

A Comparison of Methods for Transforming Belief Function Models to Probability Models

Barry R. Cobb and Prakash P. Shenoy

University of Kansas School of Business, 1300 Sunnyside Ave., Summerfield Hall
Lawrence, KS 66045-7585, USA
{brcobb, pshenoy}@ku.edu

Abstract. Recently, we proposed a new method called the plausibility transformation method to convert a belief function model to an equivalent probability model. In this paper, we compare the plausibility transformation method with the pignistic transformation method. The two transformation methods yield qualitatively different probability models. We argue that the plausibility transformation method is the correct method for translating a belief function model to an equivalent probability model that maintains belief function semantics.

1 Introduction

Bayesian probability theory and the Dempster-Shafer (D-S) theory of belief functions are two distinct calculi for modeling and reasoning with knowledge about propositions in uncertain domains. In a recent paper [1], we have argued that these two calculi have roughly the same expressive power. Also, in [2, 3], we have proposed a new method, called the plausibility transformation method, for transforming a belief function model to an equivalent probability model.

In this paper, we compare the two techniques—the pignistic transformation [11] and the plausibility transformation—for transforming a belief function model to a Bayesian probability model. In many cases, these two methods lead to radically different probability models starting from the same belief function model. We argue that the plausibility transformation method is the correct method and that it provides an equivalent probability model that is consistent with belief function semantics.

There are many different semantics of belief functions, including multivalued mapping [5], random codes [9], transferable beliefs [11], and hints [7], that are compatible with Dempster’s rule of combination. However, the semantics of belief functions as upper and lower probability bounds on some true but unknown probability function are incompatible with Dempster’s rule [12]. In this paper, we are concerned with the D-S theory of belief functions with Dempster’s rule of combination as the updating rule, and not with theories of upper and lower probabilities that admit various other rules for updating beliefs. One benefit of studying probability functions derived from D-S belief functions is a more clear understanding of D-S belief function semantics.

The remainder of this paper is organized as follows. Section 2 contains notation and definitions associated with probability theory and the Dempster–Shafer theory of belief functions. Section 3 defines the pignistic and plausibility transformation methods. Section 4 describes three examples that are studied in great detail. Section 5 contains four theorems that define the properties of the plausibility transformation. In Section 6, we summarize and conclude. Proofs of all theorems can be found in [2]. This paper is extracted from a larger unpublished working paper [2].

2 Notation and Definitions

This section establishes notation and definitions that will be used throughout the paper.

2.1 Probability Theory

A probability potential P_s for s is a function $P_s : \Omega_s \rightarrow [0, 1]$. We express our knowledge by probability potentials, which are combined to form the joint probability distribution, which is then marginalized to the relevant variables.

Projection of States. If (w, x, y, z) is a state of $\{W, X, Y, Z\}$, for example, then the projection of (w, x, y, z) to $\{W, X\}$ is simply (w, x) , which is a state of $\{W, X\}$. If s and t are sets of variables, $s \subseteq t$, and x is a state of t , then $x^{\downarrow s}$ denotes the projection of x to s .

Combination. Combination in probability theory is “pointwise” multiplication of potentials followed by normalization. Suppose P_s is a probability potential for s and P_t is a probability potential for t . Then $P_s \otimes P_t$ is a probability potential for $s \cup t$ defined as follows:

$$(P_s \otimes P_t)(x) = K^{-1} P_s(x^{\downarrow s}) P_t(x^{\downarrow t}), \quad (1)$$

for each $x \in \Omega_{s \cup t}$, where $K = \sum \{P_s(x^{\downarrow s}) P_t(x^{\downarrow t}) \mid x \in \Omega_{s \cup t}\}$ is the normalization constant.

Marginalization. Marginalization in probability theory involves addition over the state space of the variables being eliminated. Suppose P_s is a probability potential for s , and suppose $X \in s$. The marginal of P_s for $s \setminus \{X\}$, denoted by $P_s^{\downarrow (s \setminus \{X\})}$, is the probability potential for $s \setminus \{X\}$ defined as follows:

$$P_s^{\downarrow (s \setminus \{X\})}(y) = \sum \{P_s(y, x) \mid x \in \Omega_X\}, \quad (2)$$

for all $y \in \Omega_{s \setminus \{X\}}$.

