Multiplication and Comultiplication of Beliefs

Audun Jøsang and David McAnally

Distributed Systems Technology Centre
Level 7, General Purpose South, UQ Qld 4072, Australia
email: `ajosang@dstc.edu.au`, `D.McAnally@uqconnect.net`

Abstract

Multiplication and comultiplication of beliefs represent a generalisation of multiplication and comultiplication of probabilities as well as of binary logic AND and OR. Our approach follows that of subjective logic, where belief functions are expressed as opinions that are interpreted as being equivalent to beta probability distributions. We compare different types of opinion product and coproduct, and show that they represent very good approximations of the the analytical product and coproduct of beta probability distributions. We also define division and codivision of opinions, and compare our framework with other logic frameworks for combining uncertain propositions.

Key words: Belief calculus, subjective logic, Dempster Shafer, belief theory, conjunction, disjunction, unconjunction, undisjunction, three valued logic, support logic

1 Introduction

Subjective logic (Jøsang 2001 [2]) is a belief calculus based on the Dempster-Shafer belief theory (Shafer 1976 [5]). In subjective logic the term *opinion* denotes beliefs about propositions, and a set of standard and non-standard logic operators can be used to combine opinions about propositions in various ways. A particular type of multiplication and comultiplication called *propositional conjunction* and *propositional disjunction* in Jøsang 2001 [2] will be called *simple multiplication* and *simple comultiplication* here. In Jøsang 2001 [2] it was also described how every opinion can be uniquely mapped to a beta probability distribution, thereby providing a specific interpretation of belief functions in Bayesian probabilistic terms. A vacuous

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

opinion about a binary proposition is for example equivalent to a uniform probability distribution. A question left open in Jøsang 2001 [2] was why simple multiplication of two vacuous opinions produces a product opinion that when mapped to a beta distribution is slightly different from the analytical product of two uniform distributions. Below we will explain the reason for this difference, and also define alternatives to simple multiplication and simple comultiplication in the form of *normal multiplication* and *normal comultiplication* of opinions. Simple and normal multiplication and comultiplication are compared to analytical multiplication and comultiplication of beta distributions in the general case. We also define the inverse opinion operators *normal division* and *normal codivision*.

2 Fundamentals of Subjective Logic

Subjective logic is suitable for approximate reasoning in situations where there is more or less uncertainty about whether a given proposition is true or false, and this uncertainty can be expressed by a belief mass assignment [2] (BMA) where a quantity of belief mass on a given proposition can be interpreted as contributing to the probability that the proposition is true. More specifically, if a set denoted by Θ of exhaustive mutually exclusive singletons can be defined, this set is referred to as a frame of discernment. Each singleton, which will be called an *atomic element* hereafter, can be interpreted as a proposition that can be either true or false. The powerset of Θ denoted by 2^{Θ} contains all possible subsets of Θ. The set $2^{\Theta} - \{\emptyset\}$ of nonempty subsets of Θ will be called its reduced powerset. A BMA assigns belief mass to nonempty subsets of Θ (*i.e.* to elements of $2^{\Theta} - \{\emptyset\}$) without specifying any detail of how to distribute the belief mass amongst the elements of a particular subset. In this case, then for any non-atomic subset of Θ, a belief mass on that subset expresses uncertainty regarding the probability distribution over the elements of the subset. More generally, a belief mass assignment m on Θ is defined as a function from $2^{\Theta} - \{\emptyset\}$ to $[0, 1]$ satisfying:

$$\sum_{x \subseteq \Theta} m(x) = 1. \tag{1}$$

Each nonempty subset $x \subseteq \Theta$ such that $m(x) > 0$ is called a focal element of m. Special names are used to describe specific BMA classes. When $m(\Theta) = 1$ the BMA is *vacuous*. When all the focal elements are atomic elements, the BMA is Bayesian. When $m(\Theta) = 0$ the BMA is *dogmatic* [7]. Let us note, that trivially, every Bayesian belief function is dogmatic. When all the focal elements are nestable (i.e. linearly ordered by inclusion), then the BMA is *consonant*.

[2] Called *basic probability assignment* in Shafer 1976 [5].

Given a particular frame of discernment and a BMA, the Dempster-Shafer theory (Shafer 1976 [5]) defines a belief function [3] $b(x)$. In addition, subjective logic (Jøsang 2001 [2]) defines a disbelief function $d(x)$, an uncertainty function $u(x)$, a relative atomicity function $a(x/y)$ and a probability expectation $\mathrm{E}(x)$. These are all defined as follows:

$$b(x) \triangleq \sum_{\emptyset \neq y \subseteq x} m(y) \qquad \forall\, x \in 2^{\Theta}\ , \tag{2}$$

$$d(x) \triangleq \sum_{y \cap x = \emptyset} m(y) \qquad \forall\, x \in 2^{\Theta}\ , \tag{3}$$

$$u(x) \triangleq \sum_{\substack{y \cap x \neq \emptyset \\ y \not\subseteq x}} m(y) \qquad \forall\, x \in 2^{\Theta}\ , \tag{4}$$

$$a(x/y) \triangleq \frac{|x \cap y|}{|y|} \qquad \forall\, x \in 2^{\Theta}\ ,\ y \in 2^{\Theta} - \{\emptyset\}\ , \tag{5}$$

$$\mathrm{E}(x) \triangleq \sum_{y \subseteq \Theta} m(y)\ a(x/y) \qquad \forall\, x \in 2^{\Theta}\ . \tag{6}$$

The relative atomicity function of a subset x relative to the frame of discernment Θ is simply denoted by $a(x)$. It can be shown that the belief, disbelief and uncertainty functions defined above satisfy:

$$b(x) + d(x) + u(x) = 1\ , \qquad x \in 2^{\Theta} - \{\emptyset\}. \tag{7}$$

The belief, disbelief and uncertainty functions are dependent through Eq.(7) so that one is redundant. As such they represent nothing more than the traditional Bel(x) (Belief) and Pl(x) (Plausibility) pair of Shaferian belief theory, where Bel$(x) = b(x)$ and Pl$(x) = b(x) + u(x)$. However, using (Bel, Pl) instead of $(b,\ d,\ u)$ would have produced unnecessary complexity in the product and coproduct operators described in Sections 4 and 5 below. It can also be noted that our disbelief function is equivalent to the traditional Dou(x) (Doubt) of Shaferian belief theory so that Dou$(x) = d(x)$. However, the interpretation of the term “doubt” is problematic in case of e.g. “total doubt”, i.e. where Dou$(x) = 1$, whereas the term “total disbelief”, i.e. with $d(x) = 1$ leaves little room for misinterpretation. We therefore prefer to use the term “disbelief” rather than “doubt”.

Definition (6) is equivalent to the pignistic probability function described by Smets & Kennes [8], and corresponds to the principle of insufficient reason: a belief mass assigned to the union of n atomic sets is split equally among these n sets. Section 9 below describes how belief functions can be mapped to beta probability distributions, thereby making pignistic probability equivalent to expected probability. In

[3] Denoted by Bel(x) in Shafer 1976 [5].

order to reflect this equivalence and to avoid any confusion, we prefer to use the term "probability expectation", denoted by $\mathrm{E}(x)$, both for belief functions and for probability distributions, rather than to use "pignistic probability" for the former and "probability expectation" for the latter.

Subjective logic operators apply to binary frames of discernment, so in case a frame is larger than binary, a coarsening is required in order to reduce its size to binary. Coarsening in subjective logic focuses on a particular subset $x \subset \Theta$, and produces a binary frame of discernment $X = \{x, \overline{x}\}$ containing x and its complement $\overline{x}$ in Θ. The reduced powerset of X is $2^X - \{\emptyset\} = \{\{x\}, \{\overline{x}\}, X\}$ which has $2^{|X|} - 1 = 3$ elements. We will first describe *simple coarsening* and subsequently describe *normal coarsening*.

Let b_x, d_x, u_x and a_x denote the belief, disbelief, uncertainty and relative atomicity functions of x on X. According to simple coarsening which is presented in Jøsang 2001 [2], these functions are defined as:

$$
\begin{aligned}
b_x &\triangleq b(x) \ , && (8)\\
d_x &\triangleq d(x) \ , && (9)\\
u_x &\triangleq u(x) \ , && (10)\\
a_x &\triangleq [\mathrm{E}(x) - b(x)]/u(x) \ . && (11)
\end{aligned}
$$

This coarsening is called "simple" because the belief, disbelief and uncertainty functions are identical to the original functions on Θ. The relative atomicity a_x on the other hand produces a synthetic relative atomicity value which does not represent the real relative atomicity of x on Θ in general, but one that satisfies:

$$\mathrm{E}(x) = b_x + a_x u_x \qquad (12)$$

which is a special case of Eq.(6).

Next, the normal coarsening method is described. According to normal coarsening which is presented in Jøsang and Grandison 2003 [3], the belief, disbelief, uncertainty and relative atomicity functions are defined as:

For $\mathrm{E}(x) \geq b(x) + a(x)u(x)$:

$$b_x \triangleq b(x) + (\mathrm{E}(x) - b(x) - a(x)u(x))/(1 - a(x)) \ , \tag{13}$$
$$d_x \triangleq d(x) \ , \tag{14}$$
$$u_x \triangleq u(x) - (\mathrm{E}(x) - b(x) - a(x)u(x))/(1 - a(x)) \ , \tag{15}$$
$$a_x \triangleq a(x) \ . \tag{16}$$

For $\mathrm{E}(x) < b(x) + a(x)u(x)$:

$$b_x \triangleq b(x) \ , \tag{17}$$
$$d_x \triangleq d(x) + (b(x) + a(x)u(x) - \mathrm{E}(x))/a(x) \ , \tag{18}$$
$$u_x \triangleq u(x) - (b(x) + a(x)u(x) - \mathrm{E}(x))/a(x) \ , \tag{19}$$
$$a_x \triangleq a(x) \ . \tag{20}$$

This coarsening is called “normal” because the relative atomicity reflects the true relative cardinality of an element in the original frame of discernment. It is important to note the distinction between the relative cardinality of an element x in its original frame of discernment, and in the coarsened binary frame of discernment. The former is expressed by the relative atomicity, and the latter is always 0.5

With normal coarsening, the belief, disbelief and uncertainty functions on the focused frame of discernment X are in general different from the belief, disbelief and uncertainty functions on the original frame of discernment Θ, so that $b_x \geq b(x)$, $d_x \geq d(x)$, and $u_x \leq u(x)$. The interpretation of the tendency of normal coarsening to decrease the uncertainty and increase the belief and disbelief functions, is that belief masses that contribute to the uncertainty function can represent varying amounts of uncertainty relative to a given proposition. When considering for example the frame of discernment $\Theta = \{x_1, x_2, x_3\}$ and the uncertainty function u_{x_1} of normal coarsening, then the belief mass $m(\{x_1, x_2\})$ represents a smaller amount of uncertainty, and should therefore contribute less to the uncertainty function u_{x_1} than for example the belief mass $m(\Theta)$.

Simple and normal coarsening will in general produce different results, but it can be shown that simple and normal coarsening are equivalent iff

$$\mathrm{E}(x) = b(x) + a(x)u(x) \ . \tag{21}$$

A coarsening for which Eq.(21) is satisfied will be called a *Bayesian coarsening*. This will be the case when the coarsening focuses on an element in 2^Θ that can have no partly overlapping focal elements other than Θ itself. In other words, a Bayesian coarsening partitions Θ in two parts x and $\overline{x}$, where $m(\Theta)$ is the only possible belief mass that can contribute to uncertainty about x.

In the terminology of subjective logic, an opinion ω_x held by an individual about a proposition x is the ordered quadruple (b_x, d_x, u_x, a_x). Note that b_x, d_x, u_x and a_x must all fall in the closed interval $[0, 1]$, and $b_x + d_x + u_x = 1$. For both simple and normal coarsening, the expected probability for x satisfies $\mathrm{E}(\omega_x) \triangleq \mathrm{E}(x) = b_x + a_x u_x$. Although the coarsened frame of discernment X is binary, an opinion about $x \in X$ carries information about the state space size of the original frame of discernment Θ through the relative atomicity parameter a_x.

