THREE ESSAYS ON SUBJECTIVE AMBIGUITY

by

Jiankang Zhang

A thesis submitted in conformity with the requirements
for the degree of Doctor of Philosophy
Graduate Department of Economics
University of Toronto

©Copyright by Jiankang Zhang 1998
The author has granted a non-exclusive licence allowing the National Library of Canada to reproduce, loan, distribute or sell copies of this thesis in microform, paper or electronic formats.

The author retains ownership of the copyright in this thesis. Neither the thesis nor substantial extracts from it may be printed or otherwise reproduced without the author's permission.

L’auteur a accordé une licence non exclusive permettant à la Bibliothèque nationale du Canada de reproduire, prêter, distribuer ou vendre des copies de cette thèse sous la forme de microfiche/film, de reproduction sur papier ou sur format électronique.

L’auteur conserve la propriété du droit d’auteur qui protège cette thèse. Ni la thèse ni des extraits substantiels de celle-ci ne doivent être imprimés ou autrement reproduits sans son autorisation.

0-612-35378-8
THREE ESSAYS ON SUBJECTIVE AMBIGUITY

by

Jiankang Zhang

Doctor of Philosophy (1998)

Department of Economics

University of Toronto

ABSTRACT

The Subjective Expected Utility (SEU) Theory axiomatized by Savage is widely referred to as providing a subjective theory of probability because a probability measure \( p \) is derived from preference. However, the domain of \( p \) is assumed to be given exogenously, thus having an objective status in the theory. Exogeneity of the domain is not a limitation of a 'subjective theory' if it is believed that decision makers assign probabilities to all events that are relevant to the context being modeled. But choice behavior such as exhibited in the Ellsberg paradox and related evidence have demonstrated that many decision makers do not assign probabilities to all events. Decision makers may differ not only in the probabilities assigned to given events, but also in the identity of the events to which they assign probabilities. The main purpose of this thesis is to derive a fully subjective theory of probability in the sense that both the values and domain of the probability are derived from preference.

The key definition of the thesis – subjectively unambiguous events– is given in chapter 1. An implication is that the set \( \mathcal{A} \) of unambiguous events is in general not an algebra due to the fact the intersection of two unambiguous events is not necessarily unambiguous. However, \( \mathcal{A} \) is a \( \lambda \)-system, a class of sets closed with respect to complements and disjoint unions. Thus one step towards in providing a fully subjective theory of probability is to characterize those likelihood relations defined on a \( \lambda \)-system that can be represented numerically by an additive probability measure. Such a characterization is provided in chapter 3.
After imposing some axioms on preference, we derive both a probability measure on \( \mathcal{A} \) and a representation for preference on the corresponding set of 'unambiguous' acts, namely acts that are measurable with respect to \( \mathcal{A} \). The noted representation is that preference is 'based on probabilities' in the sense of probabilistic sophistication as in Machina and Schmeidler (MS). The major novelty relative to MS is that while their axioms deliver probabilistic sophistication on the domain of all acts, ours deliver probabilistic sophistication only on the domain of subjectively unambiguous acts, leaving the nature of the preference ranking of ambiguous acts unrestricted.

A well-known alternative to the Savage model is Choquet Expected Utility (CEU) Theory axiomatized by Schmeidler (1989) and Gilboa (1987). Schmeidler uses the Anscombe-Aumann set-up to deliver the CEU model, but this approach presumes the existence of objective lotteries. Gilboa avoids this drawback by using the Savage set-up, but his axioms are hard to interpret. Moreover, it is not clear from these models why and how aversion to ambiguity leads a decision maker to use a nonadditive probability measure to represent her beliefs and a CEU model to represent her preferences over acts. Adopting the Savage set up, in chapter 2 we provide a simple axiomatization of the Choquet Expected Utility model where the capacity is an inner measure. Two attractive features of the model are its specificity and the transparency of its axioms. The key axiom states that the decision-maker uses unambiguous acts to approximate ambiguous ones. The notion of 'ambiguity' is subjective and derived from preference.
ACKNOWLEDGMENTS

I especially thank Professor Larry G Epstein for pointing out this topic, and for his valuable suggestions, continuous encouragement, intellectual stimulation, time and interest which he gave me during the course of writing this thesis.

I am also grateful to Professors Massimo Marinacci, Klaus Nehring, Mike Peters, Carolyn Pitchik, Uzi Segal, Peter Wakker and Myrna Holtz Wooders for valuable discussions and suggestions.

Finally, I wish to express my sincere thanks to my wife, Jianhua Chen, who has offered continuous love and encouragement in the past five years.

This thesis is dedicated to my parents, from whom I have learned the value of an education.
CONTENTS

Chapter 1: Subjective Probabilities on Subjectively Unambiguous Events ———— 1

Chapter 2: Subjective Ambiguity, Expected Utility and Choquet Expected Utility

Chapter 3: Qualitative Probabilities on $\lambda$-Systems ———————————————————— 95
Chapter 1

Subjective Probabilities on Subjectively Unambiguous Events
1.1 INTRODUCTION

Following the Savage approach to modelling choice under uncertainty, we study preference over acts defined on a state space $S$. Our initial contribution is the definition of \textit{subjectively unambiguous events}. In the subsequent analysis, we impose some axioms on preference and derive both (1) a probability measure on the class $\mathcal{A}$ of subjectively unambiguous events and (2) a representation for preference on the corresponding set of 'unambiguous' acts, namely acts that are measurable with respect to $\mathcal{A}$. The noted representation is that preference is 'based on probabilities' in the sense of probabilistic sophistication as in [13]. The major novelty relative to [13] is that while their axioms deliver probabilistic sophistication on the domain of all acts, ours deliver probabilistic sophistication only on the domain of subjectively unambiguous acts, leaving the nature of the preference ranking of ambiguous acts unrestricted. The remainder of this introduction is devoted to motivating our analysis and clarifying its place in the literature.

Our motivation stems largely from the observation that although Savage's expected utility theory is typically referred to as providing a \textit{subjective theory of probability}, there is a sense in which it fails to be subjective. Savage assumes that the decision-maker contemplates all events, that is, all subsets of the state space are assumed admissible or measurable. Various generalizations of his analysis limit the class of measurable events to an arbitrary exogenously specified $\sigma$-algebra $\Sigma$. The subjective nature of this theory is due to the fact that a probability measure $\mathbb{P}$ on $\Sigma$ is derived from the decision-maker's preference order over the domain of measurable acts. However, the domain of the measure, either $\Sigma$ or the power set in [16], is \textit{exogeneous} to the model. In particular, and in stark contrast to the case for $\mathbb{P}$, this domain does \textit{not} depend on preference and it is \textit{not} allowed to vary with the decision-maker, except in a trivial sense. The trivial exception is where the modeler \textit{assumes} that two decision-makers $a$ and $b$ have different $\sigma$-algebras $\Sigma_a$ and $\Sigma_b$, in which case the probability measures derived for $a$ and $b$ will have different domains. But the fact that $a$ and $b$ assign probabilities to different events is an assumption - it is not derived from the preferences or choice behavior of the two decision-makers. The $\sigma$-algebra that is appropriate for $a$ or $b$ is assumed to be given and derived from other considerations. These remarks apply similarly to the generalization of Savage provided in [13].

Exogeneity of the $\sigma$-algebra is not a limitation of a 'subjective theory' if it is believed that decision-makers assign probabilities to all events that are relevant to the context being modelled. In that case, the modelling context may dictate the
appropriate specification for \( \Sigma \), independently of preference. But choice behavior such as that exhibited in the Ellsberg Paradox and related evidence have demonstrated that many decision-makers do not assign probabilities to all events. In situations where some events are ‘ambiguous’, decision-makers who are averse to such ambiguity may not assign probabilities to those events, though the likelihoods of ‘unambiguous’ events are represented in the standard probabilistic way. For example, in the case of the Ellsberg urn with balls of 3 possible colors, \( R, B \) and \( G \), where \( R + B + G = 90 \) and \( R = 30 \), all events in the class

\[
A = \{ \emptyset, \{R\}, \{B, G\}, \{R, B, G\} \}
\]

are intuitively unambiguous. Most decision-makers would presumably assign them the obvious probabilities in deciding on how to rank bets based on the color of a ball to be drawn at random. However, the common ‘ambiguity averse’ preference ranking of such bets is inconsistent with the use of probabilities for other events. On the other hand, aversion to ambiguity is not universal. Some decision-makers are indifferent to ambiguity and behave in the non-paradoxical and fully probabilistic fashion. The lesson we take from this is that decision-makers may differ not only in the probabilities assigned to given events (an aspect not well illustrated by this example), but also in the identity of the events to which they assign probabilities. Thus a subjective theory of probability should make both the domain and the values of the probability measure subjective and based on preference.

The formulation of such a fully subjective theory of probability is our ultimate objective. Naturally, following the choice-theoretic tradition of Savage, it is preference rather than probability per se that is of prime importance. Thus the desired derivation of a probability measure is as one component of a preference representation, either expected utility or probabilistic sophistication, that is to apply on an agent-specific subdomain of acts. We now clarify the nature of our contribution towards achieving the objective of a fully subjective theory.

First our terminology must be made more precise. Expressions such as ‘the event \( A \) is assigned a probability’ are meaningless; a probability measure is defined primarily by additivity, a property that refers to a class (or collection) of events.\(^1\) A meaningful statement is that \( 'B \) is a class of events where \( p \) represents the decision-maker’s likelihood relation’. Also meaningful, and our focus in this paper, is the stronger statement \( 'B \) is a class of events such that preference is probabilistically sophisticated

\(^1\)Typically, that class is taken to be an algebra or \( \sigma \)-algebra. We adopt the more neutral term ‘class’ because a different mathematical structure will be relevant here, as explained shortly.
CHAPTER 1. SUBJECTIVE PROBABILITIES

on the domain of all acts that are $B$-measurable. In general, there may be several such classes $B$ and one might expect a fully subjective theory to derive them all from the given preference. The accomplishment in this paper is more modest - we identify one particular class of events, denoted $A$, where probabilistic sophistication prevails, given suitable axioms on preference. Events in $A$ are called subjectively unambiguous. The latter term is defined by a form of separability that we feel is an intuitively necessary condition for any unambiguous event. The intuitive appeal of this definition justifies singling out $A$ amongst all the classes of events $B$ where probabilistic sophistication prevails.

Because the distinction between 'having probability' and 'being unambiguous' is central, it may be useful to elaborate here on the nature of this distinction. Ellsberg's two-urn experiment provides a clarifying example. Let $S_2$ be an urn containing 100 balls that are either red or blue in unknown proportions and $S_1$ an urn containing 50 balls of each color. Typically, $S_1$ and $S_2$ are referred to as 'unambiguous' and 'ambiguous' urns respectively. This informal terminology is consistent with the view that decision-makers dislike ambiguity and with the typical preference for betting on drawing a red ball from $S_1$ rather than from $S_2$. But such a ranking is also consistent with the use of a probability measure (presumably $(1/2, 1/2)$) for the purpose of ranking bets internal to $S_2$. In terms of the preceding discussion, if we think of $S_1 \times S_2$ as the state space and $B$ as the co-ordinate algebra $\{S_1 \times A_2 : A_2 \subset S_2\}$, then probabilistic sophistication on (acts measurable with respect to) $B$ is commonly assumed, even though events in $B$ are viewed as being ambiguous. Roughly, the existence of a probability measure that serves as the basis for the ranking of $B$-measurable acts depends exclusively on the decision-maker's view of events within $B$. On the other hand, whether or not events in $B$ are subjectively ambiguous depends also on how they are viewed relative to events outside $B$ (such as the comparison between drawing a red ball from $S_2$ as opposed to drawing it from $S_1$). See Section 1.3.3 for elaboration and for the connection with our formal definition of ambiguity.

Further comparison with the Machina-Schmeidler analysis seems worthwhile. If their axioms are imposed, then all events are subjectively unambiguous. Consequently, our Theorem 1.5.2 extends their main result by dropping the requirement that all events be unambiguous. As a further consequence, only our model is consistent with Ellsberg-type behavior; events that are deemed ambiguous by the decision-maker are excluded from the domain $A$ and are not necessarily assigned probabilities,

---

2A qualification is that our theorem delivers a countably additive probability measure, while theirs deals with the more general class of finitely additive subjective priors.
removing the source of the paradox. A ‘cost’ is that our model is silent on the nature of preference on the subset of ambiguous acts. However, from the perspective of this paper, Machina and Schmeidler are able to characterize preference over all acts only by assuming that all acts are unambiguous. In other words, they avoid the problem of ambiguous acts by assuming them away. This is not to criticize their valuable contribution; we wish merely to emphasize that our analysis extends theirs. A further extension to axiomatize the structure of preference also over ambiguous acts is an important problem for future research.

The subjective nature of unambiguous events distinguishes our model from others where a class of unambiguous events is employed or the notion of ambiguity is studied; see [7], [15], [10] and [4], for example. In all these cases, ambiguity is taken as a primitive. Here in contrast, preference is the only primitive. Zhang [21] (chapter 2) is the first to propose a definition of ‘subjectively unambiguous’. In Section 1.3.4, his definition is discussed and compared with the one proposed here. Ghirardato and Marinacci [8] define ambiguity in terms of preference in the case where the latter conforms to Choquet expected utility theory. Within the latter framework, their definition differs from ours. Rather than provide a detailed comparison of the two approaches, we prefer to acknowledge that the intuitive appeal of any definition of ‘subjectively unambiguous’ is undoubtedly subjective, and to emphasize instead the ‘practical’ case for our definition - it delivers a representation result (Theorem 1.5.2). The latter is our major contribution because, as we have argued, it constitutes a first step towards a fully subjective theory of probability. (There is no parallel result in [8].) Keeping in mind the ultimate objective of deriving a probability measure on the class \( \mathcal{A} \) of subjectively unambiguous events, it seems that a liberal definition of ‘unambiguous’ is advantageous. Our definition should be viewed from this perspective.

Beside its subjective nature, another nonstandard feature of the class of unambiguous events is that it is not a \( \sigma \)-algebra or even an algebra. Zhang [21] points out that while the class of unambiguous events is naturally taken to be closed with respect to complements and disjoint unions, it is typically not closed with respect to intersections. This point may be illustrated by borrowing Zhang's example of an Ellsberg-type urn with 4 possible colors - \( R, B, G \) and \( W \). Suppose that the total number of balls is 100 and that \( R + B = G + B = 50 \). Then it is intuitive that the class of unambiguous events is

\[
\mathcal{A} = \{ S, \emptyset, \{ B, R \}, \{ B, G \}, \{ G, W \}, \{ R, W \} \}.
\]

Observe that \( \mathcal{A} \) fails to be an algebra, because while \( \{ B, R \} \) and \( \{ B, G \} \) are unam-
biguous, their intersection \( \{B\} \) is not. As pointed out by Zhang, the appropriate mathematical structure for \( \mathcal{A} \) is a \( \lambda \)-system. This complicates the derivation of a probability measure on \( \mathcal{A} \) and, in particular, prevents us from simply invoking existing results from [16], [6] and [13]. The arguments in these studies exploit the fact that the relevant class of events is closed with respect to intersections. We rely instead on a recent representation result in [22] (chapter 3) for qualitative probabilities on \( \lambda \)-systems.

The paper proceeds as follows. The formal definition of a \( \lambda \)-system follows next. Then Section 1.3 introduces our key definition of an unambiguous event. The definition is then examined in a number of Ellsberg-type settings and also within the context of some specific models of preference. Some preference axioms are described in Section 1.4 and our main result (Theorem 1.5.2) follows in Section 1.5. Most proofs are relegated to appendices.

1.2 PRELIMINARIES

Let \((S, \Sigma)\) be a measurable space where \( S \) is the set of states and \( \Sigma \) is a \( \sigma \)-algebra. All events in this paper are assumed to lie in \( \Sigma \); we repeat this explicitly below only on occasion.

Say that a nonempty class of subsets \( \mathcal{A} \subset \Sigma \) of \( S \) is a \( \lambda \)-system if

\begin{align*}
\lambda.1 & \quad S \in \mathcal{A}; \\
\lambda.2 & \quad A \in \mathcal{A} \implies A^c \in \mathcal{A}; \text{ and} \\
\lambda.3 & \quad A_n \in \mathcal{A}, \; n = 1, 2, ... \text{ and } A_i \cap A_j = \emptyset, \; \forall i \neq j \implies \bigcup_n A_n \in \mathcal{A}.
\end{align*}

This definition and terminology appear in [3, p. 36]. A \( \lambda \)-system \( \mathcal{A} \) is closed with respect to complements and countable disjoint unions. The intuition for these properties is clear if we think of \( \mathcal{A} \) as a class of events to which the decision-maker attaches probabilities. If she can assign a probability to event \( A \), then the complementary probability is naturally assigned to \( A^c \). Similarly, if she can assign probabilities to each of the disjoint events \( A \) and \( B \), then the sum of these probabilities is naturally
assigned to $A \cup B$. On the other hand, there is no such intuition supporting closure with respect to intersections, or equivalently, with respect to arbitrary unions. Lack of closure with respect to intersections differentiates $\lambda$–systems from algebras or $\sigma$–algebras. As illustrated by the example of an Ellsberg-type urn with 4 colors, $\lambda$–systems are more appropriate for modelling families of ‘unambiguous’ events.

We have the following equivalent definition [3, p. 43]:

**Lemma 1.2.1** A nonempty class of subsets $A \subseteq \Sigma$ of $\mathcal{S}$ is a $\lambda$–system if and only if

1. $\emptyset, S \in A$;
2. $A, B \in A$ and $A \subseteq B \implies B \setminus A \in A$; and
3. $A_n \in A$ and $A_n \subseteq A_{n+1}$, $n = 1, 2, \ldots \implies \bigcup_n A_n \in A$.

Although $A$ is not an algebra, a probability measure can still be defined on $A$. Say that $p : A \to [0, 1]$ is a (finitely additive) probability measure on $A$ if:

1. $p(\emptyset) = 0$, $p(S) = 1$; and
2. $p(A \cup B) = p(A) + p(B)$, $\forall A, B \in A$, $A \cap B = \emptyset$.

Countable additivity of $p$ is defined in the usual way and will be stated explicitly where needed. Given a probability measure $p$ on $A$, call $p$ convex-ranged if for all $A \in A$ and $0 < r < 1$, there exists $B \subseteq A$, $B \in A$, such that $pB = r \left( pA \right)$.\(^3\)

As in Savage, we assume a set of outcomes $\mathcal{X}$. Prospects are modelled via (simple) acts, $\Sigma$–measurable maps from $\mathcal{S}$ to $\mathcal{X}$ having finite range. The set of acts is $\mathcal{F} = \{ ..., f, f', g, h, ... \}$. Given a $\lambda$–system $A$, define $\mathcal{F}^{ua}$ by

$$\mathcal{F}^{ua} = \{ f \in \mathcal{F} : f \text{ is } A\text{-measurable} \},$$

\(^3\)When $A$ is a $\sigma$–algebra and $p$ is countably additive, this is equivalent to non-atomicity [14, pp. 142-3] See the proof of Lemma 1.7.7 for more on convex-rangedness in the present setting.
where \( f \) is \( A \)-measurable if \{ \( s \in S : f(s) \in X \} \in A \) for any subset \( X \) of \( \mathcal{X} \). Thinking of \( A \) as the set of unambiguous events, \( \mathcal{F}^{ua} \) is naturally termed the set of unambiguous acts.

### 1.3 UNAMBIGUOUS EVENTS

#### 1.3.1 Definition

The primitives \((S, \Sigma)\) and \(\mathcal{X}\) are defined as above. The decision-maker has a preference relation \(\succeq\) on the set of acts \(\mathcal{F}\). Unambiguous events are now defined from the perspective of \(\succeq\).

**Definition 1.3.1** An event \(T\) is unambiguous if:

(a) For all disjoint subevents \(A, B\) of \(T^c\), acts \(h\) and outcomes \(x^*, x, z, z' \in \mathcal{X}\),

\[
\begin{pmatrix}
  x^* & \text{if } s \in A \\
  x & \text{if } s \in B \\
  h(s) & \text{if } s \in T^c \setminus (A \cup B) \\
  z & \text{if } s \in T
\end{pmatrix}
\begin{pmatrix}
  x & \text{if } s \in A \\
  x^* & \text{if } s \in B \\
  h(s) & \text{if } s \in T^c \setminus (A \cup B) \\
  z & \text{if } s \in T
\end{pmatrix}
\]

\[
\begin{pmatrix}
  x^* & \text{if } s \in A \\
  x & \text{if } s \in B \\
  h(s) & \text{if } s \in T^c \setminus (A \cup B) \\
  z & \text{if } s \in T
\end{pmatrix}
\begin{pmatrix}
  x & \text{if } s \in A \\
  x^* & \text{if } s \in B \\
  h(s) & \text{if } s \in T^c \setminus (A \cup B) \\
  z' & \text{if } s \in T
\end{pmatrix}
\]

\[
\begin{pmatrix}
  x^* & \text{if } s \in A \\
  x & \text{if } s \in B \\
  h(s) & \text{if } s \in T^c \setminus (A \cup B) \\
  z & \text{if } s \in T
\end{pmatrix}
\begin{pmatrix}
  x & \text{if } s \in A \\
  x^* & \text{if } s \in B \\
  h(s) & \text{if } s \in T^c \setminus (A \cup B) \\
  z' & \text{if } s \in T
\end{pmatrix}
\]

(b) The condition obtained if \(T^c\) is everywhere replaced by \(T^c\) in (a) is also satisfied. Otherwise, \(T\) is ambiguous.

The set of unambiguous events is denoted \(A\). It is nonempty because \(\emptyset\) and \(S\) are unambiguous. Observe that the defining invariance condition is required to be satisfied even if \(A\) or \(B\) is empty.

Turn to interpretation. The first two acts being compared yield identical outcomes if the true state lies in \((A \cup B)^c\). Thus the comparison is between 'bets conditional
1.3. UNAMBIGUOUS EVENTS

on \((A \cup B)'\) with stakes \(x^* \succ x\) and the outcomes shown for \((A \cup B)^c\). Paralleling the standard interpretation of Savage's axiom \(P4\), called Weak Comparative Probability in [13], the indicated ranking reveals that the decision-maker views \(A\) as conditionally more likely than \(B\). Suppose now that the outcome on event \(T\) is changed from \(z\) to \(z'\). If \(T\) is 'unambiguous', then this conditional likelihood ranking should not be affected because 'unambiguous' means or at least entails such separability or invariance. Constancy of acts on \(T\) is vital for this intuition. It is \(T\) in its entirety that is unambiguous and this does not imply anything about subsets. This leads naturally to the restriction to acts that are constant on \(T\). Finally, both \(T\) and \(T^c\) should satisfy such invariance because intuitively an event is unambiguous if and only if its complement is unambiguous.

For the exclusive purpose of the further interpretation offered in the next two paragraphs, suppose temporarily that Savage's axiom \(P4\) is satisfied, delivering a likelihood relation \(\succeq\) on all events that is derived from the ranking of bets or binary acts. Then a special case of the definition implies that if \(T\) is unambiguous, then for all \(A\) and \(B\) such that \(A \cup B = T^c\),

\[
A \succeq B \iff A \cup T \succeq B \cup T. \tag{1.5}
\]

This is shown by taking \(z = x\) and \(z = x^*\) in (1.4). Another intuitive implication of the formal definition is that the likelihood relation \(\succeq\) satisfies the additivity property of a qualitative probability when restricted to \(A\); that is, for all \(A, B \text{ and } C\) in \(\mathcal{A}\), with \(A \cap C = B \cap C = \emptyset\),

\[
B \succeq A \iff B \cup C \succeq A \cup C. \tag{1.6}
\]

(This is proven as part of the proof of Theorem 1.7.4.)

On the other hand, some readers may feel that our definition does not capture all intuitive aspects of 'unambiguous.' For example, if \(A\) and \(B\) are unambiguous, then it might be expected that (1.6) should be satisfied for all (not necessarily unambiguous) \(C\) disjoint from \(A \cup B\); this is not the case given our definition. We do not claim to have captured the complete essence of 'unambiguous'. The argument that we offer for our definition is first that it captures some intuition and second that it helps to deliver a novel representation theorem whose importance was argued in the introduction. It remains to be seen whether alternative definitions will be similarly useful.
Before turning to examples, observe that null events are unambiguous, where, as in Savage, say that an event \( T \subseteq S \) is null if for all acts \( f, g, g' \in \mathcal{F} \),

\[
\begin{pmatrix}
  f(s) & \text{if} \ s \in T^c \\
g(s) & \text{if} \ s \in T
\end{pmatrix} \sim \begin{pmatrix}
  f(s) & \text{if} \ s \in T^c \\
g'(s) & \text{if} \ s \in T
\end{pmatrix}.
\]

**Lemma 1.3.2** If \( T \) is null, then \( T \) is unambiguous.

**Proof.** Condition (1.4) is satisfied because changing the outcome on \( T \) from \( z \) to \( z' \) is a matter of indifference by the nullity of \( T \). When \( T^c \) is substituted for \( T \) in (1.4), each weak preference ranking is necessarily indifference. ■

Finally, while we have emphasized the connection, albeit imperfect, between 'unambiguous' and 'having probability', the example of state-dependent-expected-utility shows that such a connection is not delivered by our definition of unambiguous without further restrictions. Let the utility function \( U \) be given by

\[
U(f) = \int_S u_s(f) \, dp,
\]

for a suitable collection \( \{u_s\} \) of state-dependent vNM indices and some probability measure \( p \). It follows immediately from the additive separability across states that all events are unambiguous. However, there is no meaningful sense in which \( U \) is based on probabilities. It is well-known that when the vNM index varies with the state, then the probability measure \( p \) is not unique (for example, the identical \( U \) is delivered by \( v_s(\cdot) = a_s u_s(\cdot) \) and \( dq = a_s^{-1} dp \)) and its behavioral significance is unclear. We succeed later in delivering a unique probability measure for unambiguous events only by excluding such state-dependence.

### 1.3.2 Ellsberg Settings with a Single Urn

First, consider a state space consisting of only two points \( s_1 \) and \( s_2 \), corresponding to a single Ellsberg urn with balls having two possible colours. If \( \emptyset \prec_{\ell} \{s_i\} \prec_{\ell} S, i = 1, 2 \), then both singleton sets are necessarily unambiguous. Support for this designation is provided by the fact that any such likelihood relation may be represented by a probability measure; take any probability measure \( p \) such that \( p(s_1) < (\geq) 1/2 \) if \( \{s_1\} \prec_{\ell} (\geq_{\ell}) \{s_2\} \).
1.3. UNAMBIGUOUS EVENTS

Turn to the case of an urn with balls of 4 possible colors, as outlined in the introduction. The state space is

\[ S = \{ B, R, G, W \}, \]

where \( S = 100 \) and \( B + R = B + G = 50 \). The intuitive \( \lambda \)-system \( A \) of unambiguous events is described in (1.2). There is a natural probability measure \( p \) on \( A \) and it is reasonable to expect that a decision-maker would base choice on \( p \) when dealing with acts in \( \mathcal{F}^{ua} \). In particular, she might conform to expected utility theory there, with utility function

\[ U(f) = \int u(f) \, dp, \quad f \in \mathcal{F}^{ua}, \]

for some vNM index \( u \).

For the ranking of other acts, the following seems plausible: Though the decision-maker knows the probabilities of events in \( A \), no other information is available. Thus any measure in the set

\[ P = \{ m : m \text{ extends } p \text{ from } A \to \Sigma \}, \quad (1.7) \]

is admissible. Suppose that this set of measures is used as described in the multiple-priors model [9]; that is, for some vNM index \( u \), the utility of an act is computed by

\[ U(f) = \min_p \int_s u(f) \, dm. \quad (1.8) \]

Then the class of events that are subjectively unambiguous in the sense of our formal definition coincides with \( A \): Refer back to (1.7) and (1.8). For any \( x^* \succ x \),

\[
\begin{pmatrix}
    x^* & \text{if } s = W \\
    x & \text{if } s = R \\
    x & \text{if } s = G \\
    z & \text{if } s = B
\end{pmatrix}
\sim
\begin{pmatrix}
    x^* & \text{if } s = R \\
    x & \text{if } s = W \\
    x & \text{if } s = G \\
    z & \text{if } s = B
\end{pmatrix}
\]

when \( z = x \),

but indifference becomes '\( \succ \)' when \( z \) is set equal to \( x^* \). This proves that both \( \{B\} \) and its complement are ambiguous. Similarly for other singleton sets and their complements. To see that \( \{R, G\} \) (and also its complement) is ambiguous, observe that

\[
\begin{pmatrix}
    x^* & \text{if } s \in \{B, W\} \\
    z & \text{if } s \in \{R, G\}
\end{pmatrix}
\sim
\begin{pmatrix}
    x & \text{if } s \in \{B, W\} \\
    z & \text{if } s \in \{R, G\}
\end{pmatrix}
\]

when \( z = x \),

but the first act is strictly preferred when \( z = x^* \). This violates the defining property (1.4), taking \( A = \{B, W\} \) and \( B = \emptyset \). Finally, verify that each event identified in (1.2) as intuitively unambiguous is also unambiguous in the formal sense.
1.3.3 A Two-Urn Example

This example generalizes the two-urn example outlined in the introduction. It illustrates that there may exist more than one subdomain of acts where a given preference order is probabilistically sophisticated. Our representation result (Theorem 1.5.2) focuses on one of them only - the subdomain consisting of unambiguous acts. The example also demonstrates that our definition performs intuitively.

For concreteness, take outcomes $\mathcal{X} \subset \mathcal{R}^1$. Let $S = S_1 \times S_2$, where $S_1 = S_2 = \Omega$ and $\Omega$ represents the possible states in each urn. We do not insist that $\Omega$ be binary or even finite. Let $p$ be a probability measure on $\Omega$. (Implicit is a $\sigma$-algebra on $\Omega$ such that the product $\sigma$-algebra defines $\Sigma$.) The decision-maker is told that $p$ describes the distribution of states within the first urn $S_1$, but she is told less about the second urn $S_2$.

As before $\mathcal{F}$ denotes the set of acts over $S$. Denote by $\mathcal{F}_i$ the set of acts over the state space $S_i$. As a first step in defining utility over $\mathcal{F}$, let $U_2 : \mathcal{F}_2 \rightarrow \mathcal{R}^1$ be defined by

$$U_2(f) = \int_{S_2} u(f) \, d\phi(p), \quad f \in \mathcal{F}_2, \quad (1.9)$$

where $u$ is a continuous and increasing vNM index and where $\phi : [0, 1] \rightarrow [0, 1]$ is an increasing and onto map. This defines the probabilistically sophisticated subclass of Choquet expected utility functions [18]; alternatively, it corresponds to the rank-dependent-expected-utility model that has been studied in the theory of preference over lotteries or risky prospects.

To define $U$ on $\mathcal{F}$, observe that given any act $f$ over $S$ and $s_1 \in S_1$, the restriction $f(s_1, \cdot)$ can be viewed as an act over $S_2$, giving meaning to $U_2(f(s_1, \cdot))$. Thus $U$ can be defined as follows: For each $f \in \mathcal{F}$,

$$U(f) = \int_{S_1} U_2(f(s_1, \cdot)) \, dp(s_1) = \int_{S_1} \int_{S_2} u(f) \, d\phi(p(s_2)) \, dp(s_1). \quad (1.10)$$

In other words, uncertainty resolves in two stages. At the second stage and conditional on any $s_1$, $U_2$ is used to evaluate $f(s_1, \cdot)$. This evaluation can be viewed as producing the certainty equivalent outcome $u^{-1}(U_2(f(s_1, \cdot)))$, yielding the first stage act $s_1 \mapsto u^{-1}(U_2(f(s_1, \cdot)))$, which is then evaluated using expected utility theory with vNM
index u.\footnote{There is a clear parallel with the Anscombe-Aumann domain of two-stage acts that has played a large role in axiomatizations such as [18] and [9].}

This preference specification has a number of appealing features. First, preference to bet on $A_1 \times A_2$ over $B_1 \times B_2$ is independent of the stakes involved (Savage's P4), implying a complete and transitive likelihood relation $\succeq_\ell$ on all such rectangles. Moreover, $\succeq_\ell$ satisfies the following conditions for all events:

$$A_1 \times A_2 \succeq_\ell B_1 \times A_2 \iff A_1 \times A'_2 \succeq_\ell B_1 \times A'_2$$
$$A_1 \times A_2 \succeq_\ell A_1 \times B_2 \iff A'_1 \times A_2 \succeq_\ell A'_1 \times B_2.$$  

These equivalences reflect 'independence' between the two urns.

