Theory Dec. (2016) 81:153–165
DOI 10.1007/s11238-016-9534-3

Peter Fishburn's analysis of ambiguity

Mark Shattuck[1] · **Carl Wagner**[1]

Published online: 22 January 2016

Abstract In ordinary discourse the term ambiguity typically refers to vagueness or imprecision in a natural language. Among decision theorists, however, this term usually refers to imprecision in an individual's probabilistic judgments, in the sense that the available evidence is consistent with more than one probability distribution over possible states of the world. Avoiding a prior commitment to either of these interpretations, Fishburn has explored ambiguity as a primitive concept, in terms of what he calls an ambiguity measure a on the power set 2^{Ω} of a finite set Ω, characterized by five axioms. We prove, in purely set-theoretic terms, that if λ is a so-called necessity measure on 2^{Ω} and υ is its associated possibility measure, then $a = \upsilon - \lambda$ is an ambiguity measure. When Ω is construed as a set of possible exemplars of a vague predicate ϕ, then λ and υ may be regarded as arising from a fuzzy membership function f on Ω, where $f(\omega)$ designates the degree to which ϕ is applicable to ω. In this case $a(A)$ represents the degree to which the partition $\{A, A^c\}$ differentiates members of Ω with respect to the predicate ϕ. When Ω is construed as a set of possible states of the world, a necessity measure may be regarded as a very special type of lower probability known as a consonant belief function, and a possibility measure as its associated upper probability, whence $a(A)$ represents the degree of imprecision in the pair (λ, υ) with respect to the event A. Fishburn's axioms are thus consistent with an interpretation of ambiguity as linguistic vagueness, as well as (a very special sort of) probabilistic imprecision.

In Memoriam Patrick Suppes, 1922–2014.

✉ Carl Wagner
cwagner@tennessee.edu

Mark Shattuck
shattuck@math.utk.edu

1 Department of Mathematics, The University of Tennessee, Knoxville, TN 37996-1320, USA

Keywords Ambiguity · Belief function · Imprecise probability · Necessity measure · Possibility measure

1 Introduction

In ordinary discourse the term ambiguity typically refers to vagueness or imprecision in a word or phrase of some natural language. Since its introduction in the now classic paper by Ellsberg (1961), however, this term has been used by most decision theorists to refer to imprecision in an individual's probabilistic judgments, in the sense that the available evidence is consistent with more than one probability distribution over the possible states of the world. Abstaining from a prior commitment to either of these interpretations, Fishburn (1991) has explored ambiguity as a primitive concept, through an analysis of mappings $a : 2^{\Omega} \to [0, \infty)$, where Ω is a finite set of possible states of the world, and $a(A)$ denotes what he calls the ambiguity measure of A, characterized by the following axioms :

$$a(\emptyset) = a(\Omega) = 0. \tag{1.1}$$

$$a(A) = a(A^c). \tag{1.2}$$

$$a(A \cup B) \leq a(A) + a(B) - a(A \cap B). \tag{1.3}$$

$$a(A) = a(B) \Rightarrow a(A - B) = a(B - A) \quad \text{or} \quad a(A \cap B) = a(A \cup B). \tag{1.4}$$

$$a(A) > a(B) \Rightarrow a(A - B) > a(B - A) \quad \text{or} \quad a(A \cap B) > a(A \cup B). \tag{1.5}$$

These axioms arise from a subtle measurement-theoretic analysis of the binary relation $\gg$ on 2^{Ω}, where $A \gg B$ asserts that the event A is at least as ambiguous as the event B. Fishburn states and justifies six axioms for this relation and proves a representation theorem to the effect that $\gg$ satisfies these axioms if and only if there exists a mapping $a : 2^{\Omega} \to [0, \infty)$ such that $(i.)$ $A \gg B \Leftrightarrow a(A) \geq a(B)$ and $(ii.)$ axioms (1.1)–(1.5) hold.

In this paper, we investigate the extent to which axioms (1.1)–(1.5) are consistent with the notion of ambiguity as linguistic vagueness, as contrasted with probabilistic imprecision. The paper is structured as follows: In Sect. 2 we outline pertinent results from the theory of lower and upper probabilities, conceived, following Smith (1961) and Walley (1991), as threshold buying and selling prices for certain bets. We explain why it is desirable that an individual's lower probability λ and upper probability υ should coincide, respectively, with the lower envelope and upper envelope of some nonempty set $\mathbf{P}$ of probability measures[1] on 2^{Ω}, in the sense that

$$\lambda(A) = \inf\{p(A) : p \in \mathbf{P}\} \quad \text{and} \quad \upsilon(A) = \sup\{p(A) : p \in \mathbf{P}\}, \quad \text{for all } A \subseteq \Omega. \tag{1.6}$$

[1] Such a set of probability measures arises in the context of the Ellsberg (1961) urn problem, in which a ball is to be picked at random from an urn containing 30 red balls and 60 black or yellow balls in unknown proportion. Due to the ambiguity as to which probability measure on $\Omega = \{\text{red}, \text{black}, \text{yellow}\}$ is operative, subjects choosing between two possible acts in one decision problem, and two possible acts in a second decision problem, typically make choices that violate the so-called sure thing principle postulated by Savage (1972) as an axiom of rational decision making.