The opinion space can be mapped into the interior of an equal-sided triangle, where, for an opinion $\omega_x = (b_x, d_x, u_x, a_x)$, the three parameters b_x, d_x and u_x determine the position of the point in the triangle representing the opinion. Fig.1 illustrates an example where the opinion about a proposition x from a binary frame of discernment has the value $\omega_x = (0.7,\ 0.1,\ 0.2,\ 0.5)$.

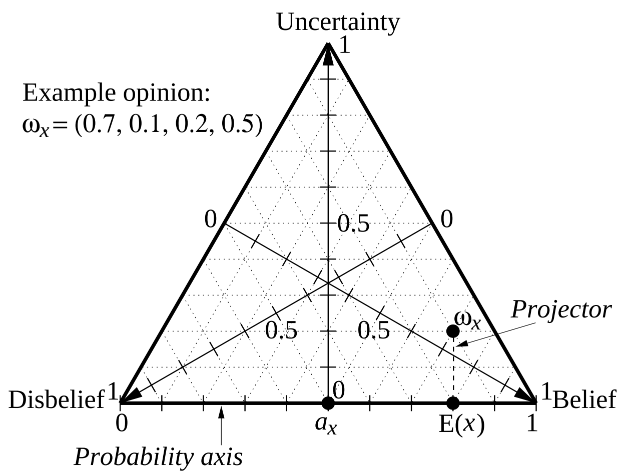


Fig. 1. Opinion triangle with example opinion

The top vertex of the triangle represents uncertainty, the bottom left vertex represents disbelief, and the bottom right vertex represents belief. The parameter b_x is the value of a linear function on the triangle which takes value 0 on the edge which joins the uncertainty and disbelief vertices and takes value 1 at the belief vertex. In other words, b_x is equal to the quotient when the perpendicular distance between the opinion point and the edge joining the uncertainty and disbelief vertices is divided by the perpendicular distance between the belief vertex and the same edge. The parameters d_x and u_x are determined similarly. The edge joining the disbelief and belief vertices is called the probability axis. The relative atomicity is indicated by a point on the probability axis, and the projector starting from the opinion point is parallel to the line that joins the uncertainty vertex and the relative atomicity point on the probability axis. The point at which the projector meets the probability axis determines the probability expectation value of the opinion, *i.e.* it coincides with the point corresponding to expectation value $b_x + a_x u_x$.

3 Products of Binary Frames of Discernment

Multiplication and comultiplication in subjective logic are binary operators that take opinions about two elements from distinct binary frames of discernment as input parameters. The product and coproduct 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.2 below.

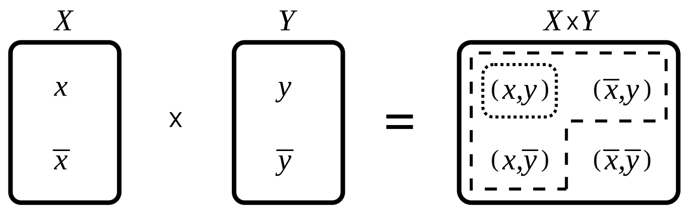


Fig. 2. Cartesian product of two binary frames of discernment

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.2. 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.2. Obviously $X \times Y$ is not binary, and coarsening is required in order to determine the product and coproduct opinions. The reduced powerset $2^{X \times Y} - \{\emptyset\}$ contains $2^{|X \times Y|} - 1 = 15$ elements. A short notation for the elements of $2^{X \times Y}$ is used below so that for example $\{(x,y),\ (x,\overline{y})\} = \{x\} \times Y$. The BMA on $X \times Y$ as a function of the opinions on x and y is defined by:

$$\begin{aligned}
m(\{(x,y)\}) = b_x b_y\ , &\quad m(\{(\overline{x},y)\}) = d_x b_y\ , \quad m(X \times \{y\}) = u_x b_y\ , \\
m(\{(x,\overline{y})\}) = b_x d_y\ , &\quad m(\{(\overline{x},\overline{y})\}) = d_x d_y\ , \quad m(X \times \{\overline{y}\}) = u_x d_y\ , \\
m(\{x\} \times Y) = b_x u_y\ , &\quad m(\{\overline{x}\} \times Y) = d_x u_y\ , \quad m(X \times Y) = u_x u_y\ .
\end{aligned} \tag{22}$$

It can be shown that the sum of the above belief masses always equals 1. The product does not produce any belief mass on the following elements:

$$\begin{aligned}
&\{(x,y),\ (\overline{x},\overline{y})\}\ , \quad (X \times \{y\}) \cup \{(x,\overline{y})\}\ , \quad (X \times \{\overline{y}\}) \cup \{(x,y)\}\ , \\
&\{(x,\overline{y}),\ (\overline{x},y)\}\ , \quad (X \times \{y\}) \cup \{(\overline{x},\overline{y})\}\ , \quad (X \times \{\overline{y}\}) \cup \{(\overline{x},y)\})\ .
\end{aligned} \tag{23}$$

The belief functions in for example $x \wedge y$ and $x \vee y$ can now be determined so that:

$$\begin{aligned} b(x \wedge y) &= m(\{(x,y)\}), \\ b(x \vee y) &= m(\{(x,y)\}) + m(\{(x,\overline{y})\}) + m(\{(\overline{x},y)\}) + \\ &\qquad m(\{x\} \times Y) + m(X \times \{y\})\ . \end{aligned} \tag{24}$$

The normal relative atomicity functions for $x \wedge y$ and $x \vee y$ can be determined by working in the respective "primitive" frames of discernment, Θ_X and Θ_Y which underlie the definitions of the sets x and y, respectively. A sample yields a value of (x, y) in the frame of discernment $X \times Y$ exactly when the sample yields an atom $\theta_X \in x$ in the frame of discernment Θ_X and an atom $\theta_Y \in y$ in the frame of discernment Θ_Y. In other words, a sample yields a value of (x, y) in the frame of discernment $X \times Y$ exactly when the sample yields an atom $(\theta_X, \theta_Y) \in x \times y$ in the frame of discernment $\Theta_X \times \Theta_Y$, so that $(x, y) \in X \times Y$ corresponds to $x \times y \subseteq \Theta_X \times \Theta_Y$ in a natural manner. Similarly, $(x, \overline{y})$ corresponds to $x \times \overline{y}$, $(\overline{x}, y)$ corresponds to $\overline{x} \times y$, and $(\overline{x}, \overline{y})$ corresponds to $\overline{x} \times \overline{y}$. The normal relative atomicity function for $x \wedge y$ is equal to:

$$a(x \wedge y) = \frac{|x \times y|}{|\Theta_X \times \Theta_Y|} = \frac{|x|\ |y|}{|\Theta_X|\ |\Theta_Y|} = \frac{|x|}{|\Theta_X|}\ \frac{|y|}{|\Theta_Y|} = a(x)a(y), \tag{25}$$

Similarly, the normal relative atomicity of $x \vee y$ is equal to

$$\begin{aligned} a(x \vee y) &= \frac{|(x \times y) \cup (x \times \overline{y}) \cup (\overline{x} \times y)|}{|\Theta_X \times \Theta_Y|} = \frac{|x \times y| + |x \times \overline{y}| + |\overline{x} \times y|}{|\Theta_X \times \Theta_Y|} \\ &= \frac{|x|\ |y| + |x|\ |\overline{y}| + |\overline{x}|\ |y|}{|\Theta_X|\ |\Theta_Y|} = a(x)a(y) + a(x)a(\overline{y}) + a(\overline{x})a(y) \\ &= a(x) + a(y) - a(x)a(y). \end{aligned}$$

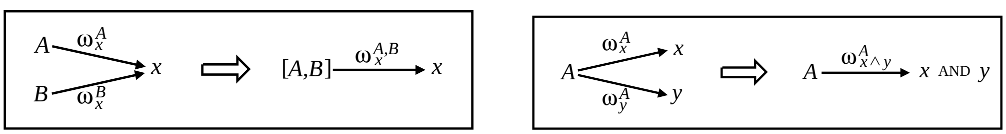
By applying simple or normal coarsening to the product frame of discernment and BMA, the simple and normal product and coproduct opinions emerge. A coarsening that focuses on $x \wedge y$ produces the product, and a coarsening that focuses on $x \vee y$ produces the coproduct. A Bayesian coarsening (i.e. when simple and normal coarsening are equivalent) is only possible in exceptional cases because some terms of Eq.(22) other than $m(X \times Y)$ will in general contribute to uncertainty about $x \wedge y$ in the case of multiplication, and to uncertainty about $x \vee y$ in the case of comultiplication. Specifically, Bayesian coarsening requires $m(X \times \{y\}) = m(\{x\} \times Y) = 0$ in case of multiplication, and $m(X \times \{\overline{y}\}) = m(\{\overline{x}\} \times Y) = 0$ in case of comultiplication. Non-Bayesian coarsenings will cause the product and coproduct of opinions to deviate from the analytically correct product and coproduct. However, the magnitude of this deviation is always small, as will be explained in Section 10.

The symbols "$\cdot$" and "$\sqcup$" will be used to denote multiplication and comultiplication of opinions respectively so that we can write:

$$\omega_{x\wedge y} \triangleq \omega_x \cdot \omega_y \tag{26}$$
$$\omega_{x\vee y} \triangleq \omega_x \sqcup \omega_y \tag{27}$$

The product of the opinions about x and y is thus the opinion about the conjunction of x and y. Similarly, the coproduct of the opinions about x and y is the opinion about the disjunction of x and y. The exact expressions for product and coproduct are given in Sections 4 and 5.

Readers might have noticed that Eq.(22) can appear to be a direct application of the non-normalised version of Dempster's rule (i.e. the conjunctive rule of combination) [5] which is a method of belief fusion. However the difference is that Dempster's rule applies to the beliefs of two different and independent observers faced with the same frame of discernment, whereas the Cartesian product of Eq.(22) applies to the beliefs of the same observer faced with two different and independent frames of discernment. Let ω_x^A and ω_x^B represent the opinions of two observers A and B about the same proposition x, and let $\omega_x^{A,B}$ represent the fusion of A and B's opinions. Let further ω_x^A and ω_y^A represent observer A's opinions about the propositions x and y, and let $\omega_{x\wedge y}^A$ represent the product of those opinions. Fig.3 below illustrates the difference between belief fusion and belief product.



(a) Belief fusion. (b) Belief product.

Fig. 3. Conceptual difference between belief fusion and belief product.

The Cartesian product as described here thus has no relationship to Dempster's rule and belief fusion other than the apparent similarity between Eq.(22) and Dempster's rule.

4 Simple Multiplication and Comultiplication

The product and coproduct of opinions held by a single individual with respect to independent propositions x and y determine the individual's opinions about their conjunction, $x \wedge y$, and disjunction, $x \vee y$, respectively. The "simple" approach to determining the product of opinions held by an individual about independent propositions x and y is to note that an outcome determines that $x \wedge y$ is certainly true if and only if the outcome determines both that x is certainly true and that y is certainly true, so that it appears natural to take $b_{x \wedge y} = b_x b_y$. Also, we can note that an outcome determines that $x \wedge y$ is false if and only if it determines that at least one of x and y is false, so that it appears natural to take $d_{x \wedge y} = d_x + d_y - d_x d_y$. Since $b_{x \wedge y} + d_{x \wedge y} + u_{x \wedge y} = 1$, then $u_{x \wedge y} = b_x u_y + u_x b_y + u_x u_y$. This makes sense since an outcome can lead to no definitive conclusion about the truth of $x \wedge y$ if and only if it does not lead to the definite conclusion that either is false and it does not lead to the definite conclusion that both are true, leaving only the three alternatives:

- x is definitely true and no definitive conclusion can be drawn about y;
- no definitive conclusion can be drawn about x and y is definitely true;
- no definitive conclusion can be drawn about x or y.

