eq.(2.3), the multinomial opinion will only have $2k-1$ degrees of freedom. A binomial opinion will for example be 3-dimensional.

The introduction of the base rate vector allows the probabilistic transformation to be independent from the internal structure of the frame. The probability expectation of multinomial opinions is a vector expressed as a function of the belief vector, the uncertainty mass and the base rate vector.

Definition 5 Probability Expectation Vector
Let $X = \{x_i | i = 1, \dots k\}$ be a frame and let ω_X be an opinion on X with belief vector $\vec{b}$ and uncertainty mass u. Let $\vec{a}$ be a base rate vector on X. The function $\vec{\mathrm{E}}_X$ from X to $[0,1]^k$ expressed as:

$$\vec{\mathrm{E}}_X(x_i) = \vec{b}(x_i) + \vec{a}(x_i)u\ . \tag{2.4}$$

is then called the probability expectation vector over X.

It can be shown that $\vec{\mathrm{E}}_X$ satisfies the additivity principle:

$$\vec{\mathrm{E}}_X(\emptyset) = 0 \quad \text{and} \quad \sum_{x \in X} \vec{\mathrm{E}}_X(x) = 1\ . \tag{2.5}$$

The base rate vector of Def.3 expresses non-informative *a priori* probability, whereas the probability expectation function of Eq.(2.4) expresses informative *a posteriori* probability.

Given a frame of cardinality k, the default base rate of for each element in the frame is $1/k$, but it is possible to define arbitrary base rates for all mutually exclusive elements of the frame, as long as the additivity constraint is satisfied.

Two different multinomial opinions on the same frame will normally share the same base rate vectors. However, it is obvious that two different observers can assign different base rates to the same frame, in addition to assigning different beliefs to the frame. This naturally reflects different views, analyses and interpretations of the same situation by different observers.

Visualising multinomial opinions is not trivial. The largest opinions that can be easily visualised are trinomial, in which case it can be represented as a point inside a tetrahedron, as shown in Fig.2.1.

In Fig.2.1, the vertical elevation of the opinion point inside the tetrahedron represents the uncertainty mass. The distances from each of the three triangular side planes to the opinion point represents the respective belief mass values. The base rate vector $\vec{a}_X$ is indicated as a point on the base plane. The line that joins the tetrahedron summit and the base rate vector point represents the director. The probability expectatioan vector poinit is geomerically determined by drawing a projection from the opinion point parallel to the director onto the base plane.

In general, the triangle and tetrahedron belong to the *simplex* family of geometrical shapes. Multinomial opinions on frames of cardinality k can in general be represented as a point in a simplex of dimension $(k+1)$. For example, binomial opinions can be represented inside a triangle which is a 3D simplex, and trinomial opinions can be represented inside a tetrahedron which is a 4D The 2D aspect of paper and computer

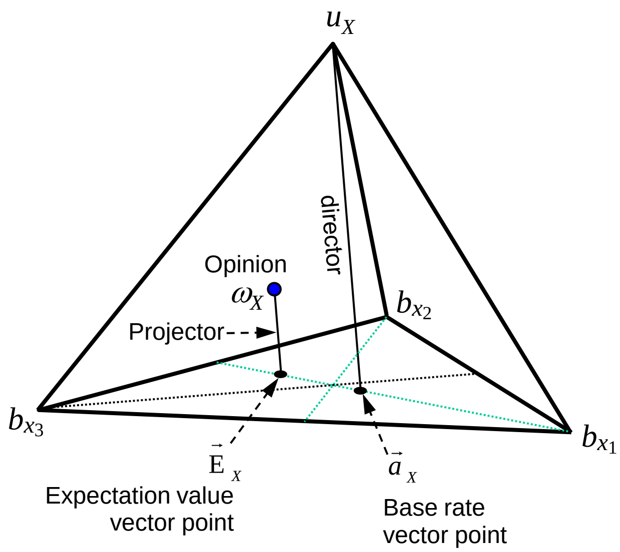


Figure 2.1: Opinion tetrahedron with example opinion

screens makes visualisation of larger multinomial opinions impractical. Opinions with dimensions larger than trinomial do not lend themselves to traditional visualisation.

2.1.2 Binomial Opinions

Opinions over binary frames are called binomial opinions, and a special notation is used for their mathematical representation. A general n-ary frame X can be considered binary when seen as a binary partitioning consisting of one of its proper subsets x and the complement $\overline{x}$.

Definition 6 (Binomial Opinion) *Let $X = \{x, \overline{x}\}$ be either a binary frame or a binary partitioning of an n-ary frame. A binomial opinion about the truth of state x is the ordered quadruple $\omega_x = (b, d, u, a)$ where:*

b	**belief**	*is the belief mass in support of x being true,*
d	**disbelief**	*is the belief mass in support of x being false,*
u	**uncertainty**	*is the amount of uncommitted belief mass,*
a	**base rate**	*is the* a priori *probability in the absence of committed belief mass.*

These components satisfy $b + d + u = 1$ and $b, d, u, a \in [0, 1]$. The characteristics of various binomial opinion classes are listed below. A binomial opinion:

where $b = 1$	is equivalent to binary logic TRUE,
where $d = 1$	is equivalent to binary logic FALSE,
where $b + d = 1$	is equivalent to a traditional probability,
where $b + d < 1$	expresses degrees of uncertainty, and
where $b + d = 0$	expresses total uncertainty.

The probability expectation value of a binomial opinion is defined as $\mathrm{E}_x = b + au$.

Binomial opinions can be represented on an equilateral triangle as shown in Fig.2.2 below. A point inside the triangle represents a (b, d, u) triple. The belief, disbelief, and uncertainty-axes run from one edge to the opposite vertex indicated by the b_x axis, d_x axis and u_x axis labels. For example, a strong positive opinion is represented by a point towards the bottom right belief vertex. The base rate[1], is shown as a point on the base line, and the probability expectation, E_x, is formed by projecting the opinion point onto the base, parallel to the base rate director line. The opinion $\omega_x = (0.2,\ 0.5,\ 0.3,\ 0.6)$ with expectation value $\mathrm{E}_x = 3.8$ is shown as an example.

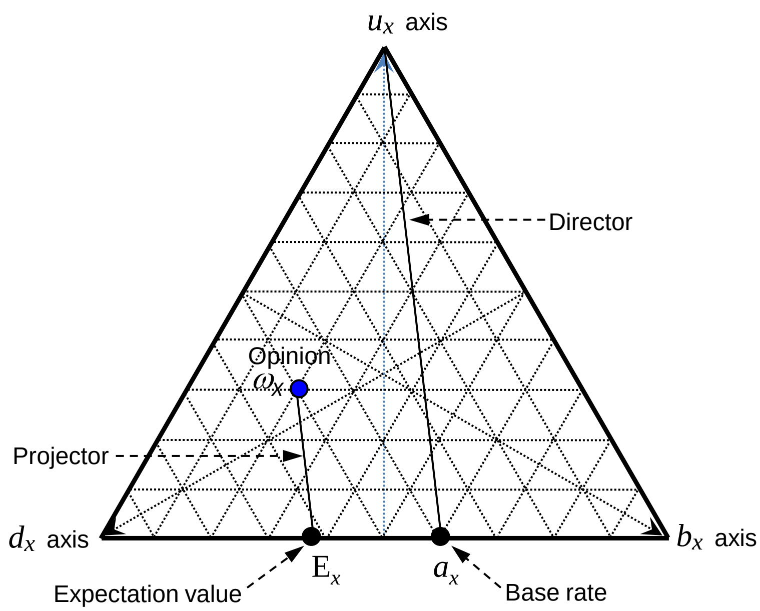


Figure 2.2: Opinion triangle with example opinion

When the opinions values are constrained to one of the three vertices, i.e. with $b = 1$, $d = 1$ or $u = 1$, subjective logic becomes a form of three-valued logic that is compatible with Kleene logic [1]. However, the three-valued arguments of Kleene logic do not contain base rates, so that the expectation value of binomial opinions can not be derived from Kleene logic arguments.

2.2 Evidence Representation

The evidence representiation of opinions is centered around the Dirichlet multinomial probability distribution. For self-containment, we briefly outline the Dirichlet multinomial model below, and refer to [2] for more details.

[1] Sometimes called relative atomicity

2.2.1 The Dirichlet Distribution

The general multinomial probability distribution over an n-ary frame is described by the multinomial Dirichlet distribution, where the special case of a probability distribution over a binary frame is described by the binomial Beta distribution. In general, the probability distribution over a k-component random variable $p(x_i),\ i = 1 \ldots k$ with sample space $[0,1]^k$, subject to the simple additivity constraint $\sum\limits_{i=1}^{k} p(x_i) = 1$. Because of this additivity requirement, the Dirichlet distribution has only $k-1$ degrees of freedom. This means that the knowledge of $k-1$ probability variables and their density determines the last probability variable and its density.

The Dirichlet distribution takes as argument a sequence of observations of the k possible outcomes represented as k positive real parameters $\vec{\alpha}(x_i),\ i = 1 \ldots k$, each corresponding to one of the possible outcomes.

Definition 7 Dirichlet Distribution

Let X be a frame consisting of k mutually disjoint elements. Let $\vec{\alpha}$ represent the evidence vector over the elements of X. In order to have a compact notation we define the vector $\vec{p} = \{\vec{p}(x_i) \mid 1 \leq i \leq k\}$ to denote the k-component random probability variable, and the vector $\vec{\alpha} = \{\vec{\alpha}(x_i) \mid 1 \leq i \leq k\}$ to denote the k-component random input argument vector $[\vec{\alpha}(x_i)]_{i=1}^{k}$. Then the multinomial Dirichlet density function over X, denoted as Dirichlet$(\vec{\alpha})$, can then be expressed as:

$$\textit{Dirichlet}(\vec{\alpha}) = f(\vec{p} \mid \vec{\alpha}) = \frac{\Gamma\left(\sum\limits_{i=1}^{k} \vec{\alpha}(x_i)\right)}{\prod\limits_{i=1}^{k} \Gamma(\vec{\alpha}(x_i))} \prod_{i=1}^{k} p(x_i)^{(\vec{\alpha}(x_i)-1)} \tag{2.6}$$

where $\vec{\alpha}(x_1), \ldots \vec{\alpha}(x_k) \geq 0$.

The vector $\vec{\alpha}$ represents the *a priori* 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 possible outcomes as a function of the base rate. It is normally assumed that $W = 2$.

The singleton elements in a frame of cardinality k can have a base rate different from the default value $1/k$. It is thereby possible to define a base rate as a vector $\vec{a}$ with arbitrary distribution over the k mutually disjoint elements x_i, as long as the simple additivity requirement expressed as $\sum_{x_i \in X} \vec{a}(x_i) = 1$ is satisfied. The total evidence $\alpha(x_i)$ for each element x_i can then be expressed as:

$$\vec{\alpha}(x_i) = \vec{r}(x_i) + W\,\vec{a}(x_i)\,, \quad \text{where} \begin{cases} \vec{r}(x_1), \ldots \vec{r}(x_k) \geq 0 \\ \vec{a}(x_1), \ldots \vec{a}(x_k) \in [0,1] \\ \sum_{i=1}^{k} \vec{a}(x_i) = 1 \\ W \geq 2 \end{cases} \tag{2.7}$$

The Dirichlelt distribution over a set of k possible states x_i can thus be represented as a function of the base rate vector $\vec{a}$, the non-informative prior weight W and the observation evidence $\vec{r}$.

The notation of Eq.(2.7) is useful, because it allows the determination of the probability distribution over frames where each element can have an arbitrary base rate. Given the Dirichlet distribution of Eq.(2.6), the probability expectation of any of the k random probability variables can now be written as:

$$\vec{\mathrm{E}}(\vec{p}(x_i) \mid \vec{r}, \vec{a}) = \frac{\vec{r}(x_i) + W\vec{a}(x_i)}{W + \sum_{i=1}^{k} \vec{r}(x_i)} \,. \tag{2.8}$$

It is normally assumed that the *a priori* distribution in case of a binary frame $X = \{x, \overline{x}\}$ is uniform. This requires that $\vec{\alpha}(x) = \vec{\alpha}(\overline{x}) = 1$, which in turn dictates that $W = 2$. Assuming an *a priori* uniform distribution over frames other than binary will require a different value for W. The non-informative prior weight W will always be equal to the cardinality of the frame over which a uniform distribution is assumed.

Selecting $W > 2$ will result in new observations having relatively less influence over the Dirichlet distribution. This could be meaningful e.g. as a representation of specific *a priori* information provided by a domain expert. It can be noted that it would be unnatural to require a uniform distribution over arbitrary large frames because it would make the sensitivity to new evidence arbitrarily small.

For example, requiring a uniform *a priori* distribution over a frame of cardinality 100, would force $W = 100$. In case an event of interest has been observed 100 times, and no other event has been observed, the derived probability expectation of the event of interest will still only be about $\frac{1}{2}$, which would be rather counterintuitive. In contrast, when a uniform distribution is assumed in the binary case, and the same evidence is analysed, the derived probability expectation of the event of interest would be close to 1, as intuition would dictate.

2.2.2 Visualising Dirichlet Distributions

Visualising Dirichlet distributions is challenging because it is a density function over $k-1$ dimensions, where k is the frame cardinality. For this reason, Dirichlet distributions over ternary frames are the largest that can be practically visualised.

With $k = 3$, the probability distribution has 2 degrees of freedom, and the additivity equation

$$\vec{p}(x_1) + \vec{p}(x_2) + \vec{p}(x_3) = 1 \tag{2.9}$$

defines a triangular plane as illustrated in Fig.2.3.

In order to visualise probability density over the triangular plane, it is convenient to lay the triangular plane horizontally in the X-Y plane, and visualise the density dimension along the Z-axis.

Let us consider the example of an urn containing balls of the three different markings: x_1, x_2 and x_3, meaning that the urn can contain multiple balls marked x_1, x_2 and x_3 respectively. Because of the three different markings, the cardinality of this frame is 3 (i.e. $k = 3$). Let us first assume that no other information than the cardinality is

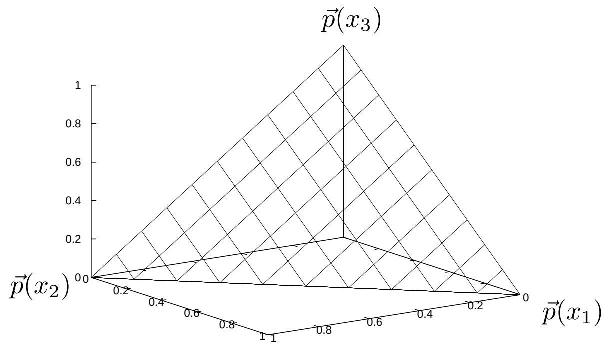
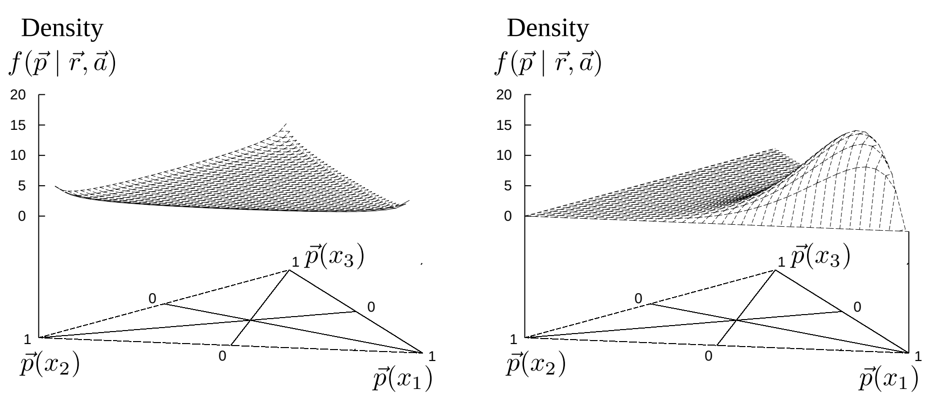


Figure 2.3: Triangular plane of trinomial probability additivity

available, meaning that the number and relative proportion of balls marked x_1, x_2 and x_3 are unknown, and that the default base rate for any of the markings is $a = 1/k = \frac{1}{3}$. Initilly, before any balls have been drawn we have $\vec{r}(x_1) = \vec{r}(x_2) = \vec{r}(x_3) = 0$. Then Eq.(2.8) dictates that the expected *a priori* probability of picking a ball of any specific marking is the default base rate probability $a = \frac{1}{3}$. The non-informative *a priori* Dirichlet density function is illustrated in Fig.2.4.a.

Let us now assume that an observer has picked (with return) 6 balls market x_1, 1 ball marked x_2 and 1 ball marked x_3, i.e. $\vec{r}(x_1) = 6,\ \vec{r}(x_2) = 1,\ \vec{r}(x_3) = 1$, then the *a posteriori* expected probability of picking a ball marked x_1 can be computed as $\vec{\mathrm{E}}(x_1) = \frac{2}{3}$. The *a posteriori* Dirichlet density function is illustrated in Fig.2.4.b.