A second attractive feature of the preference specification is that it can explain the two-urn Ellsberg Paradox; indeed, it is often invoked for that purpose. To do so, suppose that

$$\phi(t) < t, \quad t \in (0, 1). \quad (1.11)$$

This specialization implies a strict preference for betting on any event $E \subset \Omega$ when it is 'drawn' from $S_1$ rather than from $S_2$; that is, $E \times S_2 \succ_\ell S_1 \times E$.

Turn to subjectively unambiguous events and domains where preference is probabilistically sophisticated. There is at least one set of acts on which preference is probabilistically sophisticated, indeed expected utility. Define the co-ordinate $\sigma$-algebra

$$\mathcal{A}_1 = \{A_1 \times S_2 : A_1 \subset S_1\},$$

and identify $\mathcal{F}_1$ with the set of $\mathcal{A}_1$-measurable acts. (Define $\mathcal{A}_2$ similarly.) Then

$$U(f) = \int_{S_1} u(f) \, dp, \quad f \in \mathcal{F}_1.$$  

Acts in $\mathcal{F}_1$ are unambiguous in the sense of our formal definition and so the probabilistic sophistication exhibited on $\mathcal{F}_1$ is an implication of our later representation result, given that the axioms specified there are satisfied. Our definition requires some separability between an unambiguous event $T$ and other events disjoint from $T$. Here observe that if $T = T_1 \times S_2$ and if $f$ is an act yielding $g(s)$ if $s \notin T$ and $h(s)$ if $s \in T$, then $U(f) = \int_{T_1} U_2(g) \, dp + \int_{T_1} U_2(h) \, dp$. Such additive separability provides more than enough to imply that $T$ is unambiguous.
On the other hand, because \( U = U_2 \) on \( \mathcal{F}_2 \), conclude that \( U \) is probabilistically sophisticated also on \( \mathcal{F}_2 \), or more precisely, on the set of \( A_2 \)-measurable acts. But, in general, events in \( A_2 \) are ambiguous, showing that probabilistic sophistication prevails also on a subdomain of acts that are ambiguous. This claim is a special case of the next result.

**Lemma 1.3.3** In the context of the two-urn example defined by (1.10), suppose that \( \Omega = [0, 1] \), \( p \) is non-atomic and countably additive and that

\[
\phi(t) \neq t \quad \text{for all } t \in (0, 1).
\]

Then the class \( A \) of subjectively unambiguous events includes all null events and events in

\[ A_1 = \{ A_1 \times S_2 : A_1 = [0, t), 0 \leq t \leq 1 \} \cup \{ S_1 \times A_2 : A_2 \subset S_2, p(A_2) = 0 \}. \]

Moreover, \( T = S_1 \times [0, t_1) \) is ambiguous, where \( 0 < t_1 < 1 \).

**Proof.** Suppose for simplicity that \( p \) is the Lebesgue measure. In order to show that each event of the form \( T = S_1 \times [0, t_1) \) is ambiguous, where \( 0 < t_1 < 1 \), define events

\[ A = [0, t_3] \times [t_1, 1], B = [t_3, 1] \times [t_1, t_2], C = [t_3, 1] \times [t_2, 1] \]

and the corresponding acts as in the definition (1.4) of unambiguous. By routine but tedious computation, one can show that under the hypothesis (1.12), there exist \( t_2 > t_1, t_3 \) in \( (0, 1) \) and outcomes \( x^* \succ y \succ x \), such that the invariance in (1.4) is violated when \( z \sim y \) and \( z' \sim x^* \). (A detailed proof of this claim is available upon request. However, proof of the Lemma is still incomplete, as it remains to show that the other 'nonrectangular' events excluded from the specified class \( A \) are indeed ambiguous.\(^6\)

\(^6\)We conjecture that the lemma may be extended to show that for every \( t_1 \) in \( (0, 1) \), if \( \phi(t_1) \neq t_1 \), then \( S_1 \times [0, t_1] \) is ambiguous. Because it implies indifference between betting on \([0, t_1]\) from \( S_1 \) or from \( S_2 \), we think of the equality \( \phi(t_1) = t_1 \) as corresponding to a case where the decision-maker is told that the probability of \([0, t_1]\) in the second urn is \( p([0, t_1]) \), precisely as in the first urn. Then \( S_1 \times [0, t_1] \) is intuitively unambiguous.
1.3. UNAMBIGUOUS EVENTS

It is noteworthy that the class $\mathcal{A}$ is unchanged if the strict inequality in (1.11) is reversed, indicating a strict preference for betting on $S_2$. In that sense, ambiguity is logically distinct from the decision-maker's attitude towards ambiguity.

Our definition performs better intuitively when $\Omega$ contains three or more elements. We would naturally prefer that our definition perform intuitively in all settings. However, just as Savage's theory has been viewed as important in spite of its limitation to rich state spaces, we feel that a definition for 'unambiguous' that is appealing and useful in settings with rich state spaces is a worthwhile contribution, if only as a stepping stone towards a more widely appealing definition and theory.

We offer one final observation regarding this example. In the above specification, preference conforms with expected utility on $\mathcal{F}_1$, which is also the domain of unambiguous acts. But this is coincidence. For example, suppose that we reversed the 'order' of the component state spaces and defined $U$ by

$$U(f) = \int_{S_2} \int_{S_1} u(f) \, dp(s_1) \, d\phi(p(s_2)).$$

Then $u$ is an expected utility function on $\mathcal{F}_1$, but events in $S_1$ are in general ambiguous. In fact, for this preference order, neither urn is unambiguous.

1.3.4 Global Probabilistic Sophistication

If "probabilities are assigned to all events", then all events should be unambiguous. Machina and Schmeidler give precise meaning to the expression in quotation marks and call corresponding preferences probabilistically sophisticated. The precise definition of this class of preferences can be found in [13]. (See also Section 1.5 below, where we define the probabilistic sophistication of preference over any suitable domain of acts. The special case where that domain is the set of all acts $\mathcal{F}$ is the notion in [13]; it is occasionally distinguished here terminologically by the added adjective 'global'.) For the moment it suffices to note the following central axiom underlying probabilistic sophistication:

**Axiom $P4^*$ (Strong Comparative Probability):** For all disjoint events $A$ and
\[ B, \text{ outcomes } x^* \succ x \text{ and } y^* \succ y, \text{ and acts } g \text{ and } h, \]
\[
\begin{pmatrix}
  x^* & \text{if } s \in A \\
  x & \text{if } s \in B \\
  g(s) & \text{if } s \notin A \cup B
\end{pmatrix}
\preceq
\begin{pmatrix}
  x & \text{if } s \in A \\
  x^* & \text{if } s \in B \\
  g(s) & \text{if } s \notin A \cup B
\end{pmatrix}
\implies
\begin{pmatrix}
  y^* & \text{if } s \in A \\
  y & \text{if } s \in B \\
  h(s) & \text{if } s \notin A \cup B
\end{pmatrix}
\preceq
\begin{pmatrix}
  y & \text{if } s \in A \\
  y^* & \text{if } s \in B \\
  h(s) & \text{if } s \notin A \cup B
\end{pmatrix}.
\tag{1.13}
\]

It is immediate that if preference satisfies \( P4^* \), then all events (in \( \Sigma \)) are unambiguous. The converse is also true if Savage's \( P4 \) is assumed.\(^7\) We view these results as supportive of our definition. The former is noteworthy also because it ensures that our representation result Theorem 1.5.2 extends the main result in [13] (modulo the qualification mentioned in the introduction).

**1.3.5 Linearly Unambiguous**

The preceding is useful also for understanding why an alternative seemingly natural definition for 'unambiguous' was not adopted, thereby providing further perspective on our chosen definition. We motivated our definition in part by the suggestion that a necessary condition for an event to be unambiguous is that it be 'separable' from events in its complement. The following alternative definition embodies a stronger form of separability and therefore warrants some attention:

**Definition 1.3.4** An event \( T \) is **linearly unambiguous** if: (a) For all acts \( f' \) and \( f \) and all outcomes \( z \) and \( z' \),
\[
\begin{pmatrix}
  f'(s) & \text{if } s \in T^c \\
  z & \text{if } s \in T
\end{pmatrix}
\preceq
\begin{pmatrix}
  f(s) & \text{if } s \in T^c \\
  z & \text{if } s \in T
\end{pmatrix}
\implies
\tag{1.14}
\]

\(^7\)We could obtain the unqualified equivalence between \( P4^* \) and 'all events are unambiguous' if we strengthened our definition slightly to require invariance in (1.4) also if the outcomes \( x^* \) and \( x \) are replaced by \( y^* \) and \( y \), where \( y^* \succ y \). Our results below remain valid with this alternative definition. The distinction between the two alternative definitions of unambiguous seems to us to be a matter of taste.
1.3. UNAMBIGUOUS EVENTS

\[
\begin{pmatrix}
    f'(s) & \text{if } s \in T^c \\
    z' & \text{if } s \in T
\end{pmatrix} \geq
\begin{pmatrix}
    f(s) & \text{if } s \in T^c \\
    z' & \text{if } s \in T
\end{pmatrix};
\]

and (b) The condition obtained if \( T \) is everywhere replaced by \( T^c \) in (a) is also satisfied. Otherwise, say that \( T \) is linearly ambiguous.

It is apparent that if \( T \) is linearly unambiguous, then it is also unambiguous. Linear ambiguity is a stronger notion because the indicated invariance is required for all acts \( f' \) and \( f \) and not just for the subclass of ‘conditional binary acts’ as in (1.4). The economic significance of this difference is described below and is similar to the discussion in [13, Section 4.2]. In any event, it is apparent that linear ambiguity embodies a stronger form of separability than does ambiguity. So why not use it as the key notion?

One answer is that (1.14) is too demanding to correspond to the intuitive notion of ambiguity. The invariance required by (1.14) may be violated because the decision-maker views outcomes in different states as complementary or substitutable for reasons that have nothing to do with ambiguity. For example, she may be probabilistically sophisticated, thus assigning probabilities to all events and translating any act into the induced lottery over outcomes, but then she might evaluate the lottery by a risk-preference utility functional that is not linear in probabilities (violating the Independence Axiom). Decision-makers who behave as in the Allais Paradox are of this sort. Many events would be linearly ambiguous for such decision-makers, though it seems intuitively that ambiguity has nothing to do with their preferences. In contrast, for such (probabilistically sophisticated) decision-makers all events are unambiguous, as we have just observed. Roughly speaking, our formal definition of ambiguity relates to behavior in the Ellsberg Paradox but not the Allais Paradox, while linear ambiguity confounds the two.\(^8\)

A second (related) answer is that the existence of probabilistically sophisticated non-expected utility maximizers is at least a logical possibility if not an empirical fact. Such preferences are based on probabilities. A theory of subjective probability is richer if it includes them.

That is not to dispute the potential usefulness of the notion of linear ambiguity. Zhang [21], who originated the notion, shows that it can help to provide a fully

---

\(^8\) Readers who attach little importance to the Allais Paradox may feel that such a view would justify using (1.14) in place of (1.4). We feel that it argues for imposing a form of the sure-thing principle on the subdomain of unambiguous acts, rather than for changing the meaning of unambiguous.
subjective expected utility theory, in the sense of the introduction. (See Section 1.6 for further discussion.) The linearity of the expected utility function explains our choice of terminology.

1.3.6 ‘Unambiguous’ and Updating

We observed in the introduction that one would expect the set of unambiguous events to be closed with respect to complements and disjoint unions. Closure with respect to complements is built into the definition of $\mathcal{A}$. For disjoint unions, it is readily shown that if $T_1$ and $T_2$ are disjoint and each satisfies (1.4), then so does $T_1 \cup T_2$. This does not prove that the union is in $\mathcal{A}$, because that requires also the appropriate form of (1.4) for $(T_1 \cup T_2)^c$; to prove this we employ a mild axiom as described in Lemma 1.5.1 below. For the next intuitive argument, we anticipate the fact that $\mathcal{A}$ will shortly be shown to be a $\lambda$-system.

Examination of updating is supportive of our definition. Imagine that prior to the eventual realization of the true state of the world, there is an intermediate stage at which the decision-maker learns that event $T$ is true. At that point the relevant prospects are acts over $T$. Denote the set of such acts by $\mathcal{F}_T$ and by $\succeq_T$ the updated preference order over $\mathcal{F}_T$. We assume that updating takes the form prescribed by Machina [12]. Thus assume that $\succeq_T$ is defined by: For all $f', f \in \mathcal{F}_T$,

$$f' \succeq_T f \iff \begin{cases} f'(s) & \text{if } s \in T \\ h(s) & \text{if } s \in T^c \end{cases} \succeq \begin{cases} f(s) & \text{if } s \in T \\ h(s) & \text{if } s \in T^c \end{cases}.$$  

Here $h$ is the foregone unrealized alternative, an act over $T^c$. The updated order depends on $h$, but this dependence is suppressed in the notation.

We can now ask the following: Suppose that $T$ and $R \subset T$ are unambiguous with respect to the initial order $\succeq$. Then is $R$ necessarily unambiguous from the perspective of the updated order $\succeq_T$? (The latter expression is defined in the obvious way, using $T$ as the new state space.) Intuition suggests a positive answer and that is the case for our definition.

**Lemma 1.3.5** Let $T$ and $R \subset T$ be in $\mathcal{A}$, that is, unambiguous with respect to the initial order $\succeq$; update according to (1.15). Suppose that $\mathcal{A}$ is a $\lambda$-system (see Lemma 1.5.1). Then $R$ is unambiguous with respect to $\succeq_T$. 
The proof is elementary. The added hypothesis that \( \mathcal{A} \) is a \( \lambda \)-system is used only to show that \( T \setminus R = (R \cup T^c)^c \) is also in \( \mathcal{A} \).

### 1.3.7 Multiple-Priors and Choquet Expected Utility

Here we examine the nature of unambiguous events when preference is restricted to lie in the multiple-priors class [9] or the Choquet expected utility (CEU) class [18].

A multiple-priors preference order \( \succeq \) is represented by the utility function

\[
U(f) = \min_{p \in P} \int u(f) \, dp,
\]

where \( u : \mathcal{X} \rightarrow \mathcal{R}^1 \) and where \( P \) is a convex (and suitably closed) set of probability measures on \( (S, \Sigma) \). Say that all measures in \( P \) agree on an event \( T \) if \( pT = p'T \) for all \( p \) and \( p' \) in \( P \). The following result is intuitive.

**Lemma 1.3.6** The event \( T \) is unambiguous if all measures in the set \( P \) of priors agree on \( T \).

**Proof.** This follows from

\[
U(f) = \min_p \{ u(x^*) \, pA + u(x) \, pB + \int_{T^c \setminus (A \cup B)} u(h) \, dp \} + q_T u(z),
\]

where \( f \) is the first act appearing in (1.4) and \( q_T \) is the agreed probability of \( T \). ■

We have more to say about the CEU model. Let preference be represented by \( U_{\text{ceu}} \), where\(^9\)

\[
U_{\text{ceu}}(f) = \int_S u(f) \, d\nu.
\]

Here \( \nu : \Sigma \rightarrow [0, 1] \) is a capacity and \( u : \mathcal{X} \rightarrow \mathcal{R}^1 \). It is convenient to restrict attention to the special case where \( u(\mathcal{X}) \) has nonempty interior.

---

\(^9\)\( \nu \) is a capacity if it maps \( \Sigma \) into \( [0, 1] \), \( \nu(E') \geq \nu(E) \) whenever \( E' \supset E \), \( \nu(\emptyset) = 0 \) and \( \nu(S) = 1 \). The indicated integral is a Choquet integral and equals \( \sum_{i=1}^{n} u_i \left[ \nu(\bigcup_{j=1}^{n} E_j) - \nu(\bigcup_{j=m+1}^{n} E_j) \right] \) if \( E_i = \{ s : u(f(s)) = u_i \} \) and \( u_1 < \ldots < u_n \).
The CEU and multiple-priors models overlap. Say that \( \nu \) is convex if
\[
\nu(A \cup B) + \nu(A \cap B) \geq \nu(A) + \nu(B),
\]
for all events \( A \) and \( B \). The utility function \( U^{CEU} \) is also a multiple-priors model if and only if \( \nu \) is convex, in which case the relevant set of priors is \( \text{core}(\nu) \), the set of all finitely additive probability measures \( p \) on \( S \) satisfying
\[
p(A) \geq \nu(A), \quad \text{for all } A \subseteq S.
\]
A property of capacities that is weaker than convexity is exactness (see [17]): \( \nu \) is exact if
\[
\nu(A) = \inf_{m \in \text{core}(\nu)} m(A), \quad \text{for all } A \subseteq S. \tag{1.16}
\]

Two special cases of CEU preferences merit special mention because they lie at opposite extremes in terms of the extent of subjective ambiguity that they reflect. The first, corresponding to probabilistic sophistication, has
\[
\nu = \phi(p), \tag{1.17}
\]
for some probability measure \( p \) and some increasing \( \phi : [0, 1] \rightarrow [0, 1] \). As noted in Section 1.3.4, all events in \( \Sigma \) are unambiguous in this case, that is, \( \mathcal{A} = \Sigma \). At the other extreme, suppose the capacity is given by
\[
\nu(A) = \begin{cases} 
1 & \text{if } A = S \\
0 & \text{otherwise},
\end{cases}
\]
which is often described as modelling complete ignorance. Justification for this name is provided by noting that the implied utility function \( U \) satisfies \( U(f) = \inf_S u(f(s)) \); as well, the core of \( \nu \) is the set of all probability measures on \( S \). It is intuitive that complete ignorance should mean that all events other than \( \emptyset \) and \( S \) be subjectively ambiguous. That is indeed the case - it is readily verified that \( \mathcal{A} = \{ \emptyset, S \} \) for this preference order.

Returning to general capacities, define the following classes of events:
\[
\mathcal{A}_0 = \{ T \subseteq S : \nu(T + A) = \nu T + \nu A \text{ for all } A \subseteq T^c \},
\]

\footnote{Conversely, if all events are unambiguous and if \( \nu \) is convex-ranged, then (1.17) is implied for some convex-ranged probability measure \( p \). This follows from [13] because, as we have already noted, if all events are unambiguous (and if \( P4 \) is satisfied as it is by the Choquet model), then Machina and Schmeidler's axiom \( P4^* \) is implied.}
1.3. UNAMBIGUOUS EVENTS

\[ A_1 = \{ T \subset S : \nu T + \nu T^c = 1 \} \]
\[ A_2 = \cap_{m \in \text{core}(\nu)} \{ A \subset S : mA = \nu A \}, \]
consisting of those events on which all measures in the core of \( \nu \) agree. In general, \( A_0 \subset A_1 \). When \( \nu \) is exact, these three sets coincide.\(^{11}\)

**Lemma 1.3.7** If \( \nu \) is exact, then \( A_0 = A_1 = A_2 \).

A consequence is the following relation to our notion of unambiguous:

**Lemma 1.3.8** If \( A_0 \) is closed with respect to complements, then \( A_0 \subset A \). If \( \nu \) is exact, then \( A_1 \subset A \).

The conditions defining \( A_0 \) and \( A_1 \) are relatively easy to interpret, but they are only sufficient for 'unambiguous'. That necessity fails is illustrated by taking \( \nu \) to be a distortion of a probability measure as in (1.17), in which case all events are unambiguous, but \( A_1 \) may equal \( \{ 0, S \} \). The class \( A_1 \) equals the set of unambiguous events as defined in [8]. Zhang [21] shows that if \( A_0 \) is closed with respect to complements, then it coincides with the class of linearly unambiguous events. Under additional assumptions (see Corollary 1.5.4), the events termed unambiguous in [21] and [8] coincide with those in \( A \). More generally, we have the following characterization of \( A \) (see Appendix 1.7.1 for a proof). We use '+' to denote disjoint union.

**Lemma 1.3.9** \( T \) is unambiguous if and only if: For all pairwise disjoint events \( A, B, C \) and \( D \), each disjoint from \( T \),

\[
\nu(A + D) \geq \nu(B + D) \iff \nu(A + D + T) \geq \nu(B + D + T); \quad \text{and} \quad (1.18)
\]

if \( (\nu(A + D) - \nu(B + D)) (\nu(A + D + C) - \nu(B + D + C)) < 0 \), then

\[
\nu(A + D) - \nu(B + D) = \nu(A + D + T) - \nu(B + D + T) \quad \text{and} \quad (1.19)
\]

\[
\nu(A + D + C) - \nu(B + D + C) = \nu(A + D + C + T) - \nu(B + D + C + T); \quad (1.20)
\]

and the above conditions are satisfied also by \( T^e \).

\(^{11}\)This lemma appears in [8].
Observe that all of these conditions are independent of \( u \), which therefore has nothing to do with ambiguity.

This characterization admits the following incomplete interpretation: If the 'reversal' in (1.19) never occurs, then \( \nu \) is (almost) a qualitative probability within \( T^c \). In that case, the ordinal condition (1.18) alone corresponds to 'T unambiguous'. But when \( \nu \) fails to be a qualitative probability within \( T^c \), then 'T unambiguous' requires that the cardinal conditions in (1.20) and (1.21) obtain.

### 1.4 AXIOMS

Here we specify some axioms for the preference order \( \succeq \). They will deliver not only the \( \lambda \)-system properties for \( \mathcal{A} \) and a probability measure on these unambiguous events, but also the probabilistic sophistication of preference restricted to unambiguous acts. This representation result is the ultimate justification for our definition of 'unambiguous'. In particular, it confirms our rough intuition that unambiguous events are assigned probabilities, although probabilities may exist also on other collections of events.

The set \( \mathcal{F}^{ua} \) of unambiguous acts is defined by (1.3). Also useful, for any given \( A \in \mathcal{A} \), is the set of acts

\[
\mathcal{F}_A^{ua} = \{ f \in \mathcal{F} : f^{-1}(X) \cap A \in \mathcal{A} \text{ for all } X \subseteq \mathcal{X} \}. 
\]

Denote by \( x \in \mathcal{X} \) both the outcome and the constant act producing the outcome \( x \) in every state. Preference statements like \( x \preceq y \) are therefore well-defined and have the obvious meaning.

Some of the axioms for \( \succeq \) are slight variations of Savage's axioms, with names adapted from Machina and Schmeidler. Though the axioms are expressed in terms of \( \mathcal{A} \), they constitute assumptions about \( \succeq \) because \( \mathcal{A} \) is derived from \( \succeq \). A final remark is that the axioms relate primarily to \( \succeq \) restricted to \( \mathcal{F}^{ua} \).

**Axiom 1 (Monotonicity):** For all outcomes \( x \) and \( y \), non-null events \( A \in \mathcal{A} \) and

---

\(^{12}\)Excluding the case where there is a weak inequality in (1.19).
acts \( g \in \mathcal{F}^{ua}_{\mathcal{A}^c} \)

\[
\begin{pmatrix}
  x & \text{if } s \in A \\
g(s) & \text{if } s \in A^c
\end{pmatrix} \gtrless \begin{pmatrix}
y & \text{if } s \in A \\
g(s) & \text{if } s \in A^c
\end{pmatrix} \iff x \geq y.
\]

**Axiom 2 (Nondegeneracy):** There exist outcomes \( x^* \) and \( x \) such that \( x^* \succ x \).

**Axiom 3 (Weak Comparative Probability):** For all events \( A, B \in \mathcal{A} \) and outcomes \( x^* \succ x \) and \( y^* \succ y \)

\[
\begin{pmatrix}
x^* & \text{if } s \in A \\
x & \text{if } s \in A^c
\end{pmatrix} \gtrless \begin{pmatrix}
x^* & \text{if } s \in B \\
x & \text{if } s \in B^c
\end{pmatrix} \iff
\begin{pmatrix}
y^* & \text{if } s \in A \\
y & \text{if } s \in A^c
\end{pmatrix} \gtrless \begin{pmatrix}
y^* & \text{if } s \in B \\
y & \text{if } s \in B^c
\end{pmatrix}.
\]

As is well known, this axiom delivers the complete and transitive likelihood relation \( \succeq_\ell \) on \( \mathcal{A} \), where \( A \succeq_\ell B \) if \( \exists x^* \succ x \) such that

\[
\begin{pmatrix}
x^* & \text{if } s \in A \\
x & \text{if } s \in A^c
\end{pmatrix} \succeq \begin{pmatrix}
x^* & \text{if } s \in B \\
x & \text{if } s \in B^c
\end{pmatrix}.
\]

Another notable consequence of Weak Comparative Probability, that is specific to our setting, is that it immediately implies that the Machina-Schmeidler axiom \( P4^* \) is satisfied on the domain \( \mathcal{F}^{ua} \) of unambiguous acts, that is, the implication (1.13) is valid for all events \( A \) and \( B \) in \( \mathcal{A} \) and all acts \( g \) and \( h \) in \( \mathcal{F}^{ua}_{(A \cup B)^c} \). It might seem at first glance that this renders the remaining route to a suitable representation result routine, because given the key axiom \( P4^* \), the remaining and relatively uncontentious Machina-Schmeidler axioms could be assumed and their result invoked. However, their arguments (suitably translated) rely on \( \mathcal{A} \) being a \( \sigma \)-algebra, which is not generally the case when \( \mathcal{A} \) is the set of unambiguous events.

The next axiom imposes suitable richness of the set of unambiguous events. It is clear from Savage’s analysis that some richness is required to derive a probability measure on \( \mathcal{A} \). Further, Savage’s axiom \( P6 \) (suitably translated) is not adequate here because \( \mathcal{A} \) is not a \( \sigma \)-algebra. However, the spirit of Savage’s \( P6 \) is retained in the next axiom.
Axiom 4 (Small Unambiguous Event Continuity): Let \( f, g \in \mathcal{F}^{ua} \), \( f \succ g \), with \( f = (x_1, A_1; x_2, A_2; \ldots; x_n, A_n) \), \( g = (y_1, B_1; y_2, B_2; \ldots; y_m, B_m) \), where each \( A_i \) and \( B_i \) lies in \( \mathcal{A} \). Then for any outcome \( x \in \mathcal{X} \), there exist two partitions \( \{C_i\}_{i=1}^N \) and \( \{D_j\}_{j=1}^M \) of \( S \) in \( \mathcal{A} \) that refine \( \{A_i\}_{i=1}^n \) and \( \{B_j\}_{j=1}^m \) respectively, and satisfy:

\[
f \succ \begin{cases} x & \text{if } s \in D_k \\ g(s) & \text{if } s \in D_k^c \end{cases}, \quad \text{for all } k \in \{1, \ldots, N\};
\]

and

\[
\begin{cases} x & \text{if } s \in C_j \\ f(s) & \text{if } s \in C_j^c \end{cases} \succ g, \quad \text{for all } j \in \{1, \ldots, M\}.
\]

Very roughly, the axiom requires that unambiguous events can be decomposed into suitably 'small' unambiguous events. When \( \mathcal{A} \) is closed with respect to intersections, as in the standard model [16] or [13] where it is taken to be the power set, then the axiom is implied by Savage's P6, given Axioms 1-3.\(^{13}\)

The preceding axioms are familiar, at least when imposed on all of \( \mathcal{F} \), rather than just on \( \mathcal{F}^{ua} \) as here. The remaining two axioms are 'new' and are needed to accommodate the fact that \( \mathcal{A} \) may not be a \( \sigma \)-algebra.

Say that a sequence \( \{f_n\}_{n=1}^\infty \) in \( \mathcal{F}^{ua} \) converges in preference to \( f_\infty \in \mathcal{F}^{ua} \) if: For any two acts \( f_\ast, f^* \) in \( \mathcal{F}^{ua} \) satisfying \( f_\ast \prec f_\infty \prec f^* \), there exists an integer \( N \) such that

\[ f_\ast \prec f_n \prec f^*, \quad \text{whenever } n \geq N. \]

Axiom 5 (Monotone Continuity): For any \( A \in \mathcal{A} \), outcomes \( x_\ast \succ x \), act \( h \in \mathcal{F}_A^{ua} \) and decreasing sequence \( \{A_n\}_{n=1}^\infty \) in \( \mathcal{A} \) with \( A_1 \subseteq A \), define

\[
f_n = \begin{cases} x_\ast & \text{if } s \in A_n \\ x & \text{if } s \in A \setminus A_n \\ h(s) & \text{if } s \in A^c \end{cases} \quad \text{and} \quad f_\infty = \begin{cases} x_\ast & \text{if } s \in \bigcap_{n=1}^\infty A_n \\ x & \text{if } s \in A \setminus \bigcap_{n=1}^\infty A_n \\ h(s) & \text{if } s \in A^c \end{cases}.
\]

If \( f_n \in \mathcal{F}^{ua} \) for all \( n = 1, 2, \ldots \), then \( \{f_n\}_{n=1}^\infty \) converges in preference to \( f_\infty \) and \( f_\infty \in \mathcal{F}^{ua} \).

\(^{13}\)Savage's P6 applied to \( \mathcal{F}^{ua} \) would require that, given \( x \), if \( f = (x_1, A_i)_{i=1}^n \succ g = (y_1, B_i)_{i=1}^m \), where every \( A_i \) and \( B_i \) lies in \( \mathcal{A} \), then there exists a partition \( \{G_i\}_{i=1}^N \) of \( S \) in \( \mathcal{A} \) such that \( f \succ (x, G_k; g, G_k^c) \) for all \( k \). Given such a partition, and given that \( \mathcal{A} \) is closed with respect to intersections, then the collection of events \( D_{ij} = G_i \cap B_j \) satisfies the requirements in Axiom 4.
The name Monotone Continuity describes one aspect of the axiom, that requiring the indicated convergence in preference. The second component of the axiom is the requirement that the limit $f_{\infty}$ lie in $F^w$ whenever each $f_n$ is unambiguous. This will serve in particular to ensure that $\mathcal{A}$ satisfies the 'countable' closure condition $\lambda.3$ or $\lambda.3'$ required by the definition of a $\lambda$-system.

It might be felt that given the correct definition of 'unambiguous', the derivation of a probability measure on $\mathcal{A}$ should be possible with little more than some richness requirements. The axioms stated thus far can arguably be interpreted as constituting such minimal requirements. However, they do not suffice and we need one final axiom. This may reflect the fact that only some aspects of 'unambiguous' are captured in our definition. In any event, the final axiom is intuitive and arguably weak. Its statement requires some preliminaries.

A finite partition with component events from $\mathcal{A}$ is denoted $\{A_i\}$. Henceforth all partitions have unambiguous components, even where not stated explicitly. Given such a partition, use the obvious abbreviated notation $(x_i, A_i)$. For any permutation $\sigma$ of $\{1, \ldots, n\}$, $(x_{\sigma(i)}, A_i)$ denotes the act obtained by permuting outcomes between the events. Say that the finite partition $\{A_i\}$ is a uniform partition if $A_i \sim A_j$ for all $i$ and $j$ and call $\{A_i\}$ strongly uniform if in addition it satisfies: For all outcomes $(x_1, A_i)$ and for all permutations $\sigma$,

$$(x_{\sigma(i)}, A_i) \sim (x_1, A_i). \quad (1.24)$$

In particular, if $\{A_i\}_{i=1}^n$ is a strongly uniform partition, then for all index sets $I$ and $J$, subsets of $\{1, 2, \ldots, n\}$,

$$\cup_{i \in I} A_i \sim_{\mathcal{E}} \cup_{i \in J} A_i \text{ if } |I| = |J|.$$  

**Axiom 6 (Strong-Partition Neutrality):** For any two strongly uniform partitions $\{A_i\}_{i=1}^n$ and $\{B_i\}_{i=1}^n$, if $A_i \sim_{\mathcal{E}} B_i$ for all $i$, then for all $(x_i)$,

$$\begin{pmatrix} x_1 & \text{if } s \in A_1 \\ x_2 & \text{if } s \in A_2 \\ \vdots & \vdots \\ x_n & \text{if } s \in A_n \end{pmatrix} \sim \begin{pmatrix} x_1 & \text{if } s \in B_1 \\ x_2 & \text{if } s \in B_2 \\ \vdots & \vdots \\ x_n & \text{if } s \in B_n \end{pmatrix}. \quad (1.25)$$

\(^{14}\text{Machina and Schmeidler (p. 771) refer to a related axiom with the same name that is used by Arrow [2, p. 48] to deliver the countable additivity of the subjective probability measure. Here as well, countable additivity will follow from Monotone Continuity, but as an unintentional by-product.}\)
The hypothesis that the $A_i$'s and $B_i$'s satisfy (1.24) expresses another sense in which these events are unambiguous. This makes the conclusion (1.25) natural and much weaker than if the indifference in (1.25) were required for all uniform partitions. The latter axiom would go a long way towards explicitly imposing probabilistic sophistication, an unattractive feature in the present exercise where the intention is that probabilistic sophistication on unambiguous acts should result primarily from the definition of 'unambiguous'. Axiom 6 is much less vulnerable to such a criticism.