If λ and υ are an individual's lower and upper probabilities on 2^{Ω}, the mapping $a = \upsilon - \lambda$, which Walley calls the degree of imprecision of the pair (λ, υ), and which is in essence a bid-ask spread, is a natural candidate for an ambiguity measure. We prove, however, that if λ and υ are lower and upper envelopes, the mapping a, so defined, may fail to satisfy axiom (1.3), although it always satisfies axioms (1.1) and (1.2) . We further show that axiom (1.4) may fail to hold even when λ is a highly structured type of lower envelope known as a belief function and υ is its corresponding upper envelope, known as a plausibility function (Shafer 1976), and that the same is true for axiom (1.5). We conclude by describing the special class of consonant belief functions, which reappear in a different guise in Sect. 3, where we consider the numerical representation of linguistic vagueness by means of fuzzy membership functions and their associated necessity and possibility measures (Dubois et al. 2000). We show that the class of necessity measures is identical to the class of consonant belief functions, and we argue that if λ and υ are necessity and possibility measures on 2^{Ω}, then $a = \upsilon - \lambda$ is a natural candidate for a measure of linguistic vagueness. In Sect. 4, we prove that $\upsilon - \lambda$ satisfies the complete set of axioms (1.1)–(1.5) when λ is a necessity measure and υ is its corresponding possibility measure. In Sect. 5 we offer a brief summary and conclusion.

2 Probabilistic imprecision

2.1 Lower and upper subjective probabilities

Suppose, as above, that Ω is a finite set of possible states of the world, and that λ and υ are mappings from 2^{Ω} to [0, 1]. Following Smith (1961) and Walley (1981, 1991), we say that λ is your[2] lower subjective probability and υ is your upper subjective probability on 2^{Ω} if, for all $A \subseteq \Omega$,

(L) You are willing to pay any amount strictly less than $\lambda(A)$ units of linear utility (but no more than $\lambda(A)$ units) in exchange for receiving 1 such unit if the event A occurs (i.e., if the true state of the world turns out to be a member of A), and nothing otherwise. (You may or may not be willing to pay exactly $\lambda(A)$ units.)

(U) In exchange for receiving any amount strictly greater than $\upsilon(A)$ units of linear utility (but no less than $\upsilon(A)$ units), you are willing to obligate yourself to pay 1 such unit if A occurs, and nothing otherwise. (You may or may not be willing to so obligate yourself in exchange for receiving exactly $\upsilon(A)$ units.)

In the tradition of de Finetti's (1974) well-known account of subjective probability, there are no a priori structural constraints on lower or upper probabilities. It is only assumed that you announce your prices $\lambda(A)$ and $\upsilon(A)$ for all $A \subseteq \Omega$ and that you are willing to accept multiple bets on the same event. Any structural restrictions on upper and lower subjective probabilities materialize as rationality constraints, guaranteed to protect you against certain undesirable consequences of your announced buying and

[2] Our use of the second person here echoes de Finetti (1974) and Walley (1991), who explains that it is employed " to encourage you (the reader) to consider the theory as a model for your own beliefs and behaviour."

selling prices. For example, it would clearly be undesirable for your betting commitments, as expressed by λ and υ, to place you in a position of sustaining a sure loss, i.e., a net loss of utility, regardless of which state of the world turns out to be the true state. Similarly, it would be undesirable for those commitments to be incoherent, in the sense that you refuse to make a bet with outcomes uniformly better than that of a bet that you are willing to make. The following theorem is thus of substantial interest.

Theorem 2.1 *You are protected against a sure loss and also against incoherence if and only if your lower and upper probabilities λ and υ coincide, respectively, with the lower and upper envelopes of some nonempty set of probability measures on 2^{Ω}.*

Proof See Walley (1981, Theorem 2.1). □

In addition to protecting you against sure loss and incoherence, lower and upper envelopes have many pleasant structural features.

Theorem 2.2 *If λ is the lower envelope and υ is the upper envelope of a nonempty set of probability measures on 2^{Ω}, as defined by (1.6), then*

$$\lambda(\emptyset) = \upsilon(\emptyset) = 0 \quad \text{and} \quad \lambda(\Omega) = \upsilon(\Omega) = 1. \tag{2.1}$$

$$\lambda(A) \leq \upsilon(A) \quad \text{for all} \quad A \subseteq \Omega. \tag{2.2}$$

$$A \subseteq B \Rightarrow \lambda(A) \leq \lambda(B) \quad \text{and} \quad \upsilon(A) \leq \upsilon(B). \tag{2.3}$$

$$A \cap B = \emptyset \Rightarrow \lambda(A \cup B) \geq \lambda(A) + \lambda(B). \tag{2.4}$$

$$A \cap B = \emptyset \Rightarrow \upsilon(A \cup B) \leq \upsilon(A) + \upsilon(B). \tag{2.5}$$

Moreover, λ and υ are conjugates, in the sense that

$$\lambda(A) + \upsilon(A^c) = 1 \quad \textit{for all} \quad A \subseteq \Omega. \tag{2.6}$$

Proof Straightforward. □

When probabilistic imprecision is represented by the lower and upper envelopes λ and υ, then $a := \upsilon - \lambda$, which Walley (1991) calls the degree of imprecision of the pair (λ, υ), is a natural candidate for an ambiguity measure. The following result is thus somewhat disappointing.