Since the expected probability for x is $E[\omega_x] = b_x + a_x u_x$ and the expected probability for y is $E[\omega_y] = b_y + a_y u_y$, and x and y are independent, then the expected probability for $x \wedge y$ should be $E[\omega_x]E[\omega_y]$, so that the relative atomicity is given by

$$a_{x \wedge y} = \frac{b_x a_y u_y + a_x u_x b_y + a_x u_x a_y u_y}{b_x u_y + u_x b_y + u_x u_y}. \tag{28}$$

It is this simple product of opinions which is referred to in Jøsang 2001 [2] as their propositional conjunction. A numerical example of simple multiplication is visualised in Fig.4 [4] . Note that the relative atomicity $a_{x \wedge y}$ does not reflect the real relative cardinality of $x \wedge y$ in $X \times Y$.

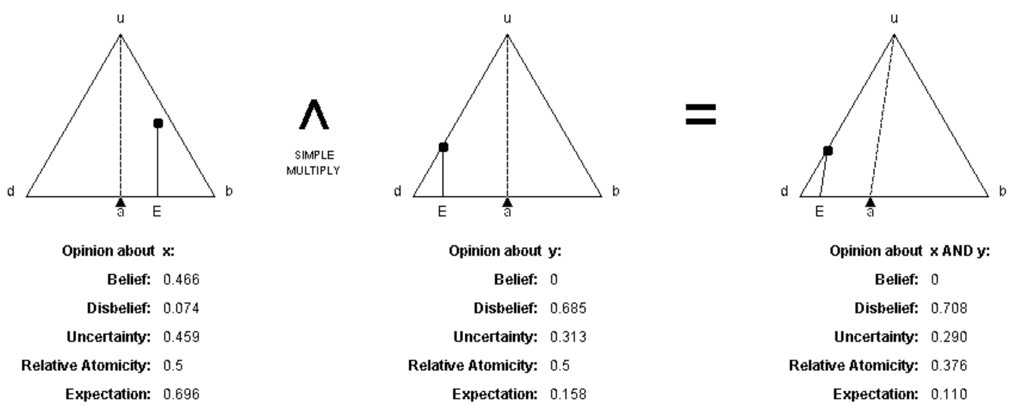


Fig. 4. Visualisation of numerical example of the simple multiplication operator

Similarly, the "simple" approach to determining the coproduct of opinions held by an individual about independent propositions x and y is to note that an outcome determines that $x \vee y$ is true if and only if it determines that at least one of x and y is true, so that it appears natural to take $b_{x \vee y} = b_x + b_y - b_x b_y$. Also, we can note that the outcome determines that $x \vee y$ is definitely false if and only if the

[4] Online demo at http://security.dstc.edu.au/spectrum/trustengine/

outcome determines both that x is definitely false and that y is definitely false, so that it appears natural to take $d_{x\vee y} = d_x d_y$. Since $b_{x\vee y} + d_{x\vee y} + u_{x\vee y} = 1$, then $u_{x\vee y} = d_x u_y + u_x d_y + u_x u_y$. This makes sense since an outcome can lead to no definitive conclusion about the truth of $x \vee y$ if and only if it does not lead to the definite conclusion that either is true and it does not lead to the definite conclusion that both are false, leaving only the three alternatives:

- x is definitely false and no definitive conclusion can be drawn about y;
- no definitive conclusion can be drawn about x and y is definitely false;
- no definitive conclusion can be drawn about x or y.

Since the expected probability for x is $E[\omega_x] = b_x + a_x u_x$ and the expected probability for y is $E[\omega_y] = b_y + a_y u_y$, and x and y are independent, then the expected probability for $x \vee y$ should be $E[\omega_x] + E[\omega_y] - E[\omega_x]E[\omega_y]$, so that the atomicity is given by

$$\begin{aligned} a_{x\vee y} &= 1 - \frac{d_x(1-a_y)u_y + (1-a_x)u_x d_y + (1-a_x)u_x(1-a_y)u_y}{d_x u_y + u_x d_y + u_x u_y} \\ &= \frac{d_x a_y u_y + a_x u_x d_y + (a_x + a_y - a_x a_y)u_x u_y}{d_x u_y + u_x d_y + u_x u_y} \\ &= \frac{a_x u_x + a_y u_y - a_x u_x b_y - b_x a_y u_y - a_x u_x a_y u_y}{u_x + u_y - b_x u_y - u_x b_y - u_x u_y}. \end{aligned}$$

It is this simple coproduct of opinions which is referred to in Jøsang 2001 [2] as their propositional disjunction. Note that the relative atomicity $a_{x\vee y}$ does not reflect the real relative cardinality of $x \vee y$ in $X \times Y$.

5 Normal Multiplication and Comultiplication

Normal conjunction and disjunction of opinions about independent propositions x and y are taken in such a way that the atomicities of $x \wedge y$ and $x \vee y$ are dependent only on the atomicities of x and y, and not on the beliefs, disbeliefs and uncertainties. By the arguments within Section 3 for justifying the relative atomicities, we can set $a_{x\wedge y} = a_x a_y$ and $a_{x\vee y} = a_x + a_y - a_x a_y$. This is in contrast to the case of "simple" conjunction and "simple" disjunction as discussed above, where atomicities of both the conjunction and the disjunction are dependent on the beliefs, disbeliefs and uncertainties of x and y. Given opinions about independent propositions, x and y, then under normal coarsening of the BMA for the Cartesian product of the binary frames of discernment, the normal opinion for the conjunction, $x \wedge y$, is given by

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

A numerical example of the normal multiplication operator is visualised in Fig.5 below. Note that in this case, the relative atomicity $a_{x\wedge y}$ is equal to the real relative cardinality of $x \wedge y$ in $X \times Y$.

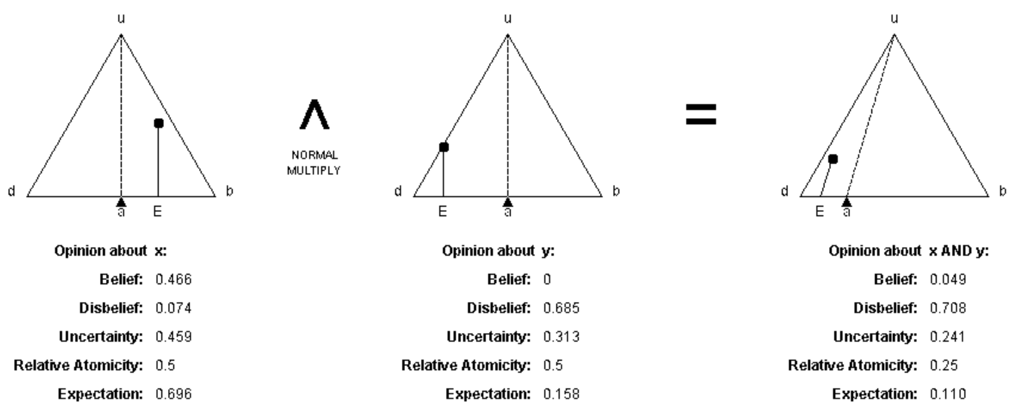


Fig. 5. Visualisation of numerical example of the normal multiplication operator

The formulae for the opinion about $x \wedge y$ are well formed unless $a_x = 1$ and $a_y = 1$, in which case the opinions ω_x and ω_y can be regarded as limiting values, and the product is determined by the relative rates of approach of a_x and a_y to 1. Specifically, if η is the limit of $\frac{1-a_x}{1-a_y}$, then

$$
\begin{aligned}
b_{x\wedge y} &= b_x b_y + \frac{\eta b_x u_y + u_x b_y}{\eta + 1}, \\
d_{x\wedge y} &= d_x + d_y - d_x d_y, \\
u_{x\wedge y} &= u_x u_y + \frac{b_x u_y + \eta u_x b_y}{\eta + 1}, \\
a_{x\wedge y} &= 1.
\end{aligned}
$$

Under normal coarsening of the BMA for the Cartesian product of the binary frames of discernment, the normal opinion for the disjunction, $x \vee y$, is given by

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

A numerical example of the normal comultiplication operator is visualised in Fig.6 below. Note that in this case, the relative atomicity $a_{x \vee y}$ is equal to the real relative cardinality of $x \vee y$ in $X \times Y$.

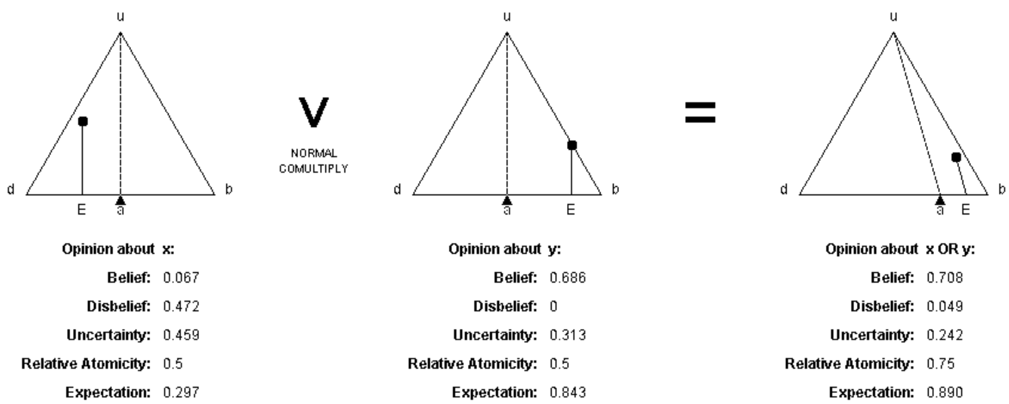


Fig. 6. Visualisation of numerical example of the normal comultiplication operator

The formulae for the opinion about $x \vee y$ are well formed unless $a_x = 0$ and $a_y = 0$. In the case that $a_x = a_y = 0$, the opinions ω_x and ω_y can be regarded as limiting values, and the product is determined by the relative rates of approach of a_x and a_y

to 0. Specifically, if ζ is the limit of $\frac{a_x}{a_y}$, then

$$
\begin{aligned}
b_{x\vee y} &= b_x + b_y - b_x b_y,\\
d_{x\vee y} &= d_x d_y + \frac{\zeta d_x u_y + u_x d_y}{\zeta + 1},\\
u_{x\vee y} &= u_x u_y + \frac{d_x u_y + \zeta u_x d_y}{\zeta + 1},\\
a_{x\vee y} &= 0.
\end{aligned}
$$

This is a self-dual system under $b \leftrightarrow d$, $u \leftrightarrow u$, $a \leftrightarrow 1 - a$, 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:

$$
\omega_{x\wedge(y\vee z)} \neq \omega_{(x\wedge y)\vee(x\wedge z)} \tag{29}
$$

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. In fact the corresponding result only holds for binary logic.