(a) Non-informative prior Dirichlet distribution (b) *A posteriori* Dirichlet distribution

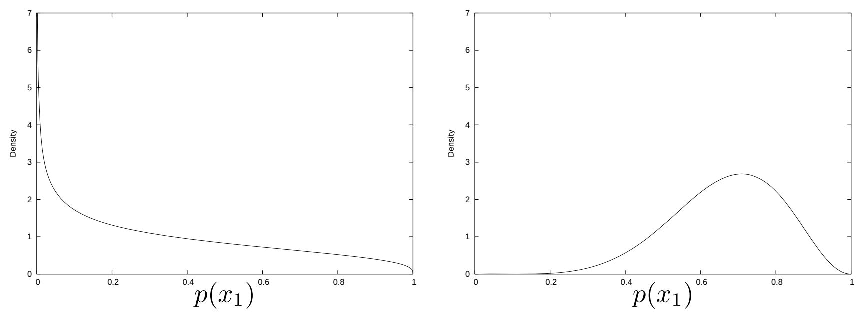
Figure 2.4: Prior and posterior Dirichlet distributions

2.2.3 Coarsening Example: From Ternary to Binary

We reuse the example of Sec.2.2.2 with the urn containing balls marked x_1, x_2 and x_3, but this time we consider a binary partition of the markings into x_1 and $\overline{x}_1 = \{x_2, x_3\}$. The base rate of picking x_1 is set to the relative atomicity of x_1, expressed as $\vec{a}(x_1) = \frac{1}{3}$. Similarly, the base rate of picking $\overline{x}_1$ is $\vec{a}(\overline{x}_1) = \vec{a}(x_2) + \vec{a}(x_3) = \frac{2}{3}$.

Let us again assume that an observer has picked (with return) 6 balls marked x_1, and 2 balls marked $\overline{x}_1$, i.e. marked x_2 or x_3. This translates into the evidence vector $\vec{r}(x_1) = 6,\ \vec{r}(\overline{x}_1) = 2$.

Since the frame has been reduced to binary, the Dirichlet distribution is reduced to a Beta distribution which is simple to visualise. The *a priori* and *a posteriori* density functions are illustrated in Fig.2.5.



(a) Non-informative *a priori* Beta distribution

(b) After 6 balls marked x_1, and 2 balls marked x_2 or x_3

Figure 2.5: Prior and posterior Beta distributions

Computing the *a posteriori* expected probability of picking ball marked x_1 with Eq.(2.8) produces $\mathrm{E}(x_1) = \frac{2}{3}$, which is the same as before the coarsening, as illustrated in Sec.2.2.2. This shows that the coarsening does not influence the probability expectation value of specific events.

2.2.4 Evidence Notation of Opinions

Dirichlet distributions translate observation evidence directly into probability density functions. The representation of the observation evidence, together with the base rate, can be used to denote opinions.

Definition 8 Evidence Notation
Let X be a frame with a Dirichlet distribution Dirichlet$(\vec{p} \mid \vec{r}, \vec{a})$. The evidence notation of opinions can then be expressed as the ordered tuple $\omega = (\vec{r}, \vec{a})$.

Let $\omega = (\vec{r}, \vec{a})$ be an opinion in evidence notation over a frame with cardinality k. Then all the k parameters of $\vec{r}$ are independent, whereas only $k-1$ parameters of $\vec{a}$ are independent because of Eq.(2.3). As expected, the evidence notation therefore has $2k-1$ dimensions of freedom, as does the belief notation.

It is possible to define a bijective mapping between the evidence notation described in Sec.2.2.1, and the belief notation of opinions described in Sec.2.1.1.

Let $X = \{x_i \mid i = 1, \cdots, k\}$ be a frame. Let $\omega_X^{\text{bn}} = (\vec{b}, u, \vec{a})$ be an opinion on X in belief notation, and let $\omega_X^{\text{en}} = (\vec{r}, \vec{a})$ be an opinion on X in evidence notation.

For the bijective mapping between ω_X^{bn} and ω_X^{en}, we require equality between the probability expectation values $\vec{\mathrm{E}}_X(x_i)$ derived from ω_X^{bn}, and those derived from ω_X^{en}. This constraint is expressed as:

For all $x_i \in X$:

$$\begin{aligned} \vec{\mathrm{E}}(\omega_X^{\text{bn}}) &= \vec{\mathrm{E}}(\omega_X^{\text{en}}) \qquad (2.10)\\ &\Updownarrow \\ \vec{b}(x_i) + \vec{a}(x_i)u &= \frac{\vec{r}(x_i)}{W + \sum_{i=1}^{k} \vec{r}(x_i)} + \frac{W\vec{a}(x_i)}{W + \sum_{i=1}^{k} \vec{r}(x_i)} \qquad (2.11) \end{aligned}$$

We also require that each belief mass $\vec{b}(x_i)$ be an increasing function of of the evidence $\vec{r}(x_i)$, and that u be a decreasing function of $\sum\limits_{i=1}^{k} \vec{r}(x_i)$. In other words, the more evidence in favour of a particular outcome x_i, the greater the belief mass on that outcome. Furthermore, the more total evidence available, the less uncertain the opinion. As already mentioned it is normally assumed that $W = 2$.

In case $u \longrightarrow 0$, then $\sum\limits_{i=1}^{k} \vec{b}(x_i) \longrightarrow 1$, and $\sum\limits_{i=1}^{k} \vec{r}(x_i) \longrightarrow \infty$, meaning that at least some, but not necessarily all, of the evidence parameters $\vec{r}(x_i)$ are infinite. We define $\vec{\eta}(x_i)$ as the the relative degree of infinity between the corresponding infinite evidence parameters $\vec{r}(x_i)$ such that $\sum\limits_{i=1}^{k} \vec{\eta}(x_i) = 1$. When infinite evidence parameters exist, any finite evidence parameter $\vec{r}(x_i)$ can be assumed to be zero in any practical situation because it will have $\vec{\eta}(x_i) = 0$, i.e. it will carry zero weight relative to the infinite evidence parameters.

These intuitive requirements together with Eq.(2.11) imply the following bijective mapping:

Theorem 1 Evidence Notation Equivalence
Let $\omega_X^{\text{bn}} = (\vec{b}, u, \vec{a})$ be an opinion expressed in belief notation, and $\omega_X^{\text{en}} = (\vec{r}, \vec{a})$ be an opinion expressed in evidence notation, both over the same frame X. The opinions ω_X^{bn} and ω_X^{en} are equivalent when the following equivalent mappings hold:

For $u \neq 0$:

$$\left\{\begin{array}{ll} \vec{b}(x_i) & = \frac{\vec{r}(x_i)}{W + \sum_{i=1}^{k} \vec{r}(x_i)} \\ \\ u & = \frac{W}{W + \sum_{i=1}^{k} \vec{r}(x_i)} \end{array}\right. \quad \Leftrightarrow \quad \left\{\begin{array}{l} \vec{r}(x_i) = \frac{W\vec{b}(x_i)}{u} \\ \\ 1 = u + \sum\limits_{i=1}^{k} \vec{b}(x_i) \end{array}\right. \qquad (2.12)$$

For $u = 0$*:*

$$\begin{cases} \vec{b}(x_i) & = \vec{\eta}(x_i) \\ \\ u & = 0 \end{cases} \quad \Leftrightarrow \quad \begin{cases} \vec{r}(x_i) & = \vec{\eta}(x_i) \sum_{i=1}^{k} \vec{r}(x_i) = \vec{\eta}(x_i)\infty \\ \\ 1 & = \sum_{i=1}^{k} \vec{b}(x_i) \end{cases} \tag{2.13}$$

In case $\vec{\eta}(x_i) = 1$ for a particular evidence parameter $\vec{r}(x_i)$, then $\vec{r}(x_i) = \infty$ and all the other evidence parameters are finite. In case $\vec{\eta}(x_i) = 1/k$ for all $i = 1 \dots k$, then all the evidence parameters are all equally infinite.

Binomial opinions correspond to Beta distributions that are normally denoted as Beta(α, β) where α and β are its two evidence parameters. Assuming that the uncertainty mass is non-zero, the Beta distribution of the binomial opinion $\omega_x = (b, d, u, a)$ is the function

$$\text{Beta}_x(\alpha, \beta) \text{ where } \begin{cases} \alpha & = Wb/u + Wa \\ \beta & = Wd/u + W(1-a) \end{cases} \tag{2.14}$$

As before, the default non-informative prior weight is $W = 2$, but larger values are also possible. A binomial opinion of a binary outcome x will in evidence notation be written as $\omega_x^{\text{en}} = (r, s, a)$, where r and s represent the amount in favour of, and against x respectively, and a represents the base rate of x.

2.3 Probabilistic Notation of Opinions

A disadvantage of the belief notation described in Sec.2.1 and the evidence notation described in Sec.2.2 is that they do not explicitly express probability expectation values of the elements in the frame. The classical probabilistic notation has the advantage that it is used in all areas of science and that people are familiar with it. The probability expectation of an opinion can easily be derived with Eq.(2.4), but this still represents a mental barrier to a direct intuitive interpretation of opinions. An intuitive representation of multinomial opinions could therefore be to represent the probability expectation value directly, together with the degree of uncertainty and the base rate function. This will be called the *probabilistic notation* of opinions:

Definition 9 Probabilistic Notation of Opinions
Let X be a frame and let ω_X^{bn} be an opinion on X in belief notation. Let $\vec{\text{E}}$ be a multinomial probability expectation function on X defined according to Def.5, let $\vec{a}$ be a multinomial base rate function on X defined according to Def.3, and let $c = 1 - u$ be the certainty function on X, where u is defined according to Def.(6). The probabilistic notation of opinions can then be expressed as the ordered tuple $\omega_X^{\text{pn}} = (\vec{\text{E}}, c, \vec{a})$.

In case $c = 1$, then $\vec{\text{E}}$ is a traditional discrete probability distribution without uncertainty. In case $c = 0$, then $\vec{\text{E}} = \vec{a}$, and no evidence has been received, so the probability distribution $\vec{\text{E}}$ is totally uncertain.

The equivalence between the belief notation and the probabilistic notation of opinions is defined below.

Theorem 2 Probabilistic Notation Equivalence
Let $\omega_X^{\mathrm{bn}} = (\vec{b}, u, \vec{a})$ be an opinion expressed in belief notation, and $\omega_X^{\mathrm{pn}} = (\vec{\mathrm{E}}, c, \vec{a})$ be an opinion expressed in probabilistic notation, both over the same frame X. The opinions ω_X^{bn} and ω_X^{pn} are equivalent when the following equivalent mappings hold:

$$\begin{cases} \vec{\mathrm{E}}(x_i) & = \vec{b}(x_i) + \vec{a}(x_i)u \\ c & = 1 - u \end{cases} \Leftrightarrow \begin{cases} \vec{b}(x_i) & = \vec{\mathrm{E}}(x_i) - \vec{a}(x_i)(1-c) \\ u & = 1 - c \end{cases} \tag{2.15}$$

Let the frame X have cardinality k. Then both the base rate vector $\vec{a}$ and the probability expectation vector $\vec{\mathrm{E}}(x_i)$ have $k-1$ degrees of freedom due to the additivity property of Eq.(2.3) and Eq.(2.5). With the addition of the independent certainty parameter c, the probabilistic notation of opinions has $2k-1$ degrees of freedom, as do the belief notation and the evidence notation of opinions.

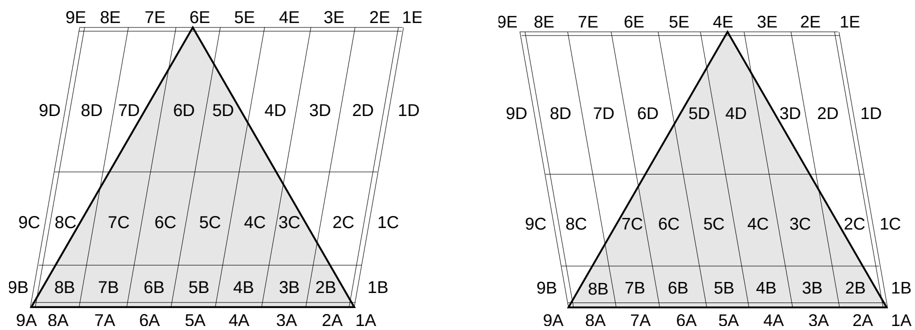
2.4 Fuzzy Category Representation

Human language provides various terms that are commonly used to express various types of likelihood and uncertainty. It is possible to express binomial opinions in terms of fuzzy verbal categories which can be specified according to the need of a particular application. An example set of fuzzy categories is provided in Table 2.1.

Likelihood Categories		Absolutely Not	Very Unlikely	Unlikely	Somewhat Unlikely	Chances about even	Somewhat Likely	Likely	Very Lkely	Absolutely
Certainty Categories		**9**	**8**	**7**	**6**	**5**	**4**	**3**	**2**	**1**
Completely Uncertain	**E**	9E	8E	7E	6E	5E	4E	3E	2E	1E
Very Uncertain	**D**	9D	8D	7D	6D	5D	4D	3D	2D	1D
Uncertain	**C**	9C	8C	7C	6C	5C	4C	3C	2C	1C
Slightly Uncertain	**B**	9B	8B	7B	6B	5B	4B	3B	2B	1B
Completely Certain	**A**	9A	8A	7A	6A	5A	4A	3A	2A	1A

Table 2.1: Fuzzy Categories

These fuzzy verbal categories can be mapped to areas in the opinion triangle as illustrated in Fig.2.6. The mapping must be defined for combinations of ranges of expectation value and uncertainty. As a result, the mapping between a specific fuzzy category from Table 2.1 and specific geometric area in the opinion triangle depends on the base rate. Without specifying the exact underlying ranges, the visualization of Fig.2.6 indicates the ranges approximately. The edge ranges are deliberately made narrow in order to have categories for near dogmatic and vacuous beliefs, as well as beliefs that express expectation values near absolute 0 or 1. The number of likelihood categories, and certainty categories, as well as the exact ranges for each, must be determined according to the need of each application, and the fuzzy categories defined here must be seen as an example. Real-world categories would likely be similar to those found in Sherman Kent's *Words of Estimated Probability* [11]; based on the *Admiralty Scale* as used within the UK National Intelligence Model[2]; or could be based on empirical results obtained from psychological experimentation.



(a) Fuzzy categories with $a = \frac{1}{3}$ (b) Fuzzy categories with $a = \frac{2}{3}$

Figure 2.6: Mapping fuzzy categories to ranges of belief as a function of the base rate

Fig.2.6 illustrates category-opinion mappings in the case of base rate $a = \frac{1}{3}$, and the case of base rate $a = \frac{2}{3}$. The mapping is determined by the overlap between category area and triangle region. Whenever a fuzzy category area overlaps, partly or completely, with the opinion triangle, that fuzzy category is a possible mapping.

Note that the fuzzy category areas overlap with different regions on the triangle depending on the base rate. For example, it can be seen that the category $7D$: "Unlikely and Very Uncertain" is possible in case $a = \frac{1}{3}$, but not in case $a = \frac{2}{3}$. This is because the expectation of a state x is defined as $\mathrm{E}(x) = b_x + a_x u_x$, so that when $a_x, u_x \longrightarrow 1$, then $\mathrm{E}(x) \longrightarrow 1$, meaning that the likelihood category "Unlikely" would be impossible.

Mapping from fuzzy categories to subjective opinions is also straight-forward. Geometrically, the process involves mapping the fuzzy adjectives to the corresponding center of the portion of the grid cell contained within the opinion triangle (see Fig.2.6). Naturally, some mappings will always be impossible for a given base rate, but these are logically inconsistent and should be excluded from selection.

It is interesting to notice that although a specific fuzzy category maps to different geometric areas in the opinion triangle depending on the base rate, it will always correspond to the same range of beta PDFs. It is simple to visualize ranges of binomial opinions with the opinion triangle, but it would not be easy to visualize ranges of beta PDFs. The mapping between binomial opinions and beta PDFs thereby provides a very powerful way of describing PDFs in terms of fuzzy categories, and vice versa.

[2] http://www.policereform.gov.uk/implementation/natintellmodel.html

Chapter 3

Properties of Subjective Logic

3.1 Probabilistic Logic

Probabilistic logic combines probability calculus and logic modelling of inference and propositions. For example, the AND operator in binary logic which is defined by the well-known truth table for AND, can simply be modelled as the product operator for probabilities. By providinig probability arguments as either $p = 1$ which is equivalent to TRUE, or as $p = 0$ which is equivalent to FALSE, the truth table of AND emerges. Similarly, many other binary logic formulas which otherwise are defined as axioms can be simply derived from probability calculus formulas.