To support the claim that Strong-Partition Neutrality is a 'weak' axiom, observe that it is satisfied by all CEU orders, proving that it falls far short of imposing probabilistic sophistication. The reason is that if $\{A_i\}$ is a strongly uniform partition, and thus satisfies (1.24), then $\nu$ is additive on the algebra generated by the partition. Thus the indifference (1.25) is implied.

We turn next to the implications of these axioms.

1.5 PROBABILISTIC SOPHISTICATION ON UNAMBIGUOUS ACTS

Define “probabilistic sophistication on unambiguous acts $F_{ua}$” by extending the definition of Machina and Schmeidler. For the convenience of the reader, the complete definition is stated here.

Some preliminary notions are required. Denote by $D(\mathcal{X})$ the set of probability distributions on $\mathcal{X}$ having finite support. A probability distribution $P = (x_1, p_1; \ldots; x_m, p_m)$ is said to first-order stochastically dominate $Q = (y_1, q_1; \ldots; y_n, q_n)$ with respect to the order $\succeq$ over the outcome set $\mathcal{X}$ if

$$\sum_{i:x_i \preceq x} p_i \leq \sum_{j:y_j \preceq x} q_j \quad \text{for all } x \in \mathcal{X}.$$  

Use the term strict dominance if the above holds with strict inequality for some $x^* \in \mathcal{X}$.

Given a real-valued function $W$ defined on a mixture subspace $dom(W)$ of $D(\mathcal{X})$, say that $W$ is mixture continuous if for any distributions $P$, $Q$ and $R$ in $dom(W)$,
1.5. PROBABILISTIC SOPHISTICATION ON UNAMBIGUOUS ACTS

the sets
\[
\{ \lambda \in [0, 1] : W(\lambda P + (1 - \lambda)Q) \geq W(R) \} \quad \text{and} \\
\{ \lambda \in [0, 1] : W(\lambda P + (1 - \lambda)Q) \leq W(R) \}
\]
are closed. Say that \( W \) is monotonic (with respect to stochastic dominance) if \( W(P)(>) \geq W(Q) \) whenever \( P \) (strictly) stochastically dominates \( Q \), \( P \) and \( Q \) in \( \text{dom}(W) \).

Given a probability measure \( p \) on \( \mathcal{A} \), denote by \( P_{f,p} \in D(\mathcal{X}) \) the distribution over outcomes induced by the act \( f \). Define
\[
D_p^{ua}(\mathcal{X}) = \{ P_{f,p} : f \in \mathcal{F}^{ua} \}.
\]
When \( p \) is convex-ranged, then \( D_p^{ua}(\mathcal{X}) \) is a mixture space.

We can finally state the desired definition. Say that \( \succeq \) is probabilistically sophisticated on \( \mathcal{F}^{ua} \) if there exists a convex-ranged probability measure \( p \) on \( \mathcal{A} \) and a real-valued, mixture continuous and monotonic function \( W \) on \( D_p^{ua}(\mathcal{X}) \) such that \( \succeq \) has utility function \( U \) of the form
\[
U(f) = W(P_{f,p}). \tag{1.26}
\]
We remind the reader that because \( \mathcal{A} \) and \( \mathcal{F}^{ua} \) are derived from the given primitive preference relation \( \succeq \) on \( \mathcal{F} \), probabilistic sophistication so-defined is a property of \( \succeq \) exclusively and does not rely on an exogenous specification of 'unambiguous acts'\(^{16}\).

Probabilistic sophistication with measure \( p \) implies that likelihood (or the ranking of unambiguous bets) is represented by \( p \); that is,
\[
A \succeq_l B \iff pA \geq pB, \quad \text{for all } A, B \in \mathcal{A}.
\]
But (1.26) is much stronger, requiring that the ranking of all (not necessarily binary) unambiguous acts be based on \( p \).

The most important special case of (1.26) is subjective expected utility, where \( W \) is an expected utility function on lotteries \( D_p^{ua}(\mathcal{X}) \) and thus \( U \) has the familiar form
\[
U(f) = \int_S u(f) \, dp. \tag{1.27}
\]
\(^{16}\)If \( \mathcal{A} \) and \( \mathcal{F}^{ua} \) are replaced by \( \Sigma \) and \( \mathcal{F} \) respectively, then one obtains the Machina-Schemidler definition of 'global' (i.e., on \( \mathcal{F} \)) probabilistic sophistication, with the minor difference that they take \( \Sigma = 2^5 \).
Turn to the implications of our axioms. A preliminary result (proven in Appendix 1.7.2) is that they imply that \( A \) is a \( \lambda \)-system.

**Lemma 1.5.1** Under Axioms 2, 4 and 5, \( A \) is a \( \lambda \)-system. In particular, if \( T_1 \) and \( T_2 \) are disjoint unambiguous events, then \( T_1 \cup T_2 \) is unambiguous.

The following is our main result:

**Theorem 1.5.2** Let \( \succeq \) be a preference order on \( F \) and \( A \) the corresponding set of unambiguous events. Then the following two statements are equivalent:

(a) \( \succeq \) satisfies axioms 1-6.

(b) \( A \) is a \( \lambda \)-system and there exists a (unique) convex-ranged and countably additive probability measure \( p \) on \( A \) such that \( \succeq \) is probabilistically sophisticated on \( \mathcal{F}^{ua} \) with underlying measure \( p \).

We would like to clarify the sense in which we have established probabilistic sophistication on an endogenous domain. One way to accomplish this is as follows: Identify from preference some class of events \( A' \) and the corresponding set of measurable acts \( \mathcal{F}' \). Then impose the Machina and Schmeidler axioms for preference \( \succeq \) restricted to \( \mathcal{F}' \), making the changes necessary to allow for the fact that \( A' \) may not be an algebra. Such an exercise would also generalize [13] and may be of some value. But this is decidedly not the nature of our contribution. For example, our representation result does not explicitly impose the key axiom \( P4^* \) used in [13]. Admittedly, \( P4^* \) on \( \mathcal{F}^{ua} \) is implied by the definition of \( \mathcal{F}^{ua} \) and by Weak Comparative Probability, but we view this as confirmation of the appropriateness of our definition of 'unambiguous.'

The bulk of the proof of Theorem 1.5.2 is found in Appendices 1.7.3 and 1.7.4. The arguments in [16], [6] and [13] must be modified because only in the present setting is the relevant class of events \( A \) not closed with respect to intersections. A key step is to show (Appendix 1.7.3) that our axioms for preference deliver the conditions for the implied likelihood relation that are used in [22] in order to obtain a representing probability measure. The proof of probabilistic sophistication is completed in Appendix 1.7.4.
We conclude this section with two corollaries (proofs are provided in Appendix 1.7.5). For the first, though we have derived a subjective probability measure on $A$, one may wonder whether this domain is maximal in this respect. We proceed to describe a result that is related to this question, though it does not provide a definitive answer to it.

Given the probability measure $p$ on $A$, one might approximate the likelihoods of ambiguous events by means of the inner and outer measures $p_*$ and $p^*$ defined as follows:\footnote{See [19] and [5]. For an application to decision theory, see [21].} For each $E$ in $\Sigma$, 
\[
p_*(E) = \sup \{p(A) : A \in A, A \subseteq E\}
\]
and 
\[
p^*(E) = \inf \{p(A) : A \in A, E \subseteq A\}.
\]
These non-additive measures provide intuitive lower and upper bounds for the likelihood assessment of $E$; in particular, $p_*(E) \leq p^*(E)$. Define 
\[\overline{A} = \{E : p_*(E) = p^*(E)\}.
\]
Events in $\overline{A}$ seem intuitively to be 'unambiguous'. The definition yields $\overline{A} \supset A$. Given our axioms, we can prove equality.

**Corollary 1.5.3** Let $\succeq$ be as in Theorem 1.5.2. Then $\overline{A} = A$.

Next we apply our theorem to the CEU model. Recall the notation in Section 1.3.7. Say that a capacity $\nu$ is **convex-ranged on** $A$ if for every $A$ in $A$,
\[\quad [0, \nu A] = \{\nu B : B \in A, B \subseteq A\}.
\]
Say that $\nu$ is **chain-continuous** if for all events in $S$,
\[
\nu A_n \searrow \nu(\cap A_n) \text{ if } A_n \searrow \quad \text{and} \\
\nu A_n \nearrow \nu(\cup A_n) \text{ if } A_n \nearrow.
\]
Corollary 1.5.4 Let \( \succeq \) be a CEU preference order with capacity \( \nu \) and subjectively unambiguous events \( \mathcal{A} \).
(a) Suppose that \( \nu \) is chain-continuous and convex-ranged on \( \mathcal{A} \). Then there exists a convex-ranged and countably additive probability measure \( p \) on \( \mathcal{A} \) and an increasing and onto map \( \phi : [0, 1] \to [0, 1] \) such that \( \nu = \phi(p) \) on \( \mathcal{A} = \overline{\mathcal{A}} \).
(b) Suppose in addition that \( \nu(\mathcal{A}_0) = [0, 1] \) and that \( \mathcal{A}_0 \) is closed with respect to complements. Then \( \nu = p \) on \( \mathcal{A} \).
(c) Suppose in addition to the assumptions in (a) and (b) that \( \nu \) is exact. Then \( \nu = p \) on \( \mathcal{A} \) and
\[
\mathcal{A} = \mathcal{A}_0 = \mathcal{A}_1 = \mathcal{A}_2.
\]
(d) Suppose that \( \nu \) is exact and satisfies the assumptions in part (a). Then \( \overline{\mathcal{A}} = \mathcal{A} \).

Under the conditions in part (a), the CEU order satisfies the axioms in Theorem 1.5.2, showing that the latter's scope extends beyond globally probabilistically sophisticated preferences.\(^\text{18}\) We repeat that the proof of existence of \( p \) is not standard because \( \mathcal{A} \) need not be a \( \sigma \)-algebra. (In (c), \( \mathcal{A} \) is a \( \sigma \)-algebra if \( \nu \) is convex, but not more generally.) Part (b) gives conditions under which the CEU preference order is expected utility (and not merely probabilistically sophisticated) on the domain of unambiguous acts. Part (c) goes further and gives conditions under which \( \mathcal{A} \) coincides with the families indicated and discussed earlier.

1.6 FURTHER RESEARCH

Improved axiomatization: We suspect that it may be possible to weaken the final axiom Strong-Partition-Neutrality.

Axiomatization of expected utility on subdomain of linearly unambiguous acts: Denote by \( \mathcal{A}^* \subset \mathcal{A} \) the class of linearly unambiguous events and by \( \mathcal{F}^{\text{tua}} \subset \mathcal{F}^{\text{ua}} \) the corresponding set of acts. (In fact, one can weaken the definition of linearly unambiguous by restricting the required invariance to hold only for acts that are \( \mathcal{A} \)-measurable.)

\(^{18}\) Under the conditions specified in Lemma 1.3.3, the two-urn example defined by (1.10)-(1.9) and the surrounding discussion, provides another example of a preference order satisfying the axioms in Theorem 1.5.2.
1.6. FURTHER RESEARCH

Assume the axioms 1-6 and add appropriate reformulations of Small Unambiguous Event Continuity and Monotone Continuity obtained by substituting $A^*$ for $A$ in the existing statements of these axioms. Then the proof of Theorem 1.5.2 may be routinely adapted to derive an expected utility representation on $F^{ua}$. This is roughly the route followed in [21].  It remains to see whether alternative expected utility representation results are possible.

Multiple domains of probabilistic sophistication: Ultimately, we would like to derive from preference all subdomains of acts where probabilistic sophistication prevails. This likely requires as a primitive notion a binary relation on events that might be thought of as ‘equally unambiguous’. For example, one might say that $A$ and $B$ are equally unambiguous if (1.6) is satisfied for all $C$ disjoint from $A \cup B$. Undoubtedly, this definition would require refinement. This is an important direction for future research.

Representation of preference on all (ambiguous) acts: Zhang [21] provides a characterization of preference on all of $F$ that has the CEU form for a special class of capacities. Much remains to be done in this direction.

For instance, Theorem 1.5.2 leaves open the question whether the axioms deliver any structure for the ranking of ambiguous acts. Conjectures along the following lines require further study.

Conjecture 1.6.1 Let $A_0$ be any $\lambda$-system of events, $F^{ua}_0$ the class of $A_0$-measurable acts and $\succeq^{ua}$ a preference relation (complete and transitive) on $F^{ua}_0$. Then the following statements are equivalent:

(a) There exists a convex-ranked and countably additive probability measure $p$ on $A_0$ such that $\succeq^{ua}$ is probabilistically sophisticated on $F^{ua}_0$ with underlying measure $p$.

(b) There exists a preference relation $\succeq$ on $F$ that agrees with $\succeq^{ua}$ on $F^{ua}_0$ and such that its class of subjectively unambiguous events, as defined in this paper, coincides with $A_0$.  

19 In order to clarify the relation between this paper and [21], we point out that the latter does not contain a separate proof of the expected utility representation result. Rather it cites the arguments in this paper, suitably adapted, for a proof.

20 A weaker requirement would replace ‘coincides with’ by ‘contains’, in which case equivalence would presumably no longer hold. Relevant questions: Does $p$ admit an extension to $\Sigma$? Is that extension unique?
The conjecture is related to the question "is any \( \lambda \)-system the family of subjectively unambiguous events for some preference order satisfying our axioms?" If not, what added properties are implied for such families? Finally, what if 'some preference order' is replaced by 'some CEU order'? In other words, the problem is to characterize the classes \( \mathcal{A} \) that coincide with the families of subjectively unambiguous events for some Choquet preference order.

\textit{Ambiguity aversion:} In the same way that risk aversion is defined by reference to riskless prospects or lotteries, any definition of ambiguity aversion requires the prior identification of unambiguous acts and events. In [4], Epstein defined the notion of aversion to ambiguity (or uncertainty) beginning with an exogenously specified \( \lambda \)-system of unambiguous events. A natural next step is to combine that paper with this one by using subjectively unambiguous events as the reference class used to define ambiguity aversion. This approach leads to the intuitive result that any CEU preference order satisfying the conditions of Corollary 1.5.4(c) is ambiguity averse. Further study of ambiguity aversion is beyond the scope of this paper.
1.7. **APPENDIX:**

### 1.7.1 **APPENDIX:** Choquet Expected Utility

**Proof of Lemma 1.3.9:** Sufficiency may be proven by routine verification. We prove necessity.

Let $T$ be unambiguous. If $x^* > x$, then $\nu(A + D) \geq \nu(B + D)$ iff

$$
\left( \begin{array}{c}
x^* \\
x \\
x \\
x^* \\
\end{array} \right) \geq \left( \begin{array}{c}
A + D \\
B \\
T^c \setminus (A + B + D) \\
T \\
\end{array} \right) \iff \\
\left( \begin{array}{c}
x^* \\
x \\
x \\
x^* \\
\end{array} \right) \not\geq \left( \begin{array}{c}
B + D \\
A \\
T^c \setminus (A + B + D) \\
T \\
\end{array} \right)
$$

$$(x^* A + D) \geq (x B + D) \iff (x^* A + T) \geq (x B + T)$$

$$\nu(A + D + T) \geq \nu(B + D + T).$$

Suppose next that (1.19) is satisfied. In fact, suppose that

$$\nu(A + D) < \nu(B + D) \quad \text{and} \quad \nu(A + D + C) > \nu(B + D + C). \quad (1.28)$$

(The other case is similar.)

The event $T$ is unambiguous only if

$$
\left( \begin{array}{c}
x^* \\
x \\
y \\
h(s) \\
z \\
\end{array} \right) \geq \left( \begin{array}{c}
x \\
x^* \\
y \\
h(s) \\
z \\
\end{array} \right) \iff \\
\left( \begin{array}{c}
x \\
x^* \\
y \\
h(s) \\
z \\
\end{array} \right) \not\geq \left( \begin{array}{c}
x \\
x^* \\
y \\
h(s) \\
z \\
\end{array} \right)
$$

(1.29)

a similar ranking obtains for the acts where $z'$ replaces $z$. Suppose that $h$ equals $\bar{y}$ on $D$ and $\bar{y}$ on $T^c \setminus (A + B + C + D)$ and that

$$\bar{y} \succ x^* \succ y \succ x \succ \bar{y}, \quad z = y \quad \text{and} \quad z' = x^*. \quad (1.30)$$
Then by the definition of Choquet integration, the above equivalence becomes

$$u(x^*) \nu(A + D) + u(y) [\nu(C + T + A + D) - \nu(A + D)] + u(x) [1 - \nu(C + T + A + D)] \geq 0$$

if and only if

$$u(x^*) \nu(T + A + D) + u(y) [\nu(C + T + A + D) - \nu(T + A + D)] + u(x) [1 - \nu(C + T + A + D)] \geq 0$$

Equivalently,

$$[u(x^*) - u(y)] (\nu(A + D) - \nu(B + D)) + [u(y) - u(x)] (\nu(C + T + A + D) - \nu(C + T + B + D)) \geq 0$$

iff

$$[u(x^*) - u(y)] (\nu(T + A + D) - \nu(T + B + D)) + [u(y) - u(x)] (\nu(C + T + A + D) - \nu(C + T + B + D)) \geq 0,$$

where this equivalence obtains for all outcomes.

By (1.28) and appropriate forms of (1.18), which has already been proven, conclude that $(\nu(A + D) - \nu(B + D))$ and $(\nu(T + A + D) - \nu(T + B + D))$ are both negative, while $(\nu(C + A + D) - \nu(C + B + D))$ and $(\nu(C + T + A + D) - \nu(C + T + B + D))$ are both positive. Because the range of $u$ has nonempty interior, we can vary the above utility values sufficiently to conclude from the preceding equivalence that

$$\nu(T + A + D) - \nu(T + B + D) = \nu(A + D) - \nu(B + D).$$

Next apply a similar argument for the case

$$\bar{y} \succ x^* \succ y \succ x \succ y, z = x \text{ and } z' = x^*,$$

in place of (1.30). One obtains the equivalence

$$[u(x^*) - u(y)] (\nu(A + D) - \nu(B + D)) + [u(y) - u(x)] (\nu(C + A + D) - \nu(C + B + D)) \geq 0$$
iff

\[ [u(x^*) - u(y)] \left( \nu(T + A + D) - \nu(T + B + D) \right) + \]
\[ [u(y) - u(x)] \left( \nu(C + T + A + D) - \nu(C + T + B + D) \right) \geq 0, \]

where this equivalence obtains for all outcomes. Apply (1.31) and conclude that

\[ \nu(C + A + D) - \nu(C + B + D) = \nu(C + T + A + D) - \nu(C + T + B + D). \]

\hfill \blacksquare

1.7. 1.7.2  \textbf{APPENDIX:} \lambda\text{-System

\textbf{Proof of Lemma 1.5.1:} Suppose that for some disjoint subsets \( A \) and \( B \) of \((T_1 \cup T_2)^c\), act \( h \) and outcomes \( x^*, x, z, z' \in \mathcal{X} \), that

\[
\begin{pmatrix}
    x^* & \text{if } s \in A \\
    x & \text{if } s \in B \\
    h & \text{if } s \in (T_1 \cup T_2)^c \setminus (A \cup B) \\
    z & \text{if } s \in T_1 \\
    z' & \text{if } s \in T_2
\end{pmatrix}
\geq
\begin{pmatrix}
    x & \text{if } s \in A \\
    x^* & \text{if } s \in B \\
    h & \text{if } s \in (T_1 \cup T_2)^c \setminus (A \cup B) \\
    h & \text{if } s \in (T_1 \cup T_2)^c \setminus (A \cup B) \\
    z & \text{if } s \in T_1 \\
    z' & \text{if } s \in T_2
\end{pmatrix}
\]

By (1.4) for \( T_2 \),

\[
\begin{pmatrix}
    x^* & \text{if } s \in A \\
    x & \text{if } s \in B \\
    h & \text{if } s \in (T_1 \cup T_2)^c \setminus (A \cup B) \\
    z & \text{if } s \in T_1 \\
    z' & \text{if } s \in T_2
\end{pmatrix}
\leq
\begin{pmatrix}
    x & \text{if } s \in A \\
    x^* & \text{if } s \in B \\
    h & \text{if } s \in (T_1 \cup T_2)^c \setminus (A \cup B) \\
    h & \text{if } s \in (T_1 \cup T_2)^c \setminus (A \cup B) \\
    z' & \text{if } s \in T_2 \\
    z & \text{if } s \in T_1
\end{pmatrix}
\]

which can be rewritten in the form

\[
\begin{pmatrix}
    x^* & \text{if } s \in A \\
    x & \text{if } s \in B \\
    h & \text{if } s \in (T_1 \cup T_2)^c \setminus (A \cup B) \\
    z' & \text{if } s \in T_2 \\
    z & \text{if } s \in T_1
\end{pmatrix}
\leq
\begin{pmatrix}
    x & \text{if } s \in A \\
    x^* & \text{if } s \in B \\
    h & \text{if } s \in (T_1 \cup T_2)^c \setminus (A \cup B) \\
    h & \text{if } s \in (T_1 \cup T_2)^c \setminus (A \cup B) \\
    z' & \text{if } s \in T_2 \\
    z & \text{if } s \in T_1
\end{pmatrix}
\]

By (1.4) for \( T_1 \),

\[
\begin{pmatrix}
    x^* & \text{if } s \in A \\
    x & \text{if } s \in B \\
    h & \text{if } s \in (T_1 \cup T_2)^c \setminus (A \cup B) \\
    z' & \text{if } s \in T_2 \\
    z & \text{if } s \in T_1
\end{pmatrix}
\leq
\begin{pmatrix}
    x & \text{if } s \in A \\
    x^* & \text{if } s \in B \\
    h & \text{if } s \in (T_1 \cup T_2)^c \setminus (A \cup B) \\
    h & \text{if } s \in (T_1 \cup T_2)^c \setminus (A \cup B) \\
    z' & \text{if } s \in T_2 \\
    z & \text{if } s \in T_1
\end{pmatrix}
\]
Therefore, \( T_2 \cup T_1 \) satisfies the appropriate form of (1.4).

It remains to prove that (1.4) is satisfied also by \((T_1 \cup T_2)^c\). By Small Unambiguous Event Continuity (Axiom 4) applied to the unambiguous events \(T_1\) and \(T_2\), there exists a partition \(\{A_i\}_{i=1}^n\) of \(S\) in \(A\) such that \((T_1 \cup T_2)^c\) equals the finite disjoint union

\[
(T_1 \cup T_2)^c = \bigcup_{A_i \subseteq (T_1 \cup T_2)^c} A_i.
\]

Thus the first part of this proof establishes (1.4) for \((T_1 \cup T_2)^c\).

To complete the proof, it suffices to show that for any \(\{A_n\}_{n=1}^\infty\), a decreasing sequence in \(A\), we have \(\bigcap_{n=1}^\infty A \in A\): By Nondegeneracy, there exist two outcomes \(x^* \succ x\). Then

\[
f_n = \begin{cases} x^* & \text{if } s \in A_n \\ x & \text{if } s \in A_n^c \end{cases} \in \mathcal{F}^{ua}, \text{ for all } n = 1, 2, \ldots
\]

By Monotone Continuity (Axiom 5),

\[
f_\infty = \begin{cases} x^* & \text{if } s \in \bigcap_{n=1}^\infty A_n \\ x & \text{if } s \in (\bigcap_{n=1}^\infty A_n)^c \end{cases} \in \mathcal{F}^{ua}.
\]

Consequently, \(\bigcap_{n=1}^\infty A_n \in A\). □

### 1.7.3 APPENDIX: Existence of Probability

The first step in proving Theorem 1.5.2 is to prove the existence of a probability measure representing \(\succeq_{\ell}\). This appendix states a theorem (proven in [22]) that delivers such a probability measure given suitable properties for \(\succeq_{\ell}\). The theorem extends [6, Theorem 14.2] to the present case of a \(\lambda\)-system of events. Next it is shown that these properties are implied by the axioms adopted for \(\succeq\), as specified in Theorem 1.5.2.

For the following theorem, \(A\) denotes any \(\lambda\)-system and \(\succeq_{\ell}\) is any binary relation on \(A\); that is, they are not necessarily derived from \(\succeq\), though the subsequent application is to that case. Denote by
1.7. APPENDIX:

\[ N(\emptyset) = \{ A \in \mathcal{A} : A \sim_\ell \emptyset \}. \]

**Theorem 1.7.1** There is a unique finitely additive, convex-ranged probability measure \( p \) on \( \mathcal{A} \) such that

\[ A \preceq_\ell B \iff p(A) \geq p(B), \quad \forall A, B \in \mathcal{A} \]

if (and only if) \( \succeq_\ell \) satisfies the following:

- **F1** \( \emptyset \preceq_\ell A \), for any \( A \in \mathcal{A} \)
- **F2** \( \emptyset \prec_\ell S \)
- **F3** \( \succeq_\ell \) is a weak order
- **F4** If \( A, B, C \in \mathcal{A} \) and \( A \cap C = B \cap C = \emptyset \), then \( A \prec_\ell B \iff A \cup C \prec_\ell B \cup C \)
- **F4'** If \( A_n \downarrow \) in \( \mathcal{A} \) and \( A_n \setminus A_{n+1} \preceq_\ell A^* \in \mathcal{A}, n = 1, 2, \ldots, \), then \( A^* \sim_\ell \emptyset \)
- **F4''** For any two uniform partitions \( \{ A_i \}_{i=1}^n \) and \( \{ B_i \}_{i=1}^n \) of \( S \) in \( \mathcal{A} \), \( \cup_{i \in I} A_i \sim_\ell \cup_{i \in J} B_i \), if \( |I| = |J| \).
- **F5** (i) If \( A \in \mathcal{A} \setminus N(\emptyset) \), then there is a finite partition \( \{ A_1, A_2, \ldots, A_n \} \) of \( S \) in \( \mathcal{A} \) such that (1) \( A_i \subset A \) or \( A_i \subset A^c \), \( i = 1, 2, \ldots, n \); (2) \( A_i \prec_\ell A \), \( i = 1, 2, \ldots, n \).

(ii) If \( A, B, C \in \mathcal{A} \setminus N(\emptyset) \) and \( A \cap C = \emptyset \), \( A \prec_\ell B \), then there is a finite partition \( \{ C_1, C_2, \ldots, C_m \} \) of \( S \) in \( \mathcal{A} \) such that (1) \( C_i \subset A \) or \( C_i \subset A^c \), and \( C_i \subset C \) or \( C_i \subset C^c \), \( i = 1, 2, \ldots, m \); (2) \( C_i \prec_\ell A \), and \( C_i \prec_\ell C \), \( i = 1, 2, \ldots, m \). and (3) \( A \cup C_i \prec_\ell B \), \( i = 1, 2, \ldots, m \).
- **F6** If \( \{ A_n \} \) is a decreasing sequence in \( \mathcal{A} \) and if \( A_* \prec_\ell \cap_{n=0}^\infty A_n \prec_\ell A^* \) for some \( A_* \) and \( A^* \) in \( \mathcal{A} \), then there exists \( N \) such that \( A_* \prec_\ell A_n \prec_\ell A^* \) for all \( n \geq N \).

Axioms F1, F2, F3 and F4 are similar to those in [6, Theorem 14.2], while F5 strengthens the corresponding axiom there. The additional axioms F4', F4'' and F6 are adopted here to compensate for the fact that \( \mathcal{A} \) is not a \( \sigma \)-algebra.

For the remainder of the appendix, \( \mathcal{A}, \succeq_\ell \) and \( \succeq \) are as specified in Theorem 1.5.2 and the axioms stated in (a) are assumed. By Lemma 1.5.1, \( \mathcal{A} \) is a \( \lambda \)-system. The objective now is to prove that conditions F1 - F6 are implied by the axioms given for \( \succeq \). Proofs that are elementary are not provided.
Lemma 1.7.2 Let \( \{A_i\} \) be a uniform partition of \( S \) in \( A \). Then for all outcomes \( \{x_i\} \) and for all permutations \( \sigma \),
\[
(x_{\sigma(i)}, A_i) \sim (x_i, A_i).
\]  
(1.32)

In other words, every uniform partition is strongly uniform.

Proof. Without loss of generality, assume \( x_1 > x_2 \) and that
\[
\begin{pmatrix}
x_1 & \text{if } s \in A_1 \\
x_2 & \text{if } s \in A_2 \\
x_3 & \text{if } s \in A_3 \\
\vdots & \vdots \\
x_n & \text{if } s \in A_n
\end{pmatrix} \succ
\begin{pmatrix}
x_2 & \text{if } s \in A_1 \\
x_1 & \text{if } s \in A_2 \\
x_3 & \text{if } s \in A_3 \\
\vdots & \vdots \\
x_n & \text{if } s \in A_n
\end{pmatrix}.
\]
Since \( \{A_i\}_{i=3}^n \) are unambiguous, the appropriate form of (1.4) implies
\[
\begin{pmatrix}
x_1 & \text{if } s \in A_1 \\
x_2 & \text{if } s \in A_1^c
\end{pmatrix} \succ
\begin{pmatrix}
x_2 & \text{if } s \in A_1^c \\
x_1 & \text{if } s \in A_2
\end{pmatrix},
\]
that is, \( A_1 \succ^\ell A_2 \), a contradiction. Similarly for the other cases. \( \blacksquare \)

Lemma 1.7.3 If \( \{A_n\}_{n=1}^{\infty} \) is a decreasing sequence in \( A \) such that \( A_n \setminus A_{n+1} \geq^\ell A^* \in A, n = 1, 2, \ldots, \) then \( A^* \sim^\ell \emptyset \).

Proof. Suppose \( A^* \succ^\ell \emptyset \). Because \( A \) is a \( \lambda \)-system (Lemma 1.5.1), \( \cap_{j=1}^{\infty} A_j \in A \) and \( \{A_n \setminus (\cap_{j=1}^{\infty} A_j)\}_{n=1}^{\infty} \) is a decreasing sequence in \( A \). Accordingly,
\[
A_n \setminus (\cap_{j=1}^{\infty} A_j) \geq^\ell A_n \setminus A_{n+1} \geq^\ell A^* \succ^\ell \emptyset.
\]
(1.33)

By Nondegeneracy, there exist outcomes \( x^* > x \). Thus,
\[
f_n = \begin{pmatrix}
x^* & \text{if } s \in A_n \setminus (\cap_{j=1}^{\infty} A_j) \\
x & \text{if } s \in A_1 \setminus (A_1 \setminus (\cap_{j=1}^{\infty} A_j)) \\
x & \text{if } s \in A_1^c
\end{pmatrix} \succ
\begin{pmatrix}
x^* & \text{if } s \in A^* \\
x & \text{if } s \in (A^*)^c
\end{pmatrix} \succ x.
\]

By Monotone Continuity, \( \{f_n\}_{n=1}^{\infty} \) converges in preference to \( f_\infty \), where
\[
f_\infty = \begin{pmatrix}
x^* & \text{if } s \in \cap_{n=1}^{\infty} (A_n \setminus (\cap_{j=1}^{\infty} A_j)) \\
x & \text{if } s \in A_1 \setminus (\cap_{n=1}^{\infty} A_1 \setminus (\cap_{j=1}^{\infty} A_j)) \\
x & \text{if } s \in A_1^c
\end{pmatrix} = x.
\]
Thus there exists an integer $N$ such that

$$f_n \prec \begin{pmatrix} x^* & \text{if } s \in A^* \\ x & \text{if } s \in (A^*)^c \end{pmatrix},$$

whenever $n \geq N$, contradicting (1.33). ■

By showing that the axioms 1-6 for $\succeq$ imply properties $F1 - F6$ for $\succeq_\ell$, we prove the following:

**Theorem 1.7.4** Let $\succeq$ be a preference order on $\mathcal{F}$ and denote by $\mathcal{A}$ the set of all unambiguous events. If $\succeq$ satisfies Axioms 1-6, then there exists a unique convex-ranged and countably additive probability measure on $\mathcal{A}$ such that

$$A \succeq_\ell B \iff pA \geq pB, \text{ for all } A, B \in \mathcal{A}.$$

**Proof.** Fix outcomes $x^* \succ x$. Properties $F1 - F3$ for $\succeq_\ell$ are immediate.

**F4:** (Note the role played here by the specific definition of $\mathcal{A}$; not any $\lambda$-system would do.) If $A \prec_\ell B$, then

$$\left( \begin{array}{c} x^* \\ x \\ h \\ x^* \\ x \end{array} \right) \prec \left( \begin{array}{c} x^* \\ x \\ h \\ x^* \\ x \end{array} \right) \quad \text{or} \quad \left( \begin{array}{c} x^* \\ x^* \\ h \\ x^* \\ x \end{array} \right) \prec \left( \begin{array}{c} x^* \\ x^* \\ h \\ x^* \\ x \end{array} \right).$$

where

$$h = \left\{ \begin{array}{l} x^* \text{ if } s \in A \cap B \\ x \text{ if } s \in C^c \setminus (A \cup B) \end{array} \right..$$

Since $C$ is unambiguous, deduce that

$$\left( \begin{array}{c} x^* \\ x \\ h \\ x^* \\ x \end{array} \right) \prec \left( \begin{array}{c} x^* \\ x \\ h \\ x^* \\ x \end{array} \right).$$

or $A \cup C \prec_\ell B \cup C$. Reverse the argument to prove the reverse implication.
F4' follows Lemma 1.7.3.

