An Axiomatic Utility Theory for Dempster-Shafer Belief Functions

Thierry Denœux THIERRY.DENOEUX@UTC.FR
Université de Technologie de Compiègne, CNRS, UMR 7253 Heudiasyc, Compiègne, France

Prakash P. Shenoy PSHENOY@KU.EDU
School of Business, University of Kansas, Lawrence, KS, 66045 USA

Abstract

The main goal of this paper is to describe an axiomatic utility theory for Dempster-Shafer belief function lotteries. The axiomatic framework used is analogous to von Neumann-Morgenstern's utility theory for probabilistic lotteries as described by Luce and Raiffa. Unlike the probabilistic case, our axiomatic framework leads to interval-valued utilities, and therefore, to a partial (incomplete) preference order on the set of all belief function lotteries. If the belief function reference lotteries we use are Bayesian belief functions, then our representation theorem coincides with Jaffray's representation theorem for his linear utility theory for belief functions. We illustrate our framework using some examples discussed in the literature. Finally, we compare our decision theory with those proposed by Jaffray and Smets.

Keywords: Dempster-Shafer theory of evidence, von Neumann-Morgenstern's utility theory, Jaffray's linear utility theory, Smets' decision theory

1. Introduction

The main goal of this paper is to propose an axiomatic utility theory for lotteries described by belief functions in the Dempster-Shafer (D-S) theory of evidence [5, 24]. The axiomatic theory is constructed similar to von Neumann-Morgenstern's (vN-M's) utility theory for probabilistic lotteries [31, 19, 18, 23, 21, 13]. Unlike the probabilistic case, our axiomatic theory leads to interval-valued utilities, and therefore to a partial (incomplete) preference order on the set of all belief function lotteries. Also, we compare our decision theory to those proposed by Jaffray [20] and Smets [29].

In the foreword to Glenn Shafer's 1976 monograph [24], Dempster writes: "... I believe that Bayesian inference will always be a basic tool for practical everyday statistics, if only because questions must be answered and decisions must be taken, so that a statistician must always stand ready to upgrade his vaguer forms of belief into precisely additive probabilities." More than 40 years after these lines were written, a lot of approaches to decision-making have been proposed (see the recent review in [6]). However, most of these methods lack a strong theoretical basis. The most important steps toward a decision theory in the D-S framework have been made by Jaffray [20], Smets [29], and Shafer [27]. However, we argue that these proposals are either not sufficiently justified from the point of view of D-S theory, or not sufficiently developed for practical use. Our goal is to propose and justify a utility theory that is in line with vN-M's utility theory, but adapted to be used with lotteries whose uncertainty is described by D-S belief functions.

In essence, the D-S theory consists of representations—basic probability assignments (also called mass functions), belief functions, plausibility functions, etc.—together with Dempster's combination rule, and a rule for marginalizing joint belief functions. The representation part of the D-S theory is also used in various other theories of belief functions. For example, in the imprecise probability community, a belief function is viewed as the lower envelope of a convex set of probability mass functions called a credal set. Using these semantics, it makes more sense to use the Fagin-Halpern combination rule [12], rather than Dempster's combination rule [17, 25, 26]. The utility theory this article proposes is designed specifically for the D-S belief function theory, and not for the other theories of belief functions. This suggests that Dempster's combination rule should be an integral part of our theory, a property that is not satisfied in the proposals by Jaffray and Smets.

There is a large literature on decision making with a (credal) set of probability mass functions motivated by Ellsberg's paradox [11]. An influential work in this area is the axiomatic framework by Gilboa-Schmeidler [15], where they use Choquet integration [3, 16] to compute expected utility. A belief function is a special case of a Choquet capacity. Jaffray's [20] work can also be regarded as belonging to the same line of research, although Jaffray works directly with belief functions without specifying a combination rule. A review of this literature can be found in, e.g., [14], where the authors propose a modification of the Gilboa-Schmeidler [15] axioms. As we said earlier, our focus here is on decision-making with D-S theory of belief functions, and not on decision-making based in belief functions with a credal set interpretation. As we will see, our interval-valued utility functions lead to intervals that

© T. Denœux & P.P. Shenoy.

are contained in the Choquet lower and upper expected utility intervals.

The remainder of this article is as follows. In Section 2, we sketch vN-M's axiomatic utility theory for probabilistic lotteries as described by Luce and Raiffa [23]. In Section 3, we describe our adaptation of vN-M's utility theory for lotteries in which uncertainty is described by D-S belief functions. Our assumptions lead to an interval-valued utility function, and consequently, to a partial (incomplete) preference order on the set of all belief function lotteries. In Section 4, we illustrate the application of our representation theorem to three examples from the literature. In Section 5, we compare our utility theory with those described by Jaffray [20], and Smets [29]. Finally, in Section 6, we summarize and conclude.

2. vN-M's Utility Theory

In this section, we describe vN-M's utility theory for decision under risk. Most of the material in this section is adapted from [23]. A decision problem can be seen as a situation in which a decision-maker (DM) has to choose a course of action (or an *act*) in some set $\mathbf{F}$. An act may have different *outcomes*, depending on the *state of nature* X. Exactly one state of nature will obtain, but this state is unknown. Let $\Omega_X = \{x_1,\dots,x_n\}$ denote the set of states of nature, and let $\mathbf{O} = \{O_1,\dots,O_r\}$ denote the set of outcomes.[1] An act can be formalized as a mapping $f : \Omega_X \to \mathbf{O}$. In this section, we assume that uncertainty about the state of nature is described by a probability mass function (PMF) p_X for X. If the DM select act f, they will get outcome O_i with probability

$$p_i = \sum_{\{x\in\Omega_X \mid f(x)=O_i\}} p_X(x). \tag{1}$$

To each act f there corresponds a PMF $\mathbf{p} = (p_1,\dots,p_r)$ for $\mathbf{O}$. We call $L = [\mathbf{O},\mathbf{p}]$ a *probabilistic lottery*. As only one state in Ω_X will be realized, a probabilistic lottery will result in exactly one outcome O_i (with probability p_i), and we assume that the lottery will not be repeated. Another natural assumption is that two acts that induce the same lottery are equivalent: the problem of expressing preference between acts then boils down to expressing preference between lotteries.

We are concerned with a DM who has preferences on $\mathscr{L}$, the set of all probabilistic lotteries on $\mathbf{O}$, and our task is to find a real-valued *utility function* $u : \mathscr{L} \to \mathbb{R}$ such that the DM strictly prefers L to L' if and only if $u(L) > u(L')$, and the DM is indifferent between L and L' if and only if $u(L) = u(L')$. We write $O_i \succ O_j$ if the DM strictly prefers O_i to O_j, write $O_i \sim O_j$ if the DM is indifferent between (or equally prefers) O_i and O_j, and write $O_i \succsim O_j$ if the DM either strictly prefers O_i to O_j or is indifferent between the two.

Of course, finding such a utility function is not always possible, unless the DM's preferences satisfy some assumptions. We can then construct a utility function that is *linear* in the sense that the utility of a lottery $L = [\mathbf{O},\mathbf{p}]$ is equal to its expected utility $\sum_{i=1}^r p_i\, u(O_i)$, where O_i is regarded as a degenerate lottery where the only possible outcome is O_i with probability 1. In the remainder of this section, we describe the assumptions that lead to the existence of such a linear utility function.

Assumption 2.1 (Weak ordering of outcomes) *For any two outcomes O_i and O_j, either $O_i \succsim O_j$ or $O_j \succsim O_i$. Also, if $O_i \succsim O_j$ and $O_j \succsim O_k$, then $O_i \succsim O_k$. Thus, the preference relation $\succsim$ over $\boldsymbol{O}$ is a weak order, i.e., it is complete and transitive.*

Given Assumption 2.1, without loss of generality, let us assume that the outcomes are labelled such that $O_1 \succsim O_2 \succsim \cdots \succsim O_r$, and to avoid trivialities, assume that $O_1 \succ O_r$.