The table below provides examples of how probabilistic generalises binary logic formulas:

Binary Logic		**Probabilistic Logic**	
AND:	$x \wedge y$	Product:	$p(x \wedge y) = p(x)p(y)$
OR;	$x \vee y$	Coproduct:	$p(x \vee y) = p(x) + p(y) - p(x)p(y)$
MP:	$\{x \rightarrow y,\ x\} \Rightarrow y$	Deduction:	$p(y\|x) = p(x)p(y\vert x) + p(\overline{x})p(y\vert\overline{x})$
MT:	$\{x \rightarrow y,\ \overline{y}\} \Rightarrow \overline{x}$	Abduction:	$p(x\vert y) = \frac{a(x)p(y\vert x)}{a(x)p(y\vert x)+a(\overline{x}p(y\vert\overline{x})}$ $p(x\vert\overline{y}) = \frac{a(x)p(\overline{y}\vert x)}{a(x)p(\overline{y}\vert x)+a(\overline{x}p(\overline{y}\vert\overline{x})}$ $p(x\overline{\|}y) = p(y)p(x\vert y) + p(\overline{y}p(x\vert\overline{y})$

Table 3.1: Correspondence between binary logic and probabilistic formulas

MP (Modus Ponens) corresponds to conditional deduction, and MT (Modus Tollens) corresponds to conditional abduction in probability calculus. The notation $p(y\|x)$ for conoditional deduction denotes the output probability of y conditionally deduced from the input conditional $p(y|x)$ and $p(y|\overline{x})$ as well as the input argument $p(x)$. Similarly, the notation $p(x\overline{\|}y)$ for conditional abduction denotes the output probability of x conditionaly abduced from the input conditional $p(y|x)$ and $p(y|\overline{x})$ as well as the input argument $p(y)$.

For example, consider the case of MT where $x \rightarrow y$ is TRUE and y is FALSE, which translates into $p(y|x) = 1$ and $p(y) = 0$. Then it can be observed from the first equation that $p(x|y) \neq 0$ because $p(y|x) = 1$. From the second equation it can be observed that $p(x|\overline{y}) = 0$ because $p(\overline{y}|x) = 1 - p(y|x) = 0$. From the third equation it can finally be seen that $p(x\overline{\|}y) = 0$ because $p(y) = 0$ and $p(x|\overline{y}) = 0$. From the probabilistic expressions it can thus be abduced that $p(x) = 1$ which translates into x being FALSE, as MT dictates.

Probabilistic logic is very powerful because it can be used to derive logic conlcusions without relying on axioms of logic, only on principles of probability calculus.

3.2 Generalising Probabilistic Logic as Subjective Logic

In case the argument opinions are equivalent to binary logic TRUE or FALSE, the result of any subjective logic operator is always equal to that of the corresponding propositional/binary logic operator. Similarly, when the argument opinions are equivalent to traditional probabilities, the result of any subjective logic operator is always equal to that of the corresponding probability operator.

In case the argument opinions contain degrees of uncertainty, the operators involving multiplication and division will produce derived opinions that always have correct expectation value but possibly with approximate variance when seen as Beta/Dirichlet probability distributions. All other operators produce opinions where the expectation value and the variance are always equal to the analytically correct values.

Different composite propositions that traditionally are equivalent in propositional logic do not necessarily have equal opinions. For example, in general

$$\omega_{x\wedge(y\vee z)} \neq \omega_{(x\wedge y)\vee(x\wedge z)} \tag{3.1}$$

although the distributivity of conjunction over disjunction, which in binary propositional logic is expressed as $x \wedge (y \vee z) \Leftrightarrow (x \wedge y) \vee (x \wedge z)$ holds. This is no surprise as the corresponding probability operator multiplication is non-distributive on comultiplication. However, multiplication is distributive over addition, as expressed by

$$\omega_{x\wedge(y\cup z)} = \omega_{(x\wedge y)\cup(x\wedge z)} \ . \tag{3.2}$$

De Morgan's laws are also satisfied as e.g. expressed by

$$\omega_{\overline{x\wedge y}} = \omega_{\overline{x}\vee\overline{y}} \ . \tag{3.3}$$

Subjective logic provides of a set of operators where input and output arguments are in the form of opinions. Opinions can be applied to frames of any cardinality, but some subjective logic operators are only defined for binomial opinions defined over binary frames. Opinion operators can be described for the belief notation, for the evidence notation, or for in the probabilistic notation, but operators defined for the belif notation of opinions are normally the simplest and most compact.

Table 3.2 provides the equivalent values and interpretation in belief notation, evidence notation, and probabilistic notation as well as in binary logic and traditional probability representation for a small set of binomial opinions.

Table 3.2: Example values with the three equivalent notations of binomial opinion, and their interpretations.

Belief (b, d, u, a)	Evidence (r, s, a)	Probabilistic (E, c, a)	Equivalent interpretation in binary logic and/or as probability value.
$(1, 0, 0, a)$	$(\infty, 0, a)$	$(1, 1, a)$	Binary logic TRUE, and probability $p = 1$
$(0, 1, 0, a)$	$(0, \infty, a)$	$(0, 1, a)$	Binary logic FALSE, and probability $p = 0$
$(0, 0, 1, a)$	$(0, 0, a)$	$(a, 0, a)$	Vacuous opinion, Beta density with prior a
$(\frac{1}{2}, \frac{1}{2}, 0, a)$	(∞, ∞, a)	$(\frac{1}{2}, 1, a)$	Dogmatic opinion, probability $p = \frac{1}{2}$, Dirac delta function with (irrelevant) prior a
$(0, 0, 1, \frac{1}{2})$	$(0, 0, \frac{1}{2})$	$(\frac{1}{2}, 0, \frac{1}{2})$	Vacuous opinion, uniform Beta distribution over binary frame
$(\frac{1}{4}, \frac{1}{4}, \frac{1}{2}, \frac{1}{2})$	$(1, 1, \frac{1}{2})$	$(\frac{1}{2}, \frac{1}{2}, \frac{1}{2})$	Symmetric Beta density after 1 positive and 1 negative observation, binary frame

It can be seen that some values correspond to binary logic and probability values, whereas other values correspond to probability density distributions. This richness of expression represents the advantage of subjective logic over other probabilistic logic frameworks. Online visualisations of subjective opinions and density functions can be accessed at `http://persons.unik.no/josang/sl/`.

Subjective logic allows extremely efficient computation of mathematically complex models. This is possible by approximating the analytical function expressions whenever needed. While it is relatively simple to analytically multiply two Beta distributions in the form of a joint distribution, anything more complex than that quickly becomes intractable. When combining two Beta distributions with some operator/connective, the analytical result is not always a Beta distribution and can involve hypergeometric series. In such cases, subjective logic always approximates the result as an opinion that is equivalent to a Beta distribution.

Chapter 4

Operators of Subjective Logic

Table 4.1 provides a brief overview of the main subjective logic operators. Additional operators exist for modeling special situations, such as when fusing opinions of multiple observers. Most of the operators correspond to well-known operators from binary logic and probability calculus, whereas others are specific to subjective logic.

Table 4.1: Correspondence between probability, set and logic operators.

Subjective logic operator	Symbol	Binary logic/ set operator	Symbol	Subjective logic notation
Addition[12]	$+$	XOR	$\cup$	$\omega_{x\cup y} = \omega_x + \omega_y$
Subtraction[12]	$-$	Difference	$\setminus$	$\omega_{x\setminus y} = \omega_x - \omega_y$
Multiplication[8]	$\cdot$	AND	$\wedge$	$\omega_{x\wedge y} = \omega_x \cdot \omega_y$
Division[8]	$/$	UN-AND	$\overline{\wedge}$	$\omega_{x\overline{\wedge}y} = \omega_x / \omega_y$
Comultiplication[8]	$\sqcup$	OR	$\vee$	$\omega_{x\vee y} = \omega_x \sqcup \omega_y$
Codivision[8]	$\overline{\sqcup}$	UN-OR	$\overline{\vee}$	$\omega_{x\overline{\vee}y} = \omega_x \overline{\sqcup} \omega_y$
Complement[3]	$\neg$	NOT	$\overline{x}$	$\omega_{\overline{x}} = \neg\omega_x$
Deduction[5, 9]	$\circledcirc$	MP	$\|$	$\omega_{Y\|X} = \omega_X \circledcirc \omega_{Y\|X}$
Abduction[5, 14]	$\overline{\circledcirc}$	MT	$\overline{\|}$	$\omega_{Y\overline{\|}X} = \omega_X \overline{\circledcirc} \omega_{X\|Y}$
Discounting[10]	$\otimes$	Transitivity	$:$	$\omega_x^{A:B} = \omega_B^A \otimes \omega_x^B$
Cumulative Fusion[10]	$\oplus$	n.a.	$\diamond$	$\omega_X^{A\diamond B} = \omega_X^A \oplus \omega_X^B$
Cumulative Unfusion[6]	$\ominus$	n.a.	$\overline{\diamond}$	$\omega_X^{A\overline{\diamond}B} = \omega_X^A \ominus \omega_X^B$
Average Fusion[10]	$\underline{\oplus}$	n.a.	$\underline{\diamond}$	$\omega_x^{A\underline{\diamond}B} = \omega_x^A \underline{\oplus} \omega_x^B$
Average Unfusion[6]	$\underline{\ominus}$	n.a.	$\underline{\overline{\diamond}}$	$\omega_X^{A\underline{\overline{\diamond}}B} = \omega_X^A \underline{\ominus} \omega_X^B$

Subjective logic is a generalisation of binary logic and probability calculus. This means that when a corresponding operator exists in binary logic, and the input parameters are equivalent to binary logic TRUE or FALSE, then the result opinion is equivalent to the result that the corresponding binary logic expression would have produced.

We will consider the case of binary logic AND which corresponds to multiplication of opinions [8]. For example, the pair of binomial opinions (in probabilistic notation) $\omega_x = (1, 1, a_x)$ and $\omega_y = (0, 1, a_y)$ produces $\omega_{x \wedge y} = (0, 1, a_x a_y)$ which is equivalent to TRUE $\wedge$ FALSE $=$ FALSE.

Similarly, when a corresponding operator exists in probability calculus, then the probability expectation value of the result opinion is equal to the result that the corresponding probability calculus expression would have produced with input arguments equal to the probability expectation values of the input opinions.

For example, the pair of argument opinions (in probabilistic notation): $\omega_x = (\mathrm{E}_x, 1, a_x)$ and $\omega_y = (\mathrm{E}_y, 1, a_y)$ produces $\omega_{x \wedge y} = (\mathrm{E}_x \mathrm{E}_y, 1, a_x a_y)$ which is equivalent to $p(x \wedge y) = p(x)p(y)$.

It is interesting to note that subjective logic represents a calculus for Dirichlet distributions when opinions are equivalent to Dirichlet distributions. Analytical manipulations of Dirichlet distributions is complex but can be done for simple operators, such as multiplication in which case it is called a joint distribution. However, this analytical method will quickly become unmanageable when applied to the more complex operators of Table 4.1 such as conditional deduction and abduction. Subjective logic therefore has the advantage of providing advanced operators for Dirichlet distributions for which no practical analytical solutions exist. It should be noted that the simplicity of some subjective logic operators comes at the cost of allowing those operators to be approximations of the analytically correct operators. This is discussed in more detail in Sec.4.2.1.

The next sections briefly describe the operators mentioned in Table 4.1. Online demonstrations of subjective logic operators can be accessed at `http://www.unik.no/people/josang/sl/`.

4.1 Addition and Subtraction

The addition of opinions in subjective logic is a binary operator that takes opinions about two mutually exclusive alternatives (*i.e.* two disjoint subsets of the same frame) as arguments, and outputs an opinion about the union of the subsets. The operator for addition first described in [12] is defined below.

Definition 10 (Addition) *Let x and y be two disjoint subsets of the same frame X, i.e. $x \cap y = \emptyset$. The opinion about $x \cup y$ as a function of the opinions about x and y is defined as:*

$$\text{Sum } \omega_{x \cup y} : \begin{cases} b_{x \cup y} & = b_x + b_y, \\ d_{x \cup y} & = \frac{a_x(d_x - b_y) + a_y(d_y - b_x)}{a_x + a_y}, \\ u_{x \cup y} & = \frac{a_x u_x + a_y u_y}{a_x + a_y}, \\ a_{x \cup y} & = a_x + a_y. \end{cases} \tag{4.1}$$

By using the symbol "+" to denote the addition operator for opinions, addition can be denoted as $\omega_{x \cup y} = \omega_x + \omega_y$.

The inverse operation to addition is subtraction. Since addition of opinions yields the opinion about $x \cup y$ from the opinions about disjoint subsets of the frame, then the difference between the opinions about x and y (*i.e.* the opinion about $x\backslash y$) can only be defined if $y \subseteq x$ where x and y are being treated as subsets of the frame X, *i.e.* the system must be in the state x whenever it is in the state y. The operator for subtraction first described in [12] is defined below.

Definition 11 (Subtraction) *Let x and y be subsets of the same frame X so that x and y, i.e. $x \cap y = y$. The opinion about $x\backslash y$ as a function of the opinions about x and y is defined as:*

The opinion about $x\backslash y$ is given by

$$\text{Difference } \omega_{x\backslash y}: \begin{cases} b_{x\backslash y} &= b_x - b_y, \\ d_{x\backslash y} &= \frac{a_x(d_x+b_y)-a_y(1+b_y-b_x-u_y)}{a_x-a_y}, \\ u_{x\backslash y} &= \frac{a_x u_x - a_y u_y}{a_x-a_y}, \\ a_{x\backslash y} &= a_x - a_y. \end{cases} \tag{4.2}$$

Since $u_{x\backslash y}$ should be nonnegative, then this requires that $a_y u_y \leq a_x u_x$, and since $d_{x\backslash y}$ should be nonnegative, then this requires that $a_x(d_x + b_y) \geq a_y(1 + b_y - b_x - u_y)$.

By using the symbol "$-$" to denote the subtraction operator for opinions, subtraction can be denoted as $\omega_{x\backslash y} = \omega_x - \omega_y$.

4.2 Binomial Multiplication and Division

This section describes the subjective logic operators that correspond to binary logic AND and OR, as well as their inverse operators. We will here describe normal multiplication and comultiplication [8] which are different from simple multiplication and comultiplication [3]. Special limit cases are described in [8].

4.2.1 Binomial Multiplication and Comultiplication

Binomial multiplication and comultiplication in subjective logic take binomial opinions about two elements from distinct binary frames of discernment as input arguments and produce a binomial opinion as result. The product and coproduct result opinions relate to subsets of the Cartesian product of the two binary frames of discernment. The Cartesian product of the two binary frames of discernment $X = \{x, \overline{x}\}$ and $Y = \{y, \overline{y}\}$ produces the quaternary set $X \times Y = \{(x,y),\ (x,\overline{y}),\ (\overline{x},y),\ (\overline{x},\overline{y})\}$ which is illustrated in Fig.4.1 below.

As will be explained below, binomial multiplication and comultiplication in subjective logic represent approximations of the analytically correct product and coproducts of Beta probability density functions. In this regard, normal multiplication and comultiplication produce the best approximations.

Let ω_x and ω_y be opinions about x and y respectively held by the same observer. Then the product opinion $\omega_{x\wedge y}$ is the observer's opinion about the conjunction $x \wedge y = \{(x,y)\}$ that is represented by the area inside the dotted line in Fig.4.1. The coproduct opinion $\omega_{x\vee y}$ is the opinion about the disjunction $x \vee y = \{(x,y),\ (x,\overline{y}),\ (\overline{x},y)\}$ that is represented by the area inside the dashed line in Fig.4.1. Obviously $X \times Y$ is

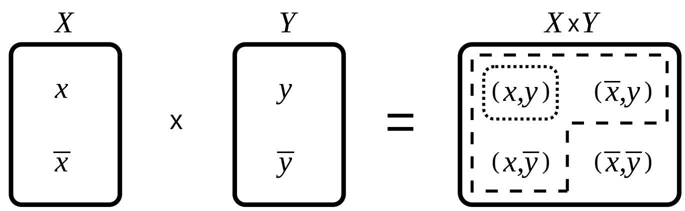


Figure 4.1: Cartesian product of two binary frames of discernment

not binary, and coarsening is required in order to determine the product and coproduct opinions as binomial opinions.

Definition 12 (Normal Binomial Multiplication) *Let $X = \{x, \overline{x}\}$ and $Y = \{y, \overline{y}\}$ be two separate frames, and let $\omega_x = (b_x, d_x, u_x, a_x)$ and $\omega_x = (b_y, d_y, u_y, a_y)$ be independent binomial opinions on x and y respectively. Given opinions about independent propositions, x and y, the binomial opinion $\omega_{x \wedge y}$ on the conjunction $(x \wedge y)$ is given by*

$$\text{Product } \omega_{x\wedge y}: \begin{cases} b_{x\wedge y} &= b_x b_y + \frac{(1-a_x)a_y b_x u_y + a_x(1-a_y)u_x b_y}{1-a_x a_y}, \\ d_{x\wedge y} &= d_x + d_y - d_x d_y, \\ u_{x\wedge y} &= u_x u_y + \frac{(1-a_y)b_x u_y + (1-a_x)u_x b_y}{1-a_x a_y}, \\ a_{x\wedge y} &= a_x a_y. \end{cases} \tag{4.3}$$

By using the symbol "$\cdot$" to denote this operator multiplication of opinions can be written as $\omega_{x\wedge y} = \omega_x \cdot \omega_y$.