6 Normal Division and Codivision

The inverse operation to multiplication is 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 $a_x \leq a_y$, $d_x \geq d_y$, and

$$
\begin{aligned}
b_x &\geq \frac{a_x(1-a_y)(1-d_x)b_y}{(1-a_x)a_y(1-d_y)},\\
u_x &\geq \frac{(1-a_y)(1-d_x)u_y}{(1-a_x)(1-d_y)}.
\end{aligned}
$$

The opinion $(b_{x\overline{\wedge}y}, d_{x\overline{\wedge}y}, u_{x\overline{\wedge}y}, a_{x\overline{\wedge}y})$, which is the quotient of the opinion about x and the opinion about y, is given by

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

if $a_x < a_y$. If $0 < a_x = a_y$, then the conditions required so that the opinion about x can be divided by the opinion about y are

$$
\begin{aligned}
b_x &= \frac{(1-d_x)b_y}{1-d_y},\\
u_x &= \frac{(1-d_x)u_y}{1-d_y},
\end{aligned}
$$

and in this case,

$$
\begin{aligned}
d_{x\overline{\wedge}y} &= \frac{d_x - d_y}{1 - d_y},\\
a_{x\overline{\wedge}y} &= 1.
\end{aligned}
$$

The only information available about $b_{x\overline{\wedge}y}$ and $u_{x\overline{\wedge}y}$ is that

$$
b_{x\overline{\wedge}y} + u_{x\overline{\wedge}y} = \frac{1-d_x}{1-d_y}.
$$

On the other hand, $b_{x\overline{\wedge}y}$ and $u_{x\overline{\wedge}y}$ can be determined if the opinion about x is considered as the limiting value of other opinions which can be divided by the opinion about y. The limiting value of the quotient of the opinions is determined by the relative rates of approach of a_x, b_x and u_x to their limits. Specifically, if γ is the limit of

$$
\frac{a_y(1-a_y)}{(a_y - a_x)(b_y + a_y u_y)} \left(\frac{(1-d_y)b_x}{1-d_x} - b_y \right) + \frac{b_y}{b_y + a_y u_y},
$$

then $0 \le \gamma \le 1$, and the limiting values of $b_{x\overline{\wedge}y}$ and $u_{x\overline{\wedge}y}$ are

$$
\begin{aligned}
b_{x\overline{\wedge}y} &= \frac{\gamma(1-d_x)}{1-d_y},\\
u_{x\overline{\wedge}y} &= \frac{(1-\gamma)(1-d_x)}{1-d_y}.
\end{aligned}
$$

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 $a_x \geq a_y$, $b_x \geq b_y$, and

$$\begin{aligned}
d_x &\geq \frac{(1-a_x)a_y(1-b_x)d_y}{a_x(1-a_y)(1-b_y)}, \\
u_x &\geq \frac{a_y(1-b_x)u_y}{a_x(1-b_y)}.
\end{aligned}$$

The opinion $(b_{x \overline{\vee} y}, d_{x \overline{\vee} y}, u_{x \overline{\vee} y}, a_{x \overline{\vee} y})$, which is the co-quotient of the opinion about x and the opinion about y, is given by

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

if $a_x > a_y$. If $a_x = a_y < 1$, then the conditions required so that the opinion about x can be codivided by the opinion about y are

$$\begin{aligned}
d_x &= \frac{(1-b_x)d_y}{1-b_y}, \\
u_x &= \frac{(1-b_x)u_y}{1-b_y},
\end{aligned}$$

and in this case,

$$\begin{aligned}
b_{x \overline{\vee} y} &= \frac{b_x - b_y}{1 - b_y}, \\
a_{x \overline{\vee} y} &= 0.
\end{aligned}$$

The only information available about $d_{x \overline{\vee} y}$ and $u_{x \overline{\vee} y}$ is that

$$d_{x \overline{\vee} y} + u_{x \overline{\vee} y} = \frac{1 - b_x}{1 - b_y}.$$

On the other hand, $d_{x \overline{\vee} y}$ and $u_{x \overline{\vee} y}$ can be determined if the opinion about x is considered as the limiting value of other opinions which can be codivided by the opinion about y. The limiting value of the co-quotient of the opinions is determined by the relative rates of approach of a_x, d_x and u_x to their limits. Specifically, if δ is the limit of

$$\frac{a_y(1-a_y)}{(a_x-a_y)(d_y+(1-a_y)u_y)}\left(\frac{(1-b_y)d_x}{1-b_x}-d_y\right)+\frac{d_y}{d_y+(1-a_y)u_y},$$

then $0 \leq \delta \leq 1$, and the limiting values of $d_{x\overline{\vee}y}$ and $u_{x\overline{\vee}y}$ are

$$\begin{aligned} d_{x\overline{\vee}y} &= \frac{\delta(1-b_x)}{1-b_y}, \\ u_{x\overline{\vee}y} &= \frac{(1-\delta)(1-b_x)}{1-b_y}. \end{aligned}$$

Given the opinion about x and the atomicity of y, it is possible to use the triangular representation of the opinion space from Fig.1 to describe geometrically the range of opinions about $x \wedge y$ and $x \vee y$.

In the case of $x \wedge y$, take the projector for ω_x, and take the intersections of the projector with the line of zero uncertainty and the line of zero belief (A and B, respectively). The intersection, A, with the line of zero uncertainty determines the probability expectation value of ω_x. Take the point, C, on the line of zero uncertainty whose distance from the disbelief vertex is a_y times the distance between the disbelief vertex and A. Take the line BC and the line through A parallel to BC. Let D and E be the intersections of these lines with the line of constant disbelief through ω_x, so that the disbelief is equal to d_x. Then $\omega_{x\wedge y}$ falls in the closed triangle determined by D, E and the disbelief vertex, and the atomicity of $x \wedge y$ is given by $a_{x\wedge y} = a_x a_y$.

In Fig. 7, this is demonstrated with an example where $\omega_x = (0.3, 0.3, 0.4, 0.6)$ and $a_y = 0.4$. The opinion ω_x has been marked with a small black circle (on the side of the shaded triangle opposite the disbelief vertex). The intersections A and B of the projector of x with the line of zero uncertainty and the line of zero belief, respectively, have been marked. The point C has been placed on the probability axis so that its distance from the disbelief vertex is 0.4 times the distance between A and the disbelief vertex (since $a_y = 0.4$). The line BC, whose direction corresponds to an atomicity of 0.24 (*i.e.* the atomicity of $x \wedge y$), has also been drawn in the triangle, and its intersection with the line of constant disbelief through ω_x (with disbelief equal to 0.3) has been marked with a white circle. This is the point D, although not marked as such in the figure. The line through A parallel to BC has also been drawn in the triangle, and its intersection with the line of constant disbelief through ω_x (the point E, although also not marked as such) has also been marked with a white circle. The triangle with vertices D, E and the disbelief vertex has been shaded, and the normal product $\omega_{x\wedge y}$ of the opinions must fall within the shaded triangle or on its boundary. In other words, the closure of the shaded triangle is the range of all possible values for the opinion $\omega_{x \wedge y}$.

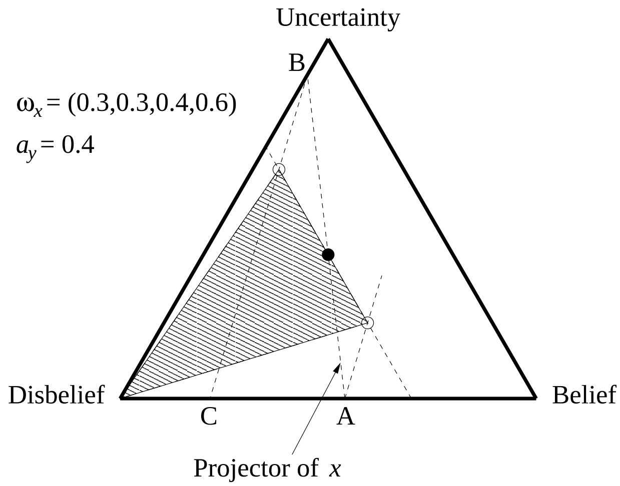


Fig. 7. Range of possible opinions for normal product

In the case of $x \vee y$, take the projector for ω_x, and take the intersections of this line with the line of zero uncertainty and the line of zero disbelief (A and B, respectively). The intersection, A, with the line of zero uncertainty determines the probability expectation value of ω_x. Take the point, C, on the line of zero uncertainty whose distance from the belief vertex is $1 - a_y$ times the distance between the belief vertex and A. Take the line BC and the line through A parallel to BC. Let D and E be the intersections of these lines with the line of constant belief through ω_x, so that the belief is equal to b_x. Then $\omega_{x \vee y}$ falls in the closed triangle determined by D, E and the belief vertex, and the atomicity of $x \vee y$ is given by $a_{x \vee y} = a_x + a_y - a_x a_y$.

The conditions required so that ω_x can be divided by ω_y can be described geometrically. Take the projector for ω_y, and take the intersections of this line with the line of zero uncertainty and with the line of zero belief. Take the lines through each of these points which are parallel to the projector for ω_x (it is required that $a_x < a_y$). Take the intersections of these lines with the line of constant disbelief through ω_y. Then ω_x can be divided by ω_y, provided ω_x falls in the closed triangle determined by these two points and the disbelief vertex.

Fig. 7 can be used to demonstrate. If $\omega_y = (0.3, 0.3, 0.4, 0.6)$ and $a_x = 0.24$, then the black circle denotes ω_y, the projector of y is the line AB, the lines through A and B parallel to the director for atomicity 0.24 are drawn in the triangle, and their intersections with the line of constant disbelief through ω_y are marked by the white circles. The closure of the shaded triangle is the range of all possible values of ω_x that allow ω_x to be divided by ω_y.

The conditions required so that ω_x can be codivided by ω_y can be described geometrically. Take the projector for ω_y, and take the intersections of this line with the line of zero uncertainty and with the line of zero disbelief. Take the lines through each of these points which are parallel to the projector for ω_x (it is required that $a_x > a_y$). Take the intersections of these lines with the line of constant belief through ω_y. Then ω_x can be codivided by ω_y, provided ω_x falls in the closed triangle determined by these two points and the belief vertex.

7 Probability Distributions over Subsets of Θ

In the previous sections, two variants of the multiplication and comultiplication operators were described. In order to interpret these operators and assess their correctness, we will define a mapping between opinions and beta probability distributions. The purpose of this is to be able to compare products of opinions with products of beta distributions, and similarly for coproducts. Ideally, they should be equivalent, but unfortunately that is not always possible.

For this analysis, we are interested in knowing the probability distribution over subsets of the frame of discernment. In the binary case it is determined by the beta distribution. In the case of exhaustive and mutually exclusive subsets, it is determined by the Dirichlet distribution which we explain in some detail in this section.

The Dirichlet distribution describes the joint distribution of k random variables $\{P_i\}_{i=1}^k$ (or equivalently, a k-component random variable $(P_i)_{i=1}^k$) with sample space $[0,1]^k$, subject to

$$\sum_{i=1}^{k} P_i = 1,$$

so that in fact, the sample space is actually

$$\{(p_i)_{i=1}^k \in [0,1]^k : \sum_{i=1}^{k} p_i = 1\},$$

of dimension $k-1$ (*i.e.* the sample space has $k-1$ degrees of freedom). Note that for any sample from a Dirichlet random variable, it is sufficient to determine values for P_i for any $k-1$ elements i of $\{1,\ldots,k\}$, as this uniquely determines the value of the other variable. The Dirichlet distribution has k positive real parameters $(\alpha_i)_{i=1}^k$, each corresponding to one of the random variables, and the probability distribution function for $(P_i)_{i=1}^{k-1}$ on the sample space

$$\{(p_i)_{i=1}^{k-1} \in [0,1]^{k-1} : \sum_{i=1}^{k-1} p_i \leq 1\}$$

is given by

$$f(p_1, \ldots, p_{k-1}) = \frac{\Gamma\left(\sum_{i=1}^{k} \alpha_i\right)}{\prod_{i=1}^{k} \Gamma(\alpha_i)} \prod_{i=1}^{k} p_i^{\alpha_i - 1},$$

where $p_k = 1 - \sum_{i=1}^{k-1} p_i$. Note that although the definition of Dirichlet random variable is symmetric, the probability density is not symmetrically defined (p_k is not an argument of the probability distribution function). The same functional form for the probability distribution function arises for any choice of $k-1$ of the component random variables (since the Jacobian of the transformations between such subsets always has absolute value 1).