F4'' follows from Lemma 1.7.2 and Axiom 6.

F5 (i): Since $A \triangleright_t 0$,

$$
\begin{pmatrix}
  x^* & \text{if } s \in A \\
  x & \text{if } s \in A^c
\end{pmatrix}
\triangleright
\begin{pmatrix}
  x^* & \text{if } s \in A \\
  x & \text{if } s \in A^c
\end{pmatrix}.
$$

By Small Unambiguous Event Continuity (Axiom 4), there is a partition $\{A_i\}_{i=1}^n$ of $S$ in $A$, refining $\{A, A^c\}$ and such that

$$
\begin{pmatrix}
  x^* & \text{if } s \in A_i \\
  x & \text{if } s \in A^c_i
\end{pmatrix}
\triangleright
\begin{pmatrix}
  x^* & \text{if } s \in A_i \\
  x & \text{if } s \in A^c_i
\end{pmatrix},
$$

That is, for each $i$, $A_i \subset A$ or $A_i \subset A^c$ and in addition, $A_i \prec_t A$.

F5 (ii): Let $A$, $B$ and $C$ be in $A \setminus \mathcal{N}(\emptyset)$, $A \cap C = \emptyset$ and $A \prec_t B$. Then

$$
f = \begin{pmatrix}
  x^* & \text{if } s \in B \\
  x & \text{if } s \in B^c
\end{pmatrix}
\triangleright
\begin{pmatrix}
  x^* & \text{if } s \in A \\
  x & \text{if } s \in A^c
\end{pmatrix} = \begin{pmatrix}
  x^* & \text{if } s \in A \\
  x & \text{if } s \in C \\
  x & \text{if } s \in A^c \setminus C
\end{pmatrix} = g.
$$

By Small Unambiguous Event Continuity, there exists a partition $\{C_1, \ldots, C_n\}$ of $S$ in $A$, refining $\{A, C, A^c \setminus C\}$ and such that

$$
f = \begin{pmatrix}
  x^* & \text{if } s \in B \\
  x & \text{if } s \in B^c
\end{pmatrix}
\triangleright
\begin{pmatrix}
  x^* & \text{if } s \in C_i \\
  g & \text{if } s \in C_i^c
\end{pmatrix},
$$

If $C_i \subset A$, then $A \cup C_i = A \implies A \cup C_i = A \prec_t B$. On the other hand, if $C_i \subset A^c$, then

$$
\begin{pmatrix}
  x^* & \text{if } s \in C_i \\
  x & \text{if } s \in A \\
  x & \text{if } s \in (A \cup C_i)^c
\end{pmatrix} = \begin{pmatrix}
  x^* & \text{if } s \in A \cup C_i \\
  x & \text{if } s \in (A \cup C_i)^c
\end{pmatrix},
$$

implying that $A \cup C_i \prec_t B$.

F6: Implied by Monotone Continuity.
1.7.4 APPENDIX: Proof of Main Result

Necessity of the axioms in Theorem 1.5.2: The necessity of Monotonicity, Nondegeneracy and Weak Comparative Probability is routine. Denote by $\succeq_D$ the order on $D^u(a(X))$ represented by $W$.

Small Unambiguous Event Continuity: Let $f \succ g$ and $x$ be as in the statement of the axiom. Denote by $P = (x_1, p_1; \ldots; x_n, p_n)$ and $Q$ the probability distributions over outcomes induced by $f$ and $g$ respectively, and let $\overline{x}$ be a least preferred outcome in $\{x\} \cup \{x_1, x_2, \ldots, x_n\}$. Since $P \succ_D Q \succeq_D \delta_{\overline{x}}$, mixture continuity and monotonicity with respect to stochastic dominance ensure there exists some sufficiently large integer $N$ such that $W \left( (1 - \frac{1}{N})P + \frac{1}{N}\delta_{\overline{x}} \right) > W(Q)$. Because $p$ is convex-ranged, we can partition each set $A_i$ into $N$ equally probable events $\{A_{ij}\}_{j=1}^N$ in $A$. Let $C_k = \bigcup_{i=1}^N A_{ik}$ for $k = 1, 2, \ldots, N$. Then $\{C_k\}_{k=1}^N$ is a partition of $S$ in $A$, $p(C_k) = 1/N$ and $p(A_i \setminus C_k) = (1 - 1/N)p_i$ for each $i$ and $k$. Consequently, $[x$ if $s \in C_k; f$ if $s \notin C_k]$ induces the probability distribution $(1/N)\delta_x + (1 - 1/N)P$ which is strictly preferred to $Q$. Combined with monotonicity with respect to first-order stochastic dominance, this yields $[x$ if $s \in C_k; f$ if $s \notin C_k] \succeq [x$ if $s \in C_k; f$ if $s \notin C_k] \succ g$. Similarly for the other part of the axiom.

Monotone Continuity: Given a decreasing sequence $\{A_n\}_{n=1}^\infty$ in $A$, $p(A_n) \searrow p(\cap_1^\infty A_i)$ by the countable additivity of $p$. The required convergence in preference is implied by mixture continuity of $W$. The limit $f_\infty$ lies in $F^u$ because we are given that $A$ is a $\lambda$-system.

Strong-Partition Neutrality: Immediate from (1.26).

Sufficiency of the axioms in Theorem 1.5.2: Let $p$ be the measure provided by Theorem 1.7.4.

Lemma 1.7.5 For unambiguous events $A$ and $B$:

(a) $A$ is null iff $A \sim_\epsilon \emptyset$.

(b) If $A \sim_\epsilon B \sim_\epsilon \emptyset$ and $A \cap B = \emptyset$, then $A \cup B \sim_\epsilon \emptyset$. 
Proof. (a) Fix \( x \succ y \). Let \( A \) be null. Then

\[
\begin{pmatrix}
    x & \text{if} s \in A \\
    y & \text{if} s \in A^c
\end{pmatrix} \sim \begin{pmatrix}
    y & \text{if} s \in A \\
    x & \text{if} s \in A^c
\end{pmatrix} = y = \begin{pmatrix}
    x & \text{if} s \in \emptyset \\
    y & \text{if} s \in S
\end{pmatrix},
\]

implying that \( A \sim_\ell \emptyset \). If \( A \) is not null, then by Monotonicity (Axiom 1),

\[
\begin{pmatrix}
    x & \text{if} s \in A \\
    y & \text{if} s \in A^c
\end{pmatrix} \succ \begin{pmatrix}
    y & \text{if} s \in A \\
    x & \text{if} s \in A^c
\end{pmatrix} = \begin{pmatrix}
    x & \text{if} s \in \emptyset \\
    y & \text{if} s \in S
\end{pmatrix},
\]

implying that \( A \succ_\ell \emptyset \).

(b) Let \( x \succ y \) and \( A \cup B \succ_\ell \emptyset \), that is,

\[
\begin{pmatrix}
    x & A \\
    x & B \\
    y & (A \cup B)^c
\end{pmatrix} \succ y.
\]

By (a), \( A \) and \( B \) are null and

\[
y = \begin{pmatrix}
    y & A \\
    y & B \\
    y & (A \cup B)^c
\end{pmatrix} \sim \begin{pmatrix}
    x & A \\
    y & B \\
    y & (A \cup B)^c
\end{pmatrix} \sim \begin{pmatrix}
    x & A \\
    x & B \\
    y & (A \cup B)^c
\end{pmatrix} \succ y.
\]

This is a contradiction. \( \blacksquare \)

For each \( f \in \mathcal{F}^{ua} \), define

\[
P_f = (x_1, p(f^{-1}(x_1)); \ldots; x_n, p(f^{-1}(x_n))).
\]

Because \( p \) is fixed, it may be suppressed in the notation. Accordingly, write

\[
P_f \in D^{ua}(\mathcal{X}) = \{P_f : f \in \mathcal{F}^{ua}\}.
\]

Define the binary relation \( \succeq_D \) on \( D^{ua}(\mathcal{X}) \) by

\[
P \succeq_D Q \text{ if } \exists f \succeq g, P = P_f \text{ and } Q = P_g.
\]

Lemma 1.7.6 If \( P_f = P_g \), then \( f \sim g \). Thus \( \succeq_D \) is complete and transitive.
1.7. APPENDIX:

Proof. We must prove that for any two partitions \( \{A_i\}_{i=1}^n \) and \( \{B_i\}_{i=1}^n \) of \( S \) in \( A \), if \( A_i \sim \top B_i, i = 1, 2, ..., n \), then for all outcomes \( \{x_i\}_{i=1}^n \),

\[
\begin{pmatrix}
  x_1 & \text{if } s \in A_1 \\
  x_2 & \text{if } s \in A_2 \\
  \vdots & \ddots \\
  x_n & \text{if } s \in A_n
\end{pmatrix}
\sim
\begin{pmatrix}
  x_1 & \text{if } s \in B_1 \\
  x_2 & \text{if } s \in B_2 \\
  \vdots & \ddots \\
  x_n & \text{if } s \in B_n
\end{pmatrix}
\tag{1.34}
\]

Case 1: \( \{A_i\}_{i=1}^n \) and \( \{B_i\}_{i=1}^n \) are uniform partitions of \( S \) in \( A \). The desired conclusion follows from Lemma 1.7.2 and Axiom 6.

Case 2: All probabilities \( \{p(A_i)\}_{i=1}^n \) and \( \{p(B_i)\}_{i=1}^n \) are rational. Because \( p \) is convex-ranged, there exist \( \{E_j\}_{j=1}^m \) and \( \{C_j\}_{j=1}^m \), two uniform partitions of \( S \) in \( A \), such that

\[
A_i = \bigcup_{E_j \subseteq A_i} E_j, \quad i = 1, 2, ..., n \quad \text{and} \quad B_i = \bigcup_{C_j \subseteq B_i} C_j, \quad i = 1, 2, ..., n.
\]

Now Case 1 may be applied.

Case 3: This is the general case where some of the probabilities \( p(A_i) \) or \( p(B_i) \) may be irrational. Suppose contrary to (1.34) that

\[
f = \begin{pmatrix}
  x_1 & \text{if } s \in A_1 \\
  x_2 & \text{if } s \in A_2 \\
  \vdots & \ddots \\
  x_n & \text{if } s \in A_n
\end{pmatrix}
\prec
\begin{pmatrix}
  x_1 & \text{if } s \in B_1 \\
  x_2 & \text{if } s \in B_2 \\
  \vdots & \ddots \\
  x_n & \text{if } s \in B_n
\end{pmatrix}
= g. \tag{1.35}
\]

Without loss of generality, assume that \( x_n \succ \cdots \succ x_2 \succ x_1 \) and \( p(A_1) = p(B_1) \) is irrational.

By the convex range of \( p \) over \( A \), there are rational numbers \( r_m \nearrow m \) and two increasing sequences \( \{A_1^m\}_{m=1}^\infty \) and \( \{B_1^m\}_{m=1}^\infty \) in \( A \) with \( A_1^m \subseteq A_1 \) and \( B_1^m \subseteq B_1 \), \( m = 1, 2, ..., \) such that \( p(A_1^m) = p(B_1^m) = r_m \nearrow p(A_1) = p(B_1) \) as \( m \to \infty \). Accordingly,

\[
p(A_1 \setminus A_1^m) = p(B_1 \setminus B_1^m) = p(A_1) - p(A_1^m) \searrow 0 \quad \text{as } m \to \infty.
\]

Thus, both \( \{A_1 \setminus A_1^m\}_{m=1}^\infty \) and \( \{B_1 \setminus B_1^m\}_{m=1}^\infty \) are decreasing sequences in \( A \) and

\[
\bigcap_{m=1}^\infty (A_1 \setminus A_1^m) \sim \top \bigcap_{m=1}^\infty (B_1 \setminus B_1^m) \sim \top 0. \tag{1.36}
\]
Define
\[
g_m = \begin{cases} x_1 & \text{if } s \in B_1^m \\ x_2 & \text{if } s \in B_1 \setminus B_1^m \\ g & \text{if } s \in B_1^c \end{cases}, \quad g_\infty = \begin{cases} x_1 & \text{if } s \in \bigcap_{m=1}^\infty B_1^m \\ x_2 & \text{if } s \in B_1 \setminus \bigcap_{m=1}^\infty B_1^m \\ g & \text{if } s \in B_1^c \end{cases}.
\]

By Lemma 1.7.5 and (1.36),
\[ g_\infty \sim g \prec g. \]

By Monotone Continuity, \( g_m \) converges to \( g_\infty \) in preference as \( m \to \infty \). Conclude that there exists an integer \( N_1 \) such that
\[ g_m \prec g \text{ whenever } m \geq N_1. \]

In particular,
\[
g_{N_1} = \begin{cases} x_1 & \text{if } s \in B_1^{N_1} \\ x_2 & \text{if } s \in B_1 \setminus B_1^{N_1} \\ g & \text{if } s \in B_1^c \end{cases} \prec g.
\]

By Monotonicity,
\[
\begin{cases} x_1 & \text{if } s \in A_1^{N_1} \\ x_2 & \text{if } s \in A_1 \setminus A_1^{N_1} \\ f & \text{if } s \in A_1^c \end{cases} \preceq \begin{cases} x_1 & \text{if } s \in A_1^{N_1} \\ f & \text{if } s \in A_1^c \end{cases} = f \succeq \begin{cases} x_1 & \text{if } s \in B_1^{N_1} \\ x_2 & \text{if } s \in B_1 \setminus B_1^{N_1} \\ g & \text{if } s \in B_1^c \end{cases}.
\]

Therefore,
\[
\begin{cases} x_1 & \text{if } s \in A_1^{N_1} \\ x_2 & \text{if } s \in A_2 \cup (A_1 \setminus A_1^{N_1}) \\ f & \text{if } s \in (A_1 \cup A_2)^c \end{cases} \succeq \begin{cases} x_1 & \text{if } s \in B_2^{N_1} \\ x_2 & \text{if } s \in B_2 \cup (B_1 \setminus B_1^{N_1}) \\ g & \text{if } s \in (B_1 \cup B_2)^c \end{cases}.
\]

Note further that \( A_2 \cup (A_1 \setminus A_1^{N_1}) \sim_\ell B_2 \cup (B_1 \setminus B_1^{N_1}) \) since \( p(A_2 \cup (A_1 \setminus A_1^{N_1})) = p(B_2 \cup (B_1 \setminus B_1^{N_1})) \). Thus a proof by induction establishes that
\[
\begin{pmatrix} x_1 & \text{if } s \in A_1^{N_1} \\ x_2 & \text{if } s \in A_2^{N_2} \\ \vdots & \vdots \\ x_n & \text{if } s \in A_n^{N_n} \end{pmatrix} \succ \begin{pmatrix} x_1 & \text{if } s \in B_1^{N_1} \\ x_2 & \text{if } s \in B_2^{N_2} \\ \vdots & \vdots \\ x_n & \text{if } s \in B_n^{N_n} \end{pmatrix},
\]

where \( A_i^{N_i} \sim_\ell B_i^{N_i}, \ i = 1, 2, \ldots, n \) and every \( p(A_i^{N_i}) = p(B_i^{N_i}) \) is rational, contradicting Case 2. \( \blacksquare \)
The rest of the proof is similar to Steps 2-6 in the proof of [13, Theorem 2]. For example, in the proof of mixture continuity of \( \geq_D \) on \( D^{wa}(\mathcal{X}) \) (Step 3), Small Unambiguous Event Continuity may be used in place of Savage's P6 in order to overcome the lack of a \( \sigma \)-algebra structure for \( A \).

### 1.7.5 APPENDIX: Proofs of Corollaries

For the proof of Corollary 1.5.3, we need the following lemma:

**Lemma 1.7.7** Let \( p_* \) be the inner measure induced from \((A, p)\). Then for any event \( E \in \Sigma \), there exists an increasing sequence \( \{A_n\} \) in \( A \) with \( A_n \subseteq A \) such that

\[
\lim_{n \to \infty} p(A_n) = p_*(A).
\]

**Proof.** Claim: For any two events \( A \subseteq B \) in \( A \), and any \( r \in (p(A), p(B)) \), there exists \( C \in A \) with \( A \subseteq C \subseteq B \) such that \( p(C) = r \). This is proven as follows: Because \( p(B \setminus A) = p(B) - p(A) > 0 \) and \( 0 < r - p(A) < p(B) - p(A) = p(B \setminus A) \), then by the convex range of \( p \), there exists \( D \in A \), with \( D \subseteq B \setminus A \) and \( p(D) = r - p(A) \). Let \( C = D \cup A \). Then \( C \in A \), \( A \subseteq C \subseteq B \) and \( p(C) = p(D) + p(A) = r \).

By the definition of \( p_* \), there exist \( \{B_n\}_{n=1}^{\infty} \) in \( A \) with \( B_n \subseteq A \), \( n = 1, 2, ... \) such that

\[
p(B_n) \geq p_*(A) - 1/n.
\]

Without loss of generality, \( p(B_n) < p(B_{n+1}) \) for all \( n \). Let \( A_1 = B_1 \). Since \( p(B_2) > p(B_1) \), the claim implies that there exists \( A_2 \in A \) with \( A_1 \subseteq A_2 \) such that \( p(A_2) = p(B_2) \). Proceed by induction to derive an increasing sequence \( \{A_n\}_{n=1}^{\infty} \) in \( A \) such that \( p(A_n) = p(B_n) \) for all \( n \).

**Proof of Corollary 1.5.3:** Let \( p_* \) and \( p^* \) agree on \( E \). By the above lemma, (and the corresponding result for \( p^* \)), there exist unambiguous events \( \{A_n\} \) and \( \{B_n\} \) such that

\[
A_n \subseteq E \subseteq B_n, \quad A_n \nsubseteq B_n, \quad B_n \setminus A_n \quad \text{and} \quad p(B_n \setminus A_n) < 1/n.
\]
Because $\mathcal{A}$ is a $\lambda$-system, $A_\infty \equiv \cup A_n$ and $B_\infty \equiv \cap B_n$ are unambiguous. By countable additivity and Lemma 1.7.5, $p(B_\infty \setminus A_\infty) = 0$ and $B_\infty \setminus A_\infty$ is null. Therefore, $E \setminus A_\infty$ is null and (by Lemma 1.3.2) unambiguous. Thus $E$, the disjoint union of $A_\infty$ and $E \setminus A_\infty$, is unambiguous. 

**Proof of Corollary 1.5.4:** (a) The assumptions on $\nu$ imply that $\succeq$ satisfies the axioms in Theorem 1.5.2. (Chain-continuity implies Monotone Continuity for $\succeq$ and convex-ranged implies Small Unambiguous Event Continuity. To verify Monotonicity, apply the special nature of unambiguous events whereby they satisfy (1.4).) Therefore, there exists a convex-ranged and countably additive $p$ representing the likelihood relation on $\mathcal{A}$ that is implicit in $\succeq$. Conclude that $p$ must be ordinally equivalent to $\nu$ on $\mathcal{A}$.

(b) From (a) and Lemma 1.3.8, $\nu = \phi(p)$ on $\mathcal{A} \supseteq \mathcal{A}_0$, where $p$ is convex-ranged on $\mathcal{A}$. Therefore, $\nu(A_0) = \phi(p(A_0)) = [0, 1]$. Because $\phi$ is (strictly) increasing and onto, conclude that $p(A_0) = [0, 1]$. Now it is straightforward to prove that $\phi$ is the identity function. (For any two $x_1, x_2 \in [0, 1]$ with $x_1 + x_2 \leq 1$, there exist $A_1 \in \mathcal{A}_0$ and $A_2 \in \mathcal{A}$ such that $p(A_1) = x_1$, $p(A_2) = x_2$ and $A_1 \cap A_2 = \emptyset$. From the definition of $\mathcal{A}_0$, $\nu(A_1 + A_2) = \nu(A_1) + \nu(A_2)$, or

$$
\nu(A_1 + A_2) = \phi(p(A_1 + A_2)) = \phi(p(A_1) + p(A_2)) \\
= \phi(x_1 + x_2) = \nu(A_1) + \nu(A_2) \\
= \phi(p(A_1)) + \phi(p(A_2)) \\
= \phi(x_1) + \phi(x_2).
$$

Since $\phi$ is continuous, $\phi$ is linear on $[0, 1]$.

(c) By (b), $\nu = p$ on $\mathcal{A}$. Let $m$ be any measure in $\text{core}(\nu)$. Then $m(\cdot) \geq \nu(\cdot) = p(\cdot)$ on $\mathcal{A}$, implying that $m$ and $p$ coincide on $\mathcal{A}$. Thus $\mathcal{A} \supseteq \mathcal{A}_2$, the class of events where all measures in the core agree. But $\mathcal{A}_1 \subset \mathcal{A}$ by Lemma 1.3.8 and $\mathcal{A}_2 = \mathcal{A}_1$ by Lemma 1.3.7.
Bibliography


Chapter 2

Subjective Ambiguity, Expected Utility and Choquet Expected Utility
2.1 INTRODUCTION

2.1.1 Motivation

Much empirical evidence, inspired by Ellsberg [4], shows that the Subjective Expected Utility (SEU) model, axiomatized by Savage [18], cannot accommodate aversion to uncertainty or ambiguity. Recently, a number of generalizations of the standard model have been developed that are both axiomatic and can accommodate the noted aversion. Most notable from the perspective of this paper is the Choquet expected utility (CEU) model, or expected utility with respect to 'nonadditive probabilities' or 'capacities,' due to Schmeidler [20] and Gilboa [9]. Schmeidler uses the Anscombe-Aumann [1] set-up to deliver the CEU model, but this approach presumes the existence of objective lotteries. Gilboa avoids this drawback by using the Savage set-up, but his axioms are hard to interpret. Furthermore, it is not clear from these models why and how aversion to ambiguity leads a decision maker (DM) to use a nonadditive probability measure to represent her beliefs about the likelihoods of events and a CEU model to represent her preferences over acts.

Following the Savage set-up, this paper provides an axiomatic generalization of SEU that is more restrictive than the general CEU model, but that can nevertheless accommodate Ellsberg type behavior. In particular, our model amounts to Choquet expected utility theory with the added restriction that the capacity is an inner measure. The greater specificity of our model is advantageous from the perspective of proper model development; one wants models that are as close as possible to SEU and can still explain the evidence at hand. The freedom within the general CEU model to choose any capacity means that one can explain 'almost anything' by a suitable choice of capacity. This embarrassment of riches is evident particularly in applications to market data, where discipline similar to that derived within the SEU framework from the rational expectations hypothesis is lacking (see [14]). Another important attractive feature of the model is the simplicity of its axioms, which facilitates understanding of the situations where they do or do not have appeal.

There are two classical extreme approaches to modeling decision making under uncertainty - the Savage probability-based model where there is only risk, and the other extreme of complete ignorance and criteria such as the maximin. (See [12, Ch. 13] for clarification of the meaning of 'complete ignorance' and for an axiomatic characterization.) Our axiomatic model provides an intermediate approach by combining
these two extremes and delivering thereby a utility function representing preferences that has two principal features. The first is that the DM has sufficient information about the set of payoff-relevant states to form probabilistic beliefs about the likelihoods of events in $A^{ua}$, a collection of events interpreted as "unambiguous". That is, there exists an additive probability measure $p$ on $A^{ua}$ representing beliefs about likelihoods and such that an expected utility function with probability measure $p$ represents preferences over unambiguous prospects; technically, over acts that are $A^{ua}$-measurable. The second principal feature of the utility function is that there is complete ignorance of the state space besides what is modeled by the measure $p$ on $A^{ua}$; a rough interpretation is that the information underlying $p$ is the only information available to DM. These two features together deliver Choquet expected utility with capacity given by the inner measure generated by $A^{ua}$ and $p$. The two classical extremes appear as special cases of our model corresponding to alternative specifications for $A^{ua}$. One obtains the Savage model of pure risk if $A^{ua}$ is the collection of all events and the other extreme of pure ignorance if $A^{ua}$ consists of only the empty set and the full state space.

An important aspect of the model is the subjective nature of $A^{ua}$; it is derived from preferences rather than being specified exogenously. For this purpose, the paper adopts a definition of 'ambiguity' and subsequent analysis that are related to but distinct from those developed in chapter 1. Clarification of the similarities and differences is provided there. We emphasize that the model in chapter 1 is silent on the nature of preferences over ambiguous acts, which is the focus and main contribution of this paper.

### 2.1.2 Two Examples

#### The Ellsberg Paradox

The Ellsberg Paradox is described here. There are 90 balls in an urn, 30 red ones and the rest of 60 either black or yellow. One ball will be drawn at random. The following preferences over acts are typical,

$$f = \begin{cases} 
$100 & \text{if } s \in R \\
$0 & \text{if } s \in B \\
$0 & \text{if } s \in Y 
\end{cases} \succ \begin{cases} 
$0 & \text{if } s \in R \\
$100 & \text{if } s \in B \\
$0 & \text{if } s \in Y 
\end{cases} = g$$
CHAPTER 2. SUBJUNCTIVE AMBIGUITY

and

\[
\begin{pmatrix}
\$100 & \text{if } s \in R \\
$0 & \text{if } s \in B \\
$100 & \text{if } s \in Y
\end{pmatrix}
\prec
\begin{pmatrix}
\$0 & \text{if } s \in R \\
$100 & \text{if } s \in B \\
$100 & \text{if } s \in Y
\end{pmatrix}
= g',
\]

where \( R, B \) and \( Y \) denote the events corresponding to the chosen ball being red or black or yellow and \( s \) refers to the ball that is drawn.

The DM picks \( f \) rather than \( g \), presumably because the chance of getting \$100 in act \( f \) is precisely known to be \( 1/3 \), but the chance of getting \$100 in \( g \) is ambiguous since any number between \( 0 \) and \( 2/3 \) is possible. Similarly, she prefers \( g' \) to \( f' \), because the chance of getting \$100 in \( g' \) is precisely known to be \( 2/3 \), whereas the chance of getting \$100 in \( f' \) is ambiguous since any number between \( 1/3 \) and \( 1 \) is possible. Now, suppose the DM’s preferences can be represented by an expected utility function \( E_p u(\cdot) \) with respect to a vNM utility index \( u \) and a probability measure \( p \). Without loss of generality, assume \( u(100) > u(0) = 0 \), then

\[
f \succ g \iff u(100)p(\{R\}) + u(0)p(\{B, Y\}) > u(100)p(\{B\}) + u(0)p(\{R, Y\})
\]

\[
\iff p(\{R\}) > p(\{B\})
\]

and

\[
f' \prec g' \iff u(100)p(\{R, Y\}) + u(0)p(\{B\}) < u(100)p(\{B, Y\}) + u(0)p(\{R\})
\]

\[
\iff p(\{R\}) < p(\{B\}).
\]

This is a contradiction.

The key point of the Ellsberg Paradox is that there are kinds of events- unambiguous and ambiguous to the DM sometimes. Intuitively, an event is unambiguous if it has precisely known probability. But the SEU model cannot distinguish them since there is only one additive probability measure defined on all events. For example, the set of unambiguous events in the Ellsberg Paradox is

\[
A^{ua} = \{\emptyset, \{R, B, Y\}, \{R\}, \{B, Y\}\}.
\]

The following probability measure \( p \) on \( A^{ua} = \{\emptyset, \{R, B, Y\}, \{R\}, \{B, Y\}\} \) is natural:

\[
p(\emptyset) = 0, \quad p(\{R, B, Y\}) = 1, \quad p(\{R\}) = 1/3 \text{ and } p(\{B, Y\}) = 2/3.
\]

(2.1)

It is also intuitive that she uses an expected utility model \( E_p u(\cdot) \) to evaluate her unambiguous acts \( F^{ua} \) which are \( A^{ua} \)-measurable (defined precisely below), where \( u \) is a vNM index.
2.1. **INTRODUCTION**

After knowing how to evaluate her unambiguous acts, a natural and simple way to evaluate her ambiguous act she uses is to approximate it by her unambiguous acts from below. Here 'below' reflects her attitude towards ambiguity. In other words, she is ambiguity averse. That is, the utility of act $h$ is $U(h) = \sup\{E_p u(h') : h \geq h' \in \mathcal{F}^{ua}\}$, where $h \geq h'$ means $h(s) \geq h'(s)$ for all $s \in \{R, B, Y\}$. Accordingly,

$$U(f) > U(g) \text{ and } U(f') < U(g').$$

That is, our model resolves the Ellsberg Paradox.

**The Four Color Example**

The set of unambiguous events in the Ellsberg Paradox is an algebra. We give an example here to show that this is not true in general, because $\mathcal{A}^{ua}$ cannot be expected to be closed with respect to intersections.

Suppose there are 100 balls in the urn and that a ball’s color may be black (B), red (R), grey (G) or white (W). The sum of black and red ball is 50 and the sum of black and grey ball is also 50. We consider the following preferences:

$$f_1 = \begin{cases} 
\$1 & \text{if } s \in B \\
\$100 & \text{if } s \in R \\
\$0 & \text{if } s \in G \\
\$0 & \text{if } s \in W 
\end{cases} \succ f_2 = \begin{cases} 
\$100 & \text{if } s \in B \\
\$0 & \text{if } s \in R \\
\$0 & \text{if } s \in G \\
\$0 & \text{if } s \in W 
\end{cases},$$

$$g_1 = \begin{cases} 
\$1 & \text{if } s \in B \\
\$100 & \text{if } s \in R \\
\$100 & \text{if } s \in G \\
\$0 & \text{if } s \in W 
\end{cases} \prec g_2 = \begin{cases} 
\$100 & \text{if } s \in B \\
\$0 & \text{if } s \in R \\
\$100 & \text{if } s \in G \\
\$0 & \text{if } s \in W 
\end{cases},$$

and

$$h_1 = \begin{cases} 
\$1 & \text{if } s \in B \\
\$100 & \text{if } s \in R \\
\$100 & \text{if } s \in G \\
\$100 & \text{if } s \in W 
\end{cases} \succ h_2 = \begin{cases} 
\$100 & \text{if } s \in B \\
\$0 & \text{if } s \in R \\
\$100 & \text{if } s \in G \\
\$100 & \text{if } s \in W 
\end{cases}.$$

The DM picks $f_1$ rather than $f_2$, mainly because the chance of getting $\$100$ in $f_1$ is the same as in $f_2$, but also with other chance to get $\$1$ in act $f_1$. The only difference

---

1This example is based on a suggestion by U. Segal.
between pairs \{g_1, g_2\} and \{f_1, f_2\} is the payoff at state \(G\). Though both events \{B\} and \{G\} are ambiguous, the combination of them, however, leads to \{B, G\} being unambiguous since it has precisely known probability of one half. Thus, picking act \(g_2\) leads her to get $100 with probability of one half, while picking act \(g_1\) leads her only to get $1 with probability of one half and no idea with how much probability to get $100. Thus, she prefers \(g_2\) to \(g_1\). Finally, the same reasons as with the pair of acts \(\{f_1, f_2\}\), she will prefer \(h_1\) to \(h_2\).