Definition 13 (Normal Binomial Comultiplication) *Let $X = \{x, \overline{x}\}$ and $Y = \{y, \overline{y}\}$ be two separate frames, and let $\omega_x = (b_x, d_x, u_x, a_x)$ and $\omega_x = (b_y, d_y, u_y, a_y)$ be independent binomial opinions on x and y respectively. The binomial opinion $\omega_{x\vee y}$ on the disjunction $x \vee y$ is given by*

$$\text{Coproduct } \omega_{x\vee y}: \begin{cases} b_{x\vee y} &= b_x + b_y - b_x b_y, \\ d_{x\vee y} &= d_x d_y + \frac{a_x(1-a_y)d_x u_y + (1-a_x)a_y u_x d_y}{a_x + a_y - a_x a_y}, \\ u_{x\vee y} &= u_x u_y + \frac{a_y d_x u_y + a_x u_x d_y}{a_x + a_y - a_x a_y}, \\ a_{x\vee y} &= a_x + a_y - a_x a_y. \end{cases} \tag{4.4}$$

By using the symbol "$\sqcup$" to denote this operator multiplication of opinions can be written as $\omega_{x\vee y} = \omega_x \sqcup \omega_y$.

Normal multiplication and comultiplication represent a self-dual system represented by $b \leftrightarrow d$, $u \leftrightarrow u$, $a \leftrightarrow 1-a$, and $\wedge \leftrightarrow \vee$, that is, for example, the expressions for $b_{x\wedge y}$ and $d_{x\vee y}$ are dual to each other, and one determines the other by the correspondence, and similarly for the other expressions. This is equivalent to the observation that the opinions satisfy de Morgan's Laws, *i.e.* $\omega_{x\wedge y} = \omega_{\overline{\overline{x}\vee\overline{y}}}$ and $\omega_{x\vee y} = \omega_{\overline{\overline{x}\wedge\overline{y}}}$. However it should be noted that multiplication and comultiplication are not distributive over each other, *i.e.* for example that:

$$\omega_{x\wedge(y\vee z)} \neq \omega_{(x\wedge y)\vee(x\wedge z)} \tag{4.5}$$

This is to be expected because if x, y and z are independent, then $x\wedge y$ and $x\wedge z$ are not generally independent in probability calculus so that distributivity does not hold. In fact distributivity of conjunction over disjunction and vice versa only holds in binary logic.

Normal multiplication and comultiplication produce very good approximations of the analytically correct products and coproducts when the arguments are Beta probabilty density functions [8]. The difference between the subjective logic product and the analytically correct product of Beta density functions is best illustrated with the example of multiplying two equal vacuous binomial opinions $\omega = (0, 0, 1, \frac{1}{2})$, that are equivalent to the uniform Beta density functions Beta$(1, 1)$.

Theorem 3 *Let Q and R be independent random probability variables with identical uniform distributions over $[0, 1]$, which for example can be described as the Beta distribution Beta$(1, 1)$. Then the probability distribution function for the product random variable $P = QR$ is given by $f(p) = -\ln p$ for $0 < p < 1$.*

The proof is given in [8]. This result applies to the case of the independent propositions x and y, where we are taking four exhaustive and mutually exclusive propositions ($x \wedge y$, $x \wedge \overline{y}$, $\overline{x} \wedge y$, $\overline{x} \wedge \overline{y}$). Specifically, this means that when the probabilities of x and y have uniform distributions, then the probability of the conjunction $x \wedge y$ has the probability distribution function $f(p) = -\ln p$ with probability expectation value $\frac{1}{4}$.

This can be contrasted with the *a priori* non-informative pobability distribution over four exhaustive and mutually exclusive propositions x_1, x_2, x_3, x_4, which can be descried by: Dirichlet $\left(\frac{1}{2}, \frac{1}{2}, \frac{1}{2}, \frac{1}{2}\right)$, so that the *a priori* probability distribution for the probability of x_1 is Beta $\left(\frac{1}{2}, \frac{3}{2}\right)$, again with probability expectation value $\frac{1}{4}$.

The difference between Beta $\left(\frac{1}{2}, \frac{3}{2}\right)$, which is derivable from Dirichlet $\left(\frac{1}{2}, \frac{1}{2}, \frac{1}{2}, \frac{1}{2}\right)$, and $-\ln p$ is illustrated in Fig.4.2 below.

The analytically correct product of two uniform distributions is represented by $-\ln p$, whereas the product produced by the multiplication operator is Beta $\left(\frac{1}{2}, \frac{3}{2}\right)$, which illustrates that multiplication and comultiplication in subjective logic produce approximate results. More specifically, it can be shown that the probability expectation value is always exact, and that the variance is approximate. The quality of the variance approximation is analysed in [8], and is very good in general. The discripencies grow with the amount of uncertainty in the arguments, so Fig.4.2 illustrates the worst case.

The advantage of the multiplication and comultiplication operators of subjective logic is their simplicity, which means that complex models can be analysed efficiently.

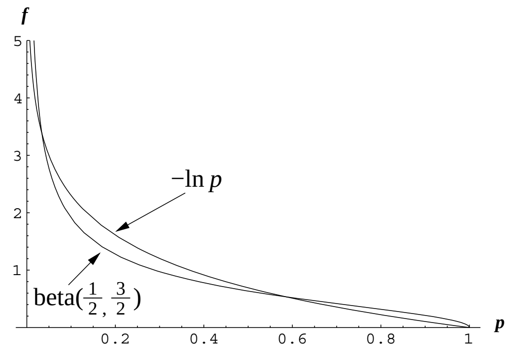


Figure 4.2: Comparison between Beta $\left(\frac{1}{2}, \frac{3}{2}\right)$ and product of uniform distributions.

The analytical result of products and coproducts of Beta distributions will in general involve the Gauss hypergeometric function [15]. The analysis of anything but the most basic models based on such functions would quickly become unmanageable.

4.2.2 Binomial Division and Codivision

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 $\omega_x = \omega_{y\wedge z}$. This requires that:

$$\begin{cases} a_x < a_y \ , \\ d_x \geq d_y \ , \\ b_x \geq \frac{a_x(1-a_y)(1-d_x)b_y}{(1-a_x)a_y(1-d_y)} \ , \\ u_x \geq \frac{(1-a_y)(1-d_x)u_y}{(1-a_x)(1-d_y)} \ . \end{cases} \tag{4.6}$$

Definition 14 (Normal Binomial Division) *Let $X = \{x, \overline{x}\}$ and $Y = \{y, \overline{y}\}$ be frames, and let $\omega_x = (b_x, d_x, u_x, a_x)$ and $\omega_y = (b_y, d_y, u_y, a_y)$ be binomial opinions on x and y satisfying Eq.(4.6). The division of ω_x by ω_y produces the quotient opinion $\omega_{x\overline{\wedge}y} = (b_{x\overline{\wedge}y}, d_{x\overline{\wedge}y}, u_{x\overline{\wedge}y}, a_{x\overline{\wedge}y})$ defined by*

$$\text{Quotient } \omega_{x\overline{\wedge}y} : \begin{cases} b_{x\overline{\wedge}y} &= \frac{a_y(b_x+a_xu_x)}{(a_y-a_x)(b_y+a_yu_y)} - \frac{a_x(1-d_x)}{(a_y-a_x)(1-d_y)}, \\ \\ d_{x\overline{\wedge}y} &= \frac{d_x-d_y}{1-d_y}, \\ \\ u_{x\overline{\wedge}y} &= \frac{a_y(1-d_x)}{(a_y-a_x)(1-d_y)} - \frac{a_y(b_x+a_xu_x)}{(a_y-a_x)(b_y+a_yu_y)}, \\ \\ a_{x\overline{\wedge}y} &= \frac{a_x}{a_y}, \end{cases} \tag{4.7}$$

By using the symbol "$/$" to denote this operator, division of opinions can be written as $\omega_{x\overline{\wedge}y} = \omega_x/\omega_y$.

The inverse operation to comultiplication is codivision. The co-quotient of opinions about propositions x and y represents the opinion about a proposition z which is independent of y such that $\omega_x = \omega_{y \vee z}$. This requires that

$$\begin{cases} a_x > a_y \ , \\ b_x \geq b_y \ , \\ d_x \geq \frac{(1-a_x)a_y(1-b_x)d_y}{a_x(1-a_y)(1-b_y)} \ , \\ u_x \geq \frac{a_y(1-b_x)u_y}{a_x(1-b_y)} \ . \end{cases} \tag{4.8}$$

Definition 15 (Normal Binomial Codivision) *Let $X = \{x, \overline{x}\}$ and $Y = \{y, \overline{y}\}$ be frames, and let $\omega_x = (b_x, d_x, u_x, a_x)$ and $\omega_y = (b_y, d_y, u_y, a_y)$ be binomial opinions on x and y satisfying Eq.(4.8). The codivision of opinion ω_x by opinion ω_x produces the co-quotient opinion $\omega_{x \overline{\vee} y} = (b_{x \overline{\vee} y}, d_{x \overline{\vee} y}, u_{x \overline{\vee} y}, a_{x \overline{\vee} y})$ defined by*

$$\text{Co-quotient } \omega_{x \overline{\vee} y}: \begin{cases} b_{x \overline{\vee} y} &= \frac{b_x - b_y}{1-b_y}, \\ \\ d_{x \overline{\vee} y} &= \frac{(1-a_y)(d_x+(1-a_x)u_x)}{(a_x-a_y)(d_y+(1-a_y)u_y)} - \frac{(1-a_x)(1-b_x)}{(a_x-a_y)(1-b_y)}, \\ \\ u_{x \overline{\vee} y} &= \frac{(1-a_y)(1-b_x)}{(a_x-a_y)(1-b_y)} - \frac{(1-a_y)(d_x+(1-a_x)u_x)}{(a_x-a_y)(d_y+(1-a_y)u_y)}, \\ \\ a_{x \overline{\vee} y} &= \frac{a_x - a_y}{1-a_y}, \end{cases} \tag{4.9}$$

By using the symbol ”$\square$” to denote this operator, codivision of opinions can be written as $\omega_{x \overline{\vee} y} = \omega_x \square \omega_y$.

4.2.3 Correspondence to Other Logic Frameworks

Multiplication, comultiplication, division and codivision of dogmatic opinions are equivalent to the corresponding probability calculus operators in Table 4.2, where e.g. $p(x)$ denotes the probability of the state variable x.

Operator name:	Probability calculus operators	
Multiplication	$p(x \wedge y)$	$= p(x)p(y)$
Division	$p(x \overline{\wedge} y)$	$= p(x)/p(y)$
Comultiplication	$p(x \vee y)$	$= p(x) + p(y) - p(x)p(y)$
Codivision	$p(x \overline{\vee} y)$	$= (p(x) - p(y))/(1 - p(y))$

Table 4.2: Probability calculus operators corresponding to opinion operators.

In the case of absolute opinions, i.e. when either $b = 1$ (absolute belief) or $d = 1$ (absolute disbelief), then the multiplication and comultiplication operators are equivalent to the AND and OR operators of binary logic.

4.3 Multinomial Multiplication

Multinomial multiplication is different from binomial multiplication in that the product opinion on the whole product frame is considered, instead of just on one element of the product frame. Fig.4.3 below illustrates the general situation.

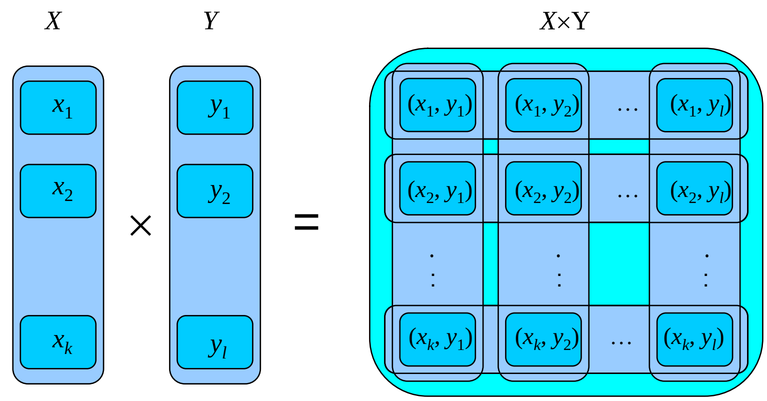


Figure 4.3: Cartesian product of two n-ary frames

Assuming the opinions ω_X and ω_Y, and the product of every combination of belief mass and uncertainty mass from the two opinions as intermediate products, there will be belief mass on all the shaded subsets of the product frame. In order to produce an opinion with only belief mass on each singleton element of $X \times Y$ as well as on $X \times Y$ itself, some of the belief mass on the row and colum subsets of $X \times Y$ must be redistributed to the singleton elements in such a way that the expectation value of each singleton element equals the product of the expecetation values of pairs of singletons from X and Y respectively.

4.3.1 General Approach

Evaluating the products of two separate multinomial opinions involves the Cartesian product of the respective frames to which the opinions apply. Let ω_X and ω_Y be two independent multinomial opinions that apply to the separate frames

$$X = \{x_1, x_2, \ldots x_k\} \text{ with cardinality } k \tag{4.10}$$

$$Y = \{y_1, y_2, \ldots y_l\} \text{ with cardinality } l\ . \tag{4.11}$$

The Cartesian product $X \times Y$ with cardinality kl is expressed as the matrix:

$$X \times Y = \begin{pmatrix} (x_1, y_1), & (x_2, y_1), & \cdots & (x_k, y_1) \\ (x_1, y_2), & (x_2, y_2), & \cdots & (x_k, y_2) \\ . & . & \cdots & . \\ . & . & \cdots & . \\ (x_1, y_l), & (x_2, y_l), & \cdots & (x_k, y_l) \end{pmatrix} \tag{4.12}$$

We now turn to the product of the multinomial opinions. The raw terms produced by $\omega_X \cdot \omega_Y$ can be separated into different groups.

1. The first group of terms consists of belief masses on singletons of $X \times Y$:

$$b^{\mathrm{I}}_{X \times Y} = \begin{cases} b_X(x_1)b_Y(y_1), & b_X(x_2)b_Y(y_1), & \ldots & b_X(x_k)b_Y(y_1) \\ b_X(x_1)b_Y(y_2), & b_X(x_2)b_Y(y_2), & \ldots & b_X(x_k)b_Y(y_2) \\ . & . & \ldots & . \\ b_X(x_1)b_Y(y_l), & b_X(x_2)b_Y(y_l), & \ldots & b_X(x_k)b_Y(y_l) \end{cases} \tag{4.13}$$

2. The second group of terms consists of belief masses on rows of $X \times Y$:

$$b^{\mathrm{Rows}}_{X \times Y} = \left(\begin{array}{cccc} u_X b_Y(y_1), & u_X b_Y(y_2), & \ldots & u_X b_Y(y_l) \end{array} \right) \tag{4.14}$$

3. The third group of terms consists of belief masses on columns of $X \times Y$:

$$b^{\mathrm{Columns}}_{X \times Y} = \left(\begin{array}{cccc} b_X(x_1)u_Y, & b_X(x_2)u_Y, & \ldots & b_X(x_k)u_Y \end{array} \right) \tag{4.15}$$

4. The last term is simply the belief mass on the whole product frame:

$$u^{\mathrm{Frame}}_{X \times Y} = u_X u_Y \tag{4.16}$$

The singleton terms of Eq.(4.13) and the term on the whole frame are unproblematic because they conform with the opinion representation of having belief mass only on singletons and on the whole frame. In contrast, the terms on rows and columns apply to overlapping subsets which is not compatible with the required opinion format, and therefore need to be reassigned. Some of it can be reassigned to singletons, and some to the whole frame. There are several possible strategies for determining the amount of uncertainty mass to be assigned to singletons and to the frame. Two methods are described below.

4.3.2 Determining Uncertainty Mass

1. **The Method of Assumed Belief Mass:** The simplest method is to assign the belief mass from the terms of Eq.(4.14) and Eq.(4.15) to singletons. Only the uncertainty mass from Eq.(4.16) is then considered as uncertainty in the product opinion, expressed as:

   $$u_{X \times Y} = u_X u_Y \; . \tag{4.17}$$

   A problem with this approach is that it in general produces less uncertainty than intuition would dictate.