Suppose that $\mathbf{L} = \{L^{(1)},\dots,L^{(s)}\}$ is a set of s lotteries, where each of the s lotteries $L^j = [\mathbf{O},\mathbf{p}^{(j)}]$ are over outcomes in $\mathbf{O}$, with PMFs $\mathbf{p}^{(j)}$ for $j = 1,\dots,s$. Suppose $\mathbf{q} = (q_1,\dots,q_s)$ is a PMF for $\mathbf{L}$ such that $q_j > 0$ for $j = 1,\dots,s$, and $\sum_{j=1}^s q_j = 1$. Then $[\mathbf{L},\mathbf{q}]$ is called a *compound* lottery whose outcome is exactly one lottery $L^{(i)}$ (with probability q_i), and lottery $L^{(i)}$ will result in one outcome O_j (with probability $p_j^{(i)}$). Notice that the PMF $\mathbf{p}^{(i)}$ is a conditional PMF for $\mathbf{O}$ in the second stage given that lottery $L^{(i)}$ is realized (with probability $q_i > 0$) in the first stage (see Figure 1). We can compute the joint PMF for $(\mathbf{L},\mathbf{O})$, and then compute the marginal $\mathbf{p}$ of the joint for $\mathbf{O}$. The following assumption states that the resulting lottery $[\mathbf{O},\mathbf{p}]$ is indifferent to the compound lottery $[\mathbf{L},\mathbf{q}]$.

Assumption 2.2 (Reduction of compound lotteries) *Any compound lottery $[\boldsymbol{L},\ \boldsymbol{q}]$, where $L^{(i)} = [\boldsymbol{O},\boldsymbol{p}^{(i)}]$, is indifferent to a simple (non-compound) lottery $[\boldsymbol{O},\ \boldsymbol{p}]$, where*

$$p_i = q_1\, p_i^{(1)} + \ldots + q_s\, p_i^{(s)} \tag{2}$$

for $i = 1,\dots,r$. PMF $(p_1,\dots,p_r)$ is the marginal for $\boldsymbol{O}$ of the joint PMF of $(\boldsymbol{L},\boldsymbol{O})$.

A simple lottery involving only outcomes O_1 and O_r with PMF $(u, 1-u)$, where $0 \le u \le 1$, is called a *reference* lottery, and is denoted by $[\{O_1,O_r\},(u,1-u)]$. Let $\mathbf{O}_2$ denote the set $\{O_1,O_r\}$.

Assumption 2.3 (Continuity) *Each outcome O_i is indifferent to a reference lottery $\widetilde{O}_i = [\boldsymbol{O}_2,(u_i,1-u_i)]$ for some u_i, where $0 \le u_i \le 1$, i.e., $O_i \sim \widetilde{O}_i$.*

1. The assumption of finiteness of the sets Ω_X and $\mathbf{O}$ is only for ease of exposition. It is unnecessary for the proof of the representation theorem in this section.

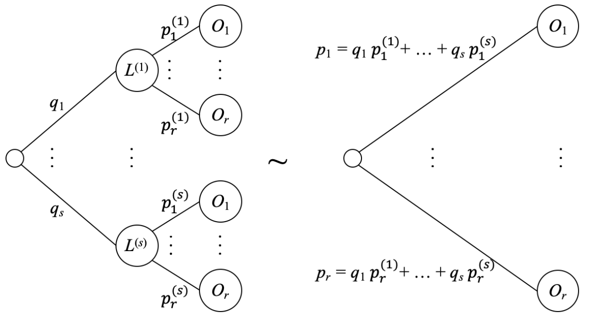

Figure 1: A two-stage compound lottery reduced to an indifferent simple lottery

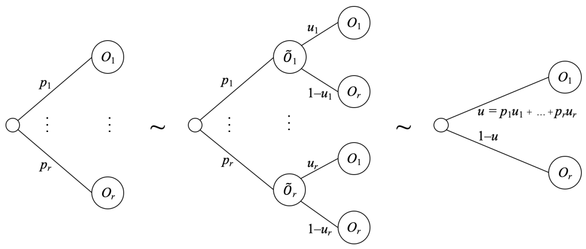

Figure 2: Reducing a lottery to an indifferent compound lottery and then to an indifferent reference lottery

Assumption 2.4 (Weak order) *The preference relation $\succsim$ for lotteries in $\mathscr{L}$ is a weak order, i.e., it is complete and transitive.*

Assumption 2.4 generalizes Assumption 2.1 for outcomes, which can be regarded as degenerate lotteries.

Assumption 2.5 (Substitutability) *In any lottery $L = [\boldsymbol{O},\boldsymbol{p}]$, if we substitute an outcome O_i by the reference lottery $\widetilde{O}_i = [\boldsymbol{O}_2,(u_i,1-u_i)]$ that is indifferent to O_i, then the result is a compound lottery that is indifferent to L.*

From Assumptions 2.1–2.5, given any lottery $L = [\mathbf{O},\mathbf{p}]$, it is possible to find a reference lottery $\widetilde{L} = [\mathbf{O}_2,(u,1-u)]$ that is indifferent to L (see Figure 2). This is expressed by Theorem 1 below. [2]

Theorem 1 ([23]) *Under Assumptions 2.1–2.5, any lottery $L = [\boldsymbol{O},\boldsymbol{p}]$ is indifferent to a reference lottery $\widetilde{L} = [\boldsymbol{O}_2,(u,1-u)]$ with*

$$u = \sum_{i=1}^{r} p_i u_i. \tag{3}$$

Assumption 2.6 (Monotonicity) *A reference lottery $L = [\boldsymbol{O}_2,(u,1-u)]$ is preferred or indifferent to reference lottery $L' = [\boldsymbol{O}_2,(u',1-u')]$ if and only if $u \geq u'$.*

2. For reasons of space, proofs of all results in this paper are omitted and can be found in [7].

As $O_1 \sim \widetilde{O}_1 = [\mathbf{O},(u_1,1-u_1)]$ and $O_r \sim \widetilde{O}_r = [\mathbf{O},(u_r,1-u_r)]$, Assumptions 2.4 and 2.6 imply that $u_1 = 1$ and $u_r = 0$. Also, from $O_1 \succsim O_2 \succsim \cdots \succsim O_r$, we can deduce that $1 = u_1 \geq u_2 \geq \cdots \geq u_r = 0$.

Assumptions 2.1–2.6 allow us to define the utility of a lottery as the probability of the best outcome O_1 in an indifferent reference lottery, and this utility function for lotteries on $\mathbf{O}$ is linear. This is stated by the following theorem.

Theorem 2 ([23]) *If the preference relation $\succsim$ on $\mathscr{L}$ satisfies Assumptions 2.1–2.6, then there are numbers u_i associated with outcomes O_i for $i = 1,\ldots,r$, such that for any two lotteries $L = [\boldsymbol{O},\boldsymbol{p}]$, and $L' = [\boldsymbol{O},\boldsymbol{p}']$, $L \succsim L'$ if and only if*

$$\sum_{i=1}^{r} p_i u_i \geq \sum_{i=1}^{r} p'_i u_i. \tag{4}$$

Thus, we can define the utility of lottery $L = [\boldsymbol{O},\boldsymbol{p}]$ as $u(L) = \sum_{i=1}^{r} p_i u_i$, where $u_i = u(O_i)$. Also, such a linear utility function is unique up to a strictly increasing affine transformation, i.e., if $u'_i = a\,u_i + b$, where $a > 0$ and b are real constants, then $u(L) = \sum_{i=1}^{r} p_i u'_i$ also qualifies as a utility function.

3. A Utility Theory for D-S Belief Function Theory

In this section, we describe a new utility theory for lotteries where the uncertainty is described by D-S belief functions.[3] These lotteries, called *belief function lotteries*,[4] will be introduced in Section 3.1. We present and discuss assumptions in Section 3.2, and state a representation theorem in Section 3.3.