Theorem 2.3 *If λ and υ are lower and upper envelopes, then $a = \upsilon - \lambda$ satisfies axioms (1.1) and (1.2), but may fail to satisfy axiom (1.3).*

Proof The proof of axiom (1.1) is straightforward, and axiom (1.2) follows from (2.6). The following example, constructed by Suppes (1974) for a different purpose, shows that a may fail to satisfy axiom (1.3): Let $\Omega = \{1, 2, 3, 4\}$ and $\mathbf{P} = \{p_1, p_2\}$, where the densities of these probability measures are, respectively, (0.25, 0.25, 0.25, 0.25) and (0.5, 0.5, 0, 0). If $A = \{1, 3\}$ and $B = \{1, 4\}$, then $a(A \cup B) = \upsilon(A \cup B) - \lambda(A \cup B) = 0.75 - 0.5 = 0.25$, $a(A) = \upsilon(A) - \lambda(A) = 0.5 - 0.5 = 0$, $a(B) = \upsilon(B) - \lambda(B) = 0.5 - 0.5 = 0$, and $a(A \cap B) = \upsilon(A \cap B) - \lambda(A \cap B) = 0.5 - 0.25 = .25$, whence $a(A \cup B) > a(A) + a(B) - a(A \cap B)$. □

Axiom (1.3) does hold, however, under slightly stronger conditions on the pair (λ, υ). Suppose that $\lambda : 2^{\Omega} \to [0,1]$ is any set function satisfying $\lambda(\emptyset) = 0$ and $\lambda(\Omega) = 1$. If, in addition, $\lambda(A \cup B) \geq \lambda(A) + \lambda(B) - \lambda(A \cap B)$, for all $A, B \subseteq \Omega$, then λ is said to be 2-monotone (or convex). The following theorem, due to the Nobel laureate (Economics, 2012) Lloyd Shapley, originated in the study of cooperative game theory.

Theorem 2.4 *If $\lambda : 2^{\Omega} \to [0,1]$ is 2-monotone, and $\upsilon(A) := 1 - \lambda(A^c)$ for all $A \subseteq \Omega$, then there exists a nonempty set* **P** *of probability measures on 2^{Ω} such that λ is the lower envelope, and υ is the upper envelope of* **P**.

Proof See Shapley (1971). □

Theorem 2.5 *Suppose that λ is 2-monotone, and υ is defined as in Theorem 2.4 above. If $a = \upsilon - \lambda$, then a satisfies Fishburn's axioms (1.1), (1.2), and (1.3).*

Proof Since, by Theorem 2.4, λ and υ are, respectively, lower and upper envelopes, it follows from Theorem 2.3 that a satisfies axioms 1.1 and 1.2.
From (i) $\lambda(A \cup B) \geq \lambda(A) + \lambda(B) - \lambda(A \cap B)$ and the definition of υ, it follows that (ii) $\upsilon(A \cup B) \leq \upsilon(A) + \upsilon(B) - \upsilon(A \cap B)$. Subtracting inequality (i) from inequality (ii) yields the desired result. □

Axioms 1.4 and 1.5 are, however, considerably more recalcitrant. Indeed, as we show in the following subsection, even a substantial strengthening of 2-monotonicity is insufficient to ensure that these axioms hold.

2.2 Belief functions

Suppose that $\lambda : 2^{\Omega} \to [0,1]$, with $\lambda(\emptyset) = 0$ and $\lambda(\Omega) = 1$, and that $t \geq 2$. If, for all $A_1, \ldots, A_t \subseteq \Omega$,

$$\lambda(A_1 \cup \cdots \cup A_t) \geq \sum_{1 \leq i \leq t} \lambda(A_i) - \sum_{1 \leq i < j \leq t} \lambda(A_i \cap A_j) + \sum_{1 \leq i < j < k \leq t} \lambda(A_i \cap A_j \cap A_k) - \cdots + (-1)^{t-1} \lambda(A_1 \cap \cdots \cap A_t), \tag{2.7}$$

λ is said to be t-monotone. Clearly, t-monotonicity implies $(t-1)$- monotonicity. If λ is t-monotone, its conjugate υ, defined by $\upsilon(A) = 1 - \lambda(A^c)$, is t-alternating, in the sense that, for all $A_1, \ldots, A_t \subseteq \Omega$,

$$\upsilon(A_1 \cap \cdots \cap A_t) \leq \sum_{1 \leq i \leq t} \upsilon(A_i) - \sum_{1 \leq i < j \leq t} \upsilon(A_i \cup A_j) + \sum_{1 \leq i < j < k \leq t} \upsilon(A_i \cup A_j \cup A_k) - \cdots + (-1)^{t-1} \upsilon(A_1 \cup \cdots \cup A_t). \tag{2.8}$$

If λ is t-monotone for all $t \geq 2$, λ is said to be infinitely monotone, or a belief function (Shafer 1976), and its conjugate υ is said to be infinitely alternating, or a plausibility function. If λ is a belief function, its Möbius transform m_λ, defined for all $E \subseteq \Omega$ by

$$m_\lambda(E) := \sum_{H \subseteq E} (-1)^{|E-H|} \lambda(H), \tag{2.9}$$

is nonnegative ($m_\lambda(E) \geq 0$ for all $E \subseteq \Omega$). Moreover, for all $A \subseteq \Omega$,

$$\lambda(A) = \sum_{E \subseteq A} m_\lambda(E), \text{ and} \tag{2.10}$$

$$\upsilon(A) = \sum_{E \cap A \neq \emptyset} m_\lambda(E). \tag{2.11}$$

Indeed, if $m : 2^\Omega \to [0, 1]$ is any mapping such that $m(\emptyset) = 0$ and $\sum_{E \subseteq \Omega} m(E) = 1$,[3] the mapping λ defined for all $A \subseteq \Omega$ by

$$\lambda(A) := \sum_{E \subseteq A} m(E) \tag{2.12}$$

is a belief function, with $m_\lambda = m$. See Shafer (1976) for full details.

