MBS Technical Report 15-02

Modeling Decisions Involving Ambiguous, Vague, or Rare Events*

Louis Narens
Department of Cognitive Sciences
University of California, Irvine

Donald Saari
Department of Mathematics
Department of Economics
University of California, Irvine

Abstract

Almost all models of decision making assume an underlying boolean space of events. This gives a logical structure to events that matches the structure of propositions of classical logic. This chapter takes a different approach, employing events that form a topology instead of a boolean algebra. This allows for new modeling concepts for judgmental heuristics, rare events, and the influence of context on decisions.

The Kolmogorov approach to probability theory, which defines probability as a normed σ-additive measure on a boolean algebra of events, has proved to be a fruitful foundation to understand issues from much of science. But there are exceptions where, for various reasons, a more flexible theory is needed. The purpose of doing so usually arises where there is a need to employ a more general form for the probability function, or to use a more general algebra of events. Both settings, for instance, occur in quantum mechanics.

This chapter describes a generalization for a normed finitely additive measure on a topology. The objectives of this extension are to present a new model of decision making that can incorporate well-documented features of human judgments of probability and to assess its "subjective rationality." Finally, the model's mathematical relationship to Chichilniski's (2009) approach for catastrophic decision making is described.

*The research for this chapter was supported by grants FA9550-08-1-0389 and FA9550-13-1-0012 from AFOSR. The chapter is based on a presentation given at the AFOSR Workshop on Catastrophic Risks, SRI Stanford, 2012.

1 Topological Event Spaces

To start by reviewing some of the basic terms, a boolean algebra of events has the form,

$$\langle \mathcal{B}, \cup, \cap, -, X, \varnothing \rangle ,$$

where $\mathcal{B}$ is a collection of subsets of the nonempty set X closed under the set-theoretic operations of $\cup$, $\cap$, and $-$ and where X and $\varnothing$ are in $\mathcal{B}$. A topology has a similar form except that it is not required to be closed under the operation of set-theoretic complementation, $-$. Thus a topology has $\emptyset$, X, finite intersections, and arbitrary unions of subsets from $\mathcal{B}$ in $\mathcal{B}$. It is useful for applications to replace $-$ with a different complementation operator called "pseudo complementation".

To be specific, let $\mathcal{T}$ be a topology. By definition, for each A in $\mathcal{T}$ the *pseudo complement of A,* denoted as $\neg A$, is the largest element B of $\mathcal{T}$ such that $A \cap B = \varnothing$. By elementary properties of "topology", $\neg A$ always exists. With this operation, a *topological algebra of events* is defined to have the form

$$\langle \mathcal{T}, \cup, \cap, \neg, X, \varnothing \rangle ,$$

where X is the universe of $\mathcal{T}$. It is not difficult to show that, with this definition, a topological algebra of events where each open set is also a closed set is a σ-boolean algebra of events.

Boolean algebras of events correspond to the classical propositional calculus in logic, where "c implies d" in a classical presentation corresponds to an expression of the form $(-C) \cup D$ in a boolean algebra of events. A topological algebra of events corresponds to a different, well-known logic called the intuitionistic propositional calculus. Similar to the boolean algebra of events and classical logic, the operation $\cup$ corresponds to disjunction "or", and the operation $\cap$ corresponds to conjunction "and" in intuitionistic logic. But rather than $-$ corresponding to the "not" operation, for intuitionistic logic, $\neg$ corresponds to negation. Unlike boolean algebra of events, the operator corresponding to intuitionistic implication cannot be defined by a simple formula involving $\cup, \cap$, and $\neg$, although it has a purely topological definition. (For details involving topological algebras of events and their relationship to intuitionistic logic, see Narens, 2014b.)

It is this difference in complementation operators that permits the logical structure of a topological algebra of events to differ from that of a boolean algebra of events. The following nine statements identify basic properties of $\neg$ for a topological event space. While the first eight remain valid by substituting $-$ for $\neg$, we call attention to Statement 9 because it becomes invalid under such a substitution. (The proofs for these statements can be found in Chapter 9 of Narens, 2007.)