3.1. Belief function lotteries

We now generalize the decision framework outline in Section 2 by assuming that uncertainty about the state of nature X with state space Ω_X is described by a BPA m_X for X. The probabilistic framework is recovered as a special case when m_X is Bayesian. As before, we define an act as a mapping $f : \Omega_X \to \mathbf{O}$. Mapping f pushes m_X forward from Ω_X to $\mathbf{O}$, transferring each mass $m_X(\mathsf{a})$ for $\mathsf{a} \in 2^{\Omega_X}$ to $\mathsf{b} = \{f(x) : x \in \mathsf{a}\}$. The resulting BPA m for $\mathbf{O}$ is then defined as follows:

$$m(\mathsf{b}) = \sum_{\{\mathsf{a}\in 2^{\Omega_X} \,|\, f[\mathsf{a}]=\mathsf{b}\}} m_X(\mathsf{a}), \tag{5}$$

for all $\mathsf{b} \subseteq \mathbf{O}$, where $f[\mathsf{a}]$ denotes the image of subset a by f [8]. Eq. (5) clearly generalizes Eq. (1). The pair $[\mathbf{O},m]$ will

3. We assume the reader is familiar with the fundamentals of D-S belief functions. A brief review appears in [7].

4. This notion was previously introduced in [6] under the name "evidential lottery."

be called a *belief function* (bf) lottery. As before, we assume that two acts can be compared from what we believe their outcomes will be, irrespective of the evidence on which we base our beliefs. This assumption is a form of what Wakker [32] calls the *principle of complete ignorance* (PCI). It implies that two acts resulting in the same bf lottery are equivalent. The problem of expressing preferences between acts becomes that of expressing preferences between bf lotteries.

Thus, we are concerned with a DM who has preferences on $\mathscr{L}_{bf}$, the set of all bf lotteries. We will define our task as finding a utility function $u : \mathscr{L}_{bf} \to [\mathbb{R}]$, where $[\mathbb{R}]$ denotes the set of closed real intervals, such that the $u(L) = [u, u+w]$ is viewed as an interval-valued utility of L. The interval-valued utility can be interpreted as follows: u and $u+w$ are, respectively, the degrees of belief and plausibility of receiving the best outcome in a bf reference lottery equivalent to L. Given two lotteries L and L', L is preferred to L' if and only if $u \geq u'$ and $u+w \geq u'+w'$. This leads to incomplete preferences on the set of all bf lotteries. If we assume $w = 0$ for all bf lotteries, then we have a real-valued utility function on $\mathscr{L}_{bf}$, and consequently, complete preferences.

Example 1 (Ellsberg's Urn) *Ellsberg [11] describes a decision problem that questions the adequacy of the vN-M axiomatic framework. Suppose we have an urn with 90 balls, of which 30 are red, and the remaining 60 are either black or yellow. We draw a ball at random from the urn. Let X denote the color of the ball drawn, with $\Omega_X = \{r, b, y\}$. Notice that the uncertainty of X can be described by a BPA m_X for X such that $m_X(\{r\}) = 1/3$, and $m_X(\{b, y\}) = 2/3$.*

First, we are offered a choice between Lottery L_1: \$100 *on red, and Lottery L_2:* \$100 *on black, i.e., in L_1, you get* \$100 *if the ball drawn is red, and* \$0 *if the ball drawn is black or yellow, and in L_2, you get* \$100 *if the ball drawn is black and* \$0 *if the ball drawn is red or yellow. Choice of L_1 can be denoted by alternative $f_1 : \Omega_X \to \{\$100, \$0\}$ such that $f_1(r) = \$100$, $f_1(b) = f_1(y) = \$0$. Similarly, choice of L_2 can be denoted by alternative $f_2 : \Omega_X \to \{\$100, \$0\}$ such that $f_2(b) = \$100$, $f_2(r) = f_2(y) = \$0$. L_1 can be represented by the BPA m_1 for $\boldsymbol{O} = \{\$0, \$100\}$ as follows: $m_1(\{\$100\}) = 1/3$, $m_1(\{\$0\}) = 2/3$. L_2 can be represented by BPA m_2 for $\boldsymbol{O}$ as follows: $m_2(\{\$0\}) = 1/3$, $m_2(\{\$0, \$100\}) = 2/3$. Notice that L_1 and L_2 are bf lotteries. Ellsberg notes that a frequent pattern of response is L_1 preferred to L_2.*

Second, we are offered a choice between L_3: \$100 *on red or yellow, and L_4:* \$100 *on black or yellow, i.e., in L_3 you get* \$100 *if the ball drawn is red or yellow, and* \$0 *if the ball drawn is black, and in L_4, you get* \$100 *if the ball drawn is black or yellow, and* \$0 *if the ball drawn is red. L_3 can be represented by BPA m_3 as follows: $m_3(\{\$100\}) = 1/3$, and $m_3(\{\$0, \$100\}) = 2/3$, and L_4 can be represented by the BPA m_4 as follows: $m_4(\{\$0\}) = 1/3$, $m_4(\{\$100\}) = 2/3$. L_3 and L_4 are also belief function lotteries. Ellsberg notes that L_4 is often strictly preferred to L_3. Also, the same subjects who prefer L_1 to L_2, prefer L_4 to L_3.*

3.2. Assumptions of our framework

As in the probabilistic case, we will assume that a DM's preferences for bf lotteries are reflexive and transitive. However, unlike the probabilistic case, we do not assume that these preferences are complete. In the probabilistic case, incomplete preferences are studied in [1], and in the case of sets of utility functions, in [9].

Our first assumption is identical to Assumption 2.1.

Assumption 3.1 (Weak ordering of outcomes) *The DM's preferences $\succsim$ for outcomes in $\boldsymbol{O} = \{O_1, \ldots, O_r\}$ are complete and transitive.*

This allows us to label the outcomes such that

$$O_1 \succsim O_2 \succsim \cdots \succsim O_r, \quad \text{and} \quad O_1 \succ O_r. \tag{6}$$

Let $\mathscr{L}_{bf}$ denote the set of all bf lotteries on $\mathbf{O} = \{O_1, \ldots O_r\}$, where the outcomes satisfy Eq. (6). As every BPA m for $\mathbf{O}$ is a bf lottery, $\mathscr{L}_{bf}$ is essentially the set of all BPAs for $\mathbf{O}$. As the set of all BPAs include Bayesian BPAs, the set $\mathscr{L}_{bf}$ is a superset of $\mathscr{L}$, i.e., every probabilistic lottery on $\mathbf{O}$ can be considered a bf lottery.

Consider a compound lottery $[\mathbf{L}, m]$, where $\mathbf{L} = \{L_1, \ldots, L_s\}$, m is a BPA for $\mathbf{L}$, and $L_j = [\mathbf{O}, m_j]$ is a bf lottery on $\mathbf{O}$, where m_j is a conditional BPA for $\mathbf{O}$ in the second stage given that lottery L_j is realized in the first stage. Assumption 3.2 posits that we can reduce the compound lottery to a simple bf lottery on $\mathbf{O}$ using the D-S calculus, and that the compound lottery is equally preferred to the reduced simple lottery on $\mathbf{O}$.

Assumption 3.2 (Reduction of compound lotteries) *Suppose $[\boldsymbol{L}, m]$ is a compound lottery as described in the previous paragraph. Then, $[\boldsymbol{L}, m] \sim [\boldsymbol{O}, m']$, where*

$$m' = \left(m \oplus \left(\bigoplus_{j=1}^{s} m_{L_j,j}\right)\right)^{\downarrow \boldsymbol{O}}, \tag{7}$$

and $m_{L_j,j}$ is a BPA for $(\boldsymbol{L}, \boldsymbol{O})$ obtained from m_j by conditional embedding, for $j = 1, \ldots, s$.