Theorem 2.6 *There exist belief function/plausibility function pairs* (λ, υ) *for which axiom* (1.4) *fails, as well as such pairs for which axiom* (1.5) *fails.*

Proof In view of the conjugacy (2.6) of λ and υ, our candidate $a(A) = \upsilon(A) - \lambda(A)$ for an ambiguity measure may be written purely in terms of λ as

$$a(A) = 1 - \lambda(A) - \lambda(A^c). \tag{2.13}$$

Let $\Omega = \{1, 2, 3, 4, 5\}$. In each case we construct a belief function λ from a mapping $m : 2^\Omega \to [0, 1]$ with $m(\emptyset) = 0$ and $\sum_{E \subseteq \Omega} m(E) = 1$, as in (2.12), denoting $m(E)$ by m_E. Let $A = \{1, 2\}$ and $B = \{2, 3, 4\}$. We first construct an m-function for which

[3] Such functions m arise on 2^Ω in a natural way in the following situation: There exists a finite set Θ, with states related to those of Ω by a mapping $T : \Theta \to 2^\Omega - \{\emptyset\}$, where $T(\theta)$ denotes the set of all states in Ω that are compatible with the state $\theta \in \Theta$. A probability measure p on 2^Θ then gives rise to m (and hence to belief and plausibility functions λ and υ) by means of the definition $m(E) := p(\{\theta \in \Theta : T(\theta) = E\})$. Strassen (1964) was the first to make a thorough study of such functions, although Dempster (1967) is typically credited as their originator in the artificial intelligence community.

axiom (1.4) fails. By (2.13),

$$\begin{aligned} a(A) &= a(B) \Leftrightarrow \lambda(A) + \lambda(A^c) = \lambda(B) + \lambda(B^c) \\ &\Leftrightarrow (m_{12} + m_1 + m_2) + (m_{345} + m_{34} + m_{35} + m_{45} + m_3 + m_4 + m_5) \\ &= (m_{234} + m_{23} + m_{24} + m_{34} + m_2 + m_3 + m_4) + (m_{15} + m_1 + m_5) \\ &\Leftrightarrow m_{345} + m_{35} + m_{45} + m_{12} = m_{234} + m_{23} + m_{24} + m_{15}; \qquad (2.14) \end{aligned}$$

$$\begin{aligned} a(A - B) &= a(B - A) \Leftrightarrow \lambda(A \cap B^c) + \lambda(A^c \cup B) = \lambda(A^c \cap B) + \lambda(A \cup B^c) \\ &\Leftrightarrow m_1 + (m_{2345} + 14 \text{ terms}) \\ &= (m_{34} + m_3 + m_4) + (m_{125} + m_{12} + m_{15} + m_{25} + m_1 + m_2 + m_5) \\ &\Leftrightarrow m_{2345} + m_{234} + m_{235} + m_{245} + m_{345} + m_{23} + m_{24} + m_{35} + m_{45} \\ &= m_{125} + m_{12} + m_{15}; \qquad (2.15) \end{aligned}$$

and

$$\begin{aligned} a(A \cap B) &= a(A \cup B) \Leftrightarrow \lambda(A \cap B) + \lambda(A^c \cup B^c) = \lambda(A \cup B) + \lambda(A^c \cap B^c) \\ &\Leftrightarrow m_2 + (m_{1345} + 14 \text{ terms}) = (m_{1234} + 14 \text{ terms}) + m_5 \\ &\Leftrightarrow m_{1345} + m_{135} + m_{145} + m_{345} + m_{15} + m_{35} + m_{45} \\ &= m_{1234} + m_{123} + m_{124} + m_{234} + m_{12} + m_{23} + m_{24}. \qquad (2.16) \end{aligned}$$

It is clearly possible to find positive real numbers m_E for each nonempty subset E of Ω such that (i) these 31 numbers sum to 1, and (ii) equation (2.14) is satisfied. If it happens that neither (2.15) nor (2.16) is satisfied, then we are done. If either (2.15) or (2.16), or both, are satisfied, choose ε such that

$$0 < \varepsilon < \min\{m_E, 1 - m_E : E \subseteq \Omega,\ E \neq \emptyset\}, \qquad (2.17)$$

and replace m_{2345} with $m^*_{2345} := m_{2345} + \varepsilon$, m_{125} with $m^*_{125} := m_{125} - \varepsilon$, m_{1345} with $m^*_{1345} := m_{1345} + \varepsilon$, and m_{1234} with $m^*_{1234} := m_{1234} - \varepsilon$, leaving all other m-values unchanged. In view of (2.17) the resulting values are all positive, and sum to 1, and continue to satisfy (2.14). Denote the left-hand side of (2.15) by l and the right-hand side by r, and denote the left-hand side of (2.16) by L and the right-hand side by R, with l^*, r^*, L^*, and R^* denoting their values when m_E is replaced by m^*_E, $E = 2345$, 125, 1345, and 1234. If $l \geq r$ and $L \geq R$, then $l^* > r^*$ and $L^* > R^*$. If $l = r$ and $L < R$, choose ε subject to (2.17), as well as the additional restriction $\varepsilon < (R - L)/2$. Then $l^* > r^*$, and $L^* < R^*$. If $l < r$ and $L = R$, choose ε subject to (2.17), as well as the additional restriction $\varepsilon < (r - l)/2$. Then $l^* < r^*$ and $L^* > R^*$. So axiom (1.4) fails to hold.