2. **The Method of Assumed Uncertainty Mass:** A method that preserves more uncertainty is to consider the belief mass from Eq.(4.14) and Eq.(4.15) as potential uncertainty mass that together with the uncertainty mass from Eq.(4.16) can be called intermediate uncertainty mass. The intermediate uncertainty mass is thus:

$$u^{\mathrm{I}}_{X\times Y} = u^{\mathrm{Rows}}_{X\times Y} + u^{\mathrm{Columns}}_{X\times Y} + u^{\mathrm{Frame}}_{X\times Y} \tag{4.18}$$

The probability expectation values of each singleton in the product frame can easily be computed as the product of the expectation values of each pair of states from X and Y, as expressed in Eq.(4.19).

$$\begin{aligned} \mathrm{E}((x_i, y_j)) &= \mathrm{E}(x_i)\mathrm{E}(y_j) \\ &= (b_X(x_i) + a_X(x_i)u_X)(b_Y(y_j) + a_Y(y_j)u_Y) \end{aligned} \tag{4.19}$$

We also require that the probability expectation values of the states in the product frame can be computed as a function of the product opinion according to Eq.(4.20).

$$\mathrm{E}((x_i, y_j)) = b_{X\times Y}((x_i, y_j)) + a_X(x_i)a_Y(y_j)u_{X\times Y} \tag{4.20}$$

In order to find the correct uncertainty mass for the product opinion, each state $(x_i, y_j) \in X\times Y$ will be investigated in turn to find the smallest uncertainty mass that satisfies both Eq.(4.20) and Eq.(4.21).

$$\frac{b^{\mathrm{I}}_{X\times Y}((x_i, y_j))}{u^{\mathrm{I}}_{X\times Y}} = \frac{b_{X\times Y}((x_i, y_j))}{u_{X\times Y}} \tag{4.21}$$

The uncertainty mass that satisfies both Eq.(4.20) and Eq.(4.21) for state (x_i, y_j) can be expressed as:

$$u^{(i,j)}_{X\times Y} = \frac{u^{\mathrm{I}}_{X\times Y}\mathrm{E}((x_i, y_j))}{b^{\mathrm{I}}_{X\times Y}((x_i, y_j)) + a_X(x_i)a_Y(y_j)u^{\mathrm{I}}_{X\times Y}} \tag{4.22}$$

The product uncertainty can now be determined as the smallest $u^{(i,j)}_{X\times Y}$ among all the states, expressed as:

$$u_{X\times Y} = \min\left\{ u^{(i,j)}_{X\times Y} \text{ where } (x_i, y_j) \in X\times Y \right\} \tag{4.23}$$

4.3.3 Determining Belief Mass

Having determined the uncertainty mass, either according to Eq.(4.17) or according to Eq.(4.23), the expression for the product expectation of Eq.(4.19) can be used to compute the belief mass on each element in the product frame, as expressed by Eq.(4.24).

$$b_{X\times Y}((x_i,y_j)) = \mathrm{E}((x_i,y_j)) - a_X(x_i)a_Y(y_j)u_{X\times Y} \tag{4.24}$$

It can be shown that the additivity property of Eq.(4.25) is preserved.

$$u_{X\times Y} + \sum_{(x_i,y_j)\in X\times Y} b_{X\times Y}((x_i,y_j)) = 1 \tag{4.25}$$

From Eq.(4.24) it follows directly that the product operator is commutative. It can also be shown that the product operator is associative.

4.3.4 Example

We consider the scenario where a GE (Genetic Engineering) process can produce Male (M) or Female (F) eggs, and that in addition, each egg can have genetical mutation S or T independently of its gender. This constitutes two binary frames $X = \{\mathrm{M},\mathrm{F}\}$ and $Y = \{\mathrm{S},\mathrm{T}\}$, or alternatively the quaternary product frame $X\times Y = \{\mathrm{MS},\mathrm{MT},\mathrm{FS},\mathrm{FT}\}$. Sensor A observes whether each egg is M or F, and Sensor B observes whether the egg has mutation S or T.

Assume that an opinion regarding the gender of a specific egg is derived from Sensor A data, and that an opinion regarding its mutation is derived from Sensor B data. Sensors A and Sensor B have thus observed different and orthogonal aspects, so their respective opinions can be combined with multiplication. This is illustrated in Fig.4.4.

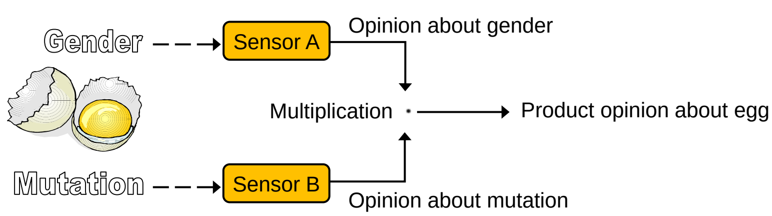


Figure 4.4: Multiplication of opinions on orthogonal aspects of GE eggs

The result of the opinion multiplication can be considered as an opinion based on a single observation where both aspects are observed at the same time. Let the observation opinions be:

$$\text{Gender } \omega_X^A : \begin{cases} \vec{b}_X^A = (0.8, 0.1) \\ u_X^A = 0.1 \\ \vec{a}_X^A = (0.5, 0.5) \end{cases} \qquad \text{Mutation } \omega_Y^B : \begin{cases} \vec{b}_Y^B = (0.7, 0.1) \\ u_Y^B = 0.2 \\ \vec{a}_Y^B = (0.2, 0.8) \end{cases} \tag{4.26}$$

The Cartesian product frame can be expressed as:

$$X \times Y = \begin{pmatrix} \text{MS}, & \text{FS} \\ \text{MT}, & \text{FT} \end{pmatrix} \tag{4.27}$$

According to Eq.(4.19) the product expectation values are:

$$\mathrm{E}(X \times Y) = \begin{pmatrix} 0.629, & 0.111 \\ 0.221, & 0.039 \end{pmatrix} \tag{4.28}$$

Below are described the results of both methods proposed in Sec.4.3.2.

1. When applying the method of *Assumed Belief Mass* where the uncertainty mass is determined according to Eq.(4.17), the product opinion is computed as:

$$b_{X \times Y} = \begin{pmatrix} 0.627, & 0.109 \\ 0.213, & 0.031 \end{pmatrix}, \quad u_{X \times Y} = 0.02, \quad a_{X \times Y} = \begin{pmatrix} 0.1, & 0.4 \\ 0.1, & 0.4 \end{pmatrix} \tag{4.29}$$

2. When applying the method of *Assumed Uncertainty* where the uncertainty mass is determined according to Eq.(4.22) and Eq.(4.23), the product opinion is computed as:

$$b_{X \times Y} = \begin{pmatrix} 0.620, & 0.102 \\ 0.185, & 0.003 \end{pmatrix}, \quad u_{X \times Y} = 0.09, \quad a_{X \times Y} = \begin{pmatrix} 0.1, & 0.4 \\ 0.1, & 0.4 \end{pmatrix} \tag{4.30}$$

The results indicate that there can be a significant difference between the two methods, and that the safest approach is to use the *assumed uncertainty* method because it preserves the most uncertainty in the product opinion.

4.4 Deduction and Abduction

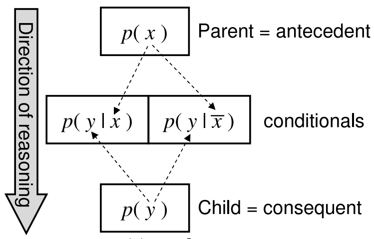
Both binary logic and probability calculus have mechanisms for conditional reasoning. In binary logic, Modus Ponens (MP) and Modus Tollens (MT) are the classical operators which are used in any field of logic that requires conditional deduction. In probability calculus, conditional probabilities together with base rates are used for analysing deductive and abductive reasoning models. Subjective logic extends the traditional probabilistic approach by allowing subjective opinions to be used as input arguments, so that deduced or abduced conclusions reflect the underlying uncertainty of the situation.

4.4.1 Probabilistic Deduction and Abduction

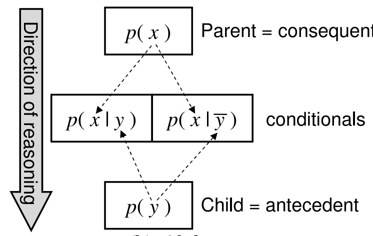
In order to clarify the principles of deduction and abduction, this section provides a brief overview of deduction and abduction in traditional probability calculus.

The notation $y\|x$, introduced in [9], denotes that the truth or probability of proposition y is derived as a function of the probability of the antecedent x together with the conditionals $p(y|x)$ and $p(y|\overline{x})$. The expression $p(y\|x)$ thus represents a derived value, whereas the expression $p(y|x)$ represents an input argument. This notational convention will also be used in subjective logic.

The deductive and abductive reasoning situations are illustrated in Fig.4.5 where x denotes the parent state and y denotes the child state of the reasoning. Conditionals are expressed as $p(\text{consequent }|\text{antecedent})$, i.e. with the consequent first, and the antecedent last.



(a) Deduction.



(b) Abduction.

Figure 4.5: Visualising deduction and abduction

It is assumed thaat the analyst has evidence about the parent, and wants to derive an opinion about the child. Defining parent and child is thus equivalent with defining the reasoning direction.

Forward conditional inference, called *deduction*, is when the parent and child states of the reasoning are the antecedent and consequent states respectively of the available conditionals.

Reverse conditional inference, called *abduction*, is when the parent state of the reasoning is the consequent of the conditionals, and the child state of the reasoning is the antecedent state of the conditionals.

Let $X = \{x_i | i = 1 \dots k\}$ be the parent frame with cardinality k, and let $Y = \{y_j | j = 1 \dots l\}$ be the child frame with cardinality l. The deductive conditional relationship between X and Y is then expressed with k vector conditionals $p(Y|x_i)$, each being of l dimensions. This is illustrated in Fig.4.6.

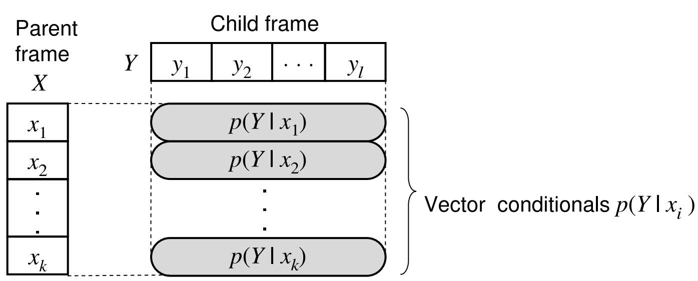


Figure 4.6: Multinomial deductive vector conditionals between parent X and child Y

The vector conditional $\vec{p}(Y|x_i)$ relates each state x_i to the frame Y. The elements of $\vec{p}(Y|x_i)$ are the scalar conditionals expressed as:

$$p(y_j|x_i), \qquad \text{where } \sum_{j=1}^{l} p(y_j|x_i) = 1\ . \tag{4.31}$$

The probabilistic expression for multinomial conditional deduction from X to Y is the vector $\vec{p}(Y\|X)$ over Y where each scalar vector element $p(y_j\|X)$ is:

$$p(y_j\|X) = \sum_{i=1}^{k} p(x_i)p(y_j|x_i)\ . \tag{4.32}$$

The multinomial probabilistic expression for inverting conditionals is:

$$p(y_j|x_i) = \frac{a(y_j)p(x_i|y_j)}{\sum_{t=1}^{l} a(y_t)p(x_i|y_t)} \tag{4.33}$$

where $a(y_j)$ represents the base rate of y_j.

By substituting the conditionals of Eq.(4.32) with inverted multinomial conditionals from Eq.(4.33), the general expression for probabilistic abduction emerges:

$$p(y_j\overline{\|}X) = \sum_{i=1}^{k} p(x_i)\left(\frac{a(y_j)p(x_i|y_j)}{\sum_{t=1}^{l} a(y_t)p(x_i|y_t)}\right)\ . \tag{4.34}$$

This will be illustrated by a numerical example below.

Example: Probabilistic Intelligence Analysis

Two countries A and B are in conflict, and intelligence analysts of country B wants to find out whether country A intends to use military aggression. The analysts of country B consider the following possible alternatives regarding country A's plans:

$$\begin{array}{ll} y_1: & \text{No military aggression from country } A \\ y_2: & \text{Minor military operations by country } A \\ y_3: & \text{Full invasion of country } B \text{ by country } A \end{array} \tag{4.35}$$

The way the analysts will determine the most likely plan of country A is by trying to observe movement of troops in country A. For this, they have spies placed inside country A. The analysts of country B consider the following possible movements of troops.

$$\begin{array}{ll} x_1: & \text{No movement of country } A\text{'s troops} \\ x_2: & \text{Minor movements of country } A\text{'s troops} \\ x_3: & \text{Full mobilisation of all country } A\text{'s troops} \end{array} \tag{4.36}$$

The analysts have defined a set of conditional probabilities of troop movements as a function of military plans, as specified by Table 4.3.

The rationale behind the conditionals are as follows. In case country A has no plans of military aggression (y_1), then there is little logistic reason for troop movements. However, even without plans of military aggression against country B it is possible

Table 4.3: Conditional probabilities $p(X|Y)$: troop movement x_i given military plan y_j

Probability vectors	Troop movements x_1 No movemt.	x_2 Minor movemt.	x_3 Full mob.
$\vec{p}(X|y_1)$:	$p(x_1|y_1) = 0.50$	$p(x_2|y_1) = 0.25$	$p(x_3|y_1) = 0.25$
$\vec{p}(X|y_2)$:	$p(x_1|y_2) = 0.00$	$p(x_2|y_2) = 0.50$	$p(x_3|y_2) = 0.50$
$\vec{p}(X|y_3)$:	$p(x_1|y_3) = 0.00$	$p(x_2|y_3) = 0.25$	$p(x_3|y_3) = 0.75$

that country A expects military aggression from country B, forcing troop movements by country A. In case country A prepares for minor military operations against country B (y_2), then necessarily troop movements are required. In case country A prepares for full invasion of country B (y_3), then significant troop movements are required.

Based on observations by spies of country B, the analysts determine the likelihoods of actual troop movements to be:

$$p(x_1) = 0.00\ , \qquad p(x_2) = 0.50\ , \qquad p(x_3) = 0.50\ . \tag{4.37}$$

The analysts are faced with an abductive reasoning situation and must first derive the conditionals $p(Y|X)$. The base rate of military plans is set to:

$$a(y_1) = 0.70\ , \qquad a(y_2) = 0.20\ , \qquad a(y_3) = 0.10\ . \tag{4.38}$$

The expression of Eq.(4.33) can now be used to derive the required conditionals, which are given in Table 4.4 below.

Table 4.4: Conditional probabilities $p(Y|X)$: military plan y_j given troop movement x_i

Military plan	Probabilities of military plans given troop movement $\vec{p}(Y|x_1)$ No movemt.	$\vec{p}(Y|x_2)$ Minor movemt.	$\vec{p}(Y|x_3)$ Full mob.
y_1: No aggr.	$p(y_1|x_1) = 1.00$	$p(y_1|x_2) = 0.58$	$p(y_1|x_3) = 0.50$
y_2: Minor ops.	$p(y_2|x_1) = 0.00$	$p(y_2|x_2) = 0.34$	$p(y_2|x_3) = 0.29$
y_3: Invasion	$p(y_3|x_1) = 0.00$	$p(y_3|x_2) = 0.08$	$p(y_3|x_3) = 0.21$

The expression of Eq.(4.32) can now be used to derive the probabilities of military plans of country A, resulting in:

$$p(y_1\|X) = 0.54\ , \qquad p(y_2\|X) = 0.31\ , \qquad p(y_3\|X) = 0.15\ . \tag{4.39}$$

Based on the results of Eq.(4.39), it seems most likely that country A does not plan any military aggression against country B. Analysing the same example with subjective logic in Sec.4.4.2 will show that these results give a misleading estimate of country A's plans because they hide the underlying uncertainty.

4.4.2 Deduction and Abduction with Subjective Opinions

Let $X = \{x_i | i = 1 \dots k\}$ and $Y = \{y_j | j = 1 \dots l\}$ be frames, where X will play the role of parent, and Y will play the role of child.

Assume the parent opinion ω_X where $|X| = k$. Assume also the conditional opinions of the form $\omega_{Y|x_i}$, where $i = 1 \dots k$. There is thus one conditional for each element x_i in the parent frame. Each of these conditionals must be interpreted as the subjective opinion on Y, given that x_i is TRUE. The subscript notation on each conditional opinion indicates not only the frame Y it applies to, but also the element x_i in the antecedent frame it is conditioned on.

By using the notation for probabilistic conditional deduction the corresponding expressions for subjective logic conditional deduction can be defined.

$$\omega_{Y\|X} = \omega_X \circledcirc \omega_{Y|X} \tag{4.40}$$

where $\circledcirc$ is the general conditional deduction operator for subjective opinions, and $\omega_{Y|X} = \{\omega_{Y|x_i} | i = 1 \dots k\}$ is a set of $k = |X|$ different opinions conditioned on each $x_i \in X$ respectively. Similarly, the expressions for subjective logic conditional abduction can be written as:

$$\omega_{Y\|X} = \omega_X \overline{\circledcirc} \omega_{X|Y} \tag{4.41}$$

where $\overline{\circledcirc}$ is the general conditional abduction operator for subjective opinions, and $\omega_{X|Y} = \{\omega_{X|y_j} | j = 1 \dots l\}$ is a set of $l = |Y|$ different multinomial opinions conditioned on each $y_j \in Y$ respectively.

In order to evaluate the expression of Eq.(4.40) and Eq.(4.41), the general deduction and abduction operators must be defined. For binomial opinions, these have been defined in [9, 14]. Deduction and abduction for multinomial opinions are described below [5].