The following proposition states that Assumption 3.2 generalizes Assumption 2.2.

Proposition 3 *Let $\boldsymbol{L} = \{L_1, \ldots, L_s\}$ be a set of bf lotteries, with $L_j = [\boldsymbol{O}, m_j]$, in which m_j is a Bayesian conditional BPA for $\boldsymbol{O}$ given L_j such that $m_j(\{O_i\}) = p_i^{(j)}$ and $\sum_{i=1}^{r} p_i^{(j)} = 1$ for $j = 1, \ldots, s$. Let $[\boldsymbol{L}, m]$ be a compound lottery in which m is a Bayesian BPA for $\boldsymbol{L}$ such that*

$m(\{L_j\} = q_j$ for $j = 1,\dots,s$ with $\sum_{j=1}^{s} q_j = 1$. Then BPA m' defined by Eq. (7) is Bayesian, and it verifies

$$m'(\{O_i\}) = \sum_{j=1}^{s} q_j p_i^{(j)} \tag{8}$$

for $i = 1,\dots,r$.

Next, we define a *bf reference* lottery $[\mathbf{O}_2, m]$ as a bf lottery on $\mathbf{O}_2 = \{O_1, O_r\}$. A bf reference lottery has three parameters $u = m(\{O_1\})$, $v = m(\{O_r\})$, and $w = m(\mathbf{O}_2)$, which are all non-negative and sum to 1. The following assumption states that any deterministic bf lottery is equally preferred to some bf reference lottery.

Assumption 3.3 (Continuity) *Any subset of outcomes $\mathsf{a} \subseteq \boldsymbol{O}$ (considered as a deterministic bf lottery) is indifferent to a bf reference lottery $\widetilde{\mathsf{a}} = [\boldsymbol{O}_2, m_{\mathsf{a}}]$ such that*

$$m_{\mathsf{a}}(\{O_1\}) = u_{\mathsf{a}}, \tag{9a}$$

$$m_{\mathsf{a}}(\{O_r\}) = v_{\mathsf{a}}, \quad \textit{and} \tag{9b}$$

$$m_{\mathsf{a}}(\boldsymbol{O}_2) = w_{\mathsf{a}}, \tag{9c}$$

where $u_{\mathsf{a}}, v_{\mathsf{a}}, w_{\mathsf{a}} \geq 0$, and $u_{\mathsf{a}} + v_{\mathsf{a}} + w_{\mathsf{a}} = 1$. Furthermore, $w_{\mathsf{a}} = 0$ if $\mathsf{a} = \{O_i\}$ is a singleton.

Notice that $Bel_{m_{\mathsf{a}}}(\{O_1\}) = u_{\mathsf{a}}$, and $Pl_{m_{\mathsf{a}}}(\{O_1\}) = u_{\mathsf{a}} + w_{\mathsf{a}} = 1 - v_{\mathsf{a}}$. For singleton subsets, the equivalent bf reference lottery is Bayesian: this ensures that Assumption 3.3 is a generalization of Assumption 2.3. For non-singleton subsets a of outcomes, we may have $w_{\mathsf{a}} > 0$, i.e., the bf reference lottery may not be Bayesian. In other words, we do not assume that ambiguity can be resolved by selecting an equivalent probabilistic reference lottery.

Example 2 *Consider lottery $L_2 = [\{\$100, \$0\}, m_2]$ in Example 1, where $m_2(\{\$0\}) = 1/3$, and $m_2(\{\$100, \$0\}) = 2/3$. Suppose we wish to assess the utility of focal set $\{\$100, \$0\}$ using a probabilistic reference lottery $[\{\$100, \$0\}, (p, 1-p)]$. A DM may have the following preferences. For any $p \leq 0.2$ she prefers $\{\$100, \$0\}$ to the probabilistic reference lottery, and for any $p \geq 0.3$, she prefers the probabilistic reference lottery to $\{\$100, \$0\}$. However, she is unable to give us a precise p such that $\{\$100, \$0\} \sim [\{\$100, \$0], (p, 1-p)]$. For such a DM, we can assess a bf reference lottery $[\{\$100, \$0\}, m_{\mathsf{a}}]$ such that $Bel_{m_{\mathsf{a}}}(\{\$100\}) = 0.2$ and $Pl_{m_{\mathsf{a}}}(\{\$100\}) = 0.3$, i.e., $u_{\{\$100,\$0\}} = 0.2$, $v_{\{\$100,\$0\}} = 0.7$, and $w_{\{\$100,\$0\}} = 0.1$.*

Assumption 3.4 (Quasi-order) *The preference relation $\succsim$ for bf lotteries on $\mathscr{L}_{bf}$ is a quasi-order, i.e., it is reflexive and transitive.*

In contrast with the probabilistic case (Assumption 2.4), we do not assume that $\succsim$ is complete. There are many reasons we may not wish to assume completeness. It is not descriptive of human behavior. Even from a normative point of view, it is questionable that a DM has complete preferences on all possible lotteries. The assumption of incomplete preferences is consistent with the D-S theory of belief functions where we have non-singleton focal sets. Several authors, such as Aumann [1], and Dubra *et al.* [9] argue why the assumption of complete preferences may not be realistic in many circumstances.

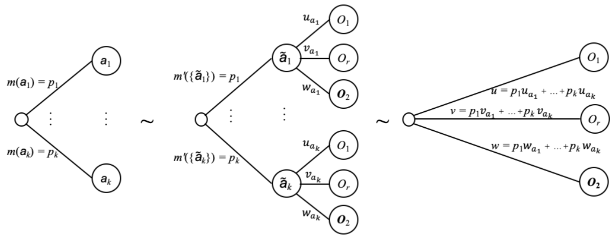

Figure 3: Reducing a bf lottery to a bf reference lottery

The substitutability assumption is similar to the probabilistic case (Assumption 2.5)—we replace an outcome in the probabilistic case by a focal set of m in the bf case.

Assumption 3.5 (Substitutability) *In any bf lottery $L = [\boldsymbol{O}, m]$, if we substitute a focal set a of m by an equally preferred bf reference lottery $\widetilde{\mathsf{a}} = [\boldsymbol{O}_2, m_{\mathsf{a}}]$, then the result is a compound lottery that is equally preferred to L.*

It follows from Assumptions 3.1–3.5 that given any bf lottery, we can reduce it to an equally preferred bf reference lottery. This is stated as Theorem 4 below.

Theorem 4 (Reducing a bf lottery) *Under Assumptions 3.1–3.5, any bf lottery $L = [\boldsymbol{O}, m]$ with focal sets $\mathsf{a}_1, \dots, \mathsf{a}_k$ is indifferent to a bf reference lottery $\widetilde{L} = [\boldsymbol{O}_2, \widetilde{m}]$, such that*

$$\widetilde{m}(\{O_1\}) = \sum_{i=1}^{k} m(\mathsf{a}_i)\, u_{\mathsf{a}_i}, \tag{10a}$$

$$\widetilde{m}(\{O_r\}) = \sum_{i=1}^{k} m(\mathsf{a}_i)\, v_{\mathsf{a}_i}, \quad \textit{and} \tag{10b}$$

$$\widetilde{m}(\boldsymbol{O}_2) = \sum_{i=1}^{k} m(\mathsf{a}_i)\, w_{\mathsf{a}_i}, \tag{10c}$$

where u_{a_i}, v_{a_i}, and w_{a_i}, are the masses assigned, respectively, to $\{O_1\}$, $\{O_r\}$, and $\boldsymbol{O}_2$, by the bf reference lottery $\widetilde{\mathsf{a}}_i$ equivalent to a_i.

Next, we formulate the monotonicity assumption. This is less obvious than it is in the probabilistic case (Assumption 2.6), as there are several ways in which intervals may be ordered. Assumption 3.6 below states that, given two bf reference lotteries L and L', the former will be preferred if and only if it assigns a higher degree of belief to the best consequence O_1, and a lower degree of belief to the worst consequence O_r.