We next construct an m-function for which axiom (1.5) fails to hold. As above we can choose positive values m_E that sum to 1 over all nonempty $E \subseteq \Omega$. Moreover, using m_{12345} and the m_i, $i = 1, \ldots, 5$, as slack variables, and linearly scaling the other m_E variables as needed, we can ensure that (2.14), as well as (2.15) and (2.16) are satisfied. Choose ε subject to (2.17), and replace m_{12} by $m^*_{12} := m_{12} + \varepsilon$, and m_{12345} by $m^*_{12345} := m_{12345} - \varepsilon$. Under these modified m-values, the equality in (2.14) is replaced by the inequality $>$, and the equalities in (2.15) and (2.16) are both replaced by the inequality $<$. □

In view of the above results, the mapping $a = \upsilon - \lambda$ does not look very promising as an ambiguity measure, at least insofar as the latter notion is characterized by Fishburn's axioms. But let us press on and consider a further restriction on the class of belief / plausibility function pairs (λ, υ).

2.3 Consonant belief functions

Suppose that λ is a belief function, denoting its Möbius transform simply by m. As noted above, $m(E) \geq 0$ for all $E \subseteq \Omega$. If $m(E) > 0$, following Shafer, we call E a focal event associated with λ. It is easy to see that a belief function λ is a probability measure if and only if all of its associated focal events are singletons.

Consider the class of belief functions whose focal events constitute a chain $\emptyset \neq E_1 \subset E_2 \subset \cdots \subset E_r \subseteq \Omega$, where $1 \leq r \leq |\Omega|$. Shafer calls such belief functions consonant belief functions. As an illustration of just how narrow this class is, we note that among belief functions that are probability measures only the so-called point masses (probability measures p for which there exists an element ω^* such that $p(A) = 1$ if $\omega^* \in A$ and $p(A) = 0$ otherwise) are consonant. Clearly, the only possible values taken on by a consonant belief function λ are the partial sums $\sum_{j=0}^{i} m(E_j)$, where $0 \leq i \leq r$ and $E_0 := \emptyset$, and $\lambda(A) = \lambda(B)$ if and only if A and B both contain E_r, or, for some $i \in \{0, \ldots, r-1\}$, both contain E_i and neither contains E_{i+1}.

Theorem 2.7 *The mapping $\lambda : 2^{\Omega} \to [0, 1]$ is a consonant belief function if and only if*

$$\lambda(\emptyset) = 0, \lambda(\Omega) = 1, \quad \text{and} \quad \lambda(A \cap B) = \min\{\lambda(A), \lambda(B)\}, \quad \text{for all } A, B \subseteq \Omega, \tag{2.18}$$

or, equivalently, if and only if the conjugate υ of λ satisfies the properties

$$\upsilon(\emptyset) = 0, \upsilon(\Omega) = 1, \quad \text{and} \quad \upsilon(A \cap B) = \max\{\upsilon(A), \upsilon(B)\}, \quad \text{for all } A, B \subseteq \Omega. \tag{2.19}$$

Proof See Shafer (1976). □

A consonant belief function λ is typically called a necessity measure [and defined by (2.18)] in the artificial intelligence community, and its conjugate υ is called a possibility measure. Note that (2.18) and (2.19) imply that, for all $A \subseteq \Omega$,

$$\lambda(A) = 0 \text{ or } \lambda(A^c) = 0, \ \text{and} \ \upsilon(A) = 1 \text{ or } \upsilon(A^c) = 1. \tag{2.20}$$

In Sect. 4 below, we shall prove that when λ is a necessity measure and υ is its conjugate possibility measure, then $a = \upsilon - \lambda$ satisfies all five of Fishburn's axioms for an ambiguity measure. Before presenting that proof, however, we describe how necessity and possibility measures also arise in the course of quantifying linguistic vagueness.

3 Linguistic vagueness

Suppose that φ is a predicate which applies with varying degrees to members of a finite set Ω. Represent these degrees by the fuzzy membership function $f_\varphi : \Omega \to [0, 1]$, where $f_\varphi(\omega)$ designates the degree to which the predicate φ is applicable to ω, with 0 = totally inapplicable and 1= totally applicable. The function f_φ might, for example, reflect the usage of a sample of speakers from a given linguistic community, with $f_\varphi(\omega)$ recording the fraction of individuals in that sample who report that they would apply the predicate φ to ω.