An important and interesting phenomenon has happened: Changing the common outcome on event \(\{G\}\) in the pair of acts \(\{f_1, f_2\}\) leads the DM to change the ranking, while changing the common outcome on event \(\{G, W\}\) in the pair of acts \(\{f_1, f_2\}\) leaves the ranking of acts unchanged. What is the ‘key’ difference between events \(\{G\}\) and \(\{G, W\}\)? Intuitively, event \(\{G\}\) is ‘ambiguous’, since the probability of event \(\{G\}\) is unknown to the DM. While event \(\{G, W\}\) is ‘unambiguous’, since the probability of event \(\{G, W\}\) is known to her. This example suggests that unambiguous and ambiguous events can be derived from preferences. Roughly, if an event \(A\) satisfies the condition that replacing a common outcome for all states in \(A\) by any other outcome does not change the ranking of the pair of acts being compared, then \(A\) is unambiguous. Otherwise, it is ambiguous (for formal expression, see section 2.3).

Similar to the Ellsberg Paradox, she has a probability measure \(p\) on her unambiguous events \(\mathcal{A}^{ua} \triangleq \{\phi, \{B, R, G, W\}, \{B, G\}, \{R, W\}, \{B, R\}, \{G, W\}\}\) with

\[
\begin{align*}
p(S) &= 1, \quad p(\phi) = 0 \quad \text{and} \\
\end{align*}
\]

If she conforms to our model, then she has an expected utility function \(E_p u(\cdot)\) to represent her preferences over \(\mathcal{F}^{ua}\) and for other acts, she uses the utility function \(U(h) = \sup \{E_p u(h') : h \geq h' \in \mathcal{F}^{ua}\}\).

Note that the set of unambiguous events \(\mathcal{A}^{ua}\) is not an algebra, since it is not intersection-closed. Both \(\{B, G\}\) and \(\{B, R\}\) are unambiguous, but \(\{B\} = \{B, G\} \cap \{B, R\}\) is ambiguous. However, \(\mathcal{A}^{ua}\) is a \(\lambda\)-system (to be defined shortly).

### 2.1.3 Related Literature

The papers in the literature that are most closely related to ours are Sarin & Wakker [17] and Jaffray & Wakker [11]. The former also employs unambiguous events, but
they do not derive them and simply assume that there exist the two kinds of events. In addition, Sarin and Wakker's key axiom differs substantially from ours; this is by necessity since the general CEU model is axiomatized in Sarin & Wakker [17], while we axiomatize the special case corresponding to an inner measure. Jaffray and Wakker exploit and adapt the well-known Dempster-Shafer representation of belief functions to show that CEU with a capacity that is a belief function captures precisely the information structure described earlier - a class of unambiguous events and complete ignorance otherwise. In comparison to the classic Savage model and the model we present, the formal model in Jaffray & Wakker [11] employs two added primitives taken from the Dempster-Shafer representation. These are an auxiliary space of states and a (multi-valued) mapping that links it with $\mathcal{S}$ and delivers a definition for 'unambiguous' events in $\mathcal{S}$. As a result, verification of the axioms requires that the modeler observe or hypothesize these auxiliary primitives; such verification is not possible through observation exclusively of choices between alternative acts over $\mathcal{S}$. This is in contrast to our model, where all axioms are expressed in terms of preferences over Savage style acts over the (payoff-relevant) state space $\mathcal{S}$, without reliance on auxiliary state spaces or ad hoc assumptions regarding $\mathcal{A}^{\text{ua}}$. Another difference from our paper is that the two papers deliver different classes of preferences. Jaffray and Wakker admit any belief function while our axioms deliver inner measures which are not necessary belief functions.

Another related paper in the decision theory literature is Mukerji [15]. At a formal level, it differs from ours substantially. In particular, it employs state spaces other than $\mathcal{S}$ that are subjective and are intended to capture the DM's limited understanding or perception of the payoff-relevant states in $\mathcal{S}$. The 'explanation' for nonadditive beliefs that it offers is based on such epistemic features; missing states or violation of logical omniscience, for example. This is in contrast to the motivation derived from the Ellsberg paradox, where presumably the DM understands fully the set of possible payoff-relevant contingencies, but where their subjective likelihoods are too vague to be expressed by probabilities. Our formal model and the appeal of the axioms are compatible with his interpretation of the source of 'ambiguity'.

The use of inner measures as a way of modeling uncertainty is due to Fagin & Halpern [6], but their notion of inner measure is more restrictive than ours elaborated below. Fagin and Halpern take the perspective of the artificial intelligence literature where beliefs are primitive rather than derived from preferences. Our contribution is to show how their analysis may be adapted to a decision theoretic framework with a more general notion of inner measures.
This paper proceeds as follows. \( \lambda \)-systems, inner measures and inner integrals are defined next. The axioms and the representation results are described in Section 2.3. Some further discussion and examples are provided in section 2.4. All proofs are collected in appendices.

### 2.2 \( \lambda \)-SYSTEMS, INNER MEASURES AND INNER INTEGRALS

Let \( S \) be a state space with associated \( \sigma \)-algebra given by the power set. The usual way of representing a decision maker's beliefs about the likelihoods of events is by means of an (additive) probability measure on \( 2^S \). This structure is too restrictive as described in the example in the section 2.1.2. In this paper, we generalize this standard approach in two ways: (i) We relax the additivity of beliefs on all events to additivity on a subset \( \mathcal{A}^{ua} \) of \( 2^S \); (ii) we require that \( \mathcal{A}^{ua} \) be a \( \lambda \)-system but not necessarily an algebra.

#### 2.2.1 \( \lambda \)-Systems

Say that a nonempty class of subsets \( \mathcal{A}^{ua} \) of \( S \) is a \( \lambda \)-system if

\[
\begin{align*}
\lambda.1 & \ S \in \mathcal{A}^{ua}; \\
\lambda.2 & \ A \in \mathcal{A}^{ua} \implies A^c \in \mathcal{A}^{ua}; \text{ and} \\
\lambda.3 & \ A_n \in \mathcal{A}^{ua}, \ n = 1, 2, \ldots \text{ and } A_i \cap A_j = \emptyset, \ \forall i \neq j \implies \bigcup_n A_n \in \mathcal{A}^{ua}.
\end{align*}
\]

This definition and terminology appear in [2, p. 36].

Think of \( \mathcal{A}^{ua} \) as the collection of DM's "unambiguous" events and that she can assign probability to each event in \( \mathcal{A}^{ua} \). \( \lambda.1 \) is a usual assumption, since the DM knows that the event \( S \) will happen without ambiguity. Thus, she will assign probability 1 to \( S \). Assumption \( \lambda.2 \) is natural. If she can assign probability to event \( A \), it is natural for her to assign the difference between the probabilities of \( S \) and \( A \) to \( A^c \). The intuition for \( \lambda.3 \) is that if she can assign probabilities to two disjoint events \( A \)
2.2. $\lambda$-SYSTEMS, INNER MEASURES AND INNER INTEGRALS

and $B$, then it is natural for her to assign the sum of probabilities of events $A$ and $B$ to the union $A \cup B$. On the other hand, even if she can assign probabilities to both events $A$ and $B$, she still cannot assign probability to the intersection $A \cap B$ when $A \cap B$ is "ambiguous"; recall the example in section 2.1.2.

If the DM can assign probabilities to events $A$ and $B$, where $A \subseteq B$, then it is natural for her to assign the difference probabilities between events $B$ and $A$ to $B \setminus A$. And this can be seen from the following Lemma (see [2, p. 43]):

**Lemma 2.2.1** A nonempty class of subsets $A$ of $S$ is a $\lambda$-system if and only if

$\lambda.1'$ $\emptyset, S \in A$;

$\lambda.2'$ $A, B \in A$ and $A \subseteq B \implies B \setminus A \in A$; and

$\lambda.3'$ $A_n \in A$ and $A_n \subseteq A_{n+1}, n = 1, 2, \ldots, \implies \bigcup_n A_n \in A$.

The key difference between a $\sigma$-algebra and $\lambda$-system is that the latter is not required to be intersection-closed. It seems evident that a $\lambda$-system is more natural for modeling unambiguous events.

2.2.2 Probabilities and Integrals

Though it is not a $\sigma$-algebra, a probability measure can still be defined on $A^{ua}$. Say that a function

$$p : A^{ua} \rightarrow [0, 1]$$

is a (finitely additive) probability measure over $A^{ua}$ if:

P.1 $p(\emptyset) = 0$, $p(S) = 1$; and

P.2 $p(A \cup B) = p(A) + p(B)$, $\forall A, B \in A^{ua}$, $A \cap B = \emptyset$.

Denote by $(S, A^{ua}, p)$ a $\lambda$-system probability space. Say that a probability $p$ is convex-ranged if for all $A \in A^{ua}$ and $0 < r < 1$, there exists $B \subseteq A$, $B \in A^{ua}$, such that $p(B) = r \cdot p(A)$. 


Given a $\lambda$-system probability space $(S, \mathcal{A}_u^a, p)$, say that a finite-ranged function $f : S \mapsto \mathbb{R}_1$ is $\mathcal{A}_u^a$-measurable,\(^2\) if for any $x \in \mathbb{R}_1$,

$$\{s \in S : f(s) \leq x\} \in \mathcal{A}_u^a.$$ 

Denote by

$$\mathcal{F}_u^a = \{f : f \text{ is } \mathcal{A}_u^a \text{- measurable and has finite range}\} \text{ and }$$

$$\mathcal{F} = \{f : f \text{ has finite range}\}.$$

The integral of $f \in \mathcal{F}_u^a$ with respect to $p$ is defined as follows:

$$\int f dp = \Sigma_i x_i p(\{s : f(s) = x_i\}). \quad (2.4)$$

### 2.2.3 Inner and Outer Measures and Integrals

Think of a $\lambda$-system $\mathcal{A}_u^a$ as consisting of all ‘unambiguous’ events and of the probability measure $p$ as representing the DM’s beliefs about the likelihoods of events in $\mathcal{A}_u^a$. Suppose that the information underlying $p$ is all that is available to her and that she is pessimistic. Then one way to assess likelihoods for ‘ambiguous’ events, that is, events not in $\mathcal{A}_u^a$, is by means of the inner measure $p_*$ corresponding to $(S, \mathcal{A}_u^a, p)$, defined by: For all events $B$,

$$p_*(B) \equiv \sup \{p(A) : A \in \mathcal{A}_u^a, A \subseteq B\}. \quad (2.5)$$

Note that $p_* = p$ is additive on $\mathcal{A}_u^a$, but generally non-additive outside $\mathcal{A}_u^a$.

Similarly, if she is optimistic, then she might assess likelihoods for ‘ambiguous’ events by means of the outer measure $p^*$ corresponding to $(S, \mathcal{A}_u^a, p)$, defined by: For all events $B$,

$$p^*(B) \equiv \inf \{p(A) : A \in \mathcal{A}_u^a, A \supseteq B\}. \quad (2.6)$$

The inner, outer measures $p_*(B), p^*(B)$ provide lower, upper bounds on the likelihood of event $B$ respectively. The use of a lower, rather than upper bound, reflects aversion to ambiguity. Similarly, the use of an upper, rather than lower bound, reflects an affinity for ambiguity. Evidently, the upper and lower bounds agree on $\mathcal{A}_u^a$.

---

\(^2\)Since the function has finite range and $\mathcal{A}_u^a$ is a $\lambda$-system, $\{s \in S : f(s) \leq x\} \in \mathcal{A}_u^a$ implies $\{s \in S : f(s) > x\} = S \setminus \{s \in S : f(s) \leq x\} \in \mathcal{A}_u^a$. 

2.2. \(\lambda\)-SYSTEMS, INNER MEASURES AND INNER INTEGRALS

The inner measure \(p\) defined in this paper is more general than that in [6], because it is generated from a probability measure on a \(\lambda\)-system, while Fagin and Halpern begin with a probability measure on a \(\sigma\)-algebra (or algebra).

Corresponding to inner and outer measures, we introduce \textit{inner and outer integrals} as follows:

\[
\int f \, dp = \sup \{ \int g \, dp : f \geq g \in \mathcal{F}^{ua} \}, \forall f \in \mathcal{F}
\]

\[
\int^* f \, dp = \inf \{ \int g \, dp : f \leq g \in \mathcal{F}^{ua} \}, \forall f \in \mathcal{F},
\]

where \(f \geq g\) means that \(f(s) \geq g(s), \forall s \in S\). Their intuitive meanings are similar to those for inner and outer measures.

2.2.4 Capacities and Choquet Integrals

Inner and outer measures are special cases of capacities. Say that \(\nu : 2^S \rightarrow [0, 1]\) is a \textit{capacity} if

\begin{enumerate}
  \item C.1 \(A \subseteq B \implies \nu(A) \leq \nu(B)\); and
  \item C.2 \(\nu(\emptyset) = 0, \nu(S) = 1\).
\end{enumerate}

Call \(\overline{\nu}\) the conjugate of \(\nu\), where \(\overline{\nu}(A) = 1 - \nu(A^c)\).

A capacity is \textit{superadditive} if: For all disjoint events \(A\) and \(B\) in \(2^S\),

\begin{enumerate}
  \item C.3 \(\nu(A \cup B) \geq \nu(A) + \nu(B)\).
\end{enumerate}

A capacity is \textit{subadditive} if: For all disjoint events \(A\) and \(B\) in \(2^S\),

\begin{enumerate}
  \item C.4 \(\nu(A \cup B) \leq \nu(A) + \nu(B)\).
\end{enumerate}

A capacity is \textit{convex} if: For all events \(A\) and \(B\) in \(2^S\),
Convexity has been interpreted by Schmeidler [20] as reflecting ambiguity aversion of the DM. But this interpretation is problematic, see Epstein [5].

It is not hard to verify the following:

**Lemma 2.2.2** Let \( p_* \) and \( p^* \) be the inner and outer measures defined in (2.5) and (2.6). Then:

(a): \( p^* \) and \( p_* \) are conjugates.

(b): \( p_* \) is superadditive. In general, the outer measure \( p^* \) is not subadditive.

**Remark 1** The inner measure \( p_* \) is not convex in general. For example, \( p_* \) defined in example 2.4.5. But \( p_* \) is convex if \( A^{wa} \) is an algebra. See Proposition 3.1 in [6] for an elementary proof for the case of \( \sigma \)-algebra.

Applications of capacities to decision theory employ the notion of Choquet integration. For any capacity \( \nu \) and integrand \( f : S \rightarrow R^1 \), the Choquet integral is defined by

\[
\int f \, d\nu \equiv \int_0^\infty \nu(\{s : f(s) \geq t\}) \, dt + \int_{-\infty}^0 \left[ \nu(\{s : f(s) \geq t\}) - 1 \right] \, dt.
\]  

When \( f \) is a finite-ranged function, we can suppose that \( f \) assumes the values \( x_1 < ... < x_n \) on the events \( E_1, ..., E_n \) respectively. In this case,

\[
\int f \, d\nu = \sum_{i=1}^n x_i [\nu(\bigcup_{j=i}^n E_j) - \nu(\bigcup_{j=i+1}^n E_j)],
\]

where \( \nu(\bigcup_{n+1}^n E_j) \equiv 0 \).

Since any inner measure is a capacity, Choquet integration is defined also for inner and outer measures. In these cases, Choquet integrals are related to inner and outer integrals, as shown next.

**Theorem 2.2.3** Let \( p_*, p^* \) be the inner, outer measures defined in (2.5) and (2.6). Then:
2.3. THE AXIOMS AND MAIN RESULTS

1. For any function \( f \in \mathcal{F} \), \( \int f \, dp \leq \int f \, dp_* \leq \int f \, dp^* \leq \int^* f \, dp \); and

2. If \( p \) is convex-ranged on \( \mathcal{A}^{ua} \) or \( p^*_*(A \cup B) = p(A) + p^*(B) \) for all \( A \in \mathcal{A}^{ua} \) and \( B \subset A^c \), then

\[
\int f dp = \int f dp_* \quad \text{and} \quad \int^* f dp = \int f dp^*. \tag{2.9}
\]

Under the condition in part 2, (2.9) and (2.10) provide a representation and novel intuition for Choquet integrals that are based on inner and outer measures. This representation is used to prove our main result, Theorem 2.3.4.

Given a capacity \( \nu \) on (the power set of) \( S \), define the core of \( \nu \) by

\[ \text{core}(\nu) = \{ m : m \text{ is a probability measure, } m(\cdot) \geq \nu(\cdot) \} \]

If \( \nu \) is convex, then its core is nonempty [19]. Furthermore, for all \( f \in \mathcal{F} \),

\[
\int f d\nu = \inf \{ \int f dm : m \in \text{core}(\nu) \} \tag{2.11}
\]

if and only if \( \nu \) is convex. This representation (2.11) is in a sense 'dual' to ours. It involves approximation by dominating measures, while ours involves approximation by dominated functions. Another aspect of the duality between these two representations is that ours considers a single measure on a \( \lambda \)-system, while the multiple-priors model has a set of measures on a single \( \sigma \)-algebra.

2.3 THE AXIOMS AND MAIN RESULTS

2.3.1 Unambiguous Events

Savage's set-up is adopted throughout. Denote by \( S \) the set of states of the world, by \( 2^S \) the set of all events and by \( \mathcal{X} \) the set of outcomes. Prospects are modeled via simple acts, mappings from \( S \) to \( \mathcal{X} \) having finite range. The set of acts is \( \mathcal{F} = \{ ..., f, ... \} \). The DM has a preference relation \( \succeq \) on the set of acts.
The following definition is central:

**Definition 2.3.1** An event \( A \) is unambiguous if: (i) For all acts \( f, f' \) and outcomes \( x, y \in \mathcal{X} \),
\[
\begin{pmatrix}
  f & \text{if } s \in A^c \\
  x & \text{if } s \in A
\end{pmatrix} \quad \text{if} \quad \begin{pmatrix}
  f' & \text{if } s \in A^c \\
  x & \text{if } s \in A
\end{pmatrix} \quad \implies \quad \begin{pmatrix}
  f & \text{if } s \in A^c \\
  x & \text{if } s \in A
\end{pmatrix} \quad \implies \quad \begin{pmatrix}
  f' & \text{if } s \in A^c \\
  x & \text{if } s \in A
\end{pmatrix}
\]
\[\text{and (ii) The condition obtained if } A \text{ is everywhere replaced by } A^c \text{ in (i) is also satisfied. Otherwise, } A \text{ is called ambiguous.}\]

Define
\[
\mathcal{A}_{ua} = \{ A \in 2^S : A \text{ is unambiguous} \} \quad \text{(2.13)}
\]
\[
\mathcal{F}_{ua} = \{ f \in \mathcal{F} : f \text{ is } \mathcal{A}_{ua}-\text{measurable} \}. \quad \text{(2.14)}
\]

Both sets are nonempty because \( \emptyset \) and \( S \) are unambiguous.

It merits emphasis that ambiguity is subjective; it is endogenously derived from the DM's preference order. This distinguishes the present analysis from that of Fishburn [8], who takes ambiguity as a primitive, and from [17], where an exogenously specified class of unambiguous events is used.

To understand definition 2.12, recall Savage's key axiom.

**Sure-Thing Principle (STP):** For all events \( A \) and all acts \( f, f', g, g' \),
\[
\begin{pmatrix}
  f & \text{if } s \in A^c \\
  g & \text{if } s \in A
\end{pmatrix} \quad \text{if} \quad \begin{pmatrix}
  f' & \text{if } s \in A^c \\
  g & \text{if } s \in A
\end{pmatrix} \quad \implies \quad \begin{pmatrix}
  f & \text{if } s \in A^c \\
  g & \text{if } s \in A
\end{pmatrix} \quad \implies \quad \begin{pmatrix}
  f' & \text{if } s \in A^c \\
  g & \text{if } s \in A
\end{pmatrix}
\]
\[\text{and (ii) The condition obtained if } A \text{ is everywhere replaced by } A^c \text{ in (i) is also satisfied. Otherwise, } A \text{ is called ambiguous.}\]
An implication is that for all events $A$ and all acts $f$, $f'$, and outcomes $x, y$,

$$
\left( \begin{array}{l}
  f & \text{if } s \in A^c \\
  x & \text{if } s \in A
\end{array} \right) \succeq \left( \begin{array}{l}
  f' & \text{if } s \in A^c \\
  x & \text{if } s \in A
\end{array} \right) \implies
\left( \begin{array}{l}
  f & \text{if } s \in A^c \\
  y & \text{if } s \in A
\end{array} \right) \succeq \left( \begin{array}{l}
  f' & \text{if } s \in A^c \\
  y & \text{if } s \in A
\end{array} \right).
$$

(2.16)

Therefore, if two acts imply different subacts ($f(\cdot)$ versus $f'(\cdot)$) over an event $A^c$, but the same outcome over the event $A$, the ranking of these acts will not depend on this common outcome. This axiom implies that preferences are separable across mutually exclusive events, which is the key property of expected utility preferences, either over objective probability distributions or over acts. However, this principle is contradicted by the typical choices in the Ellsberg Paradox. We interpret such choices as evidence that the separability required by the STP between outcome in $A$ and those in $A^c \equiv S \setminus A$ is descriptively (and perhaps even normatively) problematic when the event $A$ is “ambiguous.” In such cases, changing a common outcome in some states in $A^c$ can cast an entirely new light on the pair of acts being compared.

Though the STP is no longer appealing in similar situations, it is still of use in building an alternative formal model. If one views “ambiguity” as the only source of violation of the Sure-Thing Principle, then one is led to use STP to give formal meaning to “ambiguity.” This leads to the definition 2.12. It is immediate that if preferences satisfy STP, then all events are unambiguous.

### 2.3.2 Preferences over Unambiguous Acts

Here we specify some axioms for the preference order $\succeq$ on the set $\mathcal{F}^{ua}$ of unambiguous acts. These will deliver a probability measure $p$ on $\mathcal{A}^{ua}$ and an expected utility functional form on $\mathcal{F}^{ua}$. The axioms are similar to those in chapter 1, though, as noted earlier, the definition of ‘unambiguous’ adopted here is different. In spite of this difference, the proof of the expected utility representation (Theorem 2.3.3) is similar to that of chapter 1 and thus is deleted. The reader is referred to the cited joint paper also for further discussion of the axioms.

Following Savage, say that an event $A \in 2^S$ is null, if for all acts $f$, $g$, $g' \in \mathcal{F}$,

$$
\left( \begin{array}{l}
  f & \text{if } s \in A^c \\
  g & \text{if } s \in A
\end{array} \right) \sim \left( \begin{array}{l}
  f & \text{if } s \in A^c \\
  g' & \text{if } s \in A
\end{array} \right).
$$

Following Savage, say that an event $A \in 2^S$ is null, if for all acts $f$, $g$, $g' \in \mathcal{F}$,
Denote by \( x \in \mathcal{X} \) both the outcome and the constant act producing the outcome in every state. Preference statements like '\( x \succeq y \)' are therefore well-defined and have the obvious meaning. For any \( A \in A^{ua} \), define

\[
\mathcal{F}^{ua}_A = \{ f \in \mathcal{F} : f^{-1}(X) \cap A \in A^{ua} \text{ for all } X \subset \mathcal{X} \}.
\]

**Axiom 1 (Monotonicity)** For all outcomes \( x \) and \( y \), non-null events \( A \in A^{ua} \) and acts \( g \in \mathcal{F}^{ua}_A \),

\[
\begin{pmatrix}
    x & \text{if } s \in A \\
    g(s) & \text{if } s \in A^c
\end{pmatrix}
\succeq
\begin{pmatrix}
    y & \text{if } s \in A \\
    g(s) & \text{if } s \in A^c
\end{pmatrix}
\iff x \succeq y.
\]

**Axiom 2 (Nondegeneracy)** There exist outcomes \( x^* \) and \( x \) such that \( x^* \succ x \).

**Axiom 3 (Weak Comparative Probability)** For all events \( A, B \in A^{ua} \) and outcomes \( x^* \succ x \) and \( y^* \succ y \)

\[
\begin{pmatrix}
    x^* & \text{if } s \in A \\
    x & \text{if } s \in A^c
\end{pmatrix}
\succeq
\begin{pmatrix}
    x^* & \text{if } s \in B \\
    x & \text{if } s \in B^c
\end{pmatrix}
\iff
\begin{pmatrix}
    y^* & \text{if } s \in A \\
    y & \text{if } s \in A^c
\end{pmatrix}
\succeq
\begin{pmatrix}
    y^* & \text{if } s \in B \\
    y & \text{if } s \in B^c
\end{pmatrix}.
\]

Axiom 1 states that replacing any outcome \( y \) on a non-null unambiguous event \( A \) by a preferred outcome \( x \) always leads to a preferred act. Axiom 2 says that the relation \( \succeq \) is not trivial. When \( x^* \succ x \), the ranking \([x^* \text{ if } A; x, \text{ if } A^c] \succeq [x^* \text{ if } B; x, \text{ if } B^c]\) reveals that the DM believes \( A \) to be at least as likely as \( B \). Axiom 3 states that this revealed likelihood ranking is independent of the specific outcomes used when \( A \) and \( B \) are unambiguous. That is, the likelihood relation \( \succeq_\ell \) is complete, reflexive and transitive on \( A^{ua} \), where \( \succeq_\ell \) is defined by: \( A \succeq_\ell B \) if \( \exists x^* \succ x \) such that

\[
\begin{pmatrix}
    x^* & \text{if } s \in A \\
    x & \text{if } s \in A^c
\end{pmatrix}
\succeq
\begin{pmatrix}
    x^* & \text{if } s \in B \\
    x & \text{if } s \in B^c
\end{pmatrix}.
\]

The next axiom imposes suitable richness of the set of unambiguous events. It is clear from Savage's analysis that some richness is required to derive a probability measure on \( A^{ua} \). Further, Savage's axiom P6 (suitably translated) is not adequate here because \( A^{ua} \) is not intersection closed. Nonetheless, the spirit of Savage's P6 is retained in the next axiom.
2.3. THE AXIOMS AND MAIN RESULTS

Axiom 4 (Small Unambiguous Event Continuity) Let $f, g \in \mathcal{F}^{ua}$, $f \succ g$, with $f = (x_1, A_1; x_2, A_2; \ldots; x_n, A_n)$, $g = (y_1, B_1; y_2, B_2; \ldots; y_m, B_m)$, where each $A_i$ and $B_i$ lies in $A$. Then for any outcome $x \in \mathcal{X}$, there exist two partitions $\{C_i\}_{i=1}^N$ and $\{D_j\}_{j=1}^M$ of $S$ in $A$ that refine $\{A_i\}_{i=1}^n$ and $\{B_j\}_{j=1}^m$ respectively, and satisfy:

$$f \succ \begin{cases} x & \text{if } s \in D_k \\ g(s) & \text{if } s \in D_k^c \end{cases}, \text{ for all } k \in \{1, \ldots, N\};$$  \hspace{1cm} (2.18)

and

$$\begin{cases} x & \text{if } s \in C_j \\ f(s) & \text{if } s \in C_j^c \end{cases} \succ g, \text{ for all } j \in \{1, \ldots, M\}.$$  \hspace{1cm} (2.19)

Roughly, the axiom requires that unambiguous events can be decomposed into suitably 'small' unambiguous events. When $\mathcal{A}^{ua}$ is closed with respect to intersections, as in the standard model [18] where it is taken to be the power set, then the axiom is implied by Savage's P6, given Axioms 1-3.

Say that a sequence $\{f_n\}_{n=1}^\infty$ in $\mathcal{F}^{ua}$ converges in preference to $f_\infty \in \mathcal{F}^{ua}$ if: For any two acts $f_\ast$, $f^* \in \mathcal{F}^{ua}$ satisfying $f_\ast \prec f_\infty \prec f^*$, there exists an integer $N$ such that

$$f_\ast \prec f_n \prec f^*, \text{ whenever } n \geq N.$$  

Axiom 5 (Monotone Continuity) For any $A \in \mathcal{A}^{ua}$, outcomes $x^* \succ x$, act $h \in \mathcal{F}_A^{ua}$ and decreasing sequence $\{A_n\}_{n=1}^\infty$ in $\mathcal{A}^{ua}$ with $A_n \subseteq A$, $n = 1, 2, \ldots$, define

$$f_n = \begin{pmatrix} x^* & \text{if } s \in A_n \\ x & \text{if } s \in A \setminus A_n \\ h & \text{if } s \in A^c \end{pmatrix} \quad \text{and} \quad f_\infty = \begin{pmatrix} x^* & \text{if } s \in \cap_{n=1}^\infty A_n \\ x & \text{if } s \in A \setminus (\cap_{n=1}^\infty A_n) \\ h & \text{if } s \in A^c \end{pmatrix}.$$  

If $f_n \in \mathcal{F}^{ua}$ for all $n = 1, 2, \ldots$, then $\{f_n\}_{n=1}^\infty$ converges in preference to $f_\infty$ and $f_\infty \in \mathcal{F}^{ua}$.

The name Monotone Continuity describes one aspect of the axiom, that requiring the indicated convergence in preference. The second component of the axiom is the requirement that the limit $f_\infty$ lie in $\mathcal{F}^{ua}$ whenever each $f_n$ is unambiguous. This will serve in particular to ensure that $\mathcal{A}^{ua}$ satisfies the 'countable' closure condition $\lambda.3$ required by the definition of a $\lambda$-system.
A finite partition with component events from $A^{ua}$ is denoted $\{A_i\}$. Henceforth all partitions have unambiguous components, even where not stated explicitly. Given such a partition, use the obvious abbreviated notation $(x_i, A_i)$. For any permutation $\sigma$ of $\{1, \ldots, n\}$, $(x_{\sigma(i)}, A_i)$ denotes the act obtained by permuting outcomes between the events. Say that the finite partition $\{A_i\}$ is a uniform partition if $A_i \sim A_j$ for all $i$ and $j$ and call $\{A_i\}$ strongly uniform if in addition it satisfies: For all outcomes $\{x_i\}$ and for all permutations $\sigma$,

$$\left(x_{\sigma(i)}, A_i\right) \sim (x_i, A_i).$$

(2.20)

In particular, if $\{A_i\}_{i=1}^n$ is a strongly uniform partition, then for all index sets $I$ and $J$, subsets of $\{1, 2, \ldots, n\}$,

$$\bigcup_{i \in I} A_i \sim \bigcup_{i \in J} A_i \text{ if } |I| = |J|.$$

**Axiom 6 (Strong-Partition Neutrality)** For any two strongly uniform partitions $\{A_i\}_{i=1}^n$ and $\{B_i\}_{i=1}^n$, if $A_i \sim B_i$ for all $i$, then for all $\{x_i\}$,

$$\begin{pmatrix}
  x_1 & \text{if } s \in A_1 \\
  x_2 & \text{if } s \in A_2 \\
  \vdots & \vdots \\
  x_n & \text{if } s \in A_n
\end{pmatrix}
\sim
\begin{pmatrix}
  x_1 & \text{if } s \in B_1 \\
  x_2 & \text{if } s \in B_2 \\
  \vdots & \vdots \\
  x_n & \text{if } s \in B_n
\end{pmatrix}.

(2.21)

The hypothesis that the $A_i$'s and $B_i$'s satisfy (2.20) expresses another sense in which these events are unambiguous. This makes the conclusion (2.21) weaker than if the indifference in (2.21) were required for all uniform partitions. To support the claim that Strong-Partition Neutrality is a ‘weak’ axiom, observe that it is satisfied by all CEU orders. The reason is that if $\{A_i\}$ is a strongly uniform partition, and thus satisfies (2.20), then $\nu$ is additive on the algebra generated by the partition. Thus the indifference (2.21) is implied.

We turn next to implications of the preceding axioms. The proof of next Lemma is identical to Lemma 1.5.1 in chapter 1.