Subjective Logic Deduction

Assume that an observer perceives a conditional relationship between the two frames X and Y. Let $\omega_{Y|X}$ be the set of conditional opinions on the consequent frame Y as a function of the opinion on the antecedent frame X expressed as

$$\omega_{Y|X} : \left\{\omega_{Y|x_i},\ i = 1, \dots k\right\} \ . \tag{4.42}$$

Each conditional opinion is a tuple composed of a belief vector $\vec{b}_{Y|x_i}$, an uncertainty mass $u_{Y|x_i}$ and a base rate vector $\vec{a}_Y$ expressed as:

$$\omega_{Y|x_i} = \left(\vec{b}_{Y|x_i}, u_{Y|x_i}, \vec{a}_Y\right) \ . \tag{4.43}$$

Note the base rate vector $\vec{a}_Y$ is equal for all conditional opinions of Eq.(4.42). Let ω_X be the opinion on the antecedent frame X.

Traditional probabilistic conditional deduction can always be derived from these opinions by inserting their probability expectation values into Eq.(4.32), resulting in the expression below.

$$\mathrm{E}(y_j\|X) = \sum_{i=1}^{k} \mathrm{E}(x_i)\mathrm{E}(y_j|x_i) \quad \text{where Eq.(2.4) provides each factor.} \tag{4.44}$$

The operator for subjective logic deduction takes the uncertainty of $\omega_{Y|X}$ and ω_X into account when computing the derived opinion $\omega_{Y\|X}$ as indicated by Eq.(4.40). The method for computing the derived opinion describe below is based on a geometric analysis of the input opinions $\omega_{Y|X}$ and ω_X, and how they relate to each other.

The conditional opinions will in general define a sub-simplex inside the opinion simplex of the child frame Y. A visualisation of deduction with ternary parent and child tetrahedrons and trinomial opinions is illustrated in Fig.4.7.

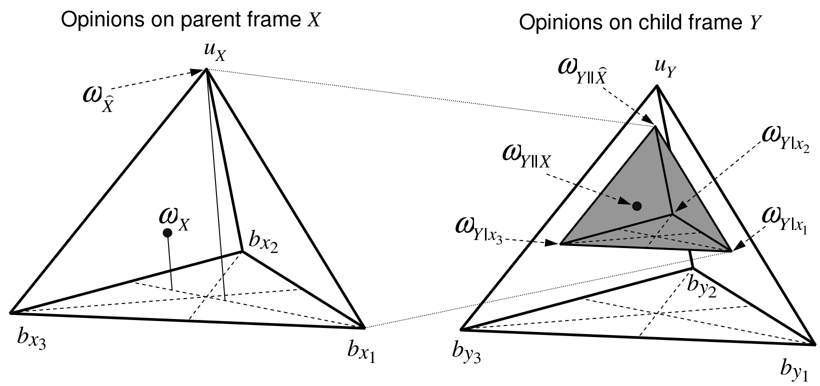


Figure 4.7: Sub-tetrahedron defined as the conditional projection of the parent tetrahedron.

The sub-simplex formed by the conditional projection of the parent simplex into the child simplex is shown as the shaded tetrahedron on the right hand side in Fig.4.7. The position of the derived opinion $\omega_{Y\|X}$ is geometrically determined by the point inside the sub-simplex that linearly corresponds to the opinion ω_X in the parent simplex.

In general, a sub-simplex will not appear as regular as in the example of Fig.4.7, and can be skewed in all possible ways. The dimensionality of the sub-simplex is equal to the smallest cardinality of X and Y. For binary frames, the sub-simplex is reduced to a triangle. Visualising a simplex larger than ternary is impractical on two-dimensional media such as paper and flat screens.

The mathematical procedure for determining the derived opinion $\omega_{Y\|X}$ is described in four steps below. The uncertainty of the sub-simplex apex will emerge from the largest sub-triangle in any dimension of Y when projected against the triangular side planes, and is derived in steps 1 to 3 below. The following expressions are needed for the computations.

$$\begin{cases} \mathrm{E}(y_t|\widehat{X}) & = \sum_{i=1}^{k} a_{x_i}\mathrm{E}(y_t|x_i)\ , \\ \mathrm{E}(y_t|(\widehat{x_r,x_s})) & = (1-a_{y_t})b_{y_t|x_s} + a_{y_t}(b_{y_t|x_r} + u_{Y|x_r})\ . \end{cases} \tag{4.45}$$

The expression $\mathrm{E}(y_t|\widehat{X})$ gives the expectation value of y_t given a vacuous opinion $\omega_{\widehat{X}}$ on X. The expression $\mathrm{E}(y_t|(\widehat{x_r,x_s}))$ gives the expectation value of y_t for the theoretical maximum uncertainty $u^{\mathrm{T}}_{y_t}$.

- **Step 1:** Determine the X-dimensions (x_r, x_s) that give the largest theoretical uncertainty $u^{\mathrm{T}}_{y_t}$ in each Y-dimension y_t, independently of the opinion on X. Each dimension's maximum uncertainty is:

$$u^{\mathrm{T}}_{y_t} = 1-\mathrm{Min}\Big[\big(1-b_{y_t|x_r}-u_{Y|x_r}+b_{y_t|x_s}\big), \forall (x_r, x_s)\in X\Big] \ . \tag{4.46}$$

The X-dimensions (x_r, x_s) are recorded for each y_t. Note that it is possible to have $x_r = x_x$.

- **Step 2:** First determine the triangle apex uncertainty $u_{y_t\|\widehat{X}}$ for each Y-dimension by assuming a vacuous opinion $\omega_{\widehat{X}}$ and the actual base rate vector $\vec{a}_X$. Assuming that $a_{y_t} \neq 0$ and $a_{y_t} \neq 1$ for all base rates on Y, each triangle apex uncertainty $u_{y_t\|\widehat{X}}$ can be computed as:

Case A: $\mathrm{E}(y_t|\widehat{X}) \leq \mathrm{E}(y_t|(\widehat{x_r, x_s}))$:

$$u_{y_t\|\widehat{X}} = \left(\frac{\mathrm{E}(y_t|\widehat{X}) - b_{y_t|x_s}}{a_{y_t}}\right) \tag{4.47}$$

Case B: $\mathrm{E}(y_t|\widehat{X}) > \mathrm{E}(y_t|(\widehat{x_r, x_s}))$:

$$u_{y_t\|\widehat{X}} = \left(\frac{b_{y_t|x_r} + u_{Y|x_r} - \mathrm{E}(y_t|\widehat{X})}{1 - a_{y_t}}\right) \tag{4.48}$$

Then determine the intermediate sub-simplex apex uncertainty $u^{\mathrm{Int}}_{Y\|\widehat{X}}$ which is equal to the largest of the triangle apex uncertainties computed above. This uncertainty is expressed as.

$$u^{\mathrm{Int}}_{Y\|\widehat{X}} = \mathrm{Max}\left[u_{y_t\|\widehat{X}}, \forall y_t \in Y\right] \ . \tag{4.49}$$

- **Step 3:** First determine the intermediate belief components $b^{\mathrm{Int}}_{y_j\|\widehat{X}}$ in case of vacuous belief on X as a function of the intermediate apex uncertainty $u^{\mathrm{Int}}_{Y\|\widehat{X}}$:

$$b^{\mathrm{Int}}_{y_j\|\widehat{X}} = \mathrm{E}(y_j\|\widehat{X}) - a_{y_j} u^{\mathrm{Int}}_{Y\|\widehat{X}} \tag{4.50}$$

For particular geometric combinations of the triangle apex uncertainties $u_{y_t\|\widehat{X}}$ it is possible that an intermediate belief component $b^{\mathrm{Int}}_{y_j\|\widehat{X}}$ becomes negative. In such cases a new adjusted apex uncertainty $u^{\mathrm{Adj}}_{y_t\|\widehat{X}}$ is computed. Otherwise the adjusted apex uncertainty is set equal to the intermediate apex uncertainty of Eq.(4.49). Thus:

Case A: $b^{\mathrm{Int}}_{y_j\|\widehat{X}} < 0: \qquad u^{\mathrm{Adj}}_{y_j\|\widehat{X}} = \mathrm{E}(y_j\|\widehat{X})/a_{y_j}$ (4.51)

Case B: $b^{\mathrm{Int}}_{y_j\|\widehat{X}} \geq 0: \qquad u^{\mathrm{Adj}}_{y_j\|\widehat{X}} = u^{\mathrm{Int}}_{Y\|\widehat{X}}$ (4.52)

Then compute the sub-simplex apex uncertainty $u_{Y\|\widehat{X}}$ as the minimum of the adjusted apex uncertainties according to:

$$u_{Y\|\widehat{X}} = \text{Min}\left[u^{\text{Adj}}_{y_j\|\widehat{X}}, \forall y_t \in Y\right] \ . \tag{4.53}$$

Note that the apex uncertainty is not necessarily the highest uncertainty of the sub-simplex. It is possible that one of the conditionals $\omega_{Y|x_i}$ actually contains a higher uncertainty, which would simply mean that the sub-simplex is skewed or tilted to the side.

- **Step 4:** Based on the sub-simplex apex uncertainty $u_{Y\|\widehat{X}}$, the actual uncertainty $u_{Y\|X}$ as a function of the opinion on X is:

$$u_{Y\|X} = u_{Y\|\widehat{X}} - \sum_{i=1}^{k}(u_{Y\|\widehat{X}} - u_{Y|x_i})b_{x_i} \ . \tag{4.54}$$

Given the actual uncertainty $u_{Y\|X}$, the actual beliefs $b_{y_j\|X}$ are:

$$b_{y_j\|X} = \text{E}(y_j\|X) - a_{y_j}u_{Y\|X} \ . \tag{4.55}$$

The belief vector $\vec{b}_{Y\|X}$ is expressed as:

$$\vec{b}_{Y\|X} = \left\{b_{y_j\|X} \mid j = 1, \dots l\right\} \ . \tag{4.56}$$

Finally, the derived opinion $\omega_{Y\|X}$ is the tuple composed of the belief vector of Eq.(4.56), the uncertainty belief mass of Eq.(4.54) and the base rate vector of Eq.(4.43) expressed as:

$$\omega_{Y\|X} = \left(\vec{b}_{Y\|X}, u_{Y\|X}, \vec{a}_Y\right) \ . \tag{4.57}$$

The method for multinomial deduction described above represents both a simplification and a generalisation of the method for binomial deduction described in [9]. In case of 2×2 deduction in particular, the methods are equivalent and produce exactly the same results.

Subjective Logic Abduction

Subjective logic abduction requires the inversion of conditional opinions of the form $\omega_{X|y_j}$ into conditional opinions of the form $\omega_{Y|x_i}$ similarly to Eq.(4.33). The inversion of probabilistic conditionals according to Eq.(4.33) uses the division operator for probabilities. While a division operator for binomial opinions is defined in [8], a division operator for multinomial opinions would be complex because it involves matrix and vector expressions. Instead we define inverted conditional opinions as an uncertainty maximised opinion.

It is natural to define base rate opinions as vacuous opinions, so that the base rate vector $\vec{a}$ alone defines their probability expectation values. The rationale for defining inversion of conditional opinions as producing maximum uncertainty is that it involves multiplication with a vacuous base rate opinion which produces an uncertainty maximised product. Let $|X| = k$ and $|Y| = l$, and assume the set of available conditionals to be:

$$\omega_{X|Y} : \ \left\{\omega_{X,y_j}, \text{ where } j = 1 \ldots l\right\} \ . \tag{4.58}$$

Assume further that the analyst requires the set of conditionals expressed as:

$$\omega_{Y|X} : \ \{\omega_{Y,x_i}, \text{ where } i = 1 \ldots k\} \ . \tag{4.59}$$

First compute the l different probability expectation values of each inverted conditional opinion $\omega_{Y|x_i}$, according to Eq.(4.33) as:

$$\mathrm{E}(y_j|x_i) = \frac{a(y_j)\mathrm{E}(\omega_{X|y_j}(x_i))}{\sum_{t=1}^{l} a(y_t)\mathrm{E}(\omega_{X|y_t}(x_i))} \tag{4.60}$$

where $a(y_j)$ denotes the base rate of y_j. Consistency requires that:

$$\mathrm{E}(\omega_{Y|x_i}(y_j)) = \mathrm{E}(y_j|x_i) \ . \tag{4.61}$$

The simplest opinions that satisfy Eq.(4.61) are the k dogmatic opinions:

$$\underline{\omega}_{Y|x_i} : \begin{cases} b_{Y|x_i}(y_j) & = \mathrm{E}(y_j|x_i), \quad \text{for } j = 1 \ldots k \ , \\ u_{Y|x_i} & = 0 \ , \\ \vec{a}_{Y|x_i} & = \vec{a}_Y \ . \end{cases} \tag{4.62}$$

Uncertainty maximisation of $\underline{\omega}_{Y|x_i}$ consists of converting as much belief mass as possible into uncertainty mass while preserving consistent probability expectation values according to Eq.(4.61). The result is the uncertainty maximised opinion denoted as $\widehat{\omega}_{Y|x_i}$. This process is illustrated in Fig.4.8.

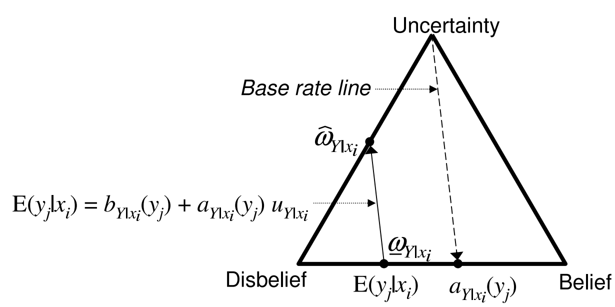


Figure 4.8: Uncertainty maximisation of dogmatic opinion

It must be noted that Fig.4.8 only represents two dimensions of the multinomial opinions on Y, namely y_j and its complement. The line defined by

$$\mathrm{E}(y_j|x_i) = b_{Y|x_i}(y_j) + a_{Y|x_i}(y_j)u_{Y|x_i} \ . \tag{4.63}$$

that is parallel to the base rate line and that joins $\underline{\omega}_{Y|x_i}$ and $\widehat{\omega}_{Y|x_i}$ in Fig.4.8, defines the opinions $\omega_{Y|x_i}$ for which the probability expectation values are consistent with Eq.(4.61). A opinion $\widehat{\omega}_{Y|x_i}$ is uncertainty maximised when Eq.(4.63) is satisfied and at least one belief mass of $\widehat{\omega}_{Y|x_i}$ is zero. In general, not all belief masses can be zero simultaneously except for vacuous opinions.

In order to find the dimension(s) that can have zero belief mass, the belief mass will be set to zero in Eq.(4.63) successively for each dimension $y_j \in Y$, resulting in l different uncertainty values defined as:

$$u^j_{Y|x_i} = \frac{\mathrm{E}(y_j|x_i)}{a_{Y|x_i}(y_j)}, \text{ where } j = 1 \ldots l \;. \tag{4.64}$$

The minimum uncertainty expressed as $\mathrm{Min}\{u^j_{Y|x_i}, \text{ for } j = 1 \ldots l\}$ determines the dimension which will have zero belief mass. Setting the belief mass to zero for any other dimension would result in negative belief mass for other dimensions. Assume that y_t is the dimension for which the uncertainty is minimum. The uncertainty maximised opinion can then be determined as:

$$\widehat{\omega}_{Y|x_i} : \begin{cases} b_{Y|x_i}(y_j) & = \mathrm{E}(y_j|x_i) - a_Y(y_j)u^t_{Y|x_i}, \text{ for } y = 1 \ldots l \\ u_{Y|x_i} & = u^t_{Y|x_i} \\ \vec{a}_{Y|x_i} & = \vec{a}_Y \end{cases} \tag{4.65}$$

By defining $\omega_{Y|x_i} = \widehat{\omega}_{Y|x_i}$, the expressions for the set of inverted conditional opinions $\omega_{Y|x_i}$ (with $i = 1 \ldots k$) emerges. Conditional abduction according to Eq.(4.41) with the original set of conditionals $\omega_{X|Y}$ is now equivalent to conditional deduction according to Eq.(4.40) as described above, where the set of inverted conditionals $\omega_{Y|X}$ is used deductively.

Example: Military Intelligence Analysis with Subjective Logic

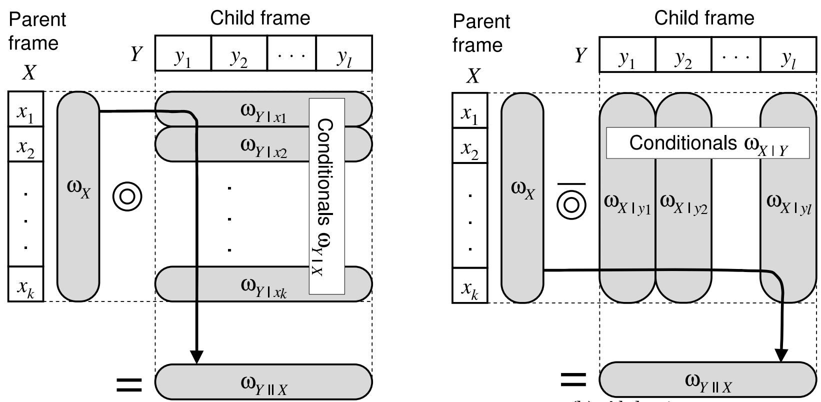
The difference between deductive and abductive reasoning with opinions is illustrated in Fig.4.9 below.