Assume that there exists at least one $\omega \in \Omega$ for which $f_\varphi(\omega) = 1$, and define $\upsilon : 2^\Omega \to [0, 1]$ by

$$\upsilon(\emptyset) = 0 \text{ and } \upsilon(A) = \max\{f_\varphi(\omega) : \omega \in A\} \text{ for all nonempty } A \subseteq \Omega, \quad (3.1)$$

and $\lambda : 2^\Omega \to [0, 1]$ by

$$\begin{aligned} \lambda(A) &= 1 - \upsilon(A^c) = 1 - \max\{f_\varphi(\omega) : \omega \in A^c\} \\ &= \min\{1 - f_\varphi(\omega) : \omega \in A^c\} \text{ for all } A \subseteq \Omega. \end{aligned} \quad (3.2)$$

Theorem 3.1 *If υ is defined by (3.1) and λ by (3.2), then*

$$\upsilon(\emptyset) = 0,\ \upsilon(\Omega) = 1,\ \text{and } \upsilon(A \cup B) = \max\{\upsilon(A), \upsilon(B)\} \text{ for all } A, B \subseteq \Omega, \quad (3.3)$$

and

$$\lambda(\emptyset) = 0, \lambda(\Omega) = 1,\ \textit{and } \lambda(A \cap B) = \min\{\lambda(A), \lambda(B)\} \text{ for all } A, B \subseteq \Omega. \quad (3.4)$$

Moreover, any set functions υ and λ satisfying (3.3) and (3.4) arise, as in (3.1) and (3.2), from some function $f : \Omega \to$ [0,1] for which there exists an $\omega \in \Omega$ such that $f(\omega) = 1$.

Proof It is straightforward to show that (3.1) implies (3.3) and (3.2) implies (3.4). Suppose that υ satisfies (3.3), and hence that $\upsilon(A_1 \cup \cdots \cup A_n) = \max\{\upsilon(A_1), \ldots, \upsilon(A_n)\}$ for any finite sequence $A_1, \ldots, A_n$ of subsets of Ω. If we define $f : \Omega \to [0, 1]$ by $f(\omega) = \upsilon(\{\omega\})$, then, for all nonempty $A \subseteq \Omega$, $\upsilon(A) = \upsilon(\cup_{\omega \in A}\{\omega\}) = \max\{\upsilon(\{\omega\}) : \omega \in A\} = \max\{f(\omega) : \omega \in A\}$. In particular, $1 = \upsilon(\Omega) = \max\{f(\omega) : \omega \in \Omega\}$, and so there exists an $\omega \in \Omega$ such that $f(\omega) = 1$. Suppose that λ satisfies (3.4). If $\overline{\lambda}(A) := 1 - \lambda(A^c)$, then $\overline{\lambda}(\emptyset) = 0$, $\overline{\lambda}(\Omega) = 1$, and $\overline{\lambda}(A \cup B) = \max\{\overline{\lambda}(A), \overline{\lambda}(B)\}$ for all $A, B \subseteq \Omega$. So by the preceding argument there exists a function $f : \Omega \to [0, 1]$ such that $f(\omega) = 1$ for some $\omega \in \Omega$ and $\overline{\lambda}(A) = \max\{f(\omega) : \omega \in A\}$ for all $A \subseteq \Omega$. In particular, $\overline{\lambda}(A^c) = \max\{f(\omega) : \omega \in A^c\}$, and so $\lambda(A) = 1 - \overline{\lambda}(A^c) = \min\{1 - f(\omega) : \omega \in A^c\}$. □

By Theorem 3.1, the set functions λ and υ arising from a fuzzy membership function, as in (3.1) and (3.2), have precisely the same properties as the necessity and possibility measures λ and υ that arose in Sect. 2.3 in an imprecise probabilistic context. In the latter context, the mapping $a = \upsilon - \lambda$ is an obvious candidate for an ambiguity measure. Can a similar case be made for a, so defined, when υ and λ arise, as in (3.1) and (3.2), from a fuzzy membership function? It is helpful here to regard a set $A \subseteq \Omega$ as a set of potential exemplars of the vague predicate φ, with $\upsilon(A)$ recording the degree to which φ is applicable to the most salient of those exemplars. In what sense does $a(A) = \upsilon(A) - \lambda(A)$ quantify linguistic vagueness? One possibility is suggested by Fishburn's (1991, p. 4) observation that when A is a proper, nonempty subset of Ω, ambiguity should really be thought of as attaching to the two-part partition $\{A, A^c\}$ of Ω. Then, since $\lambda(A) = 1 - \upsilon(A^c)$, $a(A)$ may be re-conceptualized in the form

$$a(A, A^c) = \upsilon(A) + \upsilon(A^c) - 1 = \min\{\upsilon(A), \upsilon(A^c)\}, \tag{3.5}$$

the second equality above following from the fact that $\upsilon(A) = 1$ or $\upsilon(A^c) = 1$. Note that if f_φ is simply the characteristic function of the set A (respectively, A^c), then $a(A, A^c) = 0$, as is reasonable, since all elements of A (respectively, A^c) are perfect exemplars of the predicate φ, and φ is totally inapplicable to all elements of A^c (respectively, A).

Consider now the extreme case in which $\lambda(A) = 0$ for every $A \subset \Omega$, with $\lambda(\Omega) = 1$, and $\upsilon(A) = 1$ for every nonempty subset of Ω, with $\upsilon(\emptyset) = 0$. The result $a(A, A^c) = 1$ here might strike one as counter-intuitive, since υ arises from the constant fuzzy membership function $f_\varphi(\omega) \equiv 1$. But a predicate applicable with degree 1 to every single member of the universe is essentially tautological, admitting of no differentiation between any A and A^c, and it is in this sense that assigning each partition $\{A, A^c\}$ maximal ambiguity is to be understood. To further elaborate on the preceding point, suppose that we are given just f_φ and Ω, with f_φ not uniformly equal to 1. The partitions having the least ambiguity with respect to φ arise from segregating all elements $\omega \in \Omega$ with $f_\varphi(\omega)$ minimal into one block of the partition, and placing all other elements of Ω into the other block. In other words, $a(A, A^c)$ is minimized for a given predicate φ when either A or A^c comprises the uniformly worst exemplars of φ in Ω.