We now ask what happens if instead of $(P_i)_{i=1}^{k}$, we take sums of the random variables, so we are interested in the distribution of

$$\left(\sum_{i \in J} P_i\right)_{J \in \mathbb{P}},$$

for nontrivial partitions $\mathbb{P}$ of $\{1, \ldots, k\}$ (*i.e.* any partition not consisting of the elements $\{1, \ldots, k\}$ and $\emptyset$). The distribution is given by:

Theorem 7.1

> If $(P_i)_{i=1}^{k} \sim$ Dirichlet$(\alpha_1, \ldots, \alpha_k)$, then
>
> $$\left(\sum_{i \in J} P_i\right)_{J \in \mathbb{P}} \sim \text{Dirichlet}\left(\sum_{i \in J} \alpha_i\right)_{J \in \mathbb{P}},$$
>
> *i.e.* the distribution is still a Dirichlet distribution, and the parameter corresponding to a specific sum of random variables is given by the sum of the parameters corresponding to the constituent addends.

The proof of this theorem can be found in standard textbooks, and is also given in the appendix. It follows that for any nontrivial subset J of $\{1, \ldots, k\}$ (*i.e.* any subset J not equal to $\{1, \ldots, k\}$ or $\emptyset$), $\sum_{i \in J} P_j$ is a beta distributed random variable with parameters $\sum_{i \in J} \alpha_i$ and $\sum_{i \in J'} \alpha_i$, where $J' = \{1, \ldots, k\} \backslash J$, *i.e.*

$$\sum_{i \in J} P_i \sim \text{beta}\left(\sum_{i \in J} \alpha_i, \sum_{i \in J'} \alpha_i\right).$$

This is because a random variable $P \sim$ beta(α, β) if and only if $(P, 1 - P) \sim$ Dirichlet(α, β).

In plain language this means that when a Dirichlet distribution can be defined over an exhaustive and mutually exclusive partitioning of the frame of discernment, it is possible to define a beta probability distribution over any binary coarsening of this partitioning. This corresponds to the Bayesian coarsening that was defined in Section 2.

8 *A Priori* Distribution for k Alternatives

Now, we come to the question of an *a priori* distribution function for the probabilities of k exhaustive and mutually exclusive alternatives (*e.g.* k different colours of balls in an urn). Let P_i denote the random variable describing the probability of a random sample (*e.g.* drawing a ball) yielding alternative i. Since P_i describes a probability, then the sample space for $(P_i)_{i=1}^k$ is $[0,1]^k$. Since the alternatives are exhaustive and mutually exclusive, then

$$\sum_{i=1}^{k} P_i = 1.$$

Generalising the case of 2 alternatives (with their beta distribution), we will take an *a priori* Dirichlet distribution. Since there is no reason to assume a preference for any alternative over any other alternative, then the parameters will be taken to be equal (with the result that $\mathrm{E}[P_i] = \frac{1}{k}$ for all i). In the case of 2 alternatives, a uniform distribution has been assumed (*i.e.* $\mathrm{beta}(1,1)$). The question arises as to whether this fact can be used to determine the common value of the parameters in the case of k alternatives on the grounds of consistency. It can be argued that such a determination is possible, and that the common value of the parameters is $\frac{2}{k}$. The argument goes as follows. For integers m and n, take a set of mn exhaustive mutually exclusive alternatives, and a partition of the set into m classes, each with n elements. The *a priori* distribution for the probabilities of the mn alternatives is a Dirichlet distribution with the common value of the parameters being given by $\alpha(mn)$ (here, $\alpha(k)$ denotes the common value of the parameters in the case where there are k alternatives). It follows that for the partition, the distribution for the probabilities of the m alternative classes is a Dirichlet distribution with a common value for the parameters, equal to $n\,\alpha(mn)$. Since the m classes are exhaustive mutually exclusive alternatives in their own right, with no reason for preference for any over the others, then the distribution for the probabilities should have a common value of the parameters equal to $\alpha(m)$, and so consistency requires that $\alpha(m) = n\,\alpha(mn)$. Since

$$m\,\alpha(m) = mn\,\alpha(mn) = n\,\alpha(n)$$

for all positive integers m, n, then $\alpha(n) = \frac{C}{n}$ for some constant C. Substituting $\alpha(2) = 1$ (corresponding to the uniform distribution in the case of 2 alternatives), then $C = 2$, and the common value of the parameters in the case of k alternatives is $\frac{2}{k}$.

Let J be an l-element subset of $\{1,\ldots,k\}$, then

$$\sum_{i\in J} P_i \sim \text{beta}\left(\frac{2l}{k}, \frac{2(k-l)}{k}\right).$$

This means that, in the case of variously coloured balls in an urn, if the expected *a priori* probability of picking a ball of a given colour in the absence of bias is a, then the *a priori* distribution for the probability is:

$$\text{beta}(2a, 2(1-a)) \ , \tag{30}$$

and it seems reasonable to extend this assumption to the more general case (*i.e.* in any binary event, if the expected *a priori* probability in the absence of bias is a, then an *a priori* distribution according to Eq.(30) will be assumed). Bayesian updating now allows new evidence to be added. Let r be the number of observed events of type x, and let s denote the number of observed events different from x, then the updated beta distribution can be expressed as:

$$\text{beta}(r + 2a,\ s + 2(1-a)) \ . \tag{31}$$

For example, if an observer is presented with an urn containing red and black balls, without knowing the proportion of each colour, then there is no reason to expect that the probability of picking a red ball should be greater or less than the probability of picking a black ball, so the *a priori* probability of picking a red ball is $a = 0.5$, and the *a priori* beta distribution is beta(1,1). Assume that the observer picks 8 balls of which 7 turn out to be red and only one turns out to be black. The updated beta distribution of the outcome of picking red balls is beta$(8, 2)$ which is illustrated in Fig.8.

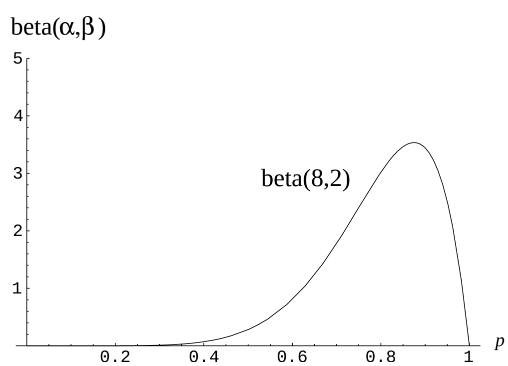


Fig. 8. Beta distribution after 7 positive and 1 negative observations

So far so good. However everything is not as simple as it seems, because there are cases where the *a priori* distribution for the probability in the absence of bias can not be determined according to the above analysis.

Take for example the following case where an event whose expected probability in the absence of bias is $a = \frac{1}{4}$, but whose *a priori* distribution for the probability is not beta $\left(\frac{1}{2}, \frac{3}{2}\right)$ as Eq.(30) would dictate.

Theorem 8.1

> Let Q and R be independent random variables, with identical uniform distributions over $[0, 1]$ (so $Q \sim$ beta$(1, 1)$ and $R \sim$ beta$(1, 1)$), then the probability distribution function for the random variable $P = QR$ is given by $f(p) = -\ln p$ for $0 < p < 1$.

The proof of this theorem is given in the appendix. Specifically, this means that if Q and R are random variables representing the probabilities of propositions x and y (which are independent), with Q and R having *a priori* uniform distributions, then P represents the probability of the conjunction $x \wedge y$, and has probability distribution function $f(p) = -\ln p$ with probability expectation value $\frac{1}{4}$.

This is the case of the independent propositions x and y, where we are taking four exhaustive and mutually exclusive propositions ($x \wedge y$, $x \wedge \overline{y}$, $\overline{x} \wedge y$, $\overline{x} \wedge \overline{y}$) with no reason for preferring any of the propositions over any of the others. As Theorem 8.1 shows, the distribution function for the probability of $x \wedge y$ is $-\ln p$ for $0 < p < 1$. Note that in the absence of bias, the probabilities of x and y are each expected to be $\frac{1}{2}$, so that the probability of $x \wedge y$ is expected to be $\frac{1}{4}$.

We will contrast this with the case of four exhaustive and mutually exclusive propositions x_1, x_2, x_3, x_4, with no reason for preferring any of these propositions over any of the others. The *a priori* distribution of the corresponding probabilities is 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.9 below.

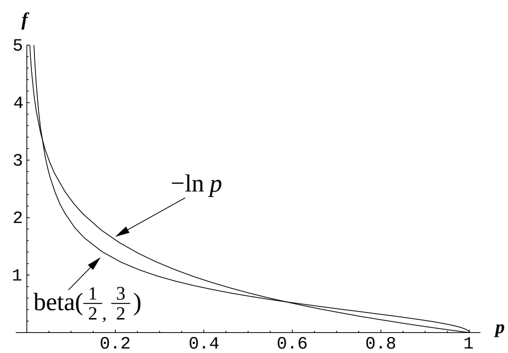


Fig. 9. Comparison between beta $\left(\frac{1}{2}, \frac{3}{2}\right)$ and product of uniform distributions.

So why the difference? The one feature that is different between the two cases is that, in the case of the conjunction $x \wedge y$, we have additional information about the probabilities. Specifically, since x and y are independent, then

$$\mathrm{P}(x \wedge y)\mathrm{P}(\overline{x} \wedge \overline{y}) = \mathrm{P}(x \wedge \overline{y})\mathrm{P}(\overline{x} \wedge y), \tag{32}$$

and we have no such relation for $\mathrm{P}(x_1)$, $\mathrm{P}(x_2)$, $\mathrm{P}(x_3)$, $\mathrm{P}(x_4)$. The result is that the two sets of circumstances are not identical. In the case of x_1, x_2, x_3, x_4, if the random variable W describes $\mathrm{P}(x_1)$, so $W \sim$ beta $\left(\frac{1}{2}, \frac{3}{2}\right)$, then $\mathrm{E}[W] = \frac{1}{4}$ and $\mathrm{E}[W^2] = \frac{1}{8}$, so that $\mathrm{Var}[W] = \frac{1}{16}$. In the case of the two independent uniform random variables (where P denotes the random variable $\mathrm{P}(x \wedge y)$), $\mathrm{E}[P] = \frac{1}{4}$ and $\mathrm{E}[P^2] = \frac{1}{9}$, so that $\mathrm{Var}[P] = \frac{7}{144}$. The fact that the variance of P is smaller than the variance of W reflects the fact that we have more information about P, and that we are therefore less uncertain about P.

9 Mapping Between Opinions and Beta Distributions

The correspondence between opinions expressed as quadruples (b_x, d_x, u_x, a_x) and beta distributions expressed as beta(α, β) is not immediately obvious. However, it is possible to fix certain requirements for the beta distribution which corresponds with a given opinion. Note that the space of opinions has three degrees of freedom (there are four variables, b_x, d_x, u_x and a_x and one relation $b_x + d_x + u_x = 1$), and the space of beta distributions has two degrees of freedom (because it has two parameters), so most of the beta distributions which correspond to an opinion can be expected to correspond to a continuum of opinions, with one degree of freedom. Since an opinion has an expectation value for the probability, *i.e.* $\mathrm{E}(x) = b_x + a_x u_x$, and the beta distribution has an expectation value for the probability, *i.e.* $\mathrm{E}(P) = \frac{\alpha}{\alpha+\beta}$, then the first requirement will be that the expectation value for the probability for the opinion be equal to the probability expectation value for the beta distribution, *i.e.*:

$$\frac{\alpha}{\alpha + \beta} = b_x + a_x u_x \ , \quad \text{or equivalently,} \quad \frac{\beta}{\alpha + \beta} = d_x + (1 - a_x) u_x \ . \tag{33}$$

Secondly, if the uncertainty decreases while the probability expectation value for the opinion and the atomicity remain constant, then that reflects a greater confidence in the individual that the probability that the system is in the state x is given by $\mathrm{E}(x) = b_x + a_x u_x$ (the size of the "interval of confidence", between b_x and $1 - d_x$ decreases as u_x decreases - meanwhile, both b_x and d_x increase as uncertainty is redistributed to belief and disbelief). The corresponding requirement for the beta distribution is that the variance for the beta distribution should decrease to reflect the greater confidence in the expectation value. Since the probability expectation value is being held constant, then $\frac{\alpha}{\beta}$ is being held constant while α and β each vary individually. The variance of the beta distribution is expressed by the formula:

$$\mathrm{Var}(P) = \frac{\alpha\beta}{(\alpha+\beta)^2(\alpha+\beta+1)} = \frac{\mathrm{E}(P)(1-\mathrm{E}(P))}{\alpha+\beta+1}. \tag{34}$$

From the above expression it can be seen that if α and β vary in such a manner that $\frac{\alpha}{\beta}$ remains constant, then the variance decreases as α and β increase, and the variance increases as α and β decrease. As a result, if the uncertainty decreases while the probability expectation value and the atomicity of the opinion remain constant, then α and β must increase in such a manner that $\frac{\alpha}{\beta}$ remains constant.