Fig.4.9 shows that deduction uses conditionals defined over the child frame, and that abduction uses conditionals defined over the parent frame.

In this example we revisit the intelligence analysis situation of Sec.4.4.1, but now with conditionals and evidence represented as subjective opinions according to Table 4.5 and Eq.(4.66).

Table 4.5: Conditional opinion $\omega_{X|Y}$: troop movement x_i given military plan y_j

Opinions $\omega_{X\|Y}$	Troop movements x_1: No movemt.	x_2: Minor movemt.	x_3: Full mob.	X Any
$\omega_{X\|y_1}$:	$b(x_1) = 0.50$	$b(x_2) = 0.25$	$b(x_3) = 0.25$	$u = 0.00$
$\omega_{X\|y_2}$:	$b(x_1) = 0.00$	$b(x_2) = 0.50$	$b(x_3) = 0.50$	$u = 0.00$
$\omega_{X\|y_3}$:	$b(x_1) = 0.00$	$b(x_2) = 0.25$	$b(x_3) = 0.75$	$u = 0.00$



(a) Deduction. (b) Abduction.

Figure 4.9: Visualising deduction and abduction with opinions

The opinion about troop movements is expressed as the opinion:

$$\omega_X = \begin{cases} b(x_1) &= 0.00, \quad a(x_1) = 0.70 \\ b(x_2) &= 0.50, \quad a(x_2) = 0.20 \\ b(x_3) &= 0.50, \quad a(x_3) = 0.10 \\ u &= 0.00 \end{cases} \tag{4.66}$$

First the conditional opinions must be inverted as expressed in Table 4.6.

Table 4.6: Conditional opinions $\omega_{Y|X}$: military plan y_j given troop movement x_i

	Opinions of military plans given troop movement		
Military plan	$\omega_{Y\|x_1}$ No movemt.	$\omega_{Y\|x_2}$ Minor movemt.	$\omega_{Y\|x_3}$ Full mob.
y_1: No aggression	$b(y_1) = 1.00$	$b(y_1) = 0.00$	$b(y_1) = 0.00$
y_2: Minor ops.	$b(y_2) = 0.00$	$b(y_2) = 0.17$	$b(y_2) = 0.14$
y_3: Invasion	$b(y_3) = 0.00$	$b(y_3) = 0.00$	$b(y_3) = 0.14$
Y: Any	$u = 0.00$	$u = 0.83$	$u = 0.72$

Then the likelihoods of country A's plans can be computed as the opinion:

$$\omega_{Y\|X} = \begin{cases} b(y_1) &= 0.00, \quad a(y_1) = 0.70, \quad \mathrm{E}(y_1) = 0.54 \\ b(y_2) &= 0.16, \quad a(y_2) = 0.20, \quad \mathrm{E}(y_2) = 0.31 \\ b(y_3) &= 0.07, \quad a(y_3) = 0.10, \quad \mathrm{E}(y_3) = 0.15 \\ u &= 0.77 \end{cases} \tag{4.67}$$

These results can be compared with those of Eq.(4.39) which were derived with probabilities only, and which are equal to the probability expectation values given in the rightmost column of Eq.(4.67). The important observation to make is that although y_1 (no aggression) seems to be country A's most likely plan in probabilistic terms, this likelihood is based on uncertainty only. The only firm evidence actually supports y_2

(minor aggression) or y_3 (full invasion), where y_2 has the strongest support ($b(y_2) = 0.16$). A likelihood expressed as a scalar probability can thus hide important aspects of the analysis, which will only come to light when uncertainty is explicitly expressed, as done in the example above.

4.5 Fusion of Multinomial Opinions

In many situations there will be multiple sources of evidence, and fusion can be used to combine evidence from different sources.

In order to provide an interpretation of fusion in subjective logic it is useful to consider a process that is observed by two sensors. A distinction can be made between two cases.

1. The two sensors observe the process during disjoint time periods. In this case the observations are independent, and it is natural to simply add the observations from the two sensors, and the resulting fusion is called *cumulative fusion*.

2. The two sensors observe the process during the same time period. In this case the observations are dependent, and it is natural to take the average of the observations by the two sensors, and the resulting fusion is called *averaging fusion*.

Fusion of binomial opinions have been described in [3, 4]. The two types of fusion for multinomial opinions are described in the following sections. When observations are partially dependent, a hybrid fusion operator can be defined [10].

4.5.1 The Cumulative Fusion Operator

The cumulative fusion rule is equivalent to *a posteriori* updating of Dirichlet distributions. Its derivation is based on the bijective mapping between the belief and evidence notations described in Sec.2.2.4.

Assume a frame X containing k elements. Assume two observers A and B who observe the outcomes of the process over two separate time periods.

Let the two observers' respective observations be expressed as $\vec{r}^A$ and $\vec{r}^B$. The evidence opinions resulting from these separate bodies of evidence can be expressed as $(\vec{r}^A, \vec{a})$ and $(\vec{r}^B, \vec{a})$

The cumulative fusion of these two bodies of evidence simply consists of vector addition of $\vec{r}^A$ and $\vec{r}^B$, expressed as:

$$(\vec{r}^A, \vec{a}) \oplus (\vec{r}^B, \vec{a}) = ((\vec{r}^A + \vec{r}^B), \vec{a}) \ . \tag{4.68}$$

The symbol "$\diamond$" denotes the fusion of two observers A and B into a single imaginary observer denoted as $A \diamond B$. All the necessary elements are now in place for presenting the cumulative rule for belief fusion.

Theorem 4 Cumulative Fusion Rule

Let ω^A and ω^B be opinions respectively held by agents A and B over the same frame $X = \{x_i \mid i = 1, \cdots, l\}$. Let $\omega^{A\diamond B}$ be the opinion such that:

Case I: For $u^A \neq 0 \;\vee\; u^B \neq 0$:

$$\begin{cases} b^{A\diamond B}(x_i) & = \frac{b^A(x_i)u^B + b^B(x_i)u^A}{u^A+u^B-u^Au^B} \\ \\ u^{A\diamond B} & = \frac{u^Au^B}{u^A+u^B-u^Au^B} \end{cases} \tag{4.69}$$

Case II: For $u^A = 0 \;\wedge\; u^B = 0$:

$$\begin{cases} b^{A\diamond B}(x_i) & = \gamma^A\, b^A(x_i) + \gamma^B b^B(x_i) \\ \\ u^{A\diamond B} & = 0 \end{cases} \quad \textit{where} \quad \begin{cases} \gamma^A = \lim\limits_{\substack{u^A\to 0\\ u^B\to 0}} \frac{u^B}{u^A+u^B} \\ \\ \gamma^B = \lim\limits_{\substack{u^A\to 0\\ u^B\to 0}} \frac{u^A}{u^A+u^B} \end{cases} \tag{4.70}$$

Then $\omega^{A\diamond B}$ is called the cumulatively fused bba of ω^A and ω^B, representing the combination of independent opinions of A and B. By using the symbol '$\oplus$' to designate this belief operator, we define $\omega^{A\diamond B} \equiv \omega^A \oplus \omega^B$.

It can be verified that the cumulative rule is commutative, associative and non-idempotent. In Case II of Theorem 4, the associativity depends on the preservation of relative weights of intermediate results, which requires the additional weight variable γ. In this case, the cumulative rule is equivalent to the weighted average of probabilities.

It is interesting to notice that the expression for the cumulative rule is independent of the *a priori* constant C. That means that the choice of a uniform Dirichlet distribution in the binary case in fact only influences the mapping between Dirichlet distributions and Dirichlet bbas, not the cumulative rule itself. This shows that the cumulative rule is firmly based on classical statistical analysis, and not dependent on arbitrary and ad hoc choices.

The cumulative rule represents a generalisation of the consensus operator [4, 3] which emerges directly from Theorem 4 by assuming a binary frame.

4.5.2 The Avaring Fusion Operator

The average rule is equivalent to averaging the evidence of Dirichlet distributions. Its derivation is based on the bijective mapping between the belief and evidence notations of Eq.(2.12) and Eq.(2.13).

Assume a frame X containing k elements. Assume two observers A and B who observe the outcomes of the process over the same time periods.

Let the two observers' respective observations be expressed as $\vec{r}^A$ and $\vec{r}^B$. The evidence opinions resulting from these separate bodies of evidence can be expressed as $(\vec{r}^A, \vec{a})$ and $(\vec{r}^B, \vec{a})$

The averaging fusion of these two bodies of evidence simply consists of averaging $\vec{r}^A$ and $\vec{r}^B$. In terms of Dirichlet distributions, this can be expressed as:

$$(\vec{r}^A, \vec{a})\underline{\oplus}(\vec{r}^B, \vec{a}) = ((\frac{\vec{r}^A + \vec{r}^B}{2}), \vec{a}) \; . \tag{4.71}$$

The symbol "$\underline{\diamond}$" denotes the averaging fusion of two observers A and B into a single imaginary observer denoted as $A \diamond B$.

Theorem 5 Averaging Fusion Rule

Let ω^A and ω^B be opinions respectively held by agents A and B over the same frame $X = \{x_i \mid i = 1, \cdots, l\}$. Let $\omega^{A\underline{\diamond}B}$ be the opinion such that:

Case I: For $u^A \neq 0 \;\; \vee \;\; u^B \neq 0$:

$$\begin{cases} b^{A\underline{\diamond}B}(x_i) = \frac{b^A(x_i)u^B + b^B(x_i)u^A}{u^A+u^B} \\ \\ u^{A\underline{\diamond}B} = \frac{2u^Au^B}{u^A+u^B} \end{cases} \tag{4.72}$$

Case II: For $u^A = 0 \;\; \wedge \;\; u^B = 0$:

$$\begin{cases} b^{A\diamond B}(x_i) = \gamma^A\, b^A(x_i) + \gamma^B b^B(x_i) \\ \\ u^{A\diamond B} = 0 \end{cases} \quad \textit{where} \quad \begin{cases} \gamma^A = \lim\limits_{\substack{u^A \to 0 \\ u^B \to 0}} \frac{u^B}{u^A+u^B} \\ \\ \gamma^B = \lim\limits_{\substack{u^A \to 0 \\ u^B \to 0}} \frac{u^A}{u^A+u^B} \end{cases} \tag{4.73}$$

Then $\omega^{A\underline{\diamond}B}$ is called the averaged opinion of ω^A and ω^B, representing the combination of the dependent opinions of A and B. By using the symbol '$\underline{\oplus}$' to designate this belief operator, we define $\omega^{A\underline{\diamond}B} \equiv \omega^A\underline{\oplus}\omega^B$.

It can be verified that the averaging fusion rule is commutative and idempotent, but not associative.

The cumulative rule represents a generalisation of the consensus rule for dependent opinions defined in [7].

4.6 Trust Transitivity

Assume two agents A and B where A trusts B, and B believes that proposition x is true. Then by transitivity, agent A will also believe that proposition x is true. This assumes that B recommends x to A. In our approach, trust and belief are formally expressed as opinions. The transitive linking of these two opinions consists of discounting B's opinion about x by A's opinion about B, in order to derive A's opinion

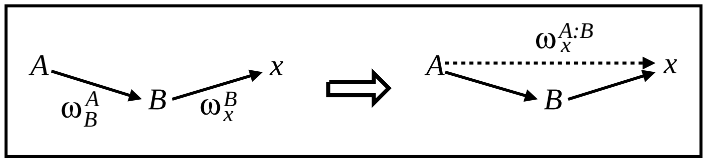


Figure 4.10: Principle of the discounting operator

about x. This principle is illustrated in Fig.4.10 below. The solid arrows represent initial direct trust, and the dotted arrow represents derived indirect trust.

Trust transitivity, as trust itself, is a human mental phenomenon, so there is no such thing as objective transitivity, and trust transitivity therefore lends itself to different interpretations. We see two main difficulties. The first is related to the effect of A disbelieving that B will give a good advice. What does this exactly mean? We will give two different interpretations and definitions. The second difficulty relates to the effect of base rate trust in a transitive path. We will briefly examine this, and provide the definition of a base rate sensitive discounting operator as an alternative to the two previous which are base rate insensitive.

4.6.1 Uncertainty Favouring Trust Transitivity

A's disbelief in the recommending agent B means that A thinks that B ignores the truth value of x. As a result A also ignores the truth value of x.

Definition 16 (Uncertainty Favouring Discounting) *Let A, B and be two agents where A's opinion about B's recommendations is expressed as $\omega^A_B = \{b^A_B, d^A_B, u^A_B, a^A_B\}$, and let x be a proposition where B's opinion about x is recommended to A with the opinion $\omega^B_x = \{b^B_x, d^B_x, u^B_x, a^B_x\}$. Let $\omega^{A:B}_x = \{b^{A:B}_x, d^{A:B}_x, u^{A:B}_x, a^{A:B}_x\}$ be the opinion such that:*

$$\begin{cases} b^{A:B}_x = b^A_B b^B_x \\ d^{A:B}_x = b^A_B d^B_x \\ u^{A:B}_x = d^A_B + u^A_B + b^A_B u^B_x \\ a^{A:B}_x = a^B_x \end{cases}$$

then $\omega^{A:B}_x$ is called the uncertainty favouring discounted opinion of A. By using the symbol $\otimes$ to designate this operation, we get $\omega^{A:B}_x = \omega^A_B \otimes \omega^B_x$. □

It is easy to prove that this operator is associative but not commutative. This means that the combination of opinions can start in either end of the path, and that the order in which opinions are combined is significant. In a path with more than one recommending entity, opinion independence must be assumed, which for example translates into not allowing the same entity to appear more than once in a transitive path. Fig.4.11 illustrates an example of applying the discounting operator for independent opinions, where $\omega^A_B = \{0.1,\ 0.8,\ 0.1\}$ discounts $\omega^B_x = \{0.8,\ 0.1,\ 0.1\}$ to produce $\omega^{A:B}_x = \{0.08,\ 0.01,\ 0.91\}$.

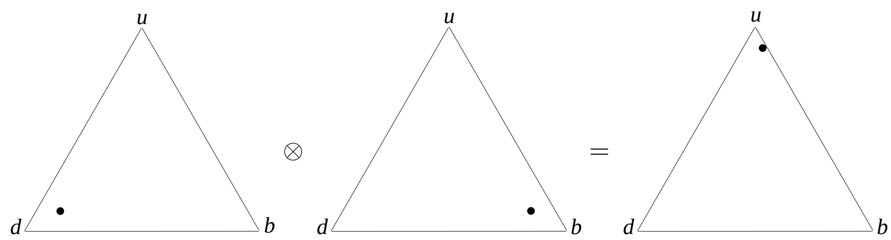


Figure 4.11: Example of applying the discounting operator for independent opinions

4.6.2 Opposite Belief Favouring

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

Definition 17 (Opposite Belief Favouring Discounting) *Let A, B and be two agents where A's opinion about B's recommendations is expressed as $\omega_B^A = \{b_B^A, d_B^A, u_B^A, a_B^A\}$, and let x be a proposition where B's opinion about x is recommended to A as the opinion $\omega_x^B = \{b_x^B, d_x^B, u_x^B, a_x^B\}$. 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:*

$$\begin{cases} b_x^{A:B} = b_B^A b_x^B + d_B^A d_x^B \\ d_x^{A:B} = b_B^A d_x^B + d_B^A b_x^B \\ u_x^{A:B} = u_B^A + (b_B^A + d_B^A) u_x^B \\ a_x^{A:B} = a_x^B \end{cases}$$

then $\omega_x^{A:B}$ is called the opposite belief favouring discounted recommendation from B to A. By using the symbol $\otimes$ to designate this operation, we get $\omega_x^{A:B} = \omega_B^A \otimes \omega_x^B$. $\square$

This operator models the principle that *"your enemy's enemy is your friend"*. That might be the case in some situations, and the operator should only be applied when the situation makes it plausible. It is doubtful whether it is meaningful to model more than two arcs in a transitive path with this principle. In other words, it is doubtful whether the enemy of your enemy's enemy necessarily is your enemy too.

4.6.3 Base Rate Sensitive Transitivity

In the transitivity operators defined in Sec.4.6.1 and Sec.4.6.2 above, a_B^A had no influence on the discounting of of the recommended (b_x^B, d_x^B, u_x^B) parameters. This can seem counterintuitive in many cases such as in the example described next.

Imagine a stranger coming to a town which is know for its citizens being honest. The stranger is looking for a car mechanic, and asks the first person he meets to direct him to a good car mechanic. The stranger receives the reply that there are two car mechanics in town, David and Eric, where David is cheap but does not always do quality work, and Eric might be a bit more expensive, but he always does a perfect job.