In the next section, we prove that if λ is a necessity measure and υ is its conjugate possibility measure, then $a = \upsilon - \lambda$ satisfies all five of Fishburn's axioms for an ambiguity measure.

4 Sufficient conditions for a Fishburn ambiguity measure

Theorem 4.1 *If λ is a necessity measure on 2^Ω and υ is its conjugate possibility measure, then $a = \upsilon - \lambda$ satisfies all five of Fishburn's axioms for an ambiguity measure.*

Proof By Theorem 2.7, λ is a belief function, and hence, 2-monotone. So by Theorem 2.5, a satisfies axioms (1.1), (1.2), and (1.3). It remains only to prove axioms (1.4) and (1.5). For ease of reference, let us reiterate the basic results on which our proofs will be based. Since λ is a necessity measure,

$$\lambda(A \cap B) = \min\{\lambda(A), \lambda(B)\}, \text{ for all } A, B \subseteq \Omega. \tag{4.1}$$

Since $\lambda(\emptyset) = 0$, (4.1) implies that, for all $A \subseteq \Omega$,

$$\lambda(A) = 0 \text{ or } \lambda(A^c) = 0. \tag{4.2}$$

Since $\upsilon(A) = 1 - \upsilon(A^c)$,

$$a(A) = 1 - \lambda(A) - \lambda(A^c), \text{ for all } A \subseteq \Omega. \tag{4.3}$$

I. We first prove (1.4). By (4.3) the hypothesis $a(A) = a(B)$ is equivalent to

$$\lambda(A) + \lambda(A^c) = \lambda(B) + \lambda(B^c), \tag{4.4}$$

and the desired conclusion, $a(A \cup B) = a(A \cap B)$ or $a(A - B) = a(B - A)$, is equivalent to

$$\lambda(A \cup B) + \lambda(A^c \cap B^c) = \lambda(A^c \cup B^c) + \lambda(A \cap B) \tag{4.5}$$

or

$$\lambda(A^c \cup B) + \lambda(A \cap B^c) = \lambda(A \cup B^c) + \lambda(A^c \cap B). \tag{4.6}$$

In what follows, we denote $\min\{x, y\}$ by $x \wedge y$ and $\max\{x, y\}$ by $x \vee y$. Suppose that $\lambda(A \cup B) = w$, $\lambda(A \cup B^c) = x$, $\lambda(A^c \cup B) = y$, and $\lambda(A^c \cup B^c) = z$. Then by (4.1), $\lambda(A) = w \wedge x$, $\lambda(A^c) = y \wedge z$, $\lambda(B) = w \wedge y$, $\lambda(B^c) = x \wedge z$, $\lambda(A \cap B) = w \wedge x \wedge y$, $\lambda(A \cap B^c) = w \wedge x \wedge z$, $\lambda(A^c \cap B) = w \wedge y \wedge z$, and $\lambda(A^c \cap B^c) = x \wedge y \wedge z$. With this notation, we must show that

$$(w \wedge x) + (y \wedge z) = (w \wedge y) + (x \wedge z) \tag{4.7}$$

implies that

$$w + (x \wedge y \wedge z) = z + (w \wedge x \wedge y) \text{ or} \tag{4.8}$$
$$y + (w \wedge x \wedge z) = x + (w \wedge y \wedge z). \tag{4.9}$$

(i.) If $w \vee x \vee y \vee z = x$, then (4.7) and (4.8) both assert that $w + (y \wedge z) = z + (w \wedge y)$.
(ii.) If $w \vee x \vee y \vee z = y$, then (4.7) and (4.8) both assert that $w + (x \wedge z) = z + (w \wedge x)$.
(iii.) If $w \vee x \vee y \vee z = w$, then (4.7) and (4.9) both assert that $x + (y \wedge z) = y + (x \wedge z)$.
(iv.) If $w \vee x \vee y \vee z = z$, then (4.7) and (4.9) both assert that $x + (w \wedge y) = y + (w \wedge x)$.

II. We next prove (1.5). The hypothesis $a(A) > a(B)$ is equivalent to

$$\lambda(B) + \lambda(B^c) > \lambda(A) + \lambda(A^c), \tag{4.10}$$

and the desired conclusion, $a(A \cap B) > a(A \cup B)$ or $a(A - B) > a(B - A)$, is equivalent to

$$\lambda(A \cup B) + \lambda(A^c \cap B^c) > \lambda(A^c \cup B^c) + \lambda(A \cap B) \text{ or} \tag{4.11}$$

$$\lambda(A \cup B^c) + \lambda(A^c \cap B) > \lambda(A^c \cup B) + \lambda(A \cap B^c). \tag{4.12}$$