Assumption 3.6 (Monotonicity) *Suppose $L = [\boldsymbol{O}_2, m]$ and $L' = [\boldsymbol{O}_2, m']$ are bf reference lotteries, with $m(\{O_1\}) = u$, $m(\boldsymbol{O}) = w$, $m'(\{O_1\}) = u'$, $m'(\boldsymbol{O}) = w'$. Then, $L \succsim L'$ if and only if $u \geq u'$ and $u + w \geq u' + w'$.*

Thus, $L \succsim L'$ if and only if $Bel_m(\{O_1\}) \geq Bel_{m'}(\{O_1\})$ and $Bel_m(\{O_r\}) \leq Bel_{m'}(\{O_r\})$, i.e., if and only if outcome O_1 is deemed both more credible and more plausible under L than it is if under L'. The corresponding indifference relation is $L \sim L'$ if and only if $u = u'$ and $w = w'$. It is clear that $\succsim$ as defined in Assumption 3.6 is reflexive and transitive. Also, the preference relation $\succsim$ on the set of all bf reference lotteries is obviously incomplete. Thus, two lotteries are *incomparable* if not $L \succsim L'$ and not $L' \succsim L$, i.e., if one of the intervals $[u, u+w]$ and $[u', u'+w']$ is strictly included in the other.

Assumptions 3.1, 3.3 and 3.6 imply the following consistency constraints between the reference bf lotteries equivalent to single outcomes:

$$1 = u_{\{O_1\}} \geq u_{\{O_2\}} \geq \ldots \geq u_{\{O_r\}} = 0. \tag{11}$$

Our final assumption has no counterpart in the vN-M theory. It states that a set a of outcomes is always at least as desirable as the worst outcome in a, and at most as desirable as the best outcome in a.

Assumption 3.7 (Consistency) *Let $\mathsf{a} \subseteq \boldsymbol{O}$, and let $\underline{O}_{\mathsf{a}}$ and $\overline{O}_{\mathsf{a}}$ denote, respectively, the worst and the best outcome in a. Then we have*

$$\mathsf{a} \succsim \underline{O}_{\mathsf{a}} \quad \textit{and} \quad \overline{O}_{\mathsf{a}} \succsim \mathsf{a}.$$

Assumptions 3.6 and 3.7 imply that, for any focal sets a of m, we have

$$u_{\mathsf{a}} \geq \min_{O_i \in \mathsf{a}} u_{\{O_i\}}, \quad \text{and} \quad u_{\mathsf{a}} + w_{\mathsf{a}} \leq \max_{O_i \in \mathsf{a}} u_{\{O_i\}}. \tag{12}$$

3.3. Representation theorem

Theorem 5 (Interval-valued utility function) *Suppose $L = [\boldsymbol{O}, m]$ and $L' = [\boldsymbol{O}, m']$ are bf lotteries on $\boldsymbol{O}$. If the preference relation $\succsim$ on $\mathscr{L}_{bf}$ satisfies Assumptions 3.1–3.6, then there are intervals $[u_{\mathsf{a}_i}, u_{\mathsf{a}_i} + w_{\mathsf{a}_i}]$ associated with subsets $\mathsf{a}_i \in 2^{\boldsymbol{O}}$ such that $L \succsim L'$ if and only if*

$$\sum_{\mathsf{a}_i \in 2^{\boldsymbol{O}}} m(\mathsf{a}_i)\, u_{\mathsf{a}_i} \geq \sum_{\mathsf{a}_i \in 2^{\boldsymbol{O}}} m'(\mathsf{a}_i)\, u_{\mathsf{a}_i} \tag{13a}$$

and

$$\sum_{\mathsf{a}_i \in 2^{\boldsymbol{O}}} m(\mathsf{a}_i)\,(u_{\mathsf{a}_i} + w_{\mathsf{a}_i}) \geq \sum_{\mathsf{a}_i \in 2^{\boldsymbol{O}}} m'(\mathsf{a}_i)\,(u_{\mathsf{a}_i} + w_{\mathsf{a}_i}). \tag{13b}$$

Thus, for a bf lottery $L = [\boldsymbol{O}, m]$, we can define

$$u(L) = [u, u + w] \tag{14}$$

as an interval-valued utility of L, with

$$u = \sum_{\mathsf{a}_i \in 2^{\boldsymbol{O}}} m(\mathsf{a}_i)\, u_{\mathsf{a}_i} \quad \textit{and} \quad w = \sum_{\mathsf{a}_i \in 2^{\boldsymbol{O}}} m(\mathsf{a}_i)\, w_{\mathsf{a}_i}. \tag{15}$$

Also, such a utility function is unique up to a strictly increasing affine transformation, i.e., if $u' = a\,u + b$, and $w' = a\,w + b$, where $a > 0$, and b are real constants, then

$$u'(L) = [u', u' + w']$$

also qualifies as an interval-valued utility function.

In the imprecise probability literature, we have lower and upper Choquet integrals as follows [15, 4]:

Definition 6 (Choquet integrals) *Suppose we have a real-valued function $u : \boldsymbol{O} \to \mathbb{R}$. The lower and upper Choquet integrals of u with respect to BPA m for $\boldsymbol{O}$, denoted by $\underline{u}_m$ and $\overline{u}_m$, are defined as follows:*

$$\underline{u}_m = \sum_{\mathsf{a} \in 2^{\boldsymbol{O}}} m(\mathsf{a}) \left(\min_{O_i \in \mathsf{a}} u(O_i) \right), \tag{16a}$$

$$\overline{u}_m = \sum_{\mathsf{a} \in 2^{\boldsymbol{O}}} m(\mathsf{a}) \left(\max_{O_i \in \mathsf{a}} u(O_i) \right). \tag{16b}$$

Thus, we can regard the interval $[\underline{u}_m, \overline{u}_m]$ as an interval-valued utility of bf lottery $[\mathbf{O}, m]$ as defined in the imprecise probability literature. It follows from Theorem 4 and Assumption 3.7 that

$$\underline{u}_m \leq u \leq u + w \leq \overline{u}_m, \tag{17}$$

where u and w are as in Eq. (15). Thus, the interval-valued utility of lottery $[\mathbf{O}, m]$ as defined in Theorem 5 is always included in the lower-upper expected utility interval. The lower and upper expectations defined by Eq. (16) can thus be seen as lower and upper bounds of the interval utility of a lottery $L = [\mathbf{O}, m]$ and could be used as conservative estimates if the equivalent bf reference lotteries $\widetilde{\mathsf{a}}_i$ cannot be elicited.

A special case of Theorem 5 is if we use Bayesian bf reference lotteries for the continuity assumption, i.e., $w_{\mathsf{a}} = 0$ for all focal sets a of m. In this case, Theorem 5 implies Corollary 7 below where we have a real-valued utility function, and consequently, a complete ordering on $\mathscr{L}_{bf}$.

Corollary 7 (Real-valued utility function) *Suppose $L = [\boldsymbol{O}, m]$ and $L' = [\boldsymbol{O}, m']$ are bf lotteries on $\boldsymbol{O}$. If the preference relation $\succsim$ on $\mathscr{L}_{bf}$ satisfies Assumptions 3.1–3.6 and if $w_{\mathsf{a}} = 0$ for all focal sets a of m and m', then there are numbers u_{a} associated with nonempty subsets $\mathsf{a} \subseteq \boldsymbol{O}$ such that $L_1 \succsim L_2$ if and only if*

$$\sum_{\mathsf{a} \in 2^{\boldsymbol{O}}} m(\mathsf{a})\, u_{\mathsf{a}} \geq \sum_{\mathsf{a} \in 2^{\boldsymbol{O}}} m'(\mathsf{a})\, u_{\mathsf{a}}.$$

Thus, for a bf lottery $L = [\boldsymbol{O}, m]$, we can define

$$u(L) = \sum_{\mathsf{a} \in 2^{\boldsymbol{O}}} m(\mathsf{a})\, u_{\mathsf{a}} \tag{18}$$

as the utility of L. Also, such a utility function is unique up to a strictly increasing affine transformation, i.e., if $u'_{\mathsf{a}} = a\, u_{\mathsf{a}} + b$, where $a > 0$, and b are real constants, then

$$u'(L) = \sum_{\mathsf{a} \in 2^{\boldsymbol{O}}} m(\mathsf{a})\, u'_{\mathsf{a}}$$

also qualifies as a utility function.