**Lemma 2.3.2** Under Axioms 1-6, $A^{ua}$ is a $\lambda$-system. In particular, if $A_1$ and $A_2$ are disjoint unambiguous events, then $A_1 \cup A_2$ is unambiguous.
Parallel to Savage’s Theorem, we have the following representation theorem for unambiguous acts $\mathcal{F}^{ua}$:

**Theorem 2.3.3** Let $\succeq$ be a preference relation on $\mathcal{F}$ and $\mathcal{A}^{ua}$ the corresponding set of unambiguous events. Then the following two statements are equivalent:

(a) $\succeq$ satisfies axioms 1–6.

(b) $\mathcal{A}^{ua}$ is a $\lambda$-system and there exist a unique countably additive convex-ranged probability measure $p$ on $\mathcal{A}^{ua}$ and a (nonconstant) utility index $u : \mathcal{X} \rightarrow \mathbb{R}^+$ such that

$$f \succeq g \iff \int u(f) dp \geq \int u(g) dp, \quad \forall f, g \in \mathcal{F}^{ua}. \quad (2.22)$$

Theorem 2.3.3 generalizes Savage’s Theorem. It is noted that there is one sense in which the Savage theory fails to be subjective. Savage assumes that the DM can figure out all events. In other words, all events are admissible and measurable to her. In our paper, not only the probability measure $p$ and the expected utility function $U(\cdot) = \int u(\cdot) dp$, but also the domain of probability measure $p$ are all endogenously derived from preference. Secondly, though the Sure-Thing Principle is not imposed on $\mathcal{F}^{ua}$, our axioms still deliver a probability measure $p$ on $\mathcal{A}^{ua}$ to represent the DM’s beliefs about the likelihoods of her unambiguous events, and a Subjective Expected Utility Model to represent preferences over $\mathcal{F}^{ua}$.

While the above theorem describes the nature of preferences over unambiguous acts, our primary concern is the nature of the ordering of ambiguous acts. We turn to this aspect of preferences.

### 2.3.3 Ambiguous Acts and Choquet Expected Utilities

The remaining task is to relate the ordering of ambiguous acts to the ordering of unambiguous ones.

The next axioms restrict both the set of outcomes $\mathcal{X}$ and $\succeq$. The following notation is adopted:

$$f \succeq g \text{ means that } f(s) \succeq g(s) \quad \forall s \in S. \quad (2.23)$$
Axiom 7 (Weak Monotonicity) For any two acts \( f, g \) with \( f \in \mathcal{F}^{\text{ua}} \) or \( g \in \mathcal{F}^{\text{ua}} \), \( f \succeq g \) if \( f \succeq g \).

Axiom 8 (Certainty Equivalent) For any \( f \in \mathcal{F} \), there exists \( x \in \mathcal{X} \) such that \( f \sim x \).

Axioms 7 and 8 are self-explanatory and common. The final axiom is the central one.

Axiom 9 (Approximation from Below) For any \( f \in \mathcal{F} \), \( x \in \mathcal{X} \), if \( f \succ x \), then there exist an act \( g \in \mathcal{F}^{\text{ua}} \) such that \( f \succeq g \) and \( g \succ x \).

The interpretation of Axiom 9 is that in order to evaluate an arbitrary act \( f \), the DM approximates \( f \) from below by an unambiguous act \( h \). Given Weak Monotonicity, any such \( h \) provides a lower bound for the utility derived from \( f \). Roughly speaking, Approximation from Below requires further that \( f \) be ranked only as highly as the most preferred of such approximating unambiguous acts \( h \). In other words, the DM facing the ambiguous act \( f \) asks “what is the worst that can happen if I choose \( f \)?” and answers this question by relying exclusively on the acts that she understands well, namely on unambiguous acts. This algorithm seems inappropriate for situations where the distinction between ‘slightly ambiguous’ and ‘highly ambiguous’ events (and acts) is important. Our DM thinks in terms of black or white rather than in terms of shades of grey.

Note that Axiom 9 does not require unambiguous acts to be rich. The DM ranks \( f \succ x \) because she can find \( h \in \mathcal{F}^{\text{ua}} \) such that \( f \succeq h \) and \( h \succ x \).

Now we can state our main result – the representation for preferences on \( \mathcal{F} \).

**Theorem 2.3.4** Let \( \succeq \) be a preference relation on \( \mathcal{F} \) and \( \mathcal{A}^{\text{ua}} \) the corresponding set of unambiguous events. Then the following two statements are equivalent:

1. \( \succeq \) satisfies axioms 1—9.
2.3. THE AXIOMS AND MAIN RESULTS

(b) $A^{ua}$ is a $\lambda$-system and there exist a unique convex-ranged, countably additive probability measure $p$ on $A^{ua}$, and a nonconstant utility index $u : \mathcal{X} \rightarrow \mathbb{R}$, having convex range such that

$$f \geq g \iff \int u(f)dp \geq \int u(g)dp, \quad f, g \in \mathcal{F}, \quad (2.24)$$

where $p_*$ is the inner measure generated from the $\lambda$-probability space $(S, A^{ua}, p)$ and integration is in the sense of Choquet.

Theorem 2.24 tells us that it is complete ignorance outside $A^{ua}$ and approximation from below that lead the DM to use a capacity $p_*$ to represent her beliefs and a Choquet Expected Utility to represent her preferences. Accordingly, complete ignorance outside $A^{ua}$ and approximation from below capture the heuristic meaning of ambiguity aversion.

The interpretation of the utility representation, particularly the subjective nature of $A^{ua}$, warrants emphasis. It cannot be argued that exclusive reliance (via $p_*$) on events in $A^{ua}$ to compute likelihoods of other events reflects an extreme or unreasonable degree of ignorance or uncertainty aversion. At a formal level, that is because $A^{ua}$ is a component of the utility representation that is inseparable from the use of inner measure. Less formally, the above argument is unsupportable because $A^{ua}$ is subjective; whether or not there is a large degree of ignorance or uncertainty aversion implied depends on the size of $A^{ua}$. Indeed, a primary role of $A^{ua}$ is to model the degree of uncertainty aversion of the DM.

We have the following natural way to model comparative uncertainty aversion. Consider two decision makers having a common vNM index $u$ but different pairs $(A^{ua}_1, p_1)$ and $(A^{ua}_2, p_2)$. Say that 2 is more uncertainty averse than 1 if $A^{ua}_2 \subset A^{ua}_1$ and $p_1$ agrees with $p_2$ on $A^{ua}_2$. Since the $\lambda$-systems and subjective probabilities are derived from preferences, this relation between two decision makers may be expressed exclusively in terms of their preferences orderings over acts.

We offer some comments here on the comparison between our axioms and others employed in literature cited in the introduction. We feel this is appropriate even though our set of axioms deliver a specialized and hence different decision model, and in spite of the inevitably subjective nature of any such discussion. Our view is that

---

3There is a parallel with the question of how to interpret the set of priors and the minimization over that set that appears in the multiple-priors model of preference in Gilboa & Schmeidler [10].
such a comparison supports our contention that the intuitive foundation available for our model serves to differentiate it from the more general forms of CEU that have been axiomatized. In his seminal paper [20], Schmeidler assumes the existence of objective randomizing devices in the form of the Anscombe-Aumann horse-race/roulette-wheel acts. Different primitives are also employed in Jaffray & Wakker [11] and Mukerji [15], as described in the introduction. Therefore, the most relevant outstanding comparison is with Gilboa [9] and Sarin & Wakker [17]. Gilboa does not rely on the existence of 'unambiguous' events, whether exogenous or derived, but his axioms are complicated and difficult to interpret. Sarin and Wakker also employ unambiguous events but they are exogenous and form a σ-algebra. In addition, the key axiom in Sarin & Wakker [17], called cumulative dominance, is simple but it builds in a concern with rank ordered outcomes that makes the CEU representation result unsurprising. In other words, the small gap between the axioms and result limits the insight into the CEU functional form that is provided by the axiomatization. We have described situations or thought processes, characterized by complete ignorance outside some λ-system, where the axiom Approximation from Below has some intuitive appeal, though admittedly it may not be intuitively appealing more generally.

Finally, there is an obvious mirror image of this analysis in which 'optimism, uncertainty affinity and Approximation From Above' replace 'pessimism, uncertainty aversion and Approximation From Below'. That is, replace axiom 9 by the following axiom:

**Axiom 10 (Approximation from Above)** For any \( f \in \mathcal{F} \), \( x \in \mathcal{X} \), if \( f < x \), then there exist an act \( g \in \mathcal{F}^{wa} \) such that \( f \leq g \) and \( g < x \).

This delivers the CEU model of preferences with capacity given by the outer measure \( p^* \).

### 2.3.4 Two Extreme Cases

Two extreme cases are discussed here. One is that if the DM's preference satisfies both approximation from below and above, intuitively, her preference should be represented by a Savage model. The following Theorem proves this is true:
2.3. THE AXIOMS AND MAIN RESULTS

Theorem 2.3.5 Let $\succeq$ be a preference relation on $\mathcal{F}$ and $\mathcal{A}^{ua}$ the corresponding set of unambiguous events. Then the following two statements are equivalent:

(a) $\succeq$ satisfies axioms 1—10.

(b) $\mathcal{A}^{ua} = 2^S$, $\mathcal{F} = \mathcal{F}^{ua}$ and there exist a unique convex-ranged, countably additive probability measure $p$ on $2^S$, and a nonconstant utility index $u : \mathcal{X} \rightarrow \mathbb{R}^1$ having convex range such that for all acts $f, g \in \mathcal{F}$,

$$f \succeq g \iff \int u(f)dp \geq \int u(g)dp. \quad (2.25)$$

Another extreme case is if the DM's information is between black and white or her attitude toward ambiguity is between pessimistic and optimistic, then her behavior does not follow approximation from below and above. This leads to the following non Choquet model:

Theorem 2.3.6 Let $\succeq$ be a preference relation on $\mathcal{F}$ and $\mathcal{A}^{ua}$ the corresponding set of unambiguous events. Then the following two statements are equivalent:

(a) $\succeq$ satisfies axioms 1—8.

(b) $\mathcal{A}^{ua}$ is a $\lambda$-system and there exist a unique convex-ranged, countably additive probability measure $p$ on $\mathcal{A}^{ua}$, and a nonconstant utility index $u : \mathcal{X} \rightarrow \mathbb{R}^1$ having convex range such that for all acts $f, g \in \mathcal{F}$, $f \succeq g \iff$

$$\alpha_f \int u(f)dp_* + (1 - \alpha_f) \int u(f)dp^* \geq \alpha_g \int u(g)dp_* + (1 - \alpha_g) \int u(g)dp^*, \quad (2.26)$$

where $p_*$, $p^*$ are the inner, outer measures generated from the $\lambda$-system probability space $(S, \mathcal{A}^{ua}, p)$, $\alpha_f$, $\alpha_g$ are constants in $[0, 1]$ and depend acts $f$, $g$ respectively, and the integrations are in the sense of Choquet.
2.4 FURTHER DISCUSSION AND EXAMPLES

2.4.1 Unambiguous Events and the Sure-Thing Principle

Given a preference order $\succeq$ on acts $\mathcal{F}$ in Savage's set-up, we provided a definition of unambiguous events in section 2.3. A natural alternative is to use STP to define unambiguous events. Denote by

$$\overline{A} = \{ A : A \text{ and } A^c \text{ satisfies (2.15) for all } f', f, g', g \}.$$  \hspace{1cm} (2.27)

But $\overline{A}$ is intersection-closed, making it unsuitable as the class of unambiguous events. In addition, $\overline{A}$ is 'too small' in general. For instance, $\overline{A} = \{ \emptyset, S \}$ in the four color example.

Lemma 2.4.1 Let $\succeq$ be a preference order on $\mathcal{F}$. Then:

1. $\overline{A}$ is intersection-closed.
2. $\overline{A} \subset A^{ua};$ But the other direction is not true.
3. If $A^{ua}$ is an algebra, then the Sure-Thing Principle is satisfied on $\mathcal{F}^{ua}$.

2.4.2 Unambiguous Events in the Choquet Model

Now suppose that the DM's preference order is represented by CEU with a monotone, superadditive capacity $\nu$. That is,

$$f \succeq g \iff \int u(f)d\nu \geq \int u(g)d\nu, \quad \forall f, g \in \mathcal{F},$$ \hspace{1cm} (2.28)

where $u : \mathcal{X} \rightarrow \mathbb{R}^1$ is a nonconstant utility index and has convex range. Since the attitude towards ambiguity is included in the capacity $\nu$, the precise relation between $\nu$ and $A^{ua}$ is of interest.

Define

$$\begin{align*}
\Sigma^0 &= \{ A \in 2^S : \nu(A_1 \cup B) = \nu(A_1) + \nu(B), \ \forall A_1 \subset A, \ B \subset A^c \} \\
\Sigma^1 &= \{ A \in 2^S : \nu(A_1 \cup B) = \nu(A_1) + \nu(B), \ \forall A_1 \subset A, \ B \subset A^c \} \\
\Sigma^2 &= \{ A \in 2^S : \nu(A \cup B) = \nu(A) + \nu(B), \ \forall B \subset A^c \} \\
\Sigma^3 &= \{ A \in 2^S : \nu(A) + \nu(A^c) = 1 \}.
\end{align*}$$
Theorem 2.4.2 Let $\geq$ on $\mathcal{F}$ be defined in 2.28 and define $A^{ua}$ and $\overline{A}$ by (2.13) and (2.27). Then:

1. $A \in \Sigma^0 \iff A_1 \in \Sigma^2$, $\forall A_1 \subseteq A$.

2. 
   \[
   \Sigma^1 = \{ A \in 2^S : \nu(B) = \nu(B \cap A) + \nu(B \cap A^c), \forall B \} \\
   = \{ A \in 2^S : \nu(B \cup A) + \nu(B \cap A) = \nu(A) + \nu(B), \forall B \} \\
   = \{ A \in 2^S : \int_A u(f(s)) d\nu = \int_A u(f) d\nu + \int_{A^c} u(f) d\nu, \forall f \in \mathcal{F} \} \\
   \subseteq \overline{A} \subseteq A^{ua}.
   \]
   Moreover, $\Sigma^1$ is an algebra.

3. If $\Sigma^2$ is symmetric in the sense that $A \in \Sigma^2 \iff A^c \in \Sigma^2$, then $\Sigma^2 = A^{ua}$. In addition, $\Sigma^2$ is closed with respect to disjoint unions.

4. If $A \in \Sigma^3$, then $A^c \in \Sigma^3$. $\Sigma^3$ is an algebra, if $\nu$ is convex, but not more generally.

5. $\Sigma^0 \subseteq \Sigma^1 \subseteq \Sigma^2 \subseteq \Sigma^3$ and they are not equal in general.

6. If $\nu$ is convex, then $\Sigma^2 = \Sigma^3$.

The collections $\Sigma^1$ and $\Sigma^3$ represent two extreme possible approaches to defining unambiguous events directly in terms of the capacity $\nu$. But $\Sigma^1$ is an algebra and if $\nu$ is not convex, then $\Sigma^3$ is not closed with respect to disjoint unions. For example:

Example 2.4.3 \footnote{This example was suggested by M. Marinacci.} Let $S = \{\omega_1, \omega_2, \omega_3\}$ and $\nu$ be defined on $(S, 2^S)$ as follows:

\[
\begin{align*}
\nu(S) &= 1, \nu(\emptyset) = 0 = \nu(\{\omega_1\}) \\
\nu(\{\omega_1, \omega_2\}) &= \nu(\{\omega_1, \omega_3\}) = \nu(\{\omega_2, \omega_3\}) = 2/3 \\
\nu(\{\omega_2\}) &= \nu(\{\omega_3\}) = 1/3.
\end{align*}
\]

Then $\nu$ is superadditive but not convex and

\[
\begin{align*}
\Sigma^0 &= \{\emptyset\}, \Sigma^1 = \{\emptyset, S\}, \\
\Sigma^3 &= \{\emptyset, S, \{\omega_1, \omega_2\}, \{\omega_1, \omega_3\}, \{\omega_2\}, \{\omega_3\}\}.
\end{align*}
\]

Our concept of unambiguous events provides an intermediate approach between the above two extreme notions, since $\Sigma^0 \subseteq \Sigma^1 \subseteq \Sigma^2 \subseteq \Sigma^3$. In the above example,

\[
\Sigma^2 = \{\emptyset, S, \{\omega_1, \omega_2\}, \{\omega_1, \omega_3\}\}.
\]
2.4.3 Examples

We provide two examples to clarify the nature of our axioms and implied utility function. They describe triples \((S, \mathcal{A}^{ua}, p)\) satisfying (with some stated exceptions) the conditions of the theorem. Assume that \(\mathcal{X} = R^1_+\) and \(u : R^1_+ \rightarrow R^1\) is strictly increasing.

Example 2.4.4 (Continuation of the Ellsberg Paradox in section 2.1.2) Let \(S = \{B, R, Y\}\) and \(\mathcal{F} = \{f : f \text{ is a finite valued function from } S \text{ to } R^1_+\}\).

Define

\[
U(f) = \int u(f) dp_*,
\]

where \(p\) is the probability measure defined in (2.1) and \(p_*\) is induced from \(p\).

Our definition of 'unambiguous' yields

\[
\mathcal{A}^{ua} = \emptyset, \{R, B, Y\}, \{R\}, \{B, Y\}
\]

Preference defined in (2.29) satisfies all Axioms except Axiom 4.

Example 2.4.5 (Continuation of the three colour example in section 2.1.2) Let \(S = \{B, R, G, W\}\) and \(\mathcal{F} = \{f : f \text{ is a finite valued function from } S \text{ to } R^1_+\}\).

Define

\[
U(f) = \int u(f) dp_*,
\]

where \(p\) is the probability measure defined in (2.2) and the inner measure \(p_*\) is induced from \(p\).

By Theorem 2.4.2,

\[
\mathcal{A}^{ua} = \{\emptyset, S, \{B, G\}, \{R, W\}, \{B, R\}, \{G, W\}\}.
\]

The preference order defined in (2.30) satisfies all Axioms except Axiom 4.
2.4. FURTHER DISCUSSION AND EXAMPLES

2.4.4 A 'Counterexample'

This example illustrates violation of approximation from below. It is a variation of the four color example in subsection 2.1.2.

Let $B + R = 80$ and $B + G = 10$. One ball is to be drawn at random. The following preferences are intuitive:

$$f_1 = \begin{pmatrix} 
$0 & \text{if } s \in B \\
$100 & \text{if } s \in R \\
$0 & \text{if } s \in G \\
$0 & \text{if } s \in W 
\end{pmatrix} \succ \begin{pmatrix} 
$100 & \text{if } s \in B \\
$0 & \text{if } s \in R \\
$100 & \text{if } s \in G \\
$100 & \text{if } s \in W 
\end{pmatrix} = f_2. $$

The DM picks $f_1$ rather than $f_2$ mainly because she knows from $B + R = 80$ and $B + G = 10$ that $R = G + 70$ and $W = B + 10$. Though the events $\{R\}$ and $R^c = \{B, G, W\}$ are both ambiguous to the DM, however, the probability of $\{R\}$ is objectively at least 0.7, and therefore greater than two times that of its complement $\{B, G, W\}$.

It is intuitive that all unambiguous events in this example are:

$$\mathcal{A}^{ua} \triangleq \{\phi, S, \{B, G\}, \{R, W\}, \{B, R\}, \{G, W\}\},$$

where $S = \{B, R, G, W\}$. And the DM's beliefs over $\mathcal{A}^{ua}$ can be represented by the following probability measure

$$p(S) = 1, \quad p(\phi) = 0 \quad \text{and} \quad p(\{B, G\}) = 0.1, \quad p(\{R, W\}) = 0.9 \quad \text{and} \quad p(\{B, R\}) = 0.8, \quad p(\{G, W\}) = 0.2$$

However,

$$\int u(f_1)dp_* < \int u(f_2)dp_*,$$

where $p_*$ is the inner measure induced from $p$ and the integral is in the sense of Choquet.

What is wrong with the inner measure model? The problem is that the decision-maker has more information than just the probabilities of events in $\mathcal{A}^{ua}$, e.g., she knows that $R = G + 70$ and $W = B + 10$, even though $\{R, G\}$ and $\{W, B\}$ are ambiguous. Extension of our model to accommodate such 'partial information' is the subject of current research.
2.5 APPENDIX

2.5.1 Appendix A:

Theorem 2.2.3 is proved here. We first prove the following Lemma:

Lemma 2.5.1 Let $(S, \mathcal{A}^u, p)$ be a $\lambda$-system probability space. If $p$ is convex-ranged or $p_*(A \cup B) = p(A) + p_*(B)$, $\forall A \in \mathcal{A}^u$ and $B \subseteq A$, then:

(i) For any $A \in 2^S$, there exist a sequence of $\{A_n\}$ in $\mathcal{A}^u$ satisfying $A_n \subseteq A_{n+1}$, where $n = 1, 2, \ldots$, such that

$$p_*(A) = \lim_{n \to \infty} p(A_n).$$

(ii) For any chain $\emptyset = A_0 \subseteq A_1 \subseteq A_2 \subseteq \cdots \subseteq A_k = S$, there exists a chain $\emptyset = B_0 \subseteq B_1 \subseteq B_2 \subseteq \cdots \subseteq B_{k-1} \subseteq B_k = S$ in $\mathcal{A}^u$ with $B_n \subseteq A_n$ for $n = 0, 1, 2, \ldots, k$ such that $p(B_n) = p_*(A_n)$ for $n = 0, 1, 2, \ldots, k$.

Proof. I: $p$ is convex-ranged.

Proof of (i): Claim: For any two events $A \subseteq B$ in $\mathcal{A}^u$, and any $r \in (p(A), p(B))$, there exists $C \in \mathcal{A}^u$ with $A \subseteq C \subseteq B$ such that $p(C) = r$. This is proven as follows: Because $p(B \setminus A) = p(B) - p(A) > 0$ and $0 < r - p(A) < p(B) - p(A) = p(B \setminus A)$, then by the convex range of $p$, there exists $D \in \mathcal{A}^u$, with $D \subseteq B \setminus A$ and $p(D) = r - p(A)$. Let $C = D \cup A$. Then $C \in \mathcal{A}^u$, $A \subseteq C \subseteq B$ and $p(C) = p(D) + p(A) = r$.

By the definition of $p_*$, there exist $\{B_n\}_{n=1}^\infty$ in $\mathcal{A}^u$ with $B_n \subseteq A$, $n = 1, 2, \ldots$ such that

$$p(B_n) \geq p_*(A) - 1/n.$$

Without loss of generality, $p(B_n) < p(B_{n+1})$ for all $n$. Let $A_1 = B_1$. Since $p(B_2) > p(B_1)$, the claim implies that there exists $A_2 \in \mathcal{A}^u$ with $A_1 \subseteq A_2$ such that $p(A_2) = p(B_2)$. Proceed by induction to derive an increasing sequence $\{A_n\}_{n=1}^\infty$ in $\mathcal{A}^u$ such that $p(A_n) = p(B_n)$ for all $n$. Let $B = \cup_n A_n$. Obviously, $p(B) = p_*(A)$.

Proof of (ii): We first prove that for each event $A \in 2^S$, there exists $B \in \mathcal{A}^u$ with $B \subseteq A$ such that $p(B) = p_*(A)$. From part (i), there exist an increasing sequence of
2.5. APPENDIX

events \{A_n\}_{n=1}^{\infty} in \mathcal{A}^{ua} such that \lim_{n \to \infty} p(A_n) = p_*(A). Let B = \bigcup_{n=1}^{\infty} A_n, then B in \mathcal{A}^{ua} with B \subset A such that p(B) = p_*(A). Next, we only need to prove that for two events A_1 \subset A_2, there exist B_1, B_2 in \mathcal{A}^{ua} with B_1 \subset B_2 such that B_i \subset A_i, p(B_i) = p(A_i), i = 1, 2. Without loss of generality, assume p_*(A_2) > p_*(A_1). Then there exists an event B_1 in \mathcal{A}^{ua} with B_1 \subset A_1 such that p(B_1) = p_*(A_1). By part (i), there exists an increasing sequence of events \{C_n\}_{n=1}^{\infty} in \mathcal{A}^{ua} with B_1 \subset C_n \subset A_2 for all n such that \lim_{n \to \infty} p(C_n) = p_*(A_2). Let B_2 = \bigcup_{n=1}^{\infty} C_n, then B_2 in \mathcal{A}^{ua} with B_1 \subset B_2 \subset A_2 such that p(B_2) = p_*(A_2).

II: p_*(A \cup B) = p(A) + p_*(B) for all A \in \mathcal{A}^{ua} and B \subseteq A^c.

Proof of (i): Case 1: p_*(A) = 0.

In this case, the conclusion is obviously true.

Case 2: p_*(A) > 0.

Take 0 < \epsilon < p_*(A). By definition of p_*, there exists A_1 \in \mathcal{A}^{ua} with A_1 \subset A such that

p(A_1) > p_*(A) - \epsilon > 0.

If p(A_1) = p_*(A), then let A_n = A_1, n = 1, 2, ..., the conclusion is proved. Now let p(A_1) < p_*(A), then p_*(A \setminus A_1) > 0 follows from p_*(A) = p_*(A_1 \cup (A \setminus A_1)) = p(A_1) + p_*(A \setminus A_1). Similarly, there exists A' \in \mathcal{A}^{ua} with A' \subset A \setminus A_1 such that

p(A') > p_*(A \setminus A_1) - \epsilon/2.

Denote A_2 = A_1 \cup A' \in \mathcal{A}^{ua}. As a result,

\[ p(A_2) = p(A_1 \cup A') = p(A_1) + p(A') = p_*(A) + (p(A') - p_*(A \setminus A_1)) > p_*(A) - \epsilon/2. \]

By induction, there exists an increasing sequence \{A_n\}_{n=1}^{\infty} in \mathcal{A}^{ua} such that

\[ p(A_n) > p_*(A) - \epsilon/n, \quad n = 1, 2, ... \quad (2.33) \]

Define B = \bigcup A_n. \ A^* \in \mathcal{A}^{ua} follows from that \mathcal{A}^{ua} is a \lambda-system. Therefore,

\[ p_*(A) - \epsilon/n < p(A_n) < p(B) \leq p_*(A), \quad \text{for all } n \]

i.e.,

\[ p_*(A) = \lim_{n \to \infty} p(A_n) = p(B). \]
Proof of (ii): We only need to prove for any two events $A_1 \subseteq A_2$, there exist two events $B_1, B_2$ in $\mathcal{A}^{ua}$ with $B_1 \subseteq B_2$ such that $B_i \subseteq A_i$ and $\mu(B_i) = \mu(A_i)$, $i = 1, 2$. By part (i), there exists event $B_1$ in $\mathcal{A}^{ua}$ with $B_1 \subseteq A_1$ such that $\mu(B_1) = \mu(A_1)$. Accordingly,

\[ \mu(A_2) = \mu(B_1 \cup (A_2 \setminus B_1)) = \mu(B_1) + \mu(A_2 \setminus B_1). \]

Again by part (i), there exists event $C_1$ in $\mathcal{A}^{ua}$ with $C_1 \subseteq A_2 \setminus B_1$ such that $\mu(C_1) = \mu(A_2 \setminus B_1)$. Let $B_2 = B_1 \cup C_1$, then $B_2 \in \mathcal{A}^{ua}$, $B_2 \subseteq A_2$, $B_1 \subseteq B_2$ and $\mu(B_2) = \mu(A_2)$.

Proof of Theorem 2.2.3:

(1): Since $\mu(A) \leq \mu(A)$, $\forall A \in \mathcal{A}^{ua}$,

\[ \int g \, dp = \int g \, d\mu \leq \int f \, d\mu, \quad f \geq g \in \mathcal{F}^{ua}. \]

And this implies

\[ \int \begin{array}{c} f \, dp = \sup \{ \int g \, dp : f \geq g \in \mathcal{F}^{ua} \} \leq \int f \, d\mu. \]

Similarly, we can prove

\[ \int f \, d\mu \leq \int f \, d\mu. \]

Finally, $\int f \, d\mu \leq \int f \, d\mu$ directly follows from $\mu(A) \leq \mu(A)$, $\forall A \in 2^S$.

(2): We first prove (2.9). By part (1), we need only prove that

\[ \sup \{ \int g \, dp : f \geq g \in \mathcal{F}^{ua} \} \geq \int f \, d\mu. \]  \hspace{1cm} \text{(2.34)}

Let

\[ f(s) = \begin{cases} x_1 & \text{if } s \in E_1 \\ x_2 & \text{if } s \in E_2 \\ \vdots & \text{...} \\ x_n & \text{if } s \in E_n. \end{cases} \]

and $x_1 < x_2 < ... < x_n$.

By definition,

\[ \int f \, d\mu = \sum_{i=1}^{n} x_i [\mu(\cup_{j=i}^{n} E_j) - \mu(\cup_{j=i+1}^{n} E_j)]. \]
2.5. APPENDIX

\[ \emptyset = \bigcup_{j=0}^{n+1} E_j \subset \bigcup_{j=0}^{n} E_j \subset \cdots \subset \bigcup_{j=2}^{n} E_j \subset \bigcup_{j=1}^{n} E_j = S \]

is a chain. By Lemma 2.5.1 (ii), there exists a chain \( \emptyset = B_0 \subset B_1 \subset \cdots \subset B_n = S \) in \( \mathcal{A}^{ua} \) with \( B_i \subset \bigcup_{j=i+1}^{n} E_j \) such that \( p_*(\bigcup_{j=i+1}^{n} E_j) = p(B_i), i = 0, 1, 2, \ldots, n \). Therefore,

\[
\int f dp_* = \sum_{i=1}^{n} x_i [p_*(\bigcup_{j=i}^{n} E_j) - p_*(\bigcup_{j=i+1}^{n} E_j)] \\
= \sum_{i=1}^{n} x_i [p(B_{n+1-i}) - p(B_{n-i})] \\
= \sum_{i=1}^{n} x_i p(B_{n+1-i} \setminus B_{n-i}) \\
= \int \hat{f} dp,
\]

where

\[
\hat{f}(s) = \begin{cases} 
    x_1 & \text{if } s \in B_n \setminus B_{n-1} \\
    x_2 & \text{if } s \in B_{n-1} \setminus B_{n-2} \\
    \vdots & \text{......} \\
    x_{n-1} & \text{if } s \in B_2 \setminus B_1 \\
    x_n & \text{if } s \in B_n
\end{cases}
\]

Obviously, \( \hat{f} \in \mathcal{F}^{ua} \) and \( \hat{f} \leq f \).

Therefore,

\[
\int f dp_* = \int \hat{f} dp \leq \sup \{ \int g dp : g \in \mathcal{F}^{ua}, g \leq f \}.
\]

The proof of (2.10) is similar to that of (2.9).

2.5.2 Appendix B:

Proof of Theorem 2.3.3: (b) \( \Longrightarrow \) (a) is routine. For the converse, argue as in Theorem 1.5.1 in chapter 1 and Theorem 3.3.1 in chapter 3 to show that \( \mathcal{A}^{ua} \) is a \( \lambda \)-system and that there is a unique convex-ranged, finitely additive probability measure \( p \) on \( \mathcal{A}^{ua} \) such that \( A \succeq B \iff p(A) \geq p(B), \forall A, B \in \mathcal{A}^{ua} \). The remainder of the proof is similar to Lemma 1.7.6 in chapter 1 and [7, pp. 203-205].

Proof of Theorem 2.3.4:

(a) \( \Longrightarrow \) (b): It is enough to prove (2.24) assuming (i) the representation (2.22) where \( \mathcal{A}^{ua} \) satisfies \( \lambda.1 \) and \( \lambda.2 \) (not necessarily \( \lambda.3 \)) and (ii) Axioms 7-9 for \( \succeq \).
Since constant acts are in $F^u$, (2.22) implies
\[ x \succeq y \iff u(x) \geq u(y). \]
Since $p$ is convex-ranged, it follows from (2.22) and Certainty Equivalent for $F^u$-measurable acts that $u$ has convex range.