Note that (4.10), as well as the disjunction of (4.11) and (4.12), are undisturbed by the substitution of B^c for B, and so it suffices to consider the following cases:

$$\lambda(B) > \lambda(A) > 0. \tag{4.13}$$

$$\lambda(B) > \lambda(A^c) > 0. \tag{4.14}$$

$$\lambda(B) > 0 \text{ and } \lambda(A) = \lambda(A^c) = 0. \tag{4.15}$$

If (4.13) holds, then $\lambda(A \cup B) \geq \lambda(B) > \lambda(A) = \min\{\lambda(A), \lambda(B)\} = \lambda(A \cap B) > 0$, and so by (4.2), $\lambda(A^c \cap B^c) = 0 = \lambda(A^c \cup B^c)$, which establishes (4.11). In both cases (4.14) and (4.15), $\lambda(A) = \lambda(B^c) = 0$, and so $\lambda(A \cap B) = \min\{\lambda(A), \lambda(B)\} = 0 = \min\{\lambda(A^c), \lambda(B^c)\} = \lambda(A^c \cap B^c)$. Also, in both cases, $\lambda(B) > \lambda(A^c)$. Suppose that the focal events of λ are $E_1 \subset \cdots \subset E_r \subseteq \Omega$, and let $E_0 := \emptyset$. Then, for some $i \in \{0, \ldots, r-1\}$, A^c contains E_i, but not E_{i+1}, and B contains E_{i+1}, whence $B^c \cap E_{i+1} = \emptyset$. So $A^c \cup B^c$ contains E_i, but not E_{i+1}, which implies that $\lambda(A^c) = m(E_0) + \cdots + m(E_i) = \lambda(A^c \cup B^c)$. Thus, $\lambda(A \cup B) \geq \lambda(B) > \lambda(A^c) = \lambda(A^c \cup B^c)$, and so (4.11) holds here as well. □

5 Summary and conclusion

The term ambiguity is, ironically, itself ambiguous, referring in some instances to vagueness in language, and in others (especially in decision theory) to probabilistic imprecision. Fishburn's axiomatization of what he termed an ambiguity measure aimed to steer a neutral course between these two interpretations. We have demonstrated that Fishburn's axioms are indeed consistent with both of these interpretations. At the core of our argument is a proof that if λ is a necessity measure and υ is its conjugate possibility measure on 2^Ω, then $a = \upsilon - \lambda$ is an ambiguity measure. The set functions λ and υ are initially construed purely abstractly as mappings from 2^Ω to [0, 1], characterized by the properties (i) $\lambda(\emptyset) = 0$, $\lambda(\Omega) = 1$, and $\lambda(A \cap B) = \min\{\lambda(A), \lambda(B)\}$; and (ii) $\upsilon(A) = 1 - \lambda(A^c)$ (whence, $\upsilon(\emptyset) = 0$, $\upsilon(\Omega) = 1$, and $\upsilon(A \cup B) = \max\{\upsilon(A), \upsilon(B)\}$). If Ω is construed as a set of possible exemplars of the vague predicate ϕ, with $f_\phi(\omega)$ denoting the degree to which ϕ is applicable to ω, then the mappings λ and υ satisfying (i) and (ii) arise here from the fuzzy membership function f_ϕ by means of the definitions $\upsilon(A) = \max\{f_\phi(\omega) : \omega \in A\}$ and $\lambda(A) = 1 - \upsilon(A^c)$, and, as we show, $a(A) = \upsilon(A) - \lambda(A) = \min\{\upsilon(A), \upsilon(A^c)\}$

measures the degree to which the partition $\{A, A^c\}$ differentiates members of Ω with respect to the predicate ϕ. If Ω is construed as a set of possible states of the world, then the mappings λ satisfying (i) coincide with a limited class of lower probabilities on 2^Ω known as consonant belief functions, and the mappings υ satisfying (ii) coincide with the upper probabilities that are conjugate to those belief functions. In this case $a(A)$ measures the degree of imprecision in the lower/upper probability pair (λ, υ) with respect to the event A. Thus Fishburn's axioms are consistent with an interpretation of ambiguity as linguistic vagueness, as well as (a very special sort of) probabilistic imprecision.

References

Dempster, A. (1967). Upper and lower probabilities induced by a multivalued mapping. *Journal of the American Statistical Association*, *69*, 325–339.

Dubois, D., Nguyen, H. T., & Prade, H. (2000). Possibility theory, probability and fuzzy sets: misunderstandings, bridges and gaps. In D. Dubois & H. Prade (Eds.), *Fundamentals of fuzzy sets* (pp. 343–438). Boston: Kluwer.

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

de Finetti, B. (1974). *Theory of probability,* vol. 1, English translation of Teoria delle Probabilità (1970). London: Wiley.

Fishburn, P. (1991). On the theory of ambiguity. *International Journal of Information and Management Sciences*, *2*, 1–16.

Fishburn, P. (1993). The axioms and algebra of ambiguity. *Theory and Decision*, *34*, 119–137.

Savage, L. (1972). *The foundations of statistics* (second revised edition). New York: Dover.

Shafer, G. (1976). *A mathematical theory of evidence*. Princeton: Princeton University Press.

Shapley, L. (1971). Cores of convex games. *International Journal of Game Theory*, *1*, 11–26.

Smith, C. (1961). Consistency in statistical inference and decision. *Journal of the Royal Statistical Society (Series B)*, *23*, 1–25.

Strassen, V. (1964). Messfehler und information. *Zeitschrift für Wahrscheinlichkeitstheorie*, *2*, 273–305.

Suppes, P. (1974). The measurement of belief. *Journal of the Royal Statistical Society (Series B)*, *36*, 160–175.

Walley, P. (1981). *Coherent lower (and upper) probabilities*, Technical Report, University of Warwick, Coventry.

Walley, P. (1991). *Statistical reasoning with imprecise probabilities*. New York: Chapman and Hall.