The utility function in Eq. (18) has exactly the same form as Jaffray's linear utility [20]. This is discussed further in Section 5.1.

4. Examples

In this section, we illustrate the application of Theorem 5 to three examples: Ellsberg's urn problem described in Example 1, the one red ball problem described in [22], and the 1,000 balls urns described in [2].

Example 3 (Ellsberg's urn) *Consider the four bf lotteries described in Example 1. Given a vacuous bf lottery $[\{\$100, \$0\}, m(\{\$100, \$0\}) = 1]$, what is an indifferent bf reference lottery? For an ambiguity-averse DM,*

$$[\{\$100, \$0\}, (1/2, 1/2)] \succ [\{\$100, \$0\}, m(\{\$100, \$0\}) = 1].$$

For such a DM, we must have $u_{\{\$100,\$0\}} + w_{\{\$100,\$0\}} < 1/2$.

For the first choice problem between L_1 (\$100 on r), and L_2 (\$100 on b), using Eq. (14), $u(L_1) = 1/3$, and

$$u(L_2) = \frac{2}{3}\left[u_{\{\$100,\$0\}}, u_{\{\$100,\$0\}} + w_{\{\$100,\$0\}}\right].$$

Thus, an ambiguity-averse DM would choose L_1. This result is valid as long as $u_{\{\$100,\$0\}} + w_{\{\$100,\$0\}} < 1/2$ and is consistent with Ellsberg's findings. For the second choice problem between L_3 (\$100 on r or y), and L_4 (\$100 on b or y),

$$u(L_3) = \frac{1}{3}(1) + \frac{2}{3}\left[u_{\{\$100,\$0\}}, u_{\{\$100,\$0\}} + w_{\{\$100,\$0\}}\right],$$

and $u(L_4) = 2/3$. An ambiguity-averse DM would choose L_4, as

$$\frac{1}{3} + \frac{2}{3} u_{\{\$100,\$0\}} + w_{\{\$100,\$0\}} < \frac{2}{3},$$

as long as $u_{\{\$100,\$0\}} + w_{\{\$100,\$0\}} < 1/2$, a result that is also consistent with Ellsberg's empirical findings.

Example 4 (One red ball) *Consider the following example called 'one red ball' in [22]. An urn possibly contains balls of six colors: red (r), blue (b), green (g), orange (o), white (w), and yellow (y). One ball is drawn at random from the urn. We are informed that the urn has a total of n balls, where n is a positive integer, and that there is exactly one red ball in the urn. Suppose random variable X denotes the color of the ball drawn from the urn. Then $\Omega_X = \{r, b, g, o, w, y\}$, and m_X is a BPA for X such that $m_X(\{r\}) = 1/n$, and $m_X(\{b, g, o, w, y\}) = (n-1)/n$. First, you pick a color, and then you draw a ball at random from the urn. You win \$100 if the color of the ball drawn from the urn matches the color you picked and you win \$0 if it doesn't. What color do you pick? In [22], the authors describe some informal experiments where all respondents chose red for $n \leq 7$, and for $n \geq 8$, several respondents preferred a color different from red.*

Suppose you pick r. The bf lottery L_r based on m_X is as follows: $[\{\$100, \$0\}, m_r]$, where $m_r(\{\$100\}) = 1/n$, and $m_r(\{\$0\}) = (n-1)/n$. If the color you pick is b, then the bf lottery L_b is $[\{\$100, \$0\}, m_b]$, where $m_b(\{\$0\}) = 1/n$, and $m_b(\{\$100, \$0\}) = (n-1)/n$. Thus, we have $u(L_r) = 1/n$, and

$$u(L_b) = \frac{n-1}{n}\left[u_{\{\$100,\$0\}}, u_{\{\$100,\$0\}} + w_{\{\$100,\$0\}}\right].$$

So, L_b is strictly preferred to L_r whenever

$$\frac{n-1}{n} u_{\{\$100,\$0\}} > \frac{1}{n},$$

i.e., whenever $u_{\{\$100,\$0\}} > 1/(n-1)$, and L_r is strictly preferred to L_b whenever

$$\frac{n-1}{n}\left(u_{\{\$100,\$0\}} + w_{\{\$100,\$0\}}\right) < \frac{1}{n},$$

i.e., whenever $u_{\{\$100,\$0\}} + w_{\{\$100,\$0\}} < 1/(n-1)$. Hence, L_b is increasingly preferred to L_r when n increases, which is consistent with the findings reported in [22]. In our model, when

$$u_{\{\$100,\$0\}} < \frac{1}{n-1} < u_{\{\$100,\$0\}} + w_{\{\$100,\$0\}},$$

the two lotteries L_r and L_b are incomparable. If forced to choose, the DM might just choose arbitrarily. As the experiment reported in [22] did not allow the respondents to express inability to choose between the two lotteries, it does not provide any evidence for or against our model.

Example 5 (Urns with 1,000 balls) *The following example is discussed in [2], where it is credited to Ellsberg in an oral conversation (with the authors of [2]). It is also discussed in [10]. There are two urns, each with 1,000 balls, numbered from $1-1{,}000$. Urn 1 has exactly one ball for each number, and there is no ambiguity. Urn 2 has unknown number of balls of each number, and there is much ambiguity. One ball is to be chosen at random from an urn of your choosing. If the number on the ball matches a specific number, e.g., 687, you win \$100, and if not, you win nothing (\$0). Which one of the two urns will you choose?*

It is reported in [2] that many respondents chose Urn 2. Why? Urn 1 has only one ball numbered 687, and therefore, the probability of winning \$100 if the choice is Urn 1 is very small, 0.001. Urn 2 could possibly have anywhere from 0 to 1,000 balls numbered 687. Thus, the choice of Urn 2, although ambiguous, is appealing. Let's analyze this problem using Theorem 5.

Let X_1 denote the number on the ball chosen from Urn 1, and let X_2 denote the number on the ball chosen from Urn 2. $\Omega_{X_1} = \Omega_{X_2} = \{1,\ldots,1000\}$. Function m_{X_1} is a BPA for X_1 as follows: $m_{X_1}(\{1\}) = \ldots = m_{X_1}(\{1000\}) = 0.001$. BPA m_{X_2} is vacuous, i.e., $m_{X_2}(\Omega_{X_2}) = 1$.

Lottery L_1 corresponding to choice of Urn 1 (say, alternative f_1) is $[\{\$100,\$0\}, m_1]$, where m_1 is a BPA for $\{\$100,\$0\}$ such that $m_1(\{\$100\}) = 0.001$, and $m_1(\{\$0\}) = 0.999$. L_1 is a bf reference lottery, and thus, $u(L_1) = 0.001$. Lottery L_2 corresponding to choice of Urn 2 (say, alternative f_2) is $[\{\$100,\$0\}, m_2]$, where m_2 is a vacuous BPA for $\{\$100,\$0\}$. The utility of L_2 is

$$u(L_2) = \left[u_{\{\$100,\$0\}}, u_{\{\$100,\$0\}} + w_{\{\$100,\$0\}}\right].$$

Consequently, L_2 is preferred to L_1 as long as

$$u_{\{\$100,\$0\}} \geq 0.001,$$

a condition that is easily satisfied. This may explain why many DMs prefer to be ambiguity-seeking in this context, i.e., prefer L_2 to L_1.