**Lemma 2.5.2** (i) For any act $f \in F$, $f \sim x$ implies that
\[ u(x) = \sup \{ \int u(g)dp : f \geq g \in F^u \}. \]
(ii) The ordering $\succeq$ on $F$ is represented by $U$ where, for any $f \in F$,
\[ U(f) = \sup \{ \int u(g)dp : f \geq g \in F^u \}. \tag{2.35} \]

**Proof.** (i) By Weak Monotonicity, $f \geq g$ implies $x \sim f \succeq g$. By (2.22),
\[ \int u(g)dp \leq u(x), \quad \forall g \in F^u, \]
this means that
\[ \sup \{ \int u(g)dp : f \geq g \in F^u \} \leq u(x). \]
For the other direction, it is true if $x$ is the worst outcome in $X$. Next, suppose that there exists $x' < x$. Since $u$ has convex range, for any sufficiently small $\varepsilon > 0$, there exists $x''$ such that
\[ u(x) - \varepsilon < u(x'') < u(x). \]
Now suppose that
\[ u(x) - \varepsilon > \sup \{ \int u(g)dp : f \geq g \in F^u \} \text{ for some } \varepsilon > 0. \]
Then $x'' \succ g$ for all $g \in F^u$. This contradicts $f \succ x''$ and Approximation from Below.

(ii) follows from (i), Certainty Equivalent and transitivity of $\succeq$. \(\blacksquare\)

It remains to show that the utility function defined in (2.35) is the required Choquet integral with respect to the inner measure $p_*$, that is,
\[ U(f) \equiv \int u(f)dp_* \tag{2.36} \]
And this directly follows from that $p$ is convex-ranged and part (ii) of Theorem 2.2.3. \(\blacksquare\)
2.5.3 Appendix C:

The proofs of Theorem 2.3.5 and 2.3.6 are provided here. First, we prove some Lemmas.

**Lemma 2.5.3** If \( p_\star(\mathcal{A} \cup \mathcal{B}) = p_\star(\mathcal{A}) + p_\star(\mathcal{B}) \) for any \( \mathcal{A} \in \mathcal{A}^u_a \) and \( \mathcal{B} \subseteq \mathcal{A}^c \), then

\[
\{ \mathcal{A} : p_\star(\mathcal{A} \cup \mathcal{B}) = p_\star(\mathcal{A}) + p_\star(\mathcal{B}), \ \forall \mathcal{B} \subseteq \mathcal{A}^c \} = \{ \mathcal{A} : p_\star(\mathcal{A}) + p_\star(\mathcal{A}^c) = 1 \}.
\]

**Proof.** We only need to prove that

\[
\Sigma^2 \triangleq \{ \mathcal{A} : p_\star(\mathcal{A} \cup \mathcal{B}) = p_\star(\mathcal{A}) + p_\star(\mathcal{B}), \ \forall \mathcal{B} \subseteq \mathcal{A}^c \} \supseteq \{ \mathcal{A} : p_\star(\mathcal{A}) + p_\star(\mathcal{A}^c) = 1 \} \triangleq \Sigma^3.
\]

First, we prove that \( \Sigma^2 \) is closed with respect to disjoint unions. Take \( \mathcal{A}_1, \mathcal{A}_2 \in \Sigma^2 \) are disjoint. For any event \( \mathcal{B} \subseteq (\mathcal{A}_1 \cup \mathcal{A}_2)^c \),

\[
p_\star((\mathcal{A}_1 \cup \mathcal{A}_2) \cup \mathcal{B}) = p_\star(\mathcal{A}_1 \cup (\mathcal{A}_2 \cup \mathcal{B})) = p_\star(\mathcal{A}_1) + p_\star(\mathcal{A}_2 \cup \mathcal{B}) = p_\star(\mathcal{A}_1) + p_\star(\mathcal{A}_2) + p_\star(\mathcal{B}) = p_\star(\mathcal{A}_1 \cup \mathcal{A}_2) + p_\star(\mathcal{B}).
\]

Next, we prove if \( \mathcal{A} \in \Sigma^2 \) and \( p_\star(\mathcal{A}) = 0 \), then \( \mathcal{A}_1 \in \Sigma^2 \) if \( \mathcal{A}_1 \subseteq \mathcal{A} \). For any \( \mathcal{B} \subseteq (\mathcal{A}_1)^c \),

\[
p_\star(\mathcal{A}_1 \cup \mathcal{B}) \leq p_\star(\mathcal{A} \cup \mathcal{B}) = p_\star(\mathcal{A} \cup (\mathcal{B} \setminus \mathcal{A})) = p_\star(\mathcal{A}) + p_\star(\mathcal{B} \setminus \mathcal{A}) \leq p_\star(\mathcal{A}_1) + p_\star(\mathcal{B}).
\]

\( p_\star(\mathcal{A}_1 \cup \mathcal{B}) \geq p_\star(\mathcal{A}_1) + p_\star(\mathcal{B}) \) directly follows from superadditivity of \( p_\star \). This completes the conclusion.

Now let \( \mathcal{A} \in \Sigma^3 \). From Lemma 2.5.1, there exist two events \( \mathcal{A}_1 \) and \( \mathcal{A}_2 \) in \( \mathcal{A}^u_a \) satisfying \( \mathcal{A}_1 \subseteq \mathcal{A} \) and \( \mathcal{A}_2 \subseteq \mathcal{A}^c \) such that

\[
p_\star(\mathcal{A}) = p_\star(\mathcal{A}_1) \text{ and } p_\star(\mathcal{A}^c) = p_\star(\mathcal{A}_2).
\]

As a result, \( (\mathcal{A}_1 \cup \mathcal{A}_2)^c \in \mathcal{A}^u_a \subseteq \Sigma^2 \) and \( p((\mathcal{A}_1 \cup \mathcal{A}_2)^c) = 0 \). Since \( \mathcal{A} \setminus \mathcal{A}_1 \subseteq (\mathcal{A}_1 \cup \mathcal{A}_2)^c \), \( (\mathcal{A}_1 \cup \mathcal{A}_2)^c \subseteq \Sigma^2 \); therefore, \( \mathcal{A} \setminus \mathcal{A}_1 \in \Sigma^2 \). Since \( \Sigma^2 \) is closed with respect to disjoint unions, \( \mathcal{A} = \mathcal{A}_1 \cup (\mathcal{A} \setminus \mathcal{A}_1) \in \Sigma^2 \). Similarly, \( \mathcal{A}^c \in \Sigma^2 \). \( \blacksquare \)

**Lemma 2.5.4** If \( p_\star(\mathcal{A}) = p^*(\mathcal{A}) \) for event \( \mathcal{A} \), then \( p_\star(\mathcal{A}) + p_\star(\mathcal{A}^c) = 1 \).
Proof. Suppose for event $A$, $p_*(A) = p^*(A)$. From $p_*(A) = p^*(A)$, we have $p_*(A^c) = 1 - p^*(A) = 1 - p_*(A) = p^*(A^c)$. If $p_*(A) + p_*(A^c) \neq 1$, then $p_*(A) + p_*(A^c) < 1$ by the superadditivity of $p_*$. Accordingly,

$$p^*(A) + p^*(A^c) = [1 - p_*(A^c)] + [1 - p_*(A)] = 2 - [p_*(A) + p_*(A^c)] > 1.$$

But this contradicts the assumption.

Proof of Theorem 2.3.5:

Obviously from Theorem 2.3.4, we have for any acts $f, g \in \mathcal{F}$

$$f \succeq g \iff \int u(f)dp_* \geq \int u(g)dp_* \iff \int u(f)dp^* \geq \int u(g)dp^*. \quad (2.37)$$

Claim 1: For any act $f$,

$$\int u(f)dp_* = \int u(f)dp^*. \quad (2.38)$$

Proof of Claim 1: If 2.38 is not true, then there exists an act $f$ such that $\int u(f)dp_* < \int u(f)dp^*$. Now pick outcome $x$ such that $\int u(f)dp_* < u(x) < \int u(f)dp^*$. But this contradicts 2.37.

Claim 2: For any event $A$, $p_*(A) = p^*(A)$.

Proof of Claim 2: Suppose this is not true, then there exists an event $A$ such that $p^*(A) > p_*(A)$. Let

$$f = \begin{cases} x^* & \text{if } s \in A \\ x & \text{if } s \in A^c \end{cases},$$

where $x^* \succ x$. Then

$$\int u(f)dp_* = u(x)[1 - p_*(A)] + u(x^*)p_*(A) = u(x) + [u(x^*) - u(x)]p_*(A) < u(x) + [u(x^*) - u(x)]p^*(A) = \int u(f)dp^*.$$

But this contradicts Claim 1.

Claim 3: If an act $f$ has the following form

$$f = \begin{cases} x & \text{If } s \in A \\ y & \text{If } s \in A^c \end{cases},$$

...
then

\[ \sup\{\int u(h)dp : h \in \mathcal{F}^{ua}, h \leq f\} = \int u(f)dp, \quad (2.39) \]

where \( x \succ y \) and the integral of left side of (2.39) is in the sense of Choquet.

Proof of Claim 3: For any integer \( n > 1 \), by the definition of inner measures, there is \( A_n \in \mathcal{A}^{ua} \) and \( A_n \subset A \) such that \( p(A_n) > p_*(A) - 1/n \). Let

\[ g_n(s) = \begin{cases} x & \text{if } s \in A_n \\ y & \text{if } s \in A_n^c. \end{cases} \]

Therefore, \( g_n \in \mathcal{F}^{ua} \) and \( g_n \leq f \). Thus,

\[ \int u(g_n)dp = u(x)p(A_n) + u(y)p(A_n^c) \]
\[ = (u(x) - u(y))p(A_n) + u(y) \]
\[ > (u(x) - u(y))p_*(A) + u(y) - (u(x) - u(y))/n. \]

Therefore,

\[ \sup\{\int u(h)dp : h \in \mathcal{F}^{ua}, h \leq f'\} \geq \lim_{n \to \infty} [(u(x) - u(y))p_*(A) + u(y) \]
\[ - (u(x) - u(y))/n] \]
\[ = (u(x) - u(y))p_*(A) + u(y) \]
\[ = \int u(f)dp. \quad (2.40) \]

The conclusion follows from (2.40) and part (1) Theorem 2.2.3.

Claim 4: If an event \( A \) is unambiguous, then

\[ p_*(A \cup B) = p(A) + p_*(B), \quad \forall B \subset A^c. \quad (2.41) \]

Proof of Claim 4: To prove (2.41), by superadditivity of \( p_* \), we only need to prove

\[ p_*(A \cup B) \leq p(A) + p_*(B), \quad \forall B \subset A^c. \quad (2.42) \]

Suppose (2.42) is not true, then there exists a unambiguous event \( A \in \mathcal{A}^{ua} \) and \( B \subset A^c \) such that

\[ p_*(A \cup B) > p(A) + p_*(B). \]

Since \( u \) is nonconstant and convex ranged, there exist outcomes satisfying

\[ u(x_1) > u(y_1) > u(y_2) > u(x_2). \quad (2.43) \]
Let
\[
f_1(s) = \begin{cases} 
  x_1 & \text{if } s \in B \\
  x_2 & \text{if } s \in A \setminus B \\
  x_1 & \text{if } s \in A 
\end{cases}, \quad 
 g_1(s) = \begin{cases} 
  y_1 & \text{if } s \in B \\
  y_2 & \text{if } s \in A \setminus B \\
  x_1 & \text{if } s \in A 
\end{cases}
\]
and
\[
f_2(s) = \begin{cases} 
  x_1 & \text{if } s \in B \\
  x_2 & \text{if } s \in A \setminus B \\
  y_1 & \text{if } s \in A 
\end{cases}, \quad 
 g_2(s) = \begin{cases} 
  y_1 & \text{if } s \in B \\
  y_2 & \text{if } s \in A \setminus B \\
  y_1 & \text{if } s \in A 
\end{cases}.
\]
(2.44)
Therefore, by (2.36) and Claim 3,
\[
f_1 \succ g_1 \iff \sup\{\int u(h)dp : h \in \mathcal{F}^{ua}, h \leq f_1\} > \sup\{\int u(h)dp : h \in \mathcal{F}^{ua}, h \leq g_1\} \iff \int u(f_1)dp_\ast > \sup\{\int u(h)dp : h \in \mathcal{F}^{ua}, h \leq g_1\}
\]
and
\[
f_2 \prec g_2 \iff \sup\{\int u(h)dp : h \in \mathcal{F}^{ua}, h \leq f_2\} < \sup\{\int u(h)dp : h \in \mathcal{F}^{ua}, h \leq g_2\} \iff \sup\{\int u(h)dp : h \in \mathcal{F}^{ua}, h \leq f_2\} < \int u(g_2)dp_\ast.
\]
If the outcomes \(\{x_1, x_2, y_1, y_2\}\) satisfy
\[
\int u(f_1)dp_\ast > \int u(g_1)dp_\ast \quad \int u(f_2)dp_\ast < \int u(g_2)dp_\ast, \quad \tag{2.45}
\]
then
\[
\sup\{\int u(h)dp : h \in \mathcal{F}^{ua}, h \leq f_1\} = \int u(f_1)dp_\ast > \int u(g_1)dp_\ast \geq \sup\{\int u(h)dp : h \in \mathcal{F}^{ua}, h \leq g_1\}
\]
and
\[
\sup\{\int u(h)dp : h \in \mathcal{F}^{ua}, h \leq f_2\} \leq \int u(f_2)dp_\ast < \int u(g_2)dp_\ast = \sup\{\int u(h)dp : h \in \mathcal{F}^{ua}, h \leq g_2\}.
\]
2.5. APPENDIX

By direct computations,
\[ \int u(f_1)dp. > \int u(g_1)dp. \iff \]
\[ [u(x_2) - u(y_2)][1 - p_*(A \cup B)] + [u(x_1) - u(y_1)][p_*(A \cup B) - p_*(A)] > 0 \quad (2.46) \]
and
\[ \int u(f_2)dp. < \int u(g_2)dp. \iff \]
\[ [u(x_2) - u(y_2)][1 - p_*(A \cup B)] + [u(x_1) - u(y_1)]p_*(B) < 0. \quad (2.47) \]
Since \( p_*(A \cup B) > p_*(A) + p_*(B) \) and the range of \( u \) is convex, there exist \( x_1, x_2, y_1, y_2 \) satisfying (2.43), (2.46) and (2.47). This implies that
\[
\begin{align*}
  f_1(s) &= \begin{cases} x_1 & \text{if } s \in B \\ x_2 & \text{if } s \in A^c \setminus B \\ x_1 & \text{if } s \in A \end{cases} > g_1(s) &= \begin{cases} y_1 & \text{if } s \in B \\ y_2 & \text{if } s \in A^c \setminus B \\ x_1 & \text{if } s \in A \end{cases} \quad \text{but} \\
  f_2(s) &= \begin{cases} x_1 & \text{if } s \in B \\ x_2 & \text{if } s \in A^c \setminus B \\ y_1 & \text{if } s \in A \end{cases} < g_2(s) &= \begin{cases} y_1 & \text{if } s \in B \\ y_2 & \text{if } s \in A^c \setminus B \\ y_1 & \text{if } s \in A. \end{cases}
\end{align*}
\]
This contradicts that \( A \) is unambiguous.

Then \( A^{ua} = 2^S \) directly follows from Lemmas 2.5.3, 2.5.4 and part 3 of Theorem 2.4.2.

**Proof of Theorem 2.3.6:** We only need to prove \((a) \implies (b)\). By Weak Monotonicity and Certainty equivalent, for any act \( f \), there exists outcome \( x_f \) such that \( x_f \sim f \) and
\[
\sup\{ \int u(g)dp : f \geq g \in \mathcal{F}^{ua} \} \leq u(x_f) \leq \inf\{ \int u(g)dp : \mathcal{F}^{ua} \in g \geq f \}.
\]
Accordingly, there exists a constant \( \alpha_f \in [0, 1] \) such that
\[
U(f) = u(x_f) = \alpha_f \sup\{ \int u(g)dp : f \geq g \in \mathcal{F}^{ua} \} + (1 - \alpha_f) \inf\{ \int u(g)dp : \mathcal{F}^{ua} \in g \geq f \}.
\]
By the convex range of \( p \) and part (ii) of theorem 2.2.3,
\[
U(f) = \alpha_f \int u(f)dp. + (1 - \alpha_f) \int u(f)dp^*.
\]
2.5.4 Appendix D:

Proof of Lemma 2.4.1 (1): we prove that $A, B \in \overline{A} \implies A \cap B \in \overline{A}$.

Suppose

$$
\begin{pmatrix}
  f(s) & \text{if } s \in A \cap B \\
  g(s) & \text{if } s \in (A \cap B)^c
\end{pmatrix}
= 
\begin{pmatrix}
  f'(s) & \text{if } s \in A \cap B \\
  g(s) & \text{if } s \in (A \cap B)^c
\end{pmatrix}.
$$

Rewrite the above in the following way

$$
\begin{pmatrix}
  f(s) & \text{if } s \in A \cap B \\
  g(s) & \text{if } s \in A \setminus B \\
  g'(s) & \text{if } s \in B \setminus A \\
  g'(s) & \text{if } s \in (A \cup B)^c
\end{pmatrix}
= 
\begin{pmatrix}
  f'(s) & \text{if } s \in A \cap B \\
  g(s) & \text{if } s \in A \setminus B \\
  g'(s) & \text{if } s \in B \setminus A \\
  g'(s) & \text{if } s \in (A \cup B)^c
\end{pmatrix}.
$$

Because $A \in \overline{A}$,

$$
\begin{pmatrix}
  f(s) & \text{if } s \in A \cap B \\
  g(s) & \text{if } s \in A \setminus B \\
  g'(s) & \text{if } s \in B \setminus A \\
  g'(s) & \text{if } s \in (A \cup B)^c
\end{pmatrix}
= 
\begin{pmatrix}
  f'(s) & \text{if } s \in A \cap B \\
  g(s) & \text{if } s \in A \setminus B \\
  g'(s) & \text{if } s \in B \setminus A \\
  g'(s) & \text{if } s \in (A \cup B)^c
\end{pmatrix}.
$$

and this is equivalent to

$$
\begin{pmatrix}
  f(s) & \text{if } s \in A \cap B \\
  g'(s) & \text{if } s \in B \setminus A \\
  g(s) & \text{if } s \in A \setminus B \\
  g'(s) & \text{if } s \in (A \cup B)^c
\end{pmatrix}
= 
\begin{pmatrix}
  f'(s) & \text{if } s \in A \cap B \\
  g'(s) & \text{if } s \in B \setminus A \\
  g(s) & \text{if } s \in A \setminus B \\
  g'(s) & \text{if } s \in (A \cup B)^c
\end{pmatrix}.
$$

By $B \in \overline{A}$,

$$
\begin{pmatrix}
  f(s) & \text{if } s \in A \cap B \\
  g'(s) & \text{if } s \in B \setminus A \\
  g'(s) & \text{if } s \in A \setminus B \\
  g'(s) & \text{if } s \in (A \cup B)^c
\end{pmatrix}
= 
\begin{pmatrix}
  f'(s) & \text{if } s \in A \cap B \\
  g'(s) & \text{if } s \in B \setminus A \\
  g'(s) & \text{if } s \in A \setminus B \\
  g'(s) & \text{if } s \in (A \cup B)^c
\end{pmatrix}.
$$

and this means

$$
\begin{pmatrix}
  f(s) & \text{if } s \in A \cap B \\
  g'(s) & \text{if } s \in (A \cap B)^c
\end{pmatrix}
\leq 
\begin{pmatrix}
  f'(s) & \text{if } s \in A \cap B \\
  g'(s) & \text{if } s \in (A \cap B)^c
\end{pmatrix}, \text{ for all } g',
$$
2.5. APPENDIX

i.e., $A \cap B \in \mathcal{A}$.

Proof of Lemma 2.4.1 (2) and (3): (2) directly follows from the definition of unambiguous events. (3) is routine. ■

Proof of Theorem 2.4.2:

Proof of 2: Step 1: We prove

$$\Sigma^1 = \{A \in 2^S : \nu(B) = \nu(B \cap A) + \nu(B \cap A^c), \forall B\}.$$  

Given $A \in \Sigma^1$, $B \cap A \subset A$ and $B \cap A^c \subset A^c$, then it follows from the definition of $\Sigma^1$ that

$$\nu(B) = \nu((B \cap A) \cup (B \cap A^c)) = \nu(B \cap A) + \nu(B \cap A^c).$$

That is, $\Sigma^1 \subset \{A \in 2^S : \nu(B) = \nu(B \cap A) + \nu(B \cap A^c), \forall B\}$.

Next, given $A \in \{A \in 2^S : \nu(B) = \nu(B \cap A) + \nu(B \cap A^c), \forall B\}$, we prove that $A \in \Sigma^1$. In fact, for any $A_1 \subset A$ and $B \subset A^c$, then

$$\nu(A_1 \cup B) = \nu((A_1 \cup B) \cap A) + \nu((A_1 \cup B) \cap A^c)$$

$$= \nu(A_1) + \nu(B).$$

That is, $\{A \in 2^S : \nu(B) = \nu(B \cap A) + \nu(B \cap A^c), \forall B\} \subset \Sigma^1$.

Step 2: We prove

$$\{A \in 2^S : \nu(B) = \nu(B \cap A) + \nu(B \cap A^c), \forall B\}$$

$$= \{A \in 2^S : \nu(B \cup A) + \nu(B \cap A) = \nu(B) + \nu(A), \forall B\}.$$  

We first prove $\{A \in 2^S : \nu(B) = \nu(B \cap A) + \nu(B \cap A^c), \forall B\}$ is an algebra.

Evidently $\emptyset \in \Sigma^1$ and $A \in \Sigma^1$ implies $A^c \in \Sigma^1$. Then $\Sigma^1$ is an algebra if

$A, B \in \Sigma^1$ implies $A \cap B \in \Sigma^1$.

Fix $C \in 2^S$. Since $A \in \Sigma^1$ and $A^c \cap C = C \setminus A$

$$\nu(C) = \nu(C \cap A) + \nu(C \setminus A),$$

since $B \in \Sigma^1$

$$\nu(C \cap A) = \nu((C \cap A) \cap B) + \nu((C \cap A) \setminus B)$$
and since $A \in \Sigma^1$

$$\nu(C \setminus (A \cap B)) = \nu((C \setminus (A \cap B)) \cap A) + \nu((C \setminus (A \cap B)) \setminus A)$$

$$= \nu((C \cap A) \setminus B) + \nu(C \setminus A).$$

Combine these equalities we derive

$$\nu(C) = \nu(C \cap (A \cap B)) + \nu(C \setminus (A \cap B)).$$

Since $C$ was arbitrary we proved $A \cap B \in \Sigma^1$, hence $\Sigma^1$ is an algebra. Finally, let $A \in \{ A \in 2^S : \nu(B) = \nu(B \cap A) + \nu(B \cap A^c), \forall B \} \text{ and any } B \in 2^S$, then

$$\nu(A \cup B) = \nu((A \cup B) \cap A) + \nu((A \cup B) \cap A^c)$$

$$= \nu(A) + \nu(B \cap A^c) = \nu(A) + [\nu(B) - \nu(B \cap A)].$$

Step 3:

$(\implies)$: Let $A \in \Sigma^1$, and $u(f(s)) = \Sigma_i u_i I_{E_i}(s)$, where $\{E_i\}$ is a finite partition of $S$, $u_1 < u_2 < \cdots < u_n$, then

$$\int u(f(s))d\nu = \sum_{i=1}^n u_i [\nu(\bigcup_{j=i}^n E_j) - \nu(\bigcup_{j=i+1}^n E_j)]$$

$$= \sum_{i=1}^n u_i [\nu(A \cap (\bigcup_{j=i}^n E_j)) + \nu(A^c \cap (\bigcup_{j=i+1}^n E_j))]$$

$$- \sum_{i=1}^n u_i [\nu(A \cap (\bigcup_{j=i+1}^n E_j)) - \nu(A^c \cap (\bigcup_{j=i+1}^n E_j))]$$

$$= \sum_{i=1}^n u_i [\nu(A \cap (\bigcup_{j=i}^n E_j)) - \nu(A \cap (\bigcup_{j=i+1}^n E_j))]$$

$$+ \sum_{i=1}^n u_i [\nu(A^c \cap (\bigcup_{j=i}^n E_j)) - \nu(A^c \cap (\bigcup_{j=i+1}^n E_j))]$$

$$= \int_A u(f)d\nu + \int_{A^c} u(f)d\nu.$$

$(\impliedby)$: Given $B \in 2^S$, let

$$f = \begin{cases} x^* & \text{if } s \in B \\ x^{**} & \text{if } s \in B^c \end{cases},$$

where $u(x^*) > u(x^{**}) = 0$ (The assumption $u(x^{**}) = 0$ is not necessary, just for convenience.)

Let $A \in 2^S$ satisfy

$$\int u(f)d\nu = \int_A u(f)d\nu + \int_{A^c} u(f)d\nu,$$
therefore,
\[
\int u(f) d\nu = u(x^*)[1 - \nu(B)] + u(x^*)\nu(B) = u(x^*)\nu(B)
\]
\[
= \int_A u(f) d\nu + \int_{A^c} u(f) d\nu
\]
\[
= u(x^*)[1 - \nu(B \cap A)] + u(x^*)\nu(B \cap A)
\]
\[
+ u(x^*)[1 - \nu(B \cap A^c)] + u(x^*)\nu(B \cap A^c)
\]
\[
= u(x^*)\nu(B \cap A) + u(x^*)\nu(B \cap A^c),
\]
\[
\Rightarrow u(x^*)\nu(B) = u(x^*)\nu(B \cap A) + u(x^*)\nu(B \cap A^c),
\]
and this means that
\[
\nu(B) = \nu(B \cap A) + \nu(B \cap A^c).
\]

Step 4: $\Sigma^1 \subset \overline{A} \subset A^{ua}$ directly follow from the definitions of $\Sigma^1$, $\overline{A}$, $A^{ua}$.

Proof of 3: ($\Rightarrow$): Let $A \in \Sigma^2$ and

\[
f(s) = \begin{cases} 
  f_1 & \text{if } s \in A^c \\
  x & \text{if } s \in A
\end{cases} = \begin{cases} 
  x_1 & \text{if } s \in E_1 \\
  x_2 & \text{if } s \in E_2 \\
  \vdots & \vdots \\
  x_n & \text{if } s \in E_n \\
  x & \text{if } s \in A
\end{cases}
\]

\[
u(x_1) < \nu(x_2) < \ldots < \nu(x_n).
\]

Case 1: $u(x) < u(x_1)$.

In this case,
\[
\int u(f) d\nu = u(x)[\nu(A \cup A^c) - \nu(A^c)] + \sum_{i=1}^n u(x_i)[\nu(\cup_{j=i}^n E_j) - \nu(\cup_{j=i+1}^n E_j)]
\]
\[
= u(x)[\nu(A) + \nu(A^c) - \nu(A^c)] + \sum_{i=1}^n u(x_i)[\nu(\cup_{j=i}^n E_j) - \nu(\cup_{j=i+1}^n E_j)]
\]
\[
= \int_{A^c} u(f_1) d\nu + u(x)\nu(A).
\]

Case 2: $u(x) = u(x_m)$ for some $m \in \{1, 2, \ldots, n\}$. 
In this case,

\[
\int u(f) \, d\nu = \sum_{i=1}^{n-1} u(x_i) \left[ \nu(\bigcup_{j=i}^{n} E_j \cup A) - \nu(\bigcup_{j=i+1}^{n} E_j \cup A) \right] \\
+ u(x) \left[ \nu(\bigcup_{j=m}^{n} E_j \cup A) - \nu(\bigcup_{j=m+1}^{n} E_j \cup A) \right] \\
+ \sum_{i=m+1}^{n} u(x_i) \left[ \nu(\bigcup_{j=m}^{i-1} E_j) - \nu(\bigcup_{j=i+1}^{n} E_j) \right] \\
= \sum_{i=1}^{n-1} u(x_i) \left[ \nu(\bigcup_{j=i}^{n} E_j) + \nu(A) - \nu(\bigcup_{j=i+1}^{n} E_j) + \nu(A) \right] \\
+ u(x) \left[ \nu(\bigcup_{j=m}^{n} E_j) + \nu(A) - \nu(\bigcup_{j=m+1}^{n} E_j) \right] \\
+ \sum_{i=m+1}^{n} u(x_i) \left[ \nu(\bigcup_{j=m}^{i-1} E_j) - \nu(\bigcup_{j=i+1}^{n} E_j) \right] \\
= \sum_{i=m+1}^{n} u(x_i) \left[ \nu(\bigcup_{j=i}^{n} E_j) - \nu(\bigcup_{j=i+1}^{n} E_j) \right] + u(x) \nu(A) \\
= \int_{A^c} u(f) \, d\nu + u(x) \nu(A).
\]

Case 3: \( u(x_m) < u(x) < u(x_{m+1}) \) for some \( m \).

In this case,

\[
\int u(f) \, d\nu = \sum_{i=1}^{n-1} u(x_i) \left[ \nu(\bigcup_{j=i}^{n} E_j \cup A) - \nu(\bigcup_{j=i+1}^{n} E_j \cup A) \right] \\
+ u(x) \left[ \nu(\bigcup_{j=m}^{n} E_j \cup A) - \nu(\bigcup_{j=m+1}^{n} E_j \cup A) \right] \\
+ \sum_{i=m+1}^{n} u(x_i) \left[ \nu(\bigcup_{j=m}^{i-1} E_j) - \nu(\bigcup_{j=i+1}^{n} E_j) \right] \\
= \sum_{i=1}^{n-1} u(x_i) \left[ \nu(\bigcup_{j=i}^{n} E_j) + \nu(A) - \nu(\bigcup_{j=i+1}^{n} E_j) - \nu(A) \right] \\
+ u(x) \left[ \nu(\bigcup_{j=m}^{n} E_j) + \nu(A) - \nu(\bigcup_{j=m+1}^{n} E_j) \right] \\
+ \sum_{i=m+1}^{n} u(x_i) \left[ \nu(\bigcup_{j=m}^{i-1} E_j) - \nu(\bigcup_{j=i+1}^{n} E_j) \right] \\
= \sum_{i=m+1}^{n} u(x_i) \left[ \nu(\bigcup_{j=i}^{n} E_j) - \nu(\bigcup_{j=i+1}^{n} E_j) \right] + u(x) \nu(A) \\
= \int_{A^c} u(f) \, d\nu + u(x) \nu(A).
\]

Case 4: \( u(x_n) < u(x) \).