If $\langle \mathcal{T}, \subseteq, \cup, \cap, \neg, X, \varnothing \rangle$ is a topological algebra of events, then the following eight statements hold for all A and B in $\mathcal{X}$:

1. $\neg X = \varnothing$ and $\neg \varnothing = X$.

2. If $A \cap B = \varnothing$, then $B \subseteq \ulcorner A$.
3. $A \cap \ulcorner A = \varnothing$.
4. If $B \subseteq A$, then $\ulcorner A \subseteq \ulcorner B$.
5. $A \subseteq \ulcorner\ulcorner A$.
6. $\ulcorner A = \ulcorner\ulcorner\ulcorner A$.
7. $\ulcorner (A \cup B) = \ulcorner A \cap \ulcorner B$.
8. $\ulcorner A \cup \ulcorner B \subseteq \ulcorner (A \cap B)$.
9. There exists a topological algebra of events $\langle \mathcal{Y}, \subseteq, \cup, \cap, \ulcorner, Y, \varnothing \rangle$ such that the following three statements hold:
   - For some A in $\mathcal{Y}$, $A \cup \ulcorner A \neq Y$.
   - For some A in $\mathcal{Y}$, $\ulcorner\ulcorner A \neq A$.
   - For some A and B in $\mathcal{Y}$, $\ulcorner (A \cap B) \neq \ulcorner A \cup \ulcorner B$.

A rich and useful concept is the definition of a "probability function," which serves as a normed measure. In part, this concept is possible because the algebra inherent in a boolean algebra of events guarantees a sufficiently abundant subset of disjoint events. In contrast, the topological algebra of events need not enjoy this property of having a sufficiently generous subset of disjoint events. Closing this gap requires altering the concept of "probability function," which is needed to provide a decent theory of probability.

The way to do so is to change the finite additivity clause in the definition of normed, finite measure to an expression that is logically equivalent for a boolean algebra of events:

For all A and B in the topology, $\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cap B)$.

With this change of definition of "normed, finitely additive measure", $\mathbb{P}$ applied to a topological algebra of events $\mathcal{T}$ is called a *probability function* on $\mathcal{T}$.

2 Boundaries of Topological Events

Many applications of Kolmogorov probability theory begin with a topological event space from which a special boolean algebra of events is selected. Along with this algebra, a measure is chosen that assigns the probability of 0 to the boundary of an event. Thus the measure allows the boundaries of events to be ignored.

In applications of topological event theory, however, we may not wish to ignore these boundaries. Instead, the boundaries, although "small," could carry substantive interpretations that cannot be ignored. In other words, when assigning probabilities, it may be important to assign positive values to some boundaries, including their parts and even isolated boundary points. As shown below, doing so allows for richer concepts to be developed in purely event terms that are not feasible when using a boolean algebra of events. In this chapter this is done for the specialized concepts of ambiguity and vagueness.

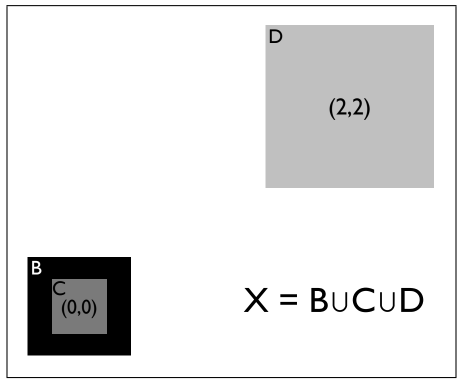

Figure 1: $\mathcal{U} = \{X, B, C, D, \varnothing\}$**, is a six element topological algebra of events with universal set** X **and an open set** C **with the "thick" boundary** $B - C$**.**

These notions are illustrated with Figure 1, which describes a topology $\mathcal{T}$ consisting of 6 open sets; each set also is an open set in the Cartesian plane with the Euclidean topology:

- The (open) sets B, C, and D have the Euclidean areas of, respectively $\frac{1}{4}$, $\frac{1}{8}$, and $\frac{3}{4}$. Geometrically, sets B and C are centered at the point (0,0), while D is centered at the point (2,2).