5. Comparison

In this section, we compare our utility theory to Jaffray's linear utility theory [20], and to Smets' two-level decision theory [29].

5.1. Comparison with Jaffray's Axiomatic Theory

Jaffray's axiomatic theory is based on considering the set of all belief functions for $\mathbf{O}$ as a mixture set as follows. Suppose m_1 and m_2 are BPAs for $\mathbf{O}$, and suppose $\lambda \in [0,1]$. Then m defined as:

$$m(\mathsf{a}) = \lambda\, m_1(\mathsf{a}) + (1-\lambda) m_2(\mathsf{a}) \tag{19}$$

for all $\mathsf{a} \in 2^{\mathbf{O}}$, is a BPA for $\mathbf{O}$. BPA m can be written as $m = \lambda\, m_1 + (1-\lambda) m_2$, and called a *mixture* of m_1 and m_2. Using the Jensen-version [21] of vN-M axiom system, Jaffray uses the following assumptions, all of which are expressed using mixture BPA functions:

Assumption 5.1 (Completeness and transitivity) *The relation $\succsim$ is complete and transitive over $\mathscr{L}_{bf}$.*

Assumption 5.2 (Independence) *For all $L_1 = [\boldsymbol{O}, m_1]$ and $L_2 = [\boldsymbol{O}, m_2]$ in $\mathscr{L}_{bf}$, and $\lambda \in (0,1)$, $L_1 \succ L_2$ implies $[\boldsymbol{O}, \lambda\, m_1 + (1-\lambda) m] \succ [\boldsymbol{O}, \lambda\, m_2 + (1-\lambda) m]$.*

Assumption 5.3 (Continuity) *For all $L_1 = [\boldsymbol{O}, m_1]$, $L_2 = [\boldsymbol{O}, m_2]$, and $L_3 = [\boldsymbol{O}, m_3]$ in $\mathscr{L}_{bf}$ such that $L_1 \succ L_2 \succ L_3$, there exists λ and μ in $(0,1)$ such that*

$$[\boldsymbol{O}, \lambda\, m_1 + (1-\lambda) m_3] \succ [\boldsymbol{O}, m_2] \succ [\boldsymbol{O}, \mu\, m_1 + (1-\mu) m_3].$$

Theorem 8 (Jaffray's representation theorem [20]) *The preference relation $\succsim$ on $\mathscr{L}_{bf}$ satisfies Assumptions 5.1–5.3 if and only if there exists a utility function $u : \mathscr{L}_{bf} \to \mathbb{R}$ such that for any lottery $L = [\boldsymbol{O}, m]$ in $\mathscr{L}_{bf}$,*

$$u(L) = \sum_{\mathsf{a} \in 2^{\boldsymbol{O}}} m(\mathsf{a})\, u_{\mathsf{a}}, \tag{20}$$

where $u_{\mathsf{a}} = u([\boldsymbol{O}, m_{\mathsf{a}}^d])$, and m_{a}^d is a deterministic BPA for $\boldsymbol{O}$ such that $m_{\mathsf{a}}^d(\mathsf{a}) = 1$.

Thus, Jaffray's axioms result in the same solution as that of Corollary 7, which is a special case of Theorem 5. As Jaffray's axioms do not use Dempster's rule explicitly, it is not clear whether Eq. (20) applies to the D-S framework or not. The mixture BPA m derived from BPAs m_1 and m_2 using Eq. (19) is not Dempster's combination rule, although Eq. (19) can be derived from a belief function model using Dempster's rule. By deriving this solution from a set of axioms making use of the basic constructs of DS theory (namely, Dempster's combination rule, marginalization, and conditional embedding), we provide additional arguments supporting Eq. (20) as a natural definition of the real-valued utility of a bf lottery in the D-S theory.

Also, there is no explicit notion of a bf reference lottery in Jaffray's framework. Thanks to our continuity axiom (Assumption 3.3), the interval-valued utility $[u_{\mathsf{a}}, u_{\mathsf{a}} + w_{\mathsf{a}}]$ in our framework receives a simple interpretation as an interval-valued probability of a best outcome O_1, in a bf reference lottery $[\mathbf{O}_2, m_{\mathsf{a}}]$ that is indifferent to a and such that $m_{\mathsf{a}}(\{O_1\}) = u_{\mathsf{a}}$, $m_{\mathsf{a}}(\{O_r\}) = 1 - (u_{\mathsf{a}} + w_{\mathsf{a}})$, and $m_{\mathsf{a}}(\mathbf{O}_2) = w_{\mathsf{a}}$. We believe that this simple interpretation can be very helpful when eliciting utilities from DMs.

5.2. Comparison with Smets' Decision Theory

Smets' decision theory [29] is a two-level framework where beliefs, represented by belief functions, are held at a credal level. When a DM has to make a decision, the marginal belief function for a variable of interest is transformed into a PMF, and the Bayesian expected utility framework is then used to make a decision.

Smets uses a transformation called the *pignistic* transform to transform belief functions into PMFs. This transform is justified in [30] using a mixture property as follows. Let T denote the belief-PMF transformation. Smets [30] argues that this transformation should be linear, i.e., we should have, for any $\lambda \in [0,1]$,

$$T(\lambda\, m_1 + (1-\lambda) m_2) = \lambda\, T(m_1) + (1-\lambda) T(m_2). \tag{21}$$

The unique transformation T verifying (21) is the pignistic transformation defined as $T(m) = BetP_m$ with

$$BetP_m(O) = \sum_{\mathsf{a} \subseteq \mathbf{O}} \frac{m(\mathsf{a})}{|\mathsf{a}|} I(O \in \mathsf{a}) \tag{22}$$

for all $O \in \mathbf{O}$, where $I(\cdot)$ is the indicator function. The pignistic PMF $BetP_m$ is mathematically identical to the Shapley value in cooperative game theory [28]. In [30], Smets attempts to derive Eq. (21) from the maximum expected utility principle. The argument, however, is quite technical and unconvincing.

Given the definition in Eq. (22), the expected utility of a bf lottery $L = [\mathbf{O}, m]$ according to the pignistic PMF is

$$u_{BetP}(L) = \sum_{O \in \mathbf{O}} BetP_m(O)\, u_{\{O\}} \tag{23a}$$

$$= \sum_{O \in \mathbf{O}} \left(\sum_{\mathsf{a} \subseteq \mathbf{O}} \frac{m(\mathsf{a})}{|\mathsf{a}|} I(O \in \mathsf{a}) \right) u_{\{O\}} \tag{23b}$$

$$= \sum_{\mathsf{a} \subseteq \mathbf{O}} m(\mathsf{a}) \left(\frac{1}{|\mathsf{a}|} \sum_{O \in \mathsf{a}} u_{\{O\}} \right). \tag{23c}$$

It is a special case of Eq. (18), with

$$u_{\mathsf{a}} = \frac{1}{|\mathsf{a}|} \sum_{O \in \mathsf{a}} u_{\{O\}}.$$

Smets' decision theory thus amounts to assuming that a DM is indifferent between a bf lottery that gives them an outcome in a for sure, and a bf reference lottery in which the probability of the best outcome is equal to the average utilities of the outcomes in a. This is consistent with our Assumptions 3.1–3.6, but it is inconsistent with Assumption 3.7. Also, this restricted model does not have a parameter to represent a DM's attitude toward ambiguity. As a result, it is unable to explain Ellsberg's paradox and the ambiguity aversion of human DMs as described, e.g., in the examples in Section 4.