2.2 Dempster-Shafer Theory of Belief Functions

A Dempster-Shafer basic probability assignment (bpa) assigns values to subsets of the state space. If Ω_s is the state space of a set of variables s , a function $m : 2^{\Omega_s} \rightarrow [0, 1]$ is a bpa for s whenever $m(\emptyset) = 0$ and

$$\sum \{m(\mathbf{a}) \mid \mathbf{a} \in 2^{\Omega_s}\} = 1. \quad (3)$$

A bpa can also be stated in terms of a corresponding plausibility function. The plausibility function Pl_m corresponding to a bpa m for s is defined as $Pl_m : 2^{\Omega_s} \rightarrow [0, 1]$ such that for all $\mathbf{a} \in 2^{\Omega_s}$,

$$Pl_m(\mathbf{a}) = \sum \{m(\mathbf{b}) \mid \mathbf{b} \cap \mathbf{a} \neq \emptyset\}. \quad (4)$$

Projection and Extension of Subsets. If r and s are sets of variables, $r \subseteq s$, and \mathbf{a} is a nonempty subset of Ω_s , then the *projection of \mathbf{a} to r* , denoted by $\mathbf{a}^{\downarrow r}$, is the subset of Ω_r given by $\mathbf{a}^{\downarrow r} = \{x^{\downarrow r} \mid x \in \mathbf{a}\}$.

By extension of a subset of a state space to a subset of a larger state space, we mean a cylinder set extension. If r and s are sets of variables, $r \subset s$, and \mathbf{a} is a subset of Ω_r , then the *extension of \mathbf{a} to s* is $\mathbf{a} \times \Omega_{s \setminus r}$. Let $\mathbf{a}^{\uparrow s}$ denote the extension of \mathbf{a} to s . For example, if \mathbf{a} is a subset of $\Omega_{\{W, X\}}$, then $\mathbf{a}^{\uparrow \{W, X, Y, Z\}} = \mathbf{a} \times \Omega_{\{Y, Z\}}$.

Combination. Calculation of a joint bpa is accomplished by using Dempster's rule of combination [5]. Consider two bpa's m_A and m_B for a and b , respectively. The combination of m_A and m_B , denoted by $m_A \oplus m_B$, is the bpa for $a \cup b$ given by

$$(m_A \oplus m_B)(\mathbf{c}) = K^{-1} \sum \{m_A(\mathbf{x})m_B(\mathbf{y}) \mid (\mathbf{x}^{\uparrow (a \cup b)}) \cap (\mathbf{y}^{\uparrow (a \cup b)}) = \mathbf{c}\} \quad (5)$$

for all nonempty $\mathbf{c} \subseteq \Omega_{a \cup b}$, where K is a normalization constant given by $K = \sum \{m_A(\mathbf{x})m_B(\mathbf{y}) \mid (\mathbf{x}^{\uparrow (a \cup b)}) \cap (\mathbf{y}^{\uparrow (a \cup b)}) \neq \emptyset\}$.

Marginalization. Suppose m is a bpa for s , and suppose $t \subset s$. The marginal of m for t , denoted $m^{\downarrow t}$, is the bpa for t defined as follows:

$$m^{\downarrow t}(\mathbf{a}) = \sum \{m(\mathbf{b}) \mid \mathbf{b}^{\downarrow t} = \mathbf{a}\} \quad (6)$$

for each $\mathbf{a} \subset \Omega_t$.

3 Transformation Methods

In this section, we define the pignistic transformation and the plausibility transformation methods for converting belief functions to probability functions.

3.1 Pignistic Transformation.

Suppose m is a bpa for s . Let $BetP_m$ denote the corresponding probability function obtained using the pignistic transformation method [10, 11]. $BetP_m$ is defined as follows:

$$BetP_m(x) = \sum_{\substack{\mathbf{a} \subseteq \Omega_s \\ x \in \mathbf{a}}} \frac{m(\mathbf{a})}{|\mathbf{a}|} \quad (7)$$

for each $x \in \Omega_s$. To simplify terminology, we will refer to the $BetP_m$ as a pignistic probability function (corresponding to bpa m).

3.2 Plausibility Transformation.

Suppose m is a bpa for s . Let Pl_m denote the plausibility function for s corresponding to bpa m . Let Pl_P_m denote the probability function for s corresponding to m obtained using the plausibility transformation method. Pl_P_m is defined as follows:

$$Pl_P_m(x) = K^{-1}Pl_m(\{x\}) \quad (8)$$

for all $x \in \Omega_s$, where $K = \sum\{Pl_m(\{x\}) \mid x \in \Omega_s\}$ is the normalization constant. To simplify terminology, we will refer to Pl_P_m as the plausibility probability function (corresponding to bpa m).