- The other open sets are $X = B \cup D = B \cup C \cup D$ (which is the universe of $\mathcal{T}$), $C \cup D$, and $\varnothing$.

Note that $\neg B = \neg C = D$ and $B = \neg\neg B = \neg\neg C$. This last expression illustrates that it is possible to have $\neg\neg C \neq C$, which is condition 9 in the above list. Also note that the $\mathcal{T}$-topological boundary of C (i.e., the set of points a of X such that each element of $\mathcal{T}$ containing a intersects C and $X - C$) is $B - C$. Further note that although $X \neq C \cup D$, it is true that $\neg\neg(C \cup D) = X$.

For each event E in $\mathcal{T}$, if $\mathbb{P}(E)$ equals the area of E, then $\mathbb{P}$ is a probability function on the topological algebra of events $\mathcal{T}$. Because $\mathbb{P}(X) = 1$ and $\neg C = D$, it follows that

$$\mathbb{P}[C \cup (\neg C)] = \frac{1}{8} + \frac{3}{4} < 1 = \mathbb{P}(X)\,. \qquad (1)$$

Equation 1 is a probabilistic form of a well-known principle of intuitionistic logic that violates the *law of the excluded middle* coming from classical logic. That is, this example violates the condition $C \cup (\neg C) = X$.

It is clear what causes the inequality in Equation 1; the boundary $B - C$ of $C \cup (\neg C)$ is ignored. This missing term has value for certain applications such as for human judgments of probability. For example, in Chapter 10 of Narens (2007), it is interpreted as potentially clear instances of C and cognitive non-instances of $\neg C$. Such instances are ignored in the cognitive calculation of the participant's subjective probability of $C \cup (\neg C)$.

Figure 1 illustrates a "thick" boundary, which is but one choice. Also the "measure method" for constructing the probability function $\mathbb{P}$ is only one way to construct probability functions for topologies. Of relevance for what follows is that a geometrically "thick" boundary is not needed in order to have a boundary, or part of a boundary, to behave as though it has a positive probability. Even individual points that are open sets can be assigned positive probabilities. In other words, a probability function on a topological algebra of events need not be produced in a usual mathematical way to derive a measure from a topology.

3 Application to Judgments of Probability

Many psychological experiments involving human judgments of uncertainty have the participants judge conditional probabilities that are of the forms $\mathbf{A}|\mathbf{Y}$ and $\mathbf{B}|\mathbf{Y}$, where $\mathbf{A}$, $\mathbf{B}$, and $\mathbf{Y}$ are descriptions, respectively, of the events A, B, and Y, with A and B being disjoint and $A \subseteq Y$ and $B \subseteq Y$. *Descriptions* are used here, because many experimental paradigms involve situations where different descriptions of the same event can lead to different results.

To provide an example, suppose the above event A is partitioned into the four events C, D, E, F with respective descriptions $\mathbf{C}$, $\mathbf{D}$, $\mathbf{E}$, $\mathbf{F}$. Suppose participants make judgments that are *sufficiently separated by time and design* so that the judgments do not influence one or another of

$$\mathbb{P}(\mathbf{A}|\mathbf{Y}),\ \mathbb{P}(\mathbf{B}|\mathbf{Y}),\ \mathbb{P}(\mathbf{C}|\mathbf{Y}),\ \mathbb{P}(\mathbf{D}|\mathbf{Y}),\ \mathbb{P}(\mathbf{E}|\mathbf{Y}),\ \mathbb{P}(\mathbf{F}|\mathbf{Y})\,.$$

The problem is that, rather than equality, many experimental studies show that

$$\mathbb{P}(\mathbf{C}|\mathbf{Y}) + \mathbb{P}(\mathbf{D}|\mathbf{Y}) + \mathbb{P}(\mathbf{E}|\mathbf{Y}) + \mathbb{P}(\mathbf{F}|\mathbf{Y}) > \mathbb{P}(\mathbf{A}|\mathbf{Y})\,, \tag{2}$$

with

$$\mathbb{P}(\mathbf{B}|\mathbf{Y}) + \mathbb{P}(\mathbf{C}|\mathbf{Y}) + \mathbb{P}(