Finally, if the uncertainty u_x is equal to zero, that represents the dogmatic opinion that the probability that the system is in the state x is b_x and the probability that the system is in the state $\overline{x}$ is d_x (since $b_x + d_x = 1$ in this case, then the laws of probability are still satisfied). Since the opinion is dogmatic, the variance of the corresponding beta distribution must be zero. This is actually impossible, so the only means of satisfying this particular requirement is to take the limit as α and β approach infinity in such a manner that $\frac{\alpha}{\beta}$ approaches $\frac{b_x}{d_x}$.

In summary, α and β must be functionally dependent on b_x, d_x, u_x and a_x in such a manner that

(1) $\alpha/\beta = (b_x + a_x u_x)/[d_x + (1-a_x)u_x]$,
(2) if a_x remains constant, and b_x, d_x and u_x vary in such a manner that $b_x + a_x u_x$ remains constant, then $\frac{\alpha}{\beta}$ remains constant (as required by the first condition) and α and β increase as u_x decreases, and α and β decrease as u_x increases,
(3) as u_x approaches zero, then α and β approach infinity in such a manner that $\frac{\alpha}{\beta}$ approaches the limiting value of $\frac{b_x}{d_x}$.

The first requirement is satisfied exactly when there exists a function γ such that $\alpha = (b_x + a_x u_x)\gamma$ and $\beta = (d_x + (1-a_x)u_x)\gamma$. The second and third requirement now reduce to the statements that as u_x decreases while a_x and $b_x + a_x u_x$ remain constant, γ must increase, and as u_x approaches zero, γ must approach infinity. After this substitution for α and β, the variance of the beta distribution is given by

$$\mathrm{Var}(P) = \frac{(b_x + a_x u_x)(d_x + (1-a_x)u_x)}{\gamma+1} = \frac{\mathrm{E}(x)(1-\mathrm{E}(x))}{\gamma+1}, \tag{35}$$

thus demonstrating explicitly that the variance decreases as γ increases.

One suggestion that would satisfy the requirements for γ is $\gamma = \delta/u_x^n$ for some function δ and some positive real number n such that δ/u_x^n increases as u_x decreases while holding $b_x + a_x u_x$ and a_x fixed, and such that δ approaches a positive function of b_x, d_x and a_x as u_x approaches zero.

One possible solution to the problem is to take the case where $\alpha = r + R$ and $\beta = s + S$, where r is the amount of evidence gathered in favour of the system being in the state x, s is the amount of evidence gathered in favour of the system being in the state $\overline{x}$, and R and S are constants to be determined. Since the belief should relate to r (i.e. the amount of evidence in favour of x) and the disbelief should relate in the same manner to s (i.e. the amount of evidence in favour of $\overline{x}$), and in an original state of ignorance, b_x and d_x should both be equal to zero in the absence of evidence, *i.e.* if $r = s = 0$, then since $r + R = (b_x + a_x u_x)\gamma$ and $s + S = (d_x + (1 - a_x)u_x)\gamma$, these conditions are satisfied when $r = b_x\gamma$, $s = d_x\gamma$, $R = a_x u_x \gamma$ and $S = (1 - a_x)u_x\gamma$, so that $\gamma = \frac{R+S}{u_x}$, and $R = Ta_x$ and $S = T(1 - a_x)$, where T could still be dependent on a_x. This falls under the previous categorisation with $\delta = T$ and $n = 1$, provided T is positive for all values of a_x. This solution yields

$$\alpha = \frac{Tb_x}{u_x} + Ta_x, \qquad \beta = \frac{Td_x}{u_x} + T(1 - a_x). \tag{36}$$

For this correspondence between opinion and beta distribution, the variance of the beta distribution is given by

$$\mathrm{Var}(P) = \frac{(b_x + a_x u_x)(d_x + (1 - a_x)u_x)u_x}{T + u_x} = \frac{\mathrm{E}(x)(1 - \mathrm{E}(x))u_x}{T + u_x}. \tag{37}$$

One can use the arguments from Section 8 to justify that it is reasonable to take T constant, and in Jøsang 2001 [2], T was taken to be constant, and set equal to 2, so that in the absence of evidence, when the atomicity is $\frac{1}{2}$, the *a priori* distribution is uniform (this requirement forces a choice of $T = 2$). This particular correspondence can be described as:

$$(b_x, d_x, u_x, a_x) \mapsto \text{beta}\left(\frac{2b_x}{u_x} + 2a_x, \frac{2d_x}{u_x} + 2(1 - a_x)\right), \tag{38}$$

As already mentioned, the beta distribution really only has two degrees of freedom, so that there will always be ranges of values in the expression for the opinion in Eq.(38) which actually produce the same beta parameters. This will be the case for the ranges of (b_x, d_x, u_x, a_x) values where u_x and $\mathrm{E}(x) = b_x + a_x u_x$ are constant.

By comparing the parameters of the beta distribution in Eq.(38) with those in Eq.(31) it can be seen that the relative atomicity in fact defines the *a priori* parameters of beta distribution expressed by:

$$\text{beta}\left(2a_x,\ 2(1 - a_x)\right) \ . \tag{39}$$

By considering the *a priori* beta parameters as separate from the evidence parameters $r = 2b_x/u_x$ and $s = 2d_x/u_x$ the expression for the beta distribution gets 3 degrees of freedom so that in fact a bijective mapping can be defined between the expression for opinions and and the augmented expression for beta distributions. Let Eq.(31) define the augmented beta distribution representation, i.e as:

$$\text{beta}(r + 2a,\ s + 2(1 - a))\ ,$$

which distinguishes between *a priori* and *a posteriori* information, then a bijective mapping between opinions and augmented beta distributions can be defined as:

$$\begin{aligned} 2b_x/u_x &\leftrightarrow r \\ 2d_x/u_x &\leftrightarrow s \\ a_x &\leftrightarrow a \end{aligned} \qquad (40)$$

It can be noted that under this correspondence the example opinion of Fig.1 and the beta distribution of Fig.8 are equivalent.

Under the correspondence of Eq.(40), as u_x becomes small, the variance is approximately proportional to u_x, so that the width of the distribution (which is characterised by the standard deviation) is approximately proportional to the square root of u_x. This means that for u_x small, there is a significant probability that the probability p that the system is in the state x will either fall below b_x or exceed $1 - d_x$, and in fact, as u_x becomes small, the probability that p will fall between b_x and $1 - d_x$ will also become small (proportional to the square root of u_x), approaching zero as u_x approaches zero. The probability that p falls below b_x approaches $\frac{1}{2}$ in the limit, and the probability that p exceeds $1 - d_x$ also approaches $\frac{1}{2}$ in the limit. This is not very satisfying since intuitively, b_x should represent the smallest practical value that p can take, and $1 - d_x$ should represent the largest practical value that p can take. Most authors in the belief theory community, including Shafer (1992) [6], reject the idea that a belief function represents a lower probability, and so from the Shaferian point of view, this objection, that b_x does not represent the smallest practical value that p can take and $1 - d_x$ does not represent the largest practical value that p can take, is not really a valid objection to the correspondence between opinion and beta distribution currently used.

While Eq.(40) provides a bijective (one-to-one) mapping from opinions to augmented beta distributions, we would also like to know the correspondence between BMAs and beta distributions. The simple and normal coarsenings described in Section 2 define two different surjective (onto) correspondences from BMAs to opinions. It was noted that simple coarsening has the drawback that the relative atomicity in general does not reflect the real relative cardinality, whereas normal coarsening has the drawback that the belief, disbelief and uncertainty parameters must be adjusted. It was also shown that these drawbacks disappear when the two coarsenings produce equal results, which is the case when Eq.(21) is satisfied. The various correspondences are illustrated in Fig.10.

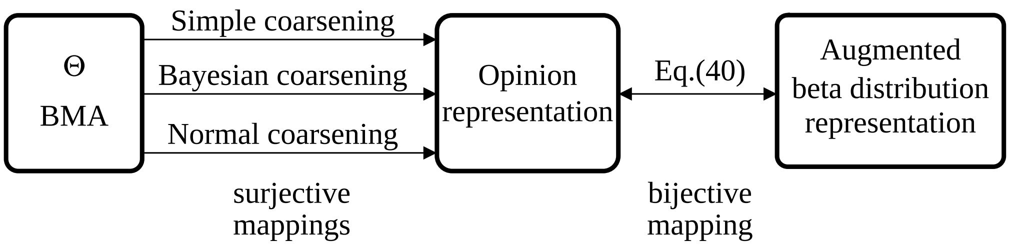


Fig. 10. Correspondence between BMAs, opinions and beta distributions.

Because the correspondence of Eq.(40) is bijective, there exists a surjective mapping from BMAs to beta distributions. The fact that there are two different mappings from BMAs to opinions can be problematic, because in practical situations, one of them must be selected. In general, normal coarsening provides the best interpretation of BMAs in terms of opinions because of the correct relative atomicity. The next section also shows that normal multiplication and comultiplication provides the best approximation of the product and coproduct of beta distributions.

10 Comparison of Multiplication and Comultiplication Operators

For the purpose of comparing simple and normal multiplication with multiplication of beta distributions, we denote by beta(ω_x) and beta(ω_y) the beta distributions corresponding to the opinions ω_x and ω_y respectively, and by beta$(\omega_x) \cdot$ beta(ω_y) the product of beta(ω_x) and beta(ω_y). Further we denote by beta$(\omega^{\mathrm{S}}_{x\wedge y})$ and beta$(\omega^{\mathrm{N}}_{x\wedge y})$ the beta distributions corresponding to the simple and normal product opinions $\omega^{\mathrm{S}}_{x\wedge y}$ and $\omega^{\mathrm{N}}_{x\wedge y}$ respectively. Similarly we denote by beta$(\omega_x) \sqcup$ beta(ω_y) the coproduct of beta(ω_x) and beta(ω_y), and by beta$(\omega^{\mathrm{S}}_{x\vee y})$ and beta$(\omega^{\mathrm{N}}_{x\vee y})$ the beta distributions corresponding to the simple and normal coproduct opinions $\omega^{\mathrm{S}}_{x\vee y}$ and $\omega^{\mathrm{N}}_{x\vee y}$respectively.