4 Three Examples

The examples in this section will highlight the differences between the pignistic and plausibility transformation methods.

4.1 Example 1: Peter, Paul, and Mary [11].

A mafia don, the Godfather, has three assassins, Peter, Paul, and Mary. Needing to assassinate an informant, Mr. Jones, the Godfather decides to first toss a fair coin to decide the sex of the assassin. If the toss results in heads, he will pick Mary for the job. If the toss results in tails, he will ask either Peter or Paul to do the job. In the case of tails, we have no knowledge of how the Godfather will select between Peter and Paul. Now suppose we find Mr. Jones assassinated. An informant in the mafia organization has informed the district attorney (DA) about the Godfather's incomplete mechanism for choosing among Peter, Paul, and Mary. The DA would like to indict Peter, Paul, or Mary (in addition to the Godfather). Who should the DA indict?

Let A denote the assassin variable with three states: Peter, Paul, and Mary. Given our knowledge of the incomplete protocol of how the assassin was selected, we can represent it by the bpa m_1 for A as follows: $m_1(\{Mary\}) = 0.5$,

$m_1(\{Peter, Paul\}) = 0.5$. The pignistic probability function corresponding to m_1 is as follows: $BetP_{m_1}(Mary) = 0.5$, $BetP_{m_1}(Peter) = BetP_{m_1}(Paul) = 0.25$. The plausibility probability function corresponding to m_1 is as follows: $Pl_P_{m_1}(Mary) = Pl_P_{m_1}(Peter) = Pl_P_{m_1}(Paul) = 1/3$. The pignistic transformation completes the Godfather’s incomplete selection protocol by dividing the 0.5 probability equally between Peter and Paul. We refer to this assignment of equal probabilities as a random choice protocol. The plausibility transformation makes no assumption about the mechanism that will be used. The mafia don may always prefer Peter to Paul, or perhaps Paul to Peter. Using standard belief function semantics, there is a 0.5 chance that Mary is not the assassin, a 0.5 chance that Peter is not the assassin, and a 0.5 chance that Paul is not the assassin. This explains the plausibility probability function $Pl_P_{m_1}$.

Clearly, the two transformation methods yield qualitatively different results starting from the same bpa m_1 . Which probability distribution can be considered as equivalent to m_1 ? In the following paragraphs, we describe one argument (flawed, in our opinion) in favor of the pignistic transformation method and two arguments (compelling, in our opinion) in favor of the plausibility transformation method.

Consider the following argument in favor of the pignistic transformation method¹.

There is exactly one “argument” for Mary and one “counter-argument” each for Mary, Peter and Paul, respectively, as follows [6]:

	<i>Arguments</i>	<i>Counter-arguments</i>	<i>Bel</i>	<i>Pl</i>
Mary	Heads	Tails	$\Rightarrow 0.5$	0.5
Peter	–	Heads	$\Rightarrow 0$	0.5
Paul	–	Heads	$\Rightarrow 0$	0.5

A transformation method should take both arguments and counter-arguments into account. The pignistic transformation method considers both in this example by averaging the weights of arguments and counter-arguments. On the other hand, the plausibility transformation method takes only counter-arguments into account (ignoring arguments).

What this argument fails to notice is that the counter-arguments for Peter and Paul are exactly the same as the argument for Mary. Thus, in averaging the weights of arguments and counter-arguments, we are selectively double-counting information, violating a fundamental tenet of uncertain reasoning. A belief function has exactly the same information as the corresponding plausibility function, $Pl(\mathbf{a}) = 1 - Bel(\Omega_A \setminus \mathbf{a})$. By ignoring arguments, the plausibility transformation method avoids double counting uncertain information.

One way to resolve the conflict between $BetP$ and Pl_P is to appeal to the property of idempotency. Suppose we have two pieces of identical, independent

¹ This argument was provided by Rolf Haenni [private communication].

evidence about the assassin, both equal to the bpa m_1 . If we use Dempster's rule to combine these two pieces of evidence, we observe that $m_1 \oplus m_1 = m_1$, i.e., m_1 is idempotent. $Pl.P_{m_1}$ is also idempotent, i.e., $Pl.P_{m_1} \otimes Pl.P_{m_1} = Pl.P_{m_1}$. However, notice that $BetP_{m_1}$ is not idempotent. Denoting $BetP_{m_1} \otimes BetP_{m_1}$ by $BetP_m$, we have $BetP_m(Mary) = 2/3$ and $BetP_m(Peter) = BetP_m(Paul) = 1/6$. Idempotency is an important qualitative property of uncertain knowledge because double-counting of idempotent information is harmless.