6. Summary and Conclusions

In this section, we summarize our proposal and sketch some future work. We start with Luce and Raiffa's version of the vN-M utility theory for probabilistic lotteries. We then consider bf lotteries, lotteries when our beliefs about the state of the world is described by DS belief functions. We use a similar set of axioms as vN-M, but first we replace each singleton outcome in a probabilistic lottery by a focal set of a BPA. Second, we replace the reduction of compound lotteries with a corresponding axiom that uses Dempster's combination rule and belief function marginalization in place of probabilistic combination (pointwise multiplication followed by normalization) and probabilistic marginalization (addition). Third, we use a bf reference lottery with two independent parameters. The axioms lead to a decision theory that involves assessing the utility of each focal element of a BPA as an interval-valued utility. Interval-valued utilities lead to a partial preference relation on the set $\mathscr{L}_{bf}$ of all bf lotteries. If we use Bayesian bf reference lotteries with a single parameter, then our axiomatic framework leads to a real-valued utility function that is exactly the same as in Jaffray's linear utility theory [20].

The decision theory that results from our axioms is more general than that proposed by Jaffray [20], which can be construed as a decision theory for belief functions interpreted as generalized probabilities. Jaffray's axiomatic theory is based on a set of mixture BPAs. A mixture of two BPAs is not the same as a Dempster's combination of two BPAs, although we could construct a belief function model where the mixture BPA is obtained by Dempster's rule. Thus, it is not clear if Jaffray's linear utility theory is applicable to D-S belief function lotteries or not. Our utility theory confirms that this is indeed the case. Our bf reference lotteries lead to interval-valued utilities, and consequently, a partial preference relation on the set of all bf lotteries.

We also compare our axiomatic theory to Smets' two-level framework [29, 30], and note that his framework is too constrained to explain ambiguity-aversion or ambiguity-seeking behavior of human DMs.

In practice, implementing the most general form of our axiomatic theory may need assessment of $2k$ parameters, where k is the number of focal sets of a bf lottery. In the worst case, k can be as large as $2^{|\mathbf{O}|} - 1$. In [7], based on additional assumptions, we propose a model based on only two parameters, which can be interpreted as reflecting both the DM's attitude to ambiguity and their indeterminacy. This model, as well as others, will have to be further studied and developed. More generally, a rigorous methodology to elicit interval-valued utilities remains to be designed.

Acknowledgments

This research was done during a visit by the second author to Université de Technologie de Compiègne (UTC) in Spring 2019. This research was funded partially by a sabbatical from the University of Kansas, and supported by the Labex MS2T, which is funded by the French Government through the program "Investments for the future" by the National Agency for Research (reference ANR-11-IDEX-0004-02). Thanks to Sébastien Destercke for many discussions on the topic of this manuscript, and to Suzanna Emelio for proofreading this manuscript.

References

[1] R. Aumann. Utility theory without the completeness axiom. *Econometrica*, 30(3):445–462, 1962.

[2] S. W Becker and F. O. Brownson. What price ambiguity? or the role of ambiguity in decision-making. *Journal of Political Economy*, 72(1):62–73, February 1964.

[3] G. Choquet. Theory of capacities. *Annales de l'Institut Fourier*, 5:131–295, 1953.

[4] G. Coletti, D. Petturiti, and B. Vantaggi. Models for pessimistic or optimistic decisions under different uncertain scenarios. *International Journal of Approximate Reasoning*, 105(2):305–326, 2019.

[5] A. P. Dempster. Upper and lower probabilities induced by a multivalued mapping. *Annals of Mathematical Statistics*, 38(2):325–339, 1967.

[6] T. Denoeux. Decision-making with belief functions: A review. *International Journal of Approximate Reasoning*, 109(6):87–110, 2019.

[7] T. Denoeux and P. P. Shenoy. An interval-valued utility theory for decision making with Dempster-Shafer belief functions. Working Paper 336, School of Business University of Kansas, Lawrence, KS, May 2019. URL http://pshenoy.faculty.ku.edu/Papers/WP336.pdf.

[8] T. Denoeux, D. Dubois, and H. Prade. Representations of uncertainty in artificial intelligence: Beyond probability and possibility. In P. Marquis, O. Papini, and H. Prade, editors, *A Guided Tour of Artificial Intelligence Research*, chapter 4. Springer Verlag, 2019.

[9] J. Dubra, F. Maccheroni, and E.A. Ok. Expected utility theory without the completeness axiom. *Journal of Economic Theory*, 115(1):118–133, 2004.

[10] H. J. Einhorn and R. M. Hogarth. Decision making under ambiguity. *Journal of Business*, 59(4):S225–S250, 1986.

[11] D. Ellsberg. Risk, ambiguity and the Savage axioms. *Quarterly Journal of Economics*, 75(2):643–669, 1961.

[12] R. Fagin and J. Y. Halpern. A new approach to updating beliefs. In P.P. Bonissone, M. Henrion, L.N. Kanal, and J.F. Lemmer, editors, *Uncertainty in Artificial Intelligence 6*, pages 347–374. North-Holland, Amsterdam, 1991.

[13] P. C. Fishburn. *The Foundations of Expected Utility*. Reidel, Dordrecht, 1982.

[14] T. Gajdos, T. Hayashi, J.-M. Tallon, and J.-C. Vergnaud. Attitude toward imprecise information. *Journal of Economic Theory*, 140(1):23–56, 2008.

[15] I. Gilboa and D. Schmeidler. Maxmin expected utility with a non-unique prior. *Journal of Mathematical Economics*, 18(2):141–153, 1989.

[16] I. Gilboa and D. Schmeidler. Additive representations of non-additive measures and the Choquet integral. *Annals of Operations Research*, 52:43–65, 1994.

[17] J. Y. Halpern and R. Fagin. Two views of belief: Belief as generalized probability and belief as evidence. *Artificial Intelligence*, 54(3):275–317, 1992.

[18] M. Hausner. Multidimensional utilities. In R. M. Thrall, C. H. Coombs, and R. L. Davis, editors, *Decision Processes*, pages 167–180. Wiley, New York, NY, 1954.

[19] I. N. Herstein and J. Milnor. An axiomatic approach to measurable utility. *Econometrica*, 21(2):291–297, 1953.

[20] J.-Y. Jaffray. Linear utility theory for belief functions. *Operations Research Letters*, 8(2):107–112, 1989.

[21] N. E. Jensen. An introduction to Bernoullian utility theory: A discussion. *Swedish Journal of Economics*, 69(3):163–183, 1967.

[22] R. Jiroušek and P. P. Shenoy. Ambiguity aversion and a decision-theoretic framework using belief functions. In *2017 IEEE Symposium Series on Computational Intelligence (SSCI) Proceedings*, pages 326–332, Piscataway, NJ, 2017. IEEE.

[23] R. D. Luce and H. Raiffa. *Games and Decisions: Introduction and Critical Survey*. Wiley, New York, NY, 1957.

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

[25] G. Shafer. Perspectives on the theory and practice of belief functions. *International Journal of Approximate Reasoning*, 4(5–6):323–362, 1990.

[26] G. Shafer. Rejoinders to comments on "Perspectives on the theory and practice of belief functions". *International Journal of Approximate Reasoning*, 6(3): 445–480, 1992.

[27] G. Shafer. Constructive decision theory. *International Journal of Approximate Reasoning*, 79(12):45–62, 2016.

[28] L. S. Shapley. A value for *n*-person games. In H. Kuhn and A. W. Tucker, editors, *Contributions to the Theory of Games*, volume 2, pages 307–317. Princeton University Press, Princeton, NJ, 1953.

[29] P. Smets. Decision making in a context where uncertainty is represented by belief functions. In R. P Srivastava and T. J. Mock, editors, *Belief Functions in Business Decisions*, pages 316–332. Physica-Verlag, Heidelberg, 2002.

[30] P. Smets. Decision making in the TBM: The necessity of the pignistic transformation. *International Journal of Approximate Reasoning*, 38(2):133–147, 2005.

[31] John von Neumann and Oskar Morgenstern. *Theory of Games and Economic Behavior*. Princeton University Press, Princeton, NJ, 2nd edition, 1947.

[32] P. P. Wakker. Dempster belief functions are based on the principle of complete ignorance. *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems*, 8(3):271–284, 2000.