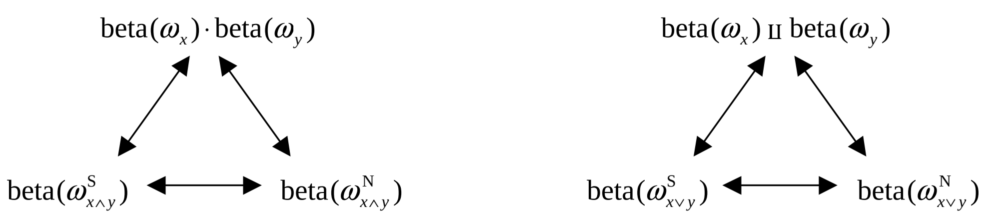
Given the interpretation of opinions as beta distributions, and assuming the product and coproduct of beta distributions to be analytically correct, it would have been desirable for multiplication and comultiplication operators of opinions to satisfy:

$$\text{beta}(\omega_{x\wedge y}) = \text{beta}(\omega_x) \cdot \text{beta}(\omega_y) \tag{41}$$

$$\text{beta}(\omega_{x\vee y}) = \text{beta}(\omega_x) \sqcup \text{beta}(\omega_y) \tag{42}$$

It is known that if the probabilities of independent propositions x and y have beta distributions, then the probabilities of $x \wedge y$ and $x \vee y$ do not have beta distributions, except under extraordinary circumstances, *i.e.* it can happen, but such a happenstance is an exception rather than the rule. It is thus impossible for the multiplication and comultiplication operators described in Sections 4 and 5 to satisfy Eq.(41) and Eq.(42) in general. The deviation between the left and right sides of Eq.(41) and Eq.(42) is partly due to non-Bayesian coarsening of $X \times Y$ as explained in Section 2, and partly due to the difference in circumstances around the *a priori* probability distributions over the conjunction and disjunction of two independent variables x and y, and four exhaustive and mutually exclusive variables (x_1, x_2, x_3, x_4), as explained in Section 8. In the following we will try to determine how close the multiplication and comultiplication operators are to satisfying Eq.(41) and Eq.(42).

First we compare the characteristics of the simple product $\text{beta}(\omega^{\text{S}}_{x\wedge y})$ with the normal product $\text{beta}(\omega^{\text{N}}_{x\wedge y})$, and then the simple coproduct $\text{beta}(\omega^{\text{S}}_{x\vee y})$ with the normal coproduct $\text{beta}(\omega^{\text{N}}_{x\vee y})$. We then compare the characteristics of the product $\text{beta}(\omega_x) \cdot \text{beta}(\omega_y)$ with the characteristics of the products $\text{beta}(\omega^{\text{S}}_{x\wedge y})$ and $\text{beta}(\omega^{\text{N}}_{x\wedge y})$. Similarly we compare the coproduct $\text{beta}(\omega_x) \sqcup \text{beta}(\omega_y)$ with the characteristics of the coproducts $\text{beta}(\omega^{\text{S}}_{x\vee y})$ and $\text{beta}(\omega^{\text{N}}_{x\vee y})$. The various comparisons are illustrated in Fig.11 below.



(a) Product comparisons (a) Coproduct comparisons

Fig. 11. Illustration of comparisons

The beta distribution corresponding to the normal opinion product $\omega^{\text{N}}_{x\wedge y}$ is given by

$$\text{beta}(\omega^{\text{N}}_{x\wedge y}) = \text{beta}\left(\frac{2FG}{H-F}, \frac{2(1-F)G}{H-F}\right), \tag{43}$$

where $F = (b_x + a_x u_x)(b_y + a_y u_y)$, $G = 1 - a_x a_y$, and $H = (1-d_x)(1-d_y)$.

Since the parameters of $\text{beta}(\omega^{\text{N}}_{x\wedge y})$ are greater than the parameters of $\text{beta}(\omega^{\text{S}}_{x\wedge y})$, then the variance of $\text{beta}(\omega^{\text{N}}_{x\wedge y})$ is less than the variance of $\text{beta}(\omega^{\text{S}}_{x\wedge y})$. This reflects the fact that the uncertainty of the normal opinion product is less than the uncertainty of the simple opinion product.

Similarly, the variance of $\text{beta}(\omega^{\text{N}}_{x\vee y})$ is less than the variance of $\text{beta}(\omega^{\text{S}}_{x\vee y})$, reflecting the fact that the uncertainty of the normal opinion coproduct is less than the uncertainty of the simple opinion coproduct.

The first and second moments of $\text{beta}(\omega^{\text{N}}_{x\wedge y})$ and $\text{beta}(\omega^{\text{S}}_{x\wedge y})$ can be compared to the first and second moments of $\text{beta}(\omega_x) \cdot \text{beta}(\omega_y)$, and similarly for the coproducts.

Since the first and second moments of a random variable $P \sim$ beta(α, β) are:

$$\mathrm{E}[P] = \frac{\alpha}{\alpha+\beta},$$
$$\mathrm{E}[P^2] = \frac{\alpha(\alpha+1)}{(\alpha+\beta)(\alpha+\beta+1)},$$

then with the map from opinion representation to beta distribution representation given by Eq.(40), the probability expectation value is given by $\mathrm{E}[P_x] = b_x + a_x u_x$, where P_x is the random variable denoting the probability of x, and the first and second moments are related by

$$\mathrm{E}[P_x^2] = \frac{\mathrm{E}[P_x](\mathrm{E}[P_x] + \frac{u_x}{2})}{1+\frac{u_x}{2}} = \frac{\mathrm{E}[P_x](2\mathrm{E}[P_x] + u_x)}{2+u_x}.$$

As a result, the first and second moments of the product and coproduct of beta distributions (i.e. beta$(\omega_x) \cdot$ beta(ω_y) and beta$(\omega_x) \sqcup$ beta(ω_y)) of independent propositions can be calculated, and these moments can be compared to the moments of the beta distributions corresponding to opinion products (i.e. beta$(\omega^{\mathrm{S}}_{x\wedge y})$ and beta$(\omega^{\mathrm{N}}_{x\wedge y})$) and coproducts (i.e. beta$(\omega^{\mathrm{S}}_{x\vee y})$ and beta$(\omega^{\mathrm{N}}_{x\vee y})$).

For the simple and normal products beta$(\omega^{\mathrm{S}}_{x\wedge y})$ and beta$(\omega^{\mathrm{N}}_{x\wedge y})$, the value of the first moment (*i.e.* the probability expectation value) is the same as the first moment of beta$(\omega_x) \cdot$ beta(ω_y), given by:

$$\mathrm{E}(x \wedge y) = (b_x + a_x u_x)(b_y + a_y u_y)\ , \tag{44}$$

Similarly for the simple and normal coproducts beta$(\omega^{\mathrm{N}}_{x\vee y})$ and beta$(\omega^{\mathrm{S}}_{x\vee y})$, the value of the first moment (*i.e.* the probability expectation value) is the same as the first moment of beta$(\omega_x) \sqcup$ beta(ω_y) which is given by:

$$\mathrm{E}(x \vee y) = b_x + a_x u_x + b_y + a_y u_y - (b_x + a_x u_x)(b_y + a_y u_y)\ , \tag{45}$$

Both of these results are to be expected since the opinions were designed specifically to yield the correct value for the first moment.

For simple product and coproduct of opinions, the second moment of beta$(\omega^{\mathrm{S}}_{x\wedge y})$ and beta$(\omega^{\mathrm{S}}_{x\vee y})$ is equal to, or greater than the second moment of beta$(\omega_x) \cdot$ beta(ω_y) and beta$(\omega_x) \sqcup$ beta(ω_y) respectively. This means that for much of the domain, the variance for beta$(\omega^{\mathrm{S}}_{x\wedge y})$ and beta$(\omega^{\mathrm{S}}_{x\vee y})$ exceeds the variance of beta$(\omega_x) \cdot$ beta(ω_y) and beta$(\omega_x) \sqcup$ beta(ω_y) respectively.

For normal product and coproduct of opinions, the second moment of beta$(\omega^{\mathrm{N}}_{x\wedge y})$ and beta$(\omega^{\mathrm{N}}_{x\vee y})$ varies between being less than, being equal to, or being greater than the second moment of beta$(\omega_x) \cdot$ beta(ω_y) and beta$(\omega_x) \sqcup$ beta(ω_y) respectively, and consequently the same relationships apply to the variance.

Based on the above analysis it seems that normal multiplication and comultiplication of opinions corresponds more closely to the analytically correct multiplication and comultiplication of beta distributions, than do simple multiplication and comultiplication of opinions. Although not perfect, normal multiplication and comultiplication are thus able to produce a good approximation of the analytically correct products and coproducts.

It is important to know how good this approximation is. This can be done by investigating the difference between the variance of beta$(\omega^{\mathrm{N}}_{x\wedge y})$ and the variance of beta$(\omega_x)\cdot$ beta(ω_y) (the case of the coproduct follows immediately from the duality, *i.e.* from de Morgan's Laws). Equivalently, the difference between the second moments can be studied since it is equal to the difference between the variances (as a consequence of the fact that the first moments are equal). The problem is difficult analytically, so a graphical approach has been adopted. Since the comparison is between the product of the beta distributions corresponding to two opinions, and the beta distribution of the normal product of the same opinions, and since each opinion has three degrees of freedom, then the problem has six degrees of freedom, which is four more than we are capable of visualising graphically (we need the third dimension for the dependent variable, *i.e.* the difference between the variances). This means for a graphical investigation into the behaviour, four independent conditions must be imposed on the opinions. In the Fig. 12, for example, we have set $b_x = d_x$, $b_y = d_y$, $u_x = u_y$ and $a_x = a_y$. The independent variables are the common value of the relative atomicity $a_x = a_y$ and the common value of the uncertainty $u_x = u_y$. The dependent variable is $V_1 - V_2$, where V_1 is the variance of beta$(\omega^{\mathrm{N}}_{x\wedge y})$, and V_2 is the variance of beta$(\omega_x)\cdot$ beta(ω_y).

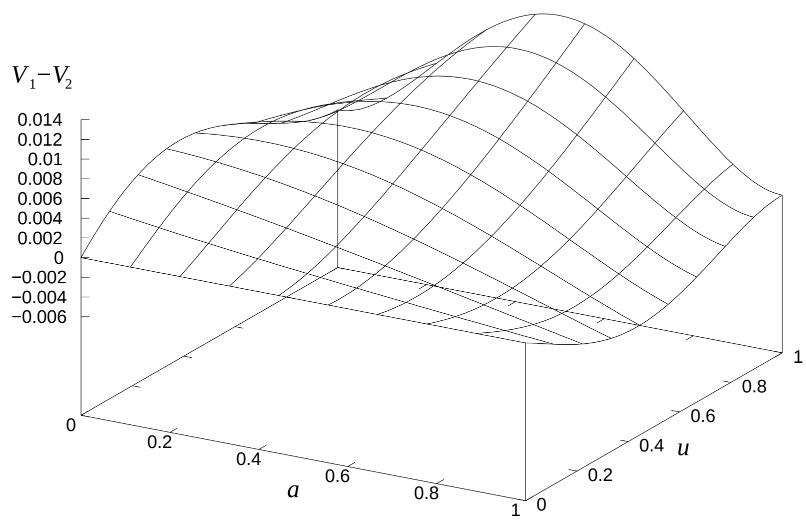


Fig. 12. Variance for beta$(\omega^{\mathrm{N}}_{x\wedge y})$ minus variance for beta$(\omega_x)\cdot$ beta(ω_y).

When the common value of the relative atomicity is 0, the difference is positive for all values of the uncertainty strictly between 0 and 1, so that the variance of beta$(\omega^{\mathrm{N}}_{x\wedge y})$ exceeds the variance of beta$(\omega_x)\cdot$beta(ω_y). When the common value of the atomicity is 1, the difference is negative for all values of the uncertainty strictly between 0 and 1, so that the variance of beta$(\omega_x) \cdot$ beta(ω_y) exceeds the variance of beta$(\omega^{\mathrm{N}}_{x\wedge y})$. When the common value of the uncertainty is 0 (*i.e.* for dogmatic opinions), the difference is zero, so that the variance of beta$(\omega^{\mathrm{N}}_{x\wedge y})$ and the variance of beta$(\omega_x) \cdot$ beta(ω_y) are equal. When the common value of the uncertainty is 1, the difference is positive for all values of the relative atomicity strictly between 0 and 1, so that the variance of beta$(\omega^{\mathrm{N}}_{x\wedge y})$ exceeds the variance of beta$(\omega_x)\cdot$beta(ω_y). The difference between the variances takes its largest magnitude when the common value for the relative atomicity is $\frac{1}{2}$ and the common value of the uncertainty is 1, and for those specific values, the variance of beta$(\omega^{\mathrm{N}}_{x\wedge y})$ exceeds the variance of beta$(\omega_x) \cdot$ beta(ω_y) by $\frac{1}{72}$. It is conjectured that the greatest difference between the variance of the beta$(\omega^{\mathrm{N}}_{x\wedge y})$ and the variance of beta$(\omega_x) \cdot$ beta(ω_y) occurs when the uncertainties of both opinions are 1 and the relative atomicities have a common value of $\frac{1}{2}$, and this greatest difference is $\frac{1}{72}$. This is precisely the difference which is illustrated in Fig.9. Because the conjectured difference is so small, then the normal product and normal coproduct can be considered to be very good approximations to the product and coproduct, respectively, of the beta distributions of the individual opinions.