Continuing the Peter, Paul or Mary saga, suppose we subsequently learn that Peter has a cast-iron alibi during the time Mr. Jones was assassinated. This piece of evidence can be represented by the bpa m_2 for A as follows: $m_2(\{Paul, Mary\}) = 1$. If we combine the two independent bpa's m_1 and m_2 , we get $(m_1 \oplus m_2)(\{Paul\}) = (m_1 \oplus m_2)(\{Mary\}) = 0.5$. Since the joint bpa has only singleton focal subsets, both the pignistic and plausibility probability functions corresponding to $m_1 \oplus m_2$ agree: $BetP_{m_1 \oplus m_2}(Paul) = Pl.P_{m_1 \oplus m_2}(Paul) = BetP_{m_1 \oplus m_2}(Mary) = Pl.P_{m_1 \oplus m_2}(Mary) = 0.5$. However, if we were using the pignistic probability distribution $BetP_{m_1}$, and we update this probability distribution (using Bayes rule) with the evidence of Peter's alibi (represented with a likelihood vector that has 0 for Peter and 1's for Paul and Mary), we end with a probability distribution for A that has probability 2/3 for Mary and 1/3 for Paul, a result that does not coincide with $BetP_{m_1 \oplus m_2}$. On the other hand, if we were using the plausibility probability distribution $Pl.P_{m_1}$, and we update this distribution with the evidence of Peter's alibi, the result is a probability distribution for A that has probability 1/2 for Paul and 1/2 for Mary, exactly the same probability distribution as $Pl.P_{m_1 \oplus m_2}$.

4.2 Example 2: Counter-Example [10].

Consider a bpa m for a variable H with state space $\Omega_H = \{h_1, \dots, h_{70}\}$ as follows: $m(\{h_1\}) = 0.30$, $m(\{h_2\}) = 0.01$, $m(\{h_2, h_3, \dots, h_{70}\}) = 0.69$. For this bpa m , the pignistic probability function $BetP_m$ is as follows: $BetP_m(h_1) = 0.30$, $BetP_m(h_2) = 0.02$, $BetP_m(h_3) = \dots = BetP_m(h_{70}) = 0.01$. The unnormalized plausibility probability function $Pl.P'_m$ is as follows: $Pl.P'_m(h_1) = 0.30$, $Pl.P'_m(h_2) = 0.70$, $Pl.P'_m(h_3) = \dots = Pl.P'_m(h_{70}) = 0.69$.

Clearly, the two probability functions are very different. The pignistic probability function has h_1 15 times more likely than h_2 whereas the plausibility probability function has h_2 2.33 times more likely than h_1 . Our interpretation is that the pignistic transformation uses a random protocol where the probability of 0.69 is divided equally amongst the 69 states h_2, \dots, h_{70} . Smets [?] argues that the originality of Shafer's model is that—unlike probabilistic models—it does not resort to an argument of symmetry to arbitrarily split belief assigned to non-singleton subsets into equal parts; however, we interpret the pignistic transformation as performing this very allocation.

Shafer [8] states that $m(A)$ should be interpreted as the probability mass that is “confined to A but can move freely to every point of A ” (p. 40). In this example, we have belief of 0.70 against h_1 , a belief of 0.30 against h_2 , and a belief of 0.31 against h_3, \dots, h_{70} . Rather than use a random choice protocol, the

plausibility transformation assumes that all mass can move freely to any state in the focal element of the belief function, which is consistent with belief function semantics.

Another compelling argument for the plausibility transformation method is as follows. Consider an hypothetical situation where we have n independent pieces of evidence, all exactly equal to m . Combining these n pieces of evidence by Dempster's rule yields m^n . For $n \geq 500$, we observe that $m^n(\{h_2\}) \approx 1$, so the result is more consistent with $Pl.P_m$ (that has h_2 as the most probable state) than with $BetP_m$ (that has h_1 as the most probable state). Notice that if we combine $Pl.P_m$ n times using Bayes rule (or pointwise multiplication) and denote the result by $(Pl.P_m)^n$, for large n we get the result that $(Pl.P_m)^n(h_2) \approx 1$.

4.3 Example 3: Target Identification Problem [4].

A target identification system is composed of 30 sensors, $S_i, i = 1, \dots, 30$. Each sensor S_i is in one of two states x_i or y_i . The state of the sensors depends on an unknown target that is assumed to be in one of two states: t_1 denoting friend, or t_2 denoting foe. The state of each sensor also depends on whether it is working or not. When in working condition, a sensor reading of x_i correctly identifies a target of type t_1 and a sensor reading of y_i correctly identifies a target of type t_2 . When the sensors are not in working condition, nothing is known about what the sensor readings mean. The first 11 sensors S_1, \dots, S_{11} are high quality sensors, and the remaining 19 sensors S_{12}, \dots, S_{30} are low quality sensors. A high quality sensor has a 99% probability of being in working condition whereas a low quality sensor has only a 90% probability of being in working condition. Data in the form of sensor readings is collected as follows: $x_1, \dots, x_{10}, y_{11}, x_{12}, y_{13}, \dots, y_{30}$. What conclusions can we draw about the actual target type?

First, we will represent the evidence from the 30 sensors by bpa's and compute the joint belief function for T . Subsequently, we will represent the evidence by probability functions using the pignistic transformation and the plausibility transformation, in each case computing the joint probability function for T .

Table 1. Bpa Encoding of Sensor Readings

Sensor $S_i = x_i$ $i = 1, \dots, 10$	Sensor $S_{11} = y_{11}$	Sensor $S_{12} = x_{12}$	Sensor $S_i = y_i$ $i = 13, \dots, 30$
$\mathbf{a} \subseteq 2^{\Omega_T} \quad m_i(\mathbf{a})$	$\mathbf{a} \subseteq 2^{\Omega_T} \quad m_{11}(\mathbf{a})$	$\mathbf{a} \subseteq 2^{\Omega_T} \quad m_{12}(\mathbf{a})$	$\mathbf{a} \subseteq 2^{\Omega_T} \quad m_i(\mathbf{a})$
$\{t_1\}$ 0.99	$\{t_2\}$ 0.99	$\{t_1\}$ 0.90	$\{t_2\}$ 0.90
$\{t_1, t_2\}$ 0.01	$\{t_1, t_2\}$ 0.01	$\{t_1, t_2\}$ 0.10	$\{t_1, t_2\}$ 0.10

Table 1 shows the data collected from the sensors represented as evidence in bpa's. We can reach a conclusion about the target identity by calculating the joint bpa for the 30 sensors. Using Dempster's rule, the joint bpa m is given by

$m = m_1 \oplus \dots \oplus m_{30}$. The results are presented in Table 2. Thus, as per the belief function model, the target is approximately 10 times more likely to be a friend than a foe.

Next, we will model this problem using probabilities from pignistic transformations of the 30 belief functions. The probability functions are shown in Table 3. The results of combining the 30 probability functions using pointwise multiplication and normalizing the resulting probability function are presented in Table 4.

Table 2. Joint Bpa and Plausibility Functions for 30 Sensors

$a \in 2^{\Omega_T}$	Un-normalized bpa	Normalized bpa (m)	Plausibility (Pl_m)
\emptyset	≈ 1	0	0
$\{t_1\}$	$\approx 1.00 \times 10^{-20}$	0.9091	0.9091
$\{t_2\}$	$\approx 1.00 \times 10^{-21}$	0.0909	0.0909
$\{t_1, t_2\}$	$\approx 1.00 \times 10^{-41}$	0.0000	1

Table 3. Pignistic Probability Function Encoding of Sensor Readings

Sensor $S_i = x_i$ $i = 1, \dots, 10$		Sensor $S_{11} = y_{11}$		Sensor $S_{12} = x_{12}$		Sensor $S_i = y_i$ $i = 13, \dots, 30$	
$x \in \Omega_T$	$BetP_{m_i}(x)$	$x \in \Omega_T$	$BetP_{m_{11}}(x)$	$x \in \Omega_T$	$BetP_{m_{12}}(x)$	$x \in \Omega_T$	$BetP_{m_i}(x)$
t_1	0.995	t_1	0.005	t_1	0.95	t_1	0.05
t_2	0.005	t_2	0.995	t_2	0.05	t_2	0.95

Table 4. The Joint Pignistic Probability Model for the Target Identification Problem

$x \in \Omega_T$	Un-normalized Probability	Normalized Probability
t_1	$\approx 1.723E - 26$	≈ 0.0820
t_2	$\approx 1.930E - 25$	≈ 0.9180
Sum	$\approx 2.102E - 25$	1

Notice that the pignistic probability model of the target identification problem is qualitatively different from the belief function model. As per the pignistic probability model, the probability that the target is a foe is approximately 11 times more likely than the probability that the target is a friend. In general,

if m_1 and m_2 are two bpa's on the same domain, then $(BetP_{m_1} \otimes BetP_{m_2}) \neq BetP_{m_1 \oplus m_2}$.

Next, consider the probability model for the target identification problem obtained from the belief function model using the plausibility transformation. This model is shown in Table 5. If we combine the 30 plausibility probability functions using pointwise multiplication and normalize the resulting probability function, we obtain the results in Table 6.

Table 5. Plausibility Probability Function Encoding of Sensor Readings

Sensor $S_i = x_i$ $i = 1, \dots, 10$		Sensor $S_{11} = y_{11}$		Sensor $S_{12} = x_{12}$		Sensor $S_i = y_i$ $i = 13, \dots, 30$	
$x \in \Omega_T$	$PLP_{m_i}(x)$	$x \in \Omega_T$	$PLP_{m_{11}}(x)$	$x \in \Omega_T$	$PLP_{m_{12}}(x)$	$x \in \Omega_T$	$PLP_{m_i}(x)$
t_1	0.9901	t_1	0.0099	t_1	0.9091	t_1	0.0909
t_2	0.0099	t_2	0.9901	t_2	0.0909	t_2	0.9091

Table 6. The Joint Plausibility Probability Model

$x \in \Omega_T$	Un-normalized Probability	Normalized Probability
t_1	$\approx 1.4656E - 21$	≈ 0.9091
t_2	$\approx 1.4656E - 22$	≈ 0.0909
Sum	$\approx 1.6121E - 21$	1

Notice that the conclusion is similar to the result obtained in the belief function model. In the next section, we will show that this equivalence between the belief function model conclusion and plausibility probability function is always true.

5 Justification and Properties of the Plausibility Transformation

In all three examples described in the previous section, there is a discrepancy between the pignistic probability function(s) obtained before and after combining all evidence. Smets [10] resolves this apparent discrepancy of the pignistic transformation by stating that beliefs are held at the credal level and one only descends to the probability space for decision making at the time a decision has to be made. However, we view decision-making as a dynamic activity.

Probability theory and belief function theory are two uncertainty calculi with roughly the same expressive power [1]. One should get roughly the same results regardless of the calculi one is using to represent knowledge if the models built

using the calculi are equivalent. An appropriate transformation method can allow a model of an uncertain domain in one calculus to be translated into the other. Thus we can exploit the advantages of both calculi.

The pignistic transformation is justified based on a so-called “rationality” requirement, which implies a mathematical requirement of linearity. Other justifications for the pignistic transformation are given in [10, 11]. Some intuitive justifications for the plausibility transformation are given in [2, 3]. Here we will state four theorems that demonstrate that the plausibility transformation is consistent with belief functions semantics.

Theorem 1. Suppose m_1, \dots, m_k are k bpa’s. Suppose $Pl_{m_1}, \dots, Pl_{m_k}$ are the associated plausibility functions, and suppose $Pl_{P_{m_1}}, \dots, Pl_{P_{m_k}}$ are the corresponding probability functions. If $m = m_1 \oplus \dots \oplus m_k$ is the joint bpa, Pl_m is the associated plausibility function, and Pl_{P_m} is the corresponding plausibility probability function, then $Pl_{P_{m_1}} \otimes \dots \otimes Pl_{P_{m_k}} = Pl_{P_m}$.

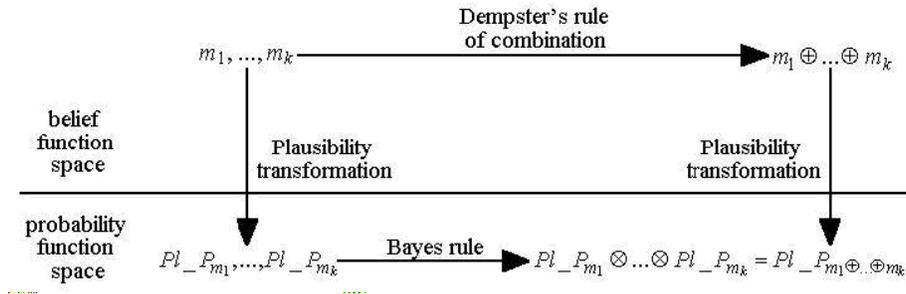


Fig. 1. A Pictorial Depiction of Theorem 1.

Theorem 1 is depicted pictorially in Figure 1. Notice that from a computational perspective, it is much faster to compute $Pl_{P_{m_1}} \otimes \dots \otimes Pl_{P_{m_k}}$ than it is to compute Pl_{P_m} (since the latter involves Dempster’s rule of combination and the former involves Bayes rule).

Given bpa m , we don’t view Pl_{P_m} as an approximation of m . Instead, we view Pl_{P_m} as an equivalent probability encoding of the information in m . Thus if we have a belief function model consisting of $\{m_1, \dots, m_k\}$, then we view $\{Pl_{P_{m_1}}, \dots, Pl_{P_{m_k}}\}$ as an equivalent probability model. Theorem 1 can be viewed as a regularity condition for any transformation method. As demonstrated in the Peter, Paul, and Mary, and the target identification problems, the pignistic transformation does not satisfy this condition.

If a unique state x exists in a bpa m such that $\lim_{n \rightarrow \infty} m^n(\{x\}) = 1$ (where $m^n = m \oplus \dots \oplus m$, n times), an equivalent probability function should have x as its most probable state. This property is satisfied for the plausibility transformation, as stated in the following theorem.

Theorem 2. Consider a bpa m for s (with corresponding plausibility function Pl_m) such that $x \in \Omega_s$ is the most plausible state, i.e., $Pl_m(\{x\}) > Pl_m(\{y\})$, for all $y \in \Omega_s \setminus \{x\}$. If Pl_{m^∞} denotes the plausibility function corresponding to m^∞ , then $Pl_{m^\infty}(\{x\}) = 1$, and $Pl_{m^\infty}(\{y\}) = 0$ for all $y \in \Omega_s \setminus \{x\}$.

In Example 2 presented in Section 4, $m^{500}(\{h_2\}) \approx 1$, so the most plausible hypothesis in m is h_2 , consistent with $Pl_P m$ and not $Bet P m$.

If a bpa function has a subset of most plausible states, all with equal plausibility, the following theorem applies.

Theorem 3. Consider a bpa m for s (with corresponding plausibility function Pl_m) such that $\mathbf{t} \subseteq \Omega_s$ is a subset of most plausible states, i.e., $Pl_m(\{x\}) = Pl_m(\{y\})$ for all $x, y \in \mathbf{t}$, and $Pl_m(\{x\}) > Pl_m(\{z\})$ for all $x \in \mathbf{t}$, and $z \in \Omega_s \setminus \mathbf{t}$. Then there exists a partition $\{\mathbf{a}_1, \dots, \mathbf{a}_k\}$ of \mathbf{t} such that $m^\infty(\mathbf{a}_i) = 1/k$ for $i = 1, \dots, k$, i.e., $Pl_{m^\infty}(x) = Pl_{m^\infty}(y) = 1/k$ for all $x, y \in \mathbf{t}$, and $Pl_{m^\infty}(z) = 0$ for all $z \in \Omega_s \setminus \mathbf{t}$.

In the Peter, Paul, and Mary saga described earlier, the initial belief function m_1 has a corresponding plausibility function where each state has equal plausibilities. Theorem 3 applies with $a_1 = \{Mary\}$, $a_2 = \{Peter, Paul\}$, and $k = 2$. The next theorem states that $Pl_P m$ is idempotent if m is idempotent.

Theorem 4. If m is idempotent with respect to Dempster's rule, i.e., $m \oplus m = m$, then $Pl_P m$ is idempotent with respect to Bayes rule, i.e., $Pl_P m \otimes Pl_P m = Pl_P m$.

As demonstrated in the Peter, Paul, and Mary example, $Bet P m$ does not satisfy this property.

6 Conclusions and Summary

In summary, if T transforms a bpa m in a belief function model to an equivalent probability function $T(m)$, T should satisfy four basic properties:

- 1) *Invariance with respect to combination:* $T(m_1 \oplus \dots \oplus m_n) = T(m_1) \otimes \dots \otimes T(m_n)$, which is satisfied for the plausibility transformation, according to Theorem 1;
- 2) *Unique most plausible state:* $\text{Lim}_{n \rightarrow \infty} T^n(m)(h_i) = 1$ if $\text{Lim}_{n \rightarrow \infty} m^n(h_i) = 1$, which is satisfied for the plausibility transformation according to Theorem 2;
- 3) *Non-unique most plausible states:* If $\text{Lim}_{n \rightarrow \infty} Pl_{m^n}(x) = \text{Lim}_{n \rightarrow \infty} Pl_{m^n}(y)$ for all $x, y \in \mathbf{t} \subseteq \Omega_s$ and $\text{Lim}_{n \rightarrow \infty} Pl_{m^n}(z) = 0$ for all $z \in \Omega_s \setminus \mathbf{t}$, then $\text{Lim}_{n \rightarrow \infty} T^n(m)(x) = \text{Lim}_{n \rightarrow \infty} T^n(m)(y)$ for all $x, y \in \mathbf{t}$, and $\text{Lim}_{n \rightarrow \infty} T^n(m)(z) = 0$ for all $z \in \Omega_s \setminus \mathbf{t}$; this property is satisfied for the plausibility transformation according to Theorem 3; and
- 4) *Idempotency:* $T(m)$ is idempotent if m is idempotent, which is satisfied by the plausibility probability transformation according to Theorem 4.

The main goal of this paper is to compare the pignistic and plausibility transformation methods for transforming belief function models to probability models. Until now, most of the literature on belief functions has used the pignistic method. The pignistic transformation method does not satisfy the invariance with respect to combination, most plausible, and idempotency axioms. On the other hand, the plausibility transformation satisfies all intuitively acceptable axioms we have postulated for an acceptable transformation method. We conjecture that the plausibility transformation method is the only method that satisfies these axioms, but we don't have a proof of this claim.

Acknowledgements

The research was partly funded by a contract from Sparta, Inc., to the second author. We are grateful for extensive comments from Philippe Smets, Rolf Haenni, and Yang Chen on earlier drafts of this paper.

References

1. Cobb, B.R., Shenoy, P.P.: A Comparison of Bayesian and Belief Function Reasoning. Working Paper No. 292, University of Kansas School of Business (2002)
2. Cobb, B.R., Shenoy, P.P.: On Transformations of Belief Function Models to Probability Models. Working Paper No. 293, University of Kansas School of Business (2003)
3. Cobb, B.R., Shenoy, P.P.: Converting Belief Function Models to Probability Models Using the Plausibility Transformation. Working Paper, University of Kansas School of Business, submitted to UAI-03 (2003)
4. Delmotte, F., Smets, P.: Target Identification Based on the Transferable Belief Model Interpretation of the Dempster-Shafer Model, Part II: Applications. Working Paper, LAMIH-Université de Valenciennes, IRIDIA-Université Libre de Bruxelles (2001)
5. Dempster, A.P.: New Methods of Reasoning Toward Posterior Distributions Based on Sample Data. *Annals of Mathematical Statistics*. **37** (1987) 355–374
6. Haenni, R., Lehmann, N.: Probabilistic Argumentation Systems: A New Perspective on Dempster-Shafer Theory. *International Journal of Intelligent Systems*, in press (2003)
7. Kohlas, J., Monney, P.A.: A Mathematical Theory of Hints. In: *Lecture Notes in Economics and Mathematical Systems*, vol. 425. Springer-Verlag, Berlin (1995)
8. Shafer, G.: *A Mathematical Theory of Evidence*. Princeton University Press, Princeton, N.J. (1976)
9. Shafer, G.: Belief Functions and Possibility Measures. In: Bezdek, J. (ed.): *The Analysis of Fuzzy Information*. CRC Press, Boca Raton, FL (1987) 51–84
10. Smets, P.: Decision Making in a Context Where Uncertainty is Represented by Belief Functions. In: Srivastava, R.P., Mock, T.J. (eds.): *Belief Functions in Business Decisions*. Physica-Verlag, Heidelberg (2002) 17–61
11. Smets, P., Kennes, R.: The Transferable Belief Model. *Artificial Intelligence*. **66** (1994) 191–234
12. Walley, P.: *Statistical Reasoning with Imprecise Probabilities*. Chapman and Hall, London (1991)