In this case,

\[
\int u(f) \, d\nu = \sum_{i=1}^{n} u(x_i) \left[ \nu(\bigcup_{j=i}^{n} E_j \cup A) - \nu(\bigcup_{j=i+1}^{n} E_j \cup A) \right] + u(x) \nu(A) \\
= \sum_{i=1}^{n} u(x_i) \left[ \nu(\bigcup_{j=i}^{n} E_j) + \nu(A) - \nu(\bigcup_{j=i+1}^{n} E_j) - \nu(A) \right] + u(x) \nu(A) \\
= \sum_{i=1}^{n} u(x_i) \left[ \nu(\bigcup_{j=i}^{n} E_j) - \nu(\bigcup_{j=i+1}^{n} E_j) \right] + u(x) \nu(A) \\
= \int_{A^c} u(f_1) \, d\nu + u(x) \nu(A).
\]

In summary, if \( A \in \Sigma^2 \), then

\[
\int u(f) \, d\nu = \int_{A^c} u(f_1) \, d\nu + u(x) \nu(A)
\]
and this implies that

\[
\begin{pmatrix}
  f_1 & \text{if } s \in A^c \\
  x & \text{if } s \in A
\end{pmatrix} \succ
\begin{pmatrix}
  g_1 & \text{if } s \in A^c \\
  x & \text{if } s \in A
\end{pmatrix} \quad \iff
\begin{pmatrix}
  f_1 & \text{if } s \in A^c \\
  y & \text{if } s \in A
\end{pmatrix} \succ
\begin{pmatrix}
  g_1 & \text{if } s \in A^c \\
  y & \text{if } s \in A
\end{pmatrix}, \quad \forall y \in \mathcal{X}.
\]

\((\iff)\): Suppose it is not true. Without loss of generality, then there exists a unambiguous event \( A \in \mathcal{A}^{ua} \) and \( B \subseteq A^c \) such that

\[\nu(A \cup B) > \nu(A) + \nu(B).\]

Since \( u \) is nonconstant and convex ranged, there exist outcomes satisfying

\[u(x_1) > u(y_1) > u(y_2) > u(x_2).\] \quad (2.48)

Let

\[
\begin{align*}
f_1(s) &= \begin{pmatrix}
  x_1 & \text{if } s \in B \\
  x_2 & \text{if } s \in A^c \setminus B \\
  x_1 & \text{if } s \in A
\end{pmatrix}, &
g_1(s) &= \begin{pmatrix}
  y_1 & \text{if } s \in B \\
  y_2 & \text{if } s \in A^c \setminus B \\
  x_1 & \text{if } s \in A
\end{pmatrix} \quad \text{and} \\
f_2(s) &= \begin{pmatrix}
  x_1 & \text{if } s \in B \\
  x_2 & \text{if } s \in A^c \setminus B \\
  y_1 & \text{if } s \in A
\end{pmatrix}, &
g_2(s) &= \begin{pmatrix}
  y_1 & \text{if } s \in B \\
  y_2 & \text{if } s \in A^c \setminus B \\
  y_1 & \text{if } s \in A
\end{pmatrix}.
\end{align*}
\]

By direct computations,

\[
\int u(f_1) d\nu > \int u(g_1) d\nu \quad \iff \\
[u(x_2) - u(y_2)[1 - \nu(A \cup B)] + [u(x_1) - u(y_1)][\nu(A \cup B) - \nu(A)] > 0 \quad (2.49)
\]

and

\[
\int u(f_2) d\nu < \int u(g_2) d\nu \quad \iff \\
[u(x_2) - u(y_2)[1 - \nu(A \cup B)] + [u(x_1) - u(y_1)]\nu(B) < 0. \quad (2.50)
\]
Since $\nu(A \cup B) > \nu(A) + \nu(B)$ and the range of $u$ is convex, there exist $x_1, x_2, y_1, y_2$ satisfying (2.48), (2.49) and (2.50). This implies that $f_1 \succ g_1$ and $f_2 \prec g_2$. This contradicts that $A$ is unambiguous. \hfill \blacksquare

**Proof of 4:** Obviously $A \in \Sigma^3 \implies A^c \in \Sigma^3$ and $\Omega, \phi$ are in $\Sigma^3$. To prove that $\Sigma^3$ is an algebra when $\nu$ is convex, it suffices to show that if $A_1, A_2$ are in $\Sigma^3$, then $A_1 \cap A_2$ and $A_1 \cup A_2$ are also in $\Sigma^3$. That $\nu$ is convex implies that $\overline{\nu}$ is concave. Therefore,

\[
\nu(A_1) + \nu(A_2) \leq \nu(A_1 \cap A_2) + \nu(A_1 \cup A_2) \leq \overline{\nu}(A_1 \cap A_2) + \overline{\nu}(A_1 \cup A_2)
\]

Because

\[
\nu(A_1 \cup A_2) \leq \overline{\nu}(A_1 \cup A_2) \quad \text{and} \quad \nu(A_1 \cap A_2) \leq \overline{\nu}(A_1 \cap A_2),
\]

then

\[
\nu(A_1 \cap A_2) = \overline{\nu}(A_1 \cap A_2) \quad \text{and} \quad \nu(A_1 \cup A_2) = \overline{\nu}(A_1 \cup A_2).
\]

**Proof of 5:** The set inclusions follow from the definitions of $\Sigma^i$. Example 2.4.3 shows that the sets differ in general.

**Proof of 6:** We only need to prove that $\Sigma^3 \subseteq \Sigma^2$. Since $\nu$ is convex,

\[
\nu(A) = \inf\{p(A) : p \text{ is in the core of } \nu\}.
\]

Let $\nu(A) = 1 - \nu(A^c)$. Then we prove $p(A) = q(A)$, for any $p, q$ in the core of $\nu$. Suppose this is not true. That is, there are $p, q$ in the core of $\nu$ such that $p(A) > q(A)$ and $p(A^c) < q(A^c)$. Thus, $\nu(A) \leq q(A)$, $\nu(A^c) \leq p(A^c)$ and $\nu(A) + \nu(A^c) \leq q(A) + p(A^c) < q(A) + q(A^c) = 1$, a contradiction. If $p(A) = q(A)$ for any $p, q$ in the core of $\nu$, then $A$ is in $\Sigma^2$. \hfill \blacksquare
Bibliography


Chapter 3

Qualitative Probabilities on λ-Systems
3.1 Introduction

Given a state space $S$ and a binary relation $\succeq$ on a class $\mathcal{A}$ of events or subsets of $S$, a number of papers have described necessary and sufficient conditions in order that $\succeq$ admit numerical representation by a (finitely additive) probability measure. For a finite state space see [9] and [12]; for an infinite state space, see [11], [14] and [6], where the representing probability measure is convex-ranged. The cited results help to axiomatize decision theories in which preference is based on probabilities and these represent beliefs about likelihoods of events (see also [10]).

In all of the above studies, it is assumed that $\mathcal{A}$ is a $\sigma$-algebra. This a priori restriction is problematic for the following reason: The Ellsberg Paradox and related evidence show that many decision-makers do not attach probabilities to some events, namely to ‘ambiguous’ ones. In other words, probabilities are assigned only to ‘unambiguous’ events. But the collection of such events is typically not a $\sigma$-algebra because, as illustrated shortly, it is not closed with respect to intersections. On the other hand, it is intuitive that the collection of unambiguous events be closed with respect to complements and disjoint unions. Thus, as argued in [13] (chapter 2) and chapter 3, a $\lambda$-system is the more appropriate mathematical structure for modeling the collection of unambiguous events.

This paper provides necessary and sufficient conditions such that the binary relation $\succeq$ on a $\lambda$-system can be represented by a convex-ranged, finitely additive probability measure.

To illustrate the failure of $\mathcal{A}$ to be an algebra, consider the following example taken from [13]: There are 100 balls in an urn and a ball’s color may be black ($B$), red ($R$), grey ($G$) or white ($W$). The sum of black and red ball is 50 and the sum of black and grey ball is also 50. One ball will be drawn at random. It is intuitive that the unambiguous events are given by

$$\mathcal{A} \triangleq \{\emptyset, \{B, R, G, W\}, \{B, G\}, \{R, W\}, \{B, R\}, \{G, W\}\},$$

where each of these events has the obvious probability. Observe that $\{B, G\}$ and $\{B, R\}$ are in $\mathcal{A}$, but that $\{B\} = \{B, G\} \cap \{B, R\}$ is not in $\mathcal{A}$. 
3.2. PRELIMINARIES

3.2 Preliminaries

3.2.1 λ-Systems

Let $S$ be a nonempty set. Say that a nonempty class of subsets $\mathcal{A}$ of $S$ is a $\lambda$-system, if

\begin{align*}
\lambda.1 & \quad S \in \mathcal{A}; \\
\lambda.2 & \quad A \in \mathcal{A} \implies A^c \in \mathcal{A}; \text{ and} \\
\lambda.3 & \quad A_n \in \mathcal{A}, n = 1, 2, \ldots \text{ and } A_i \cap A_j = \emptyset, \forall i \neq j \implies \bigcup_n A_n \in \mathcal{A}.
\end{align*}

These properties are intuitive if we interpret $\mathcal{A}$ as the collection of all events that are assigned probability by the decision-maker. For example, if she can assign probability to $A$, then it is natural for her to assign the complementary probability to $A^c$. For $\lambda.3$, if she can assign probabilities to the disjoint events $A$ and $B$, then it is natural for her to assign the sum of these probabilities to $A \cup B$. On the other hand, when these events have a nonempty intersection, the union may very well be ambiguous. Equivalently, as we have seen, the intersection of two unambiguous events may be ambiguous.

We have the following Lemma (see [2]):

**Lemma 3.2.1** A nonempty class of subsets $\mathcal{A}$ of $S$ is a $\lambda$-system if and only if

\begin{align*}
\lambda.1' & \quad \emptyset, S \in \mathcal{A}; \\
\lambda.2' & \quad A, B \in \mathcal{A} \text{ and } A \subseteq B \implies B \setminus A \in \mathcal{A}; \text{ and} \\
\lambda.3' & \quad A_n \in \mathcal{A} \text{ and } A_n \subseteq A_{n+1}, n = 1, 2, \ldots \implies \bigcup_n A_n \in \mathcal{A}.
\end{align*}

Say that a function

\[ p : \mathcal{A} \rightarrow [0, 1] \quad (3.1) \]

is a countably additive probability measure over $\mathcal{A}$ if:
P.1 \( p(\emptyset) = 0, p(S) = 1; \) and

P.2 \( p(\bigcup_n A_n) = \sum_n p(A_n), \forall A_i \cap A_j = \emptyset, \) for any \( i \neq j. \)

Denote by \( (S, \mathcal{A}, p) \) a \( \lambda \)-system probability space. Say that a probability \( p \) is *convex-ranged* if for all \( A \in \mathcal{A} \) and \( 0 < r < 1, \) there exists \( B \subset A, B \in \mathcal{A}, \) such that \( p(B) = r p(A). \)

### 3.2.2 Qualitative Probabilities

Let \( \succeq \) be a binary relation \( \mathcal{A}. \) Say that \( \succeq \) is a *qualitative probability* if

Q.1 \( \succeq \) is a weak order (complete and transitive);
Q.2 \( A \succeq \emptyset \) for all \( A \in \mathcal{A}; \)
Q.3 \( S \succ \emptyset; \) and
Q.4 \( A \cap C = B \cap C = \emptyset \) implies \( [A \succeq B \iff A \cup C \succeq B \cup C]. \)

Say that a countably additive probability measure \( p \) on \( \mathcal{A} \) represents \( \succeq \) if

\[
A \succeq B \iff p(A) \geq p(B),
\]

for all \( A \) and \( B \) in \( \mathcal{A}. \)

Say that a sequence \( \{A_n\}_{n=1}^{\infty} \) in \( \mathcal{A} \) *converges to* \( A \in \mathcal{A} \) if for any events \( A_*, A^* \) in \( \mathcal{A} \) with \( A_* \prec A \prec A^* \), there exists an integer \( N \) such that

\[
A_* \prec A_n \prec A^*, \quad \text{whenever } n \geq N.
\]

### 3.3 Main Theorem

Given a qualitative probability \( \succeq \) on a \( \lambda \)-system \( \mathcal{A}, \) can we find a representing convex-ranged probability measure? Necessary and sufficient conditions are given next.
3.4. THE PROOF OF THEOREM 3.1

Denote by

\[ \mathcal{N}(\emptyset) = \{ A \in \mathcal{A} : A \sim \emptyset \}. \]

A partition \( \{ A_i \}_{i=1}^n \) of \( S \) in \( \mathcal{A} \) is uniform if \( A_1 \sim A_2 \sim \cdots \sim A_n \).

**Theorem 3.3.1** Let \( \mathcal{A} \) be a \( \lambda \)-system and \( \succeq \) a qualitative probability on \( \mathcal{A} \). Then there exists a convex-ranged, countably additive probability measure \( p \) on \( \mathcal{A} \) representing \( \succeq \) if and only if \( \succeq \) satisfies the following:

(a) (i) If \( A \in \mathcal{A} \setminus \mathcal{N}(\emptyset) \), then there is a finite partition \( \{ A_1, A_2, \ldots, A_n \} \) of \( S \) in \( \mathcal{A} \) such that (1) \( A_i \subset A \) or \( A_i \subset A^c \), \( i = 1, 2, \ldots, n \); (2) \( A_i \prec A \), \( i = 1, 2, \ldots, n \).

(ii) If \( A, B, C \in \mathcal{A} \setminus \mathcal{N}(\emptyset) \) and \( A \cap C = \emptyset \), \( A \prec B \), then there is a finite partition \( \{ C_1, C_2, \ldots, C_m \} \) of \( S \) in \( \mathcal{A} \) such that (1) \( C_i \subset A \) or \( C_i \subset A^c \), and \( C_i \subset C \) or \( C_i \subset C^c \), \( i = 1, 2, \ldots, m \); (2) \( C_i \prec A \), and \( C_i \prec C \), \( i = 1, 2, \ldots, m \). and (3) \( A \cup C_i \prec B \), \( i = 1, 2, \ldots, m \).

(b) If \( \{ A_n \}_{n=1}^{\infty} \) is a decreasing sequence of events in \( \mathcal{A} \), then \( \{ A_n \}_{n=1}^{\infty} \succeq \)-converges to \( \cap_{n=1}^{\infty} A_n \) and \( \{ A_n \setminus A_{n+1} \}_{n=1}^{\infty} \succeq \)-converges to \( \emptyset \).

(c) For any two uniform partitions \( \{ A_i \}_{i=1}^n \) and \( \{ B_i \}_{i=1}^n \) of \( S \) in \( \mathcal{A} \), \( \cup_{i \in I} A_i \sim \cup_{i \in J} B_i \) if \( |I| = |J| \).

Moreover, under the conditions (a) – (c), the representing measure \( p \) is unique.

(a) is similar to F5 in Fishburn [6]. The additional axioms (b) and (c) are adopted here to compensate for the fact that \( \mathcal{A} \) is not a \( \sigma \)-algebra.

### 3.4 The Proof of Theorem 3.1

We prove Theorem 3.1 in this section. Because necessity is routine, we provide only the proof of sufficiency. Throughout, \( \succeq \) is a qualitative probability on the \( \lambda \)-system \( \mathcal{A} \).
Lemma 3.4.1 If $A \sim B \sim \emptyset$ and $A \cap B = \emptyset$, then $A \cup B \sim \emptyset$.

Proof. Suppose $A \cup B \succ \emptyset$. Then $A \cup B \succ \emptyset \cup B \sim \emptyset$. By Q.4, $A \succ \emptyset$, a contradiction.

Lemma 3.4.2 If events $A, B \in \mathcal{A}$ with $A \cup B \in \mathcal{A}$, then

$$A \succeq_\ell B \iff A^c \succeq_\ell B^c.$$ 

Proof. Because $\mathcal{A}$ is a $\lambda$-system and $A \cup B \in \mathcal{A}$,

$$A \setminus B, B \setminus A, A \cap B \text{ and } (A \cup B)^c \text{ are all in } \mathcal{A}.$$ 

By Q.4,

$$A \succeq_\ell B \iff (A \setminus B) \cup (A \cap B) \succeq_\ell (B \setminus A) \cup (A \cap B) \iff A \setminus B \succeq_\ell B \setminus A \iff (A \setminus B) \cup (A \cup B)^c \succeq_\ell (B \setminus A) \cup (A \cup B)^c \iff B^c \succeq_\ell A^c.$$ 

Lemma 3.4.3 Let $\succeq$ satisfy (a)-(c) and let $\{A_n\}_{n=1}^{\infty}$ be an increasing sequence of events in $\mathcal{A}$.

1. $\{A_{n+1} \setminus A_n\}_{n=1}^{\infty}$ converges in $\succeq$ to $\emptyset$; and

2. If $\bigcup_{n} A_n \succ B \in \mathcal{A}$ and $\bigcup_{n} A_n \subseteq B^c$, then there exists an integer $N$ such that $A_n \succ B$, whenever $n \geq N$.

Proof. For part 1, by (b) and the fact that $\mathcal{A}$ is a $\lambda$-system, $\{A_n^c\}_{n=1}^{\infty}$ is a decreasing sequence of events in $\mathcal{A}$ and $\{A_n^c \setminus A_{n+1}^c\}_{n=1}^{\infty}$ $\succeq$-converges to $\emptyset$. It follows that $\{A_{n+1} \setminus A_n\}_{n=1}^{\infty}$ $\succeq$-converges to $\emptyset$, because $A_n^c \setminus A_{n+1}^c = A_{n+1} \setminus A_n$.

For part 2, by Lemma 3.4.2, $\cap_{n} A_n^c = (\bigcup_{n} A_n)^c \prec B^c$. By (b), $\{A_n^c\}_n$ is a decreasing sequence of events in $\mathcal{A}$ and $\succeq$-converges to $\cap_{n} A_n^c$. So, there exists an integer $N$ such
that $A_n^c < B^c$, whenever $n \geq N$. Again by Lemma 3.4.2, $A_n \succ B$, whenever $n \geq N$. ■

The following lemma is adapted from [6, pp. 195-8].

**Lemma 3.4.4** Let $\succeq$ satisfy (a)-(c). Then it also satisfies the following:

**C1** If $B, C \in A$ and $B \subseteq C$, then $\emptyset \preceq B \preceq C \preceq S$.

**C4** ($\preceq$) If $A, B, C, D, B \cup C \in A$, $A \cap C = B \cap D = \emptyset$ and $A \preceq B, C \preceq D$, then $A \cup C \preceq B \cup D$.

**C4** ($\sim$) If $A, B, C, D, B \cup C \in A$, $A \cap C = B \cap D = \emptyset$, $A \sim B$ and $C \sim D$, then $A \cup C \sim B \cup D$.

**C5** If $\emptyset \prec A \in A$, then $A$ can be partitioned into two events $B$ and $C$ in $A$ such that $\emptyset \preceq B$ and $\emptyset \preceq C$.

**C6** If $A, B, C \in A \setminus N(\emptyset)$ are pairwise disjoint, $A \preceq B$ and $B \prec A \cup C$, then there exists $D \in A$, $D \subset C$ such that $\emptyset \prec D$ and $B \cup D \prec A \cup (C \setminus D)$.

**C7** If $A, B \in A \setminus N(\emptyset)$ and $A \cap B = \emptyset$, then $B$ can be partitioned into $C$ and $D$ in $A$ such that $C \preceq D \preceq A \cup C$.

**C8** If $\emptyset \prec A \in A$, then $A$ can be partitioned into $B$ and $C$ in $A$ such that $B \sim C$.

**C9** If $\emptyset \prec A \in A$, then for any positive integer $n$ there is a $2^n$ part partition of $A$ in $A$ such that $\sim$ holds between each two events in the partition.

**Proof.**

**C1**: Let $B, C \in A$ and $B \subseteq C$. By $\lambda.2'$, $C \setminus B \in A$. By $Q.2$, $\emptyset \preceq C \setminus B$. By $Q.4$, $B = \emptyset \cup B \preceq (C \setminus B) \cup B = C$. Clearly, $\emptyset \preceq B$ and $C \preceq S$.

**C4($\preceq$)**: $B \cup C \in A \implies B \setminus C, C \setminus B \in A$. $A \preceq B$ and $Q.4 \implies$

$$A \cup (C \setminus B) \preceq B \cup (C \setminus B) = B \cup C.$$

---

1The labelling of these properties follows Fishburn. $C2$ is not needed, while $C3$ is not valid in our setting where $A$ is only a $\lambda$-system.
CHAPTER 3. QUALITATIVE PROBABILITIES ON $\lambda$-SYSTEMS

$C \leq D$ and Q.4 $\implies$ 

\[ B \cup C = C \cup (B \setminus C) \leq D \cup (B \setminus C). \]

Thus, 

\[ A \cup (C \setminus B) \leq D \cup (B \setminus C). \]

Because $B \cap C$ is disjoint from $A \cup D$, 

\[ A \cup C = A \cup (C \setminus B) \cup (B \cap C) \leq D \cup (B \setminus C) \cup (B \cap C) = D \cup B. \]

C4(\sim): Implied by C4(\leq).

C5: By (a) (i), there is a partition \{D_i : i = 1, 2, \ldots, n\} of $S$ in $A$ such that $D_i \prec A$, $i = 1, 2, \ldots, n$ and $D_i \subset A$ or $D_i \subset A^c$. Therefore, $D_i \cap A = D_i$ or $\emptyset$ and $A = \cup_i (D_i \cap A)$. There exists at least one $D_i \subseteq A$ such that $\emptyset < D_i$, for otherwise, Lemma 3.4.1 would imply $A \sim \emptyset$, contradicting the hypothesis. Let $D_1 \prec A$. Now prove that $\emptyset < A \setminus D_1$. If $A \setminus D_1 \leq \emptyset$, then by Q.4, $A = (A \setminus D_1) \cup D_1 \leq \emptyset \cup D_1 = D_1$, a contradiction.

C6: Let $A, B, C \in A$ be pairwise disjoint, $A \leq B$, $B \prec A \cup C$ and $A, B \in A \setminus N(\emptyset)$. We are to prove that there exists $D \in A$, $\emptyset < D \subset C$ and such that $B \cup D < A \cup (C \setminus D)$. If $C \leq \emptyset$, then Q.4 and $A \cap C = \emptyset$ imply $C \cup A \leq \emptyset \cup A = A \leq B$, contradicting the hypothesis $B < A \cup C$. Therefore, $\emptyset < C$. By (a) (ii), there exists a partition \{D_1, D_2, \ldots, D_n\} of $S$ in $A$ such that (a) $D_i \subset C$ or $D_i \subset C^c$ and $D_i \subset B$ or $D_i \subset B^c$, $i = 1, 2, \ldots, n$; (b) $D_i \prec C$ and $D_i \prec B$, $i = 1, 2, \ldots, n$; (c) $B \cup D_i \prec A \cup C$, $i = 1, 2, \ldots, n$. If all $D_i \subset C$ are in $N(\emptyset)$, then $C = \cup D_i \subset C$, $D_i$ is also in $N(\emptyset)$ by Lemma 3.4.1, a contradiction. Therefore, there exists $D_i \subset C$ such that $\emptyset < D_i$ and $B \cup D_i \prec A \cup C$. By C5, there exist a partition \{D', D''\} of $D_i$ in $A$ such that $\emptyset < D' \leq D''$. Thus,

\[ B \cup D' \cup D'' = B \cup D_i \prec A \cup C = A \cup (C \setminus D') \cup D'. \]

By Q.4, $B \cup D'' \prec A \cup (C \setminus D')$ and, using also $D' \leq D''$ and $B \cap D' = B \cap D'' = \emptyset$,

\[ B \cup D' \leq B \cup D'' \prec A \cup (C \setminus D'). \]

C7: Let $A, B$ be two disjoint events in $A$, $\emptyset < A$ and $\emptyset < B$. If $B \leq A$, the conclusion follows easily from C5; take any partition of $B$ into $C$ and $D$ with $\emptyset < C \leq D$. 

Assume that $A < B$. By (a) (ii), there is a partition \{$G_1, G_2, \ldots, G_n$\} of $B$ in $A$ such that $G_i < A$, $i = 1, 2, \ldots, n$. Proceed precisely as in [6, p. 196].

**Proof of CS:** By the arguments in [6, pp. 196-7], there exists a sequence \{$B_1, C_1, D_1$\}, \{$B_2, C_2, D_2$\}, \ldots, \{$B_n, C_n, D_n$\}, \ldots of three part partitions of $A$ in $A$, such that for each $n \geq 1$,

(a) $B_n < C_n \cup D_n$ and $C_n < B_n \cup D_n$,

(b) $B_n \subseteq B_{n+1}$, $C_n \subseteq C_{n+1}$, $D_{n+1} \subseteq D_n$,

(c) $D_{n+1} \subseteq D_n \setminus D_{n+1}$,

It follows that $\cap_n D_n \in A$, $\emptyset < D_n$ and that $D_n$ contains the two disjoint and equally likely events $D_{n+1}$ and $D_n \setminus D_{n+1}$. In addition, $\cap_n D_n \sim \emptyset$. (If not, then by C5 and Lemma 3.4.3, there exists $\emptyset < A^* \in A$ such that $\cap_n D_n \supset A^*$. Conclude that $D_n \setminus D_{n+1} \supset D_{n+1} \supset A^*$, $n = 1, 2, \ldots$ and hence (by (b)) $A^* \sim \emptyset$, a contradiction.)

Let

$$B = \cup_{n=1}^\infty B_n \quad \text{and} \quad C = (\cup_{n=1}^\infty C_n) \cup (\cap_{n=1}^\infty D_n).$$

Then $B \cup C$ is a partition of $A$ in $A$. We show that\(^2\)

$$B \sim C.$$

If not, then $B < C$ (the other case is handled similarly). By Q.4, $C \sim \cup_{n=1}^\infty C_n$ because $\cap_n D_n \sim \emptyset$. Thus $B < \cup_{n=1}^\infty C_n$ and by (2) of Lemma 3.4.3, there exists $N$ such that

$$B < C_n \text{ whenever } n \geq N.$$

By C5, there exists $G$ in $A$ with $\emptyset < G \subseteq C_N$ and $B \cup G < C_N \setminus G$, implying

$$B \cup G < (\cup_{n=1}^\infty C_n) \setminus G. \quad (3.2)$$

Because $B \cap G = \emptyset$ and $B_n \subseteq B$, we have

$$B_n \cup G \subseteq B \cup G. \quad (3.3)$$

\(^2\)Fishburn’s proof cannot be adopted unaltered, because it exploits the $\sigma$-algebra structure for $A$. 

3.4. **THE PROOF OF THEOREM 3.1**

103
By Q.4 and (b), $D_n < G$ for large $n$, so that, again by Q.4,

$$B_n \cup D_n < B_n \cup G.$$  

(3.4)

**Claim:** $(\cup_{i=1}^{\infty} C_i) \setminus D_n = C_n$. Argue as follows: Because for each $n$, $\{C_n, B_n, D_n\}$ is a partition of $A$ in $\mathcal{A}$ and $C_n \nsubseteq B_n \setminus D_n$, then

$$C_{n+i} \setminus D_n = [(C_{n+i} \setminus C_n) \cup C_n] \setminus D_n = C_n, \quad i = 1, 2, ...$$

and $C_{n+i} \setminus C_n \subseteq D_n$ for each $i$. Thus

$$(\cup_{i=1}^{\infty} C_i) \setminus D_n = \cup_{i=1}^{\infty} (C_i \setminus D_n) = \lim_{m \to \infty} \cup_{i=1}^{m} (C_i \setminus D_n) = C_n.$$ 

It follows that $(\cup_{i=1}^{\infty} C_i) \setminus D_n \in \mathcal{A}$ and therefore that

$$(\cup_{i=1}^{\infty} C_i) \cap D_n = (\cup_{i=1}^{\infty} C_i) \setminus ((\cup_{i=1}^{\infty} C_i) \setminus D_n) \in \mathcal{A}.$$ 

We also have $(\cup_{i=1}^{\infty} C_i) \cap D_n \subseteq D_n < G$ for large $n$ and

$$\cup_{i=1}^{\infty} C_i = (\cup_{i=1}^{\infty} C_i \setminus G) \cup G$$

$$= (\cup_{i=1}^{\infty} C_i \setminus D_n) \cup ((\cup_{i=1}^{\infty} C_i) \cap D_n),$$

$$= C_n \cup ((\cup_{i=1}^{\infty} C_i) \cap D_n).$$

Because $C_n \cup G = C_n \in \mathcal{A}$, it follows by C4($\leq$) that for $n \geq N$

$$\cup_{i=1}^{\infty} C_i \setminus G \subseteq (\cup_{i=1}^{\infty} C_i \setminus D_n).$$  

(3.5)

Conclude from the Claim and (3.2)-(3.5) that

$$B_n \cup D_n < B_n \cup G \leq B \cup G < (\cup_{n=1}^{\infty} C_n) \setminus G \leq (\cup_{i=1}^{\infty} C_i \setminus D_n) < C_n.$$ 

This contradicts the construction $C_n < B_n \cup D_n$ from (a).

**C9:** Let $\emptyset < A \in \mathcal{A}$. By C8, there is a partition $\{A_1, A_2\}$ of $A$ in $\mathcal{A}$ such that $A_1 \sim A_2 \succ \emptyset$. Again by C8, there exist partitions $\{A_{11}, A_{12}\}$ of $A_1$ in $\mathcal{A}$ and $\{A_{21}, A_{22}\}$ of $A_2$ in $\mathcal{A}$ such that $A_{11} \sim A_{12} \succ \emptyset$ and $A_{21} \sim A_{22} \succ \emptyset$. Next prove that $A_{11} \sim A_{12} \sim A_{21} \sim A_{22}$. If not, then, for example, $A_{11} \succ A_{21}$, implying by
3.4. **THE PROOF OF THEOREM 3.1**

C4 that $A_{12} \succ A_{22}$. Therefore, again by C4, $A_1 = A_{11} \cup A_{12} \succ A_{21} \cup A_{12} = A_2$, a contradiction. Proceed by induction. ■

**Proof of Sufficiency in Theorem 3.3.1:** Let (a)-(b) hold. Call a partition \( \{A_1, A_2, \ldots, A_m\} \) of \( A \) in \( \mathcal{A} \) a u.p. (uniform partition) when \( \emptyset < A \) and \( A_1 \sim A_i \) for all \( i \). Let

\[
C(r, 2^n) = \{ A \in \mathcal{A} : A \text{ is the union of } r \text{ events in some } 2^n \text{ part u.p. of } S \}.
\]

Prove sufficiency through a series of steps following Fishburn. The proofs of Steps 3-5 are as in [6, pp. 198-9].

**Step 1:** If \( A, B \in C(r, 2^n) \), then \( A \sim B \): Implied by (c).

**Step 2:** If \( A \in C(r, 2^n) \) and \( B \in C(r2^m, 2^{n+m}) \), then \( A \sim B \): Implied by (c).

**Step 3:** If \( A \in C(r, 2^n) \), \( B \in C(t, 2^m) \), then \( A \preceq B \iff r/2^n \leq t/2^m \).

**Step 4:** For \( A \in \mathcal{A} \), let \( k(A, 2^n) \) be the largest integer \( r \) (possibly zero) such that \( B \preceq A \) when \( B \in C(r, 2^n) \), and define

\[
p(A) = \sup \{ \frac{k(A, 2^n)}{2^n} : n = 0, 1, 2, \ldots \}.
\]

Then \( p(\emptyset) = 0 \), \( p(S) = 1 \), \( p(A) \geq 0 \) for all \( A \in \mathcal{A} \) and

\[
A \in C(r, 2^n) \implies p(A) = r/2^n.
\]

**Step 5:** \( A \preceq B \implies p(A) \leq p(B) \).

**Step 6:** \( p \) is finitely additive: Let \( A, B \in \mathcal{A} \) and \( A \cap B = \emptyset \). For each \( n \), there is a \( 2^n \) part u.p. of \( S \) for which there exist \( A_n \) and \( B_n \), unions of elements of this partition, such that \( A_n \cap B_n = \emptyset \), \( A_n \in C(k(A, 2^n), 2^n) \), \( B_n \in C(k(B, 2^n), 2^n) \), \( A_n \preceq A \) and \( B_n \preceq B \). The key is to show that

\[
A_n \cup B_n \preceq A \cup B.
\]

Given this, the proof may be completed as in Fishburn. To prove (3.8), one might try to invoke C4, but it is inapplicable because it is not certain that \( A \cup B_n \in \mathcal{A} \). This added hypothesis, absent in the standard case where \( \mathcal{A} \) is a \( \sigma \)-algebra, requires a slight variation of the argument in Fishburn.
Prove that there exists another $2^n$ uniform partition $\{C_i\}_{i=1}^{2^n}$ such that

$$\bigcup_{i \in I} C_i \sim A_n, \quad \bigcup_{i \in J} C_i \sim B_n \quad \text{and} \quad \bigcup_{i \in J} C_i \cup A_n \in A,$$

where $I$ and $J$ are arbitrary disjoint subsets of $\{1, 2, \ldots, 2^n\}$ with $|I| = k(A, 2^n)$ and $|J| = k(B, 2^n)$.

Assume that $A \succ \emptyset$, $B \succ \emptyset$ and $(A \cup B)^c \succ \emptyset$. (The other cases may be handled similarly.) By C9, there are three uniform partitions $\{C_i^1\}_{i=1}^{2^n}, \{C_i^2\}_{i=1}^{2^n}$ and $\{C_i^3\}_{i=1}^{2^n}$ of $A$, $B$, $(A \cup B)^c$ in $A$ respectively. Let

$$C_i = C_i^1 \cup C_i^2 \cup C_i^3, \quad i = 1, 2, \ldots, 2^n.$$ 

Then $\{C_i\}_{i=1}^{2^n}$ is a uniform partition of $S$ in $A$ and $\bigcup_{i \in I} C_i \cup A \in A$. By (c), $\bigcup_{i \in I} C_i \sim A_n$ and $\bigcup_{i \in J} C_i \sim B_n$. (Recall that $A_n$ is a disjoint union of $|I|$ elements of a uniform partition of $S$; and similarly for $B_n$.) Now C4 may be applied to deduce (3.8).

Step 7: $p$ is countably additive: It follows from condition (b) and the convex range of $p$ on $A$.

Step 8: $p$ is convex-ranged: The proof is identical to [6, p. 199].
Bibliography