11 Correspondence to Other Logic Frameworks

The subjective logic operators described above represent generalisations of classical probability and logic operators in the context of belief theory. In the case of dogmatic opinions, i.e. when $u_x = 0$, opinions are equivalent to probabilities through the correspondence $\mathrm{P}(x) = b_x$. Opinions thus represent a generalisation of probabilities. Furthermore, probabilities represent a generalisation of truth values in binary logic, where TRUE is equivalent to the special case $\mathrm{P}(x) = 1$ and FALSE is equivalent to the special case $\mathrm{P}(x) = 0$. In subjective logic these cases are expressed by the opinions $(1, 0, 0, a_x)$ for TRUE, and by $(0, 1, 0, a_x)$ for FALSE, where $0 < a_x < 1$.

Multiplication, comultiplication, division and codivision of dogmatic opinions are equivalent to the corresponding probability operators in Table 1.

Operator name:	Operator expression
Multiplication	$\mathrm{P}(x)\mathrm{P}(y)$
Division	$\mathrm{P}(x)/\mathrm{P}(y)$
Comultiplication	$\mathrm{P}(x) + \mathrm{P}(y) - \mathrm{P}(x)\mathrm{P}(y)$
Codivision	$(\mathrm{P}(x) - \mathrm{P}(y))/(1 - \mathrm{P}(y))$

Table 1
Probability operators resulting from opinion operators.

The correspondence between binary logic operators and probability/opinion operators is given in Table 2 below. Some of the operators are not widely used, and new names and symbols had to be defined.

Opinion operator	Symbol	Set operator	Logic operator	Symbol
Multiplication	$\cdot$	Conjunction	AND	$\wedge$
Division	$/$	Unconjunction	UN-AND	$\overline{\wedge}$
Comultiplication	$\sqcup$	Disjunction	OR	$\vee$
Codivision	$\overline{\sqcup}$	Undisjunction	UN-OR	$\overline{\vee}$

Table 2
Correspondence between probability, set and logic operators.

It can be shown that the multiplication and comultiplication operators produce the classical truth tables of AND and OR for the special cases where $\mathrm{P}(x) = 1$ or $\mathrm{P}(x) = 0$. Similarly the truth tables of UN-AND and UN-OR can be determined through the division and codivision operators. Table 3, which e.g. can be derived from the expressions in Table 1, defines the complete truth table.

		AND	OR	UN-AND	UN-OR
x	y	$x \wedge y$	$x \vee y$	$x \overline{\wedge} y$	$x \overline{\vee} y$
F	F	F	F	T or F	F
F	T	F	T	F	undefined
T	F	F	T	undefined	T
T	T	T	T	T	T or F

Table 3
Truth tables for AND, OR, UN-AND and UN-OR.

It can be shown that simple multiplication and comultiplication represent a generalisation of the $\wedge$ and $\vee$ operators in Kleene's (1950) three valued logic [4], where opinions with $u = 1$ can be interpreted as *undefined* or *undetermined* in Kleene's terminology. In *weak* Kleene logic, $x \wedge y$ is undefined if for example x is undefined and y is FALSE, because not all arguments have defined values. In *strong* Kleene logic, $x \wedge y$ is FALSE if x is undefined and y is FALSE, because as long as one argument is FALSE the value of the other argument is irrelevant. Similarly for OR when one of the arguments has value TRUE. When applied to opinions where either $b = 1$, $d = 1$ or $u = 1$, simple multiplication and comultiplication produce the truth tables of Kleene's strong $\wedge$ and $\vee$ operators.

It can also be mentioned that simple multiplication and comultiplication are equivalent to the "AND" and "OR" operators of Baldwin's support logic (Baldwin 1986 [1]) except for the relative atomicity parameter which is absent in Baldwin's logic. In Baldwin's logic, each proposition has a support pair $[S_l, S_u]$ where S_l represents the lower or necessary support, and S_u represents the upper or possible support. This is in fact the same as the [*Belief*, *Plausibility*] pair of classical belief theory, but instead of focusing on frames of discernment, Baldwin's theory focuses on individual propositions. A support pair $[0, 1]$ is equivalent to a vacuous BMA, and would correspond to the undefined truth value in Kleene's logic. The support pairs $[0, 0]$ and $[1, 1]$ correspond to false and true propositions respectively.

Having established the correspondence to Kleene's three valued logic, and to Baldwin's support logic, it can be useful to illustrate what subjective logic can do in addition. Let for example S be a set of independent propositions with undefined truth value in Kleene's terminology, or with support pairs $[0, 1]$ in Baldwin's terminology. The conjunction of all the statements $s \in S$ would always produce a proposition with undefined truth value in Kleene's theory, and a support pair $[0, 1]$ in Baldwin's theory. However, the intuitive interpretation of the conjunction of undefined/uncertain propositions is that the likelihood of the conjunctive proposition being true decreases as a function of $|S|$. Similarly, the likelihood of a disjunctive proposition of S being true increases as a function of $|S|$. This intuitive observation can not be derived by Kleene's or Baldwin's frameworks, but is explicitly reflected in subjective logic through the relative atomicity and the probability expectation value.

12 Conclusion

The two coarsening methods described in Sec.2 describe two different surjective mappings from a generalised frame of discernment and BMA to the opinion space. Eq.(40) defines a bijective mapping between opinions and the sub-class of beta probability distribution functions beta(α, β) where $\alpha + \beta \geq 2$.

The coarsening process together with the bijective correspondence between opinions and beta distributions provides a specific interpretation of belief functions in terms of Bayesian probabilities. Two opinions that correspond to beta distributions can be multiplied or comultiplied to produce a new product or coproduct opinion that also corresponds to a beta distribution. Under this interpretation our analysis of multiplication and comultiplication of opinions has led us to the conclusion that these operators only provide an approximation of the analytical multiplication and comultiplication of beta distributions.

In general the product of two beta distributions is not a beta distribution, and the analytical expressions for products and coproducts of probability distributions quickly become exceedingly complex, whereas the expressions for products and coproducts of opinions are very simple. The advantage of doing calculations with opinions rather that with probability distributions, is a dramatic reduction in complexity.

Since it appears that the variance of the beta distribution for the normal product of opinions differs from the variance of the product of the beta distributions for the individual opinions by no more than about 0.014 (and a similar result holds for the coproduct by de Morgan's Laws), then the approximation of the product of the beta distributions by the normal product of the opinions is very good.

References

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

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

[3] A. Jøsang and T. Grandison. Conditional Inference in Subjective Logic. In Xuezhi Wang, editor, *Proceedings of the 6th International Conference on Information Fusion*, 2003.

[4] S.C. Kleene. *Introduction to Metamathematics*. D. Van Nostrand, Princeton, N.J., 1950.

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

[6] G. Shafer. Response to the discussion of belief functions. *International Journal of Approximate Reasoning*, 6(3):445–480, 1992.

[7] Ph. Smets. The transferable belief model for quantified belief representation. In D.M. Gabbay and Ph. Smets, editors, *Handbook of Defeasible Reasoning and Uncertainty Management Systems, Vol.1*, pages 267–301. Kluwer, Doordrecht, 1998.

[8] Ph. Smets and R. Kennes. The transferable belief model. *Artificial Intelligence*, 66:191–234, 1994.

Appendix

Proof of theorem 7.1

Because of the symmetry between the random variables, and because of iteration, it suffices to prove that

$$(P_1, \ldots, P_{k-2}, P_{k-1} + P_k) \sim \text{Dirichlet}(\alpha_1, \ldots, \alpha_{k-2}, \alpha_{k-1} + \alpha_k).$$

This can be proven by evaluating the marginal probability density function

$$
\begin{aligned}
& f(P_1 = p_1, \ldots, P_{k-2} = p_{k-2}) \\
= & \int_0^{1-\sum_{i=1}^{k-2} p_i} f(p_1, \ldots, p_{k-1})\, dp_{k-1} \\
= & \frac{\Gamma\left(\sum_{i=1}^{k} \alpha_i\right)}{\prod_{i=1}^{k} \Gamma(\alpha_i)} \prod_{i=1}^{k-2} p_i^{\alpha_i - 1} \int_0^{1-\sum_{i=1}^{k-2} p_i} p_{k-1}^{\alpha_{k-1}-1} p_k^{\alpha_k - 1}\, dp_{k-1},
\end{aligned}
$$

where $p_k = 1 - \sum_{i=1}^{k-1} p_i$. Changing the variable of integration to

$$
q = \frac{p_{k-1}}{1 - \sum_{i=1}^{k-2} p_i},
$$

then the integral becomes

$$
\begin{aligned}
& f(P_1 = p_1, \ldots, P_{k-2} = p_{k-2}) \\
= & \frac{\Gamma\left(\sum_{i=1}^{k} \alpha_i\right)}{\prod_{i=1}^{k} \Gamma(\alpha_i)} \prod_{i=1}^{k-2} p_i^{\alpha_i - 1} \int_0^{1-\sum_{i=1}^{k-2} p_i} p_{k-1}^{\alpha_{k-1}-1} p_k^{\alpha_k - 1}\, dp_{k-1} \\
= & \frac{\Gamma\left(\sum_{i=1}^{k} \alpha_i\right)}{\prod_{i=1}^{k} \Gamma(\alpha_i)} \prod_{i=1}^{k-2} p_i^{\alpha_i - 1} \left(1 - \sum_{i=1}^{k-2} p_i\right)^{\alpha_{k-1}+\alpha_k - 1} \int_0^1 q^{\alpha_{k-1}-1} (1-q)^{\alpha_k - 1}\, dq \\
= & \frac{\Gamma\left(\sum_{i=1}^{k} \alpha_i\right)}{\left(\prod_{i=1}^{k-2} \Gamma(\alpha_i)\right) \Gamma(\alpha_{k-1} + \alpha_k)} \left(\prod_{i=1}^{k-2} p_i^{\alpha_i - 1}\right) \left(1 - \sum_{i=1}^{k-2} p_i\right)^{\alpha_{k-1}+\alpha_k - 1},
\end{aligned}
$$

thus demonstrating that $(P_1, \ldots, P_{k-2}, P_{k-1}+P_k) \sim \text{Dirichlet}(\alpha_1, \ldots, \alpha_{k-2}, \alpha_{k-1}+\alpha_k)$, and completing the proof.

$\square$

Since Q and R are independent uniformly distributed random variables and $P = QR$, then for all $p \in (0,1)$,

$$\begin{aligned}
\mathrm{P}(p < P < 1) &= \mathrm{P}(p < QR < 1) \\
&= \int_p^1 dq\, \mathrm{P}\left(\frac{p}{q} < R < 1\right) \\
&= \int_p^1 dq\, \int_{\frac{p}{q}}^1 dr \\
&= \int_p^1 dq\, \left(1 - \frac{p}{q}\right) \\
&= [q - p\ln q]_p^1 \\
&= 1 - p + p\ln p.
\end{aligned}$$

Since $f(p) = -d\mathrm{P}(p < P < 1)/dp$, then $f(p) = -\ln p$ for $0 < p < 1$.

$\square$