Translated into the formalism of subjective logic, the stranger has no other info about the person he asks than the base rate that the citizens in the town are honest. The stranger is thus ignorant, but the expectation value of a good advice is still very high. Without taking a^A_B into account, the result of the definitions above would be that the stranger is completely ignorant about which if the mechanics is the best.

An intuitive approach would then be to let the expectation value of the stranger's trust in the recommender be the discounting factor for the recommended (b^B_x, d^B_x) parameters.

Definition 18 (Base Rate Sensitive Discounting) *The base rate sensitive discounting of a belief* $\omega^B_x = (b^B_x, d^B_x, u^B_x, a^B_x)$ *by a belief* $\omega^A_B = (b^A_B, d^A_B, u^A_B, a^A_B)$ *produces the transitive belief* $\omega^{A:B}_x = (b^{A:B}_x, d^{A:B}_x, u^{A:B}_x, a^{A:B}_x)$ *where*

$$\begin{cases} b^{A:B}_x = & \mathrm{E}(\omega^A_B) b^B_x \\ d^{A:B}_x = & \mathrm{E}(\omega^A_B) d^B_x \\ u^{A:B}_x = & 1 - \mathrm{E}(\omega^A_B)(b^B_x + d^B_x) \\ a^{A:B}_x = & a^B_x \end{cases} \tag{4.74}$$

where the probability expectation value $\mathrm{E}(\omega^A_B) = b^A_B + a^A_B u^A_B$.

However this operator must be applied with care. Assume again the town of honest citizens, and let let the stranger A have the opinion $\omega^A_B = (0,\ 0,\ 1,\ 0.99)$ about the first person B she meets, i.e. the opinion has no basis in evidence other than a very high base rate defined by $a^A_B = 0.99$. If the person B now recommends to A the opinion $\omega^B_x = (1,\ 0,\ 0,\ a)$, then, according to the base rate sensitive discounting operator of Def.18, A will have the belief $\omega^{A:B}_x = (0.99,\ 0,\ 0.01,\ a)$ in x. In other words, the highly certain belief $\omega^{A:B}_x$ is derived on the basis of the highly uncertain belief ω^A_B, which can seem counterintuitive. This potential problem could be amplified as the trust path gets longer. A safety principle could therefore be to only apply the base rate sensitive discounting to the last transitive link.

There might be other principles that better reflect human intuition for trust transitivity, but we will leave this question to future research. It would be fair to say that the base rate insensitive discounting operator of Def.16 is safe and conservative, and that the base rate sensitive discounting operator of Def.18 can be more intuitive in some situations, but must be applied with care.

Chapter 5

Applications

Subjective logic represents a generalisation of probability calculus and logic under uncertainty. Subjective logic will always be equivalent to traditional probability calculus when applied to traditional probabilities, and will be equivalent to binary logic when applied to TRUE and FALSE statements.

While subjective logic has traditionally been applied to binary frames, we have shown that it can easily be extended and be applicable to frames larger than binary. The input and output parameters of subjective logic are beliefs in the form of opinions. We have described three different equivalent notations of opinions which provides rich interpretations of opinions. This also allows the analyst to choose the opinion representation that best suits a particular situation.

5.1 Fusion of Opinions

The cumulative and averaging rules of belief fusion make it possible to use the theory of belief functions for modelling situations where evidence is combined in a cumulative or averaging fashion. Such situations could previously not be correctly modelled within the framework of belief theory. It is worth noticing that the cumulative, averaging rules and Dempster's rule apply to different types of belief fusion, and that, strictly speaking, is meaningless to compare their performance in the same examples. The notion of cumulative and averaging belief fusion as opposed to conjunctive belief fusion has therefore been introduced in order to make this distinction explicit.

The following scenario will illustrate using the cumulative and the averaging fusion operators, as well as multiplication. Assume that a GE (Genetical Engineering) process can produce Male (M) or Female (F) eggs, and that in addition, each egg can have genetical mutation X or Y independently of its gender. This constitutes the quaternary frame $\Theta = \{\mathrm{MX}, \mathrm{MY}, \mathrm{FX}, \mathrm{FY}\}$. Sensors IA and IB simultaneously observe whether each egg is M or F, and Sensor II observes whether the egg has mutation X or Y.

Assume that Sensors IA and IB have derived two separate opinions regarding the gender of a specific egg, and that Sensor II has produced an opinion regarding its mutation. Because Sensors IA and IB have observed the same aspect simultaneously, the opinions should be fused with averaging fusion. Sensor II has observed a different and orthogonal aspect, so the output of the averaging fusion and the opinion of Sensor II should be combined with multiplication. This is illustrated in Fig.5.1.

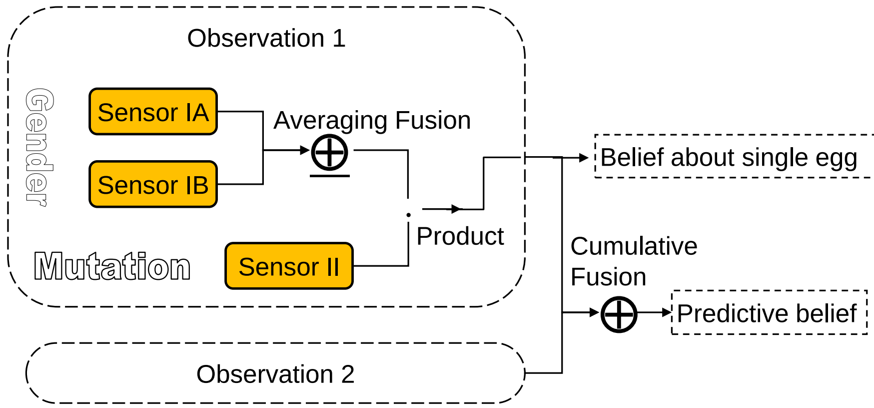


Figure 5.1: Applying different types of belief fusion according to the situation

This result from fusing the two orthogonal opinions with multiplication can now be considered as a single observation. By combining opinions from multiple observations it is possible to express the most likely status of future eggs as a predictive opinion. We are now dealing with two different situations which must be considered separately. The first situation relates to the state of a given egg that the sensors have already observed. The second situation relates to the possible state of eggs that will be produced in the future. An opinion in the first situation is based on the sensors as illustrated inside Observation 1 in Fig.5.1. The second situation relates to combining multiple observations, as illustrated by fusing the opinions from Observation 1 and Observation 2 in Fig.5.1.

5.2 Bayesian Networks with Subjective Logic

A Bayesian network is a graphical model for conditional relationships. Specifically, a Bayesian network is normally defined as a directed acyclic graph of nodes representing variables and arcs representing conditional dependence relations among the variables.

Equipped with the operators for conditional deduction and abduction, it is possible to analyse Bayesian networks with subjective logic. For example, the simple Bayesian network:

$$X \longrightarrow Y \longrightarrow Z \tag{5.1}$$

can be modelled by defining conditional opinions between the three frames. In case conditionals can be obtained with X as antecedent and Y as consequent, then deductive reasoning can be applied to the edge $[X : Y]$. In case there are available conditionals with Y as antecedent and X as consequent, then abductive reasoning must be applied.

In the example illustrated in Fig.5.2 it is assumed that deductive reasoning can be applied to both $[X : Y]$ and $[Y : Z]$.

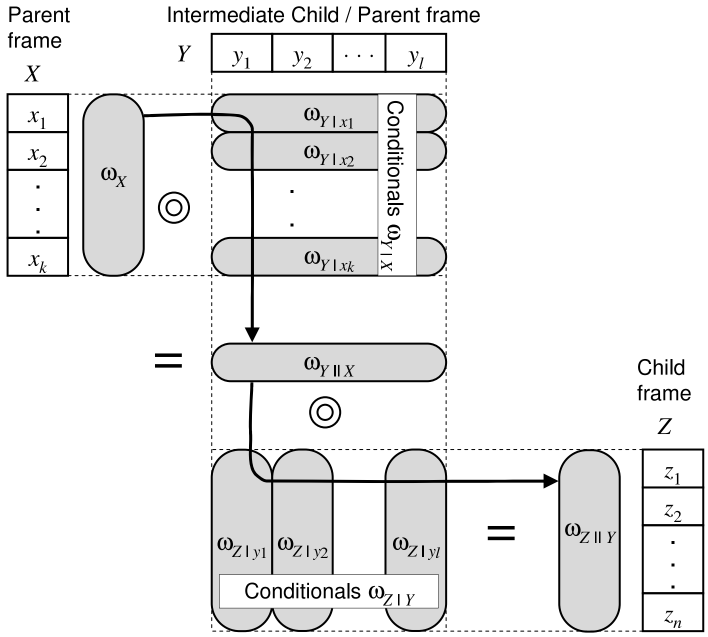


Figure 5.2: Deductive opinion structure for the Bayesian network of Eq.(5.1)

The frames X and Y thus represent parent and child of the first conditional edge, and the frames Y and Z represent parent and child of the second conditional edge respectively.

This chaining of conditional reasoning is possible because of the symmetry between the parent and child frames. They both consist of sets of mutually exclusive elements, and subjective opinions can be applied to both. In general it is arbitrary which frame is the antecedent and which frame is the consequent in a given conditional edge. Conditional reasoning is possible in either case, by applying the deductive or the abductive operator.

Frame pairs to consider as parent-child relationships must have a degree of relevance to each other. The relevance between two nodes can be formally expressed as a relevance measure, and is a direct function of the conditionals. For probabilistic conditional deduction, the relevance denoted as $\mathrm{R}(y, x)$ between two states y and x can be defined as:

$$\mathrm{R}(y, x) = |p(y|x) - p(y|\overline{x})| \; . \tag{5.2}$$

It can be seen that $\mathrm{R}(y, x) \in [0, 1]$, where $\mathrm{R}(y, x) = 0$ expresses total irrelevance, and $\mathrm{R}(y, x) = 1$ expresses total relevance between y and x.

For conditionals expressed as opinions, the same type of relevance between a given state $y_j \in Y$ and a given state $x_i \in X$ can be defined as:

$$\mathrm{R}(y_j, x_i) = |\mathrm{E}(\omega_{Y|x_i}(y_j) - \mathrm{E}(\omega_{Y|\overline{x_j}}(y_j))| \ . \tag{5.3}$$

The relevance between a child frame Y and a given state $x_i \in X$ of a parent frame can be defined as:

$$\mathrm{R}(Y, x_i) = \sum_{j=1}^{l} \mathrm{R}(y_j, x_i)/l \ . \tag{5.4}$$

Finally, the relevance between a child frame Y and a parent frame X can be defined as:

$$\mathrm{R}(Y, X) = \sum_{i=1}^{k} \mathrm{R}(Y, x_i)/k \ . \tag{5.5}$$

In our model, the relevance measure of Eq.(5.5) is the most general.

In many situations there can be multiple parents for the same child, which requires fusion of the separate child opinions into a single opinion. The question then arises which type of fusion is most appropriate. The two most typical situations to modelled are the cumulative case and the averaging case.

Cumulative fusion is applicable when independent evidence is accumulated over time such as by continuing observation of outcomes of a process. Averaging fusion is applicable when two sources provide different but independent opinions so that each opinion is weighed as a function of its certainty. The fusion operators are described in Sec.4.5.

By fusing child opinions resulting from multiple parents, arbitrarily large Bayesian networks can be constructed. Depending on the situation it must be decided whether the cumulative or the averaging operator is applicable. An example with three grand-parent frames X_1, X_2, X_3, two parent parent frames Y_1, Y_2 and one child frame Z is illustrated in Fig.5.3 below.

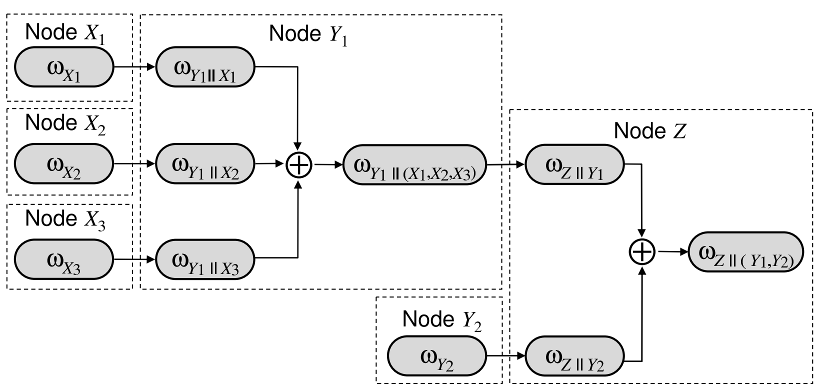


Figure 5.3: Bayesian network with multiple parent evidence nodes

The nodes X_1, X_2, X_3 and Y_2 represent initial parent frames because they do not themselves have parents in the model. Opinions about the initial parent nodes represent the input evidence to the model.

When representing Bayesian networks as graphs, the structure of conditionals is hidden in the edges, and only the nodes consisting of parent and children frames are shown.

When multiple parents can be identified for the same child, there are two important considerations. Firstly, the relative relevance between the child and each parent, and secondly the relevance or dependence between the parents.

Strong relevance between child and parents is desirable, and models should include the strongest child-parent relationships that can be identified, and for which there is evidence directly or potentially available.

Dependence between parents should be avoided as far as possible. A more subtle and hard to detect dependence can originate from hidden parent nodes outside the Bayesian network model itself. In this case the parent nodes have a hidden common grand parent node which makes them dependent. Philosophically speaking everything depends on everything in some way, so absolute independence is never achievable. There will often be some degree of dependence between evidence sources, but which from a practical perspective can be ignored. When building Bayesian network models it is important to be aware of possible dependencies, and try to select parent evidence nodes that have the lowest possible degree of dependence.

As an alternative method for managing dependence it could be possible to allow different children to share the same parent by fissioning the parent opinion, or alternatively taking dependence into account during the fusion operation. The latter option can be implemented by applying the averaging fusion operator.

It is also possible that evidence opinions provided by experts need to be discounted due to the analysts doubt in their reliability. This can be done with the trust transitivity operator[1] described in Sec.4.6

[1] Also called discounting operator

Bibliography

[1] M. Fitting. Kleene's three-valued logics and their children. *Fundamenta Informaticae*, 20:113–131, 1994.

[2] A. Gelman et al. *Bayesian Data Analysis, 2nd ed*. Chapman and Hall/CRC, Florida, USA, 2004.

[3] A. Jøsang. A Logic for Uncertain Probabilities. *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems*, 9(3):279–311, June 2001.

[4] A. Jøsang. The Consensus Operator for Combining Beliefs. *Artificial Intelligence Journal*, 142(1–2):157–170, October 2002.

[5] A. Jøsang. Conditional Reasoning with Subjective Logic. *Journal of Multiple-Valued Logic and Soft Computing*, 15(1):5–38, 2008.

[6] A. Jøsang. Cumulative and Averaging Unfusion of Beliefs. In *The Proceedings of the International Conference on Information Processing and Management of Uncertainty (IPMU2008)*, Malaga, June 2008.

[7] A. Jøsang and S.J. Knapskog. A Metric for Trusted Systems (full paper). In *Proceedings of the 21st National Information Systems Security Conference*. NSA, October 1998.

[8] A. Jøsang and D. McAnally. Multiplication and Comultiplication of Beliefs. *International Journal of Approximate Reasoning*, 38(1):19–51, 2004.

[9] A. Jøsang, S. Pope, and M. Daniel. Conditional deduction under uncertainty. In *Proceedings of the 8th European Conference on Symbolic and Quantitative Approaches to Reasoning with Uncertainty (ECSQARU 2005)*, 2005.

[10] A. Jøsang, S. Pope, and S. Marsh. Exploring Different Types of Trust Propagation. In *Proceedings of the 4th International Conference on Trust Management (iTrust)*, Pisa, May 2006.

[11] Sherman Kent. Words of estimated probability. In Donald P. Steury, editor, *Sherman Kent and the Board of National Estimates: Collected Essays*. CIA, Center for the Study of Intelligence, 1994.

[12] D. McAnally and A. Jøsang. Addition and Subtraction of Beliefs. In *Proceedings of Information Processing and Management of Uncertainty in Knowledge-Based Systems (IPMU 2004)*, Perugia, July 2004.

[13] N. J. Nilsson. Probabilistic logic. *Artificial Intelligence*, 28(1):71–87, 1986.

[14] Simon Pope and Audun Jøsang. Analsysis of Competing Hypotheses using Subjective Logic. In *Proceedings of the 10th International Command and Control Research and Technology Symposium (ICCRTS)*. United States Department of Defense Command and Control Research Program (DoDCCRP), 2005.

[15] A.P. Prudnikov, Yu.A. Brychkov, and O.I. Marichev. *Integrals and Series*, (translated from Russian), volume 1–3. Gordon and Breach Science Publishers, Amsterdam, New York, 1986.

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

[17] M. Smithson. *Ignorance and Uncertainty: Emerging Paradigms*. Springer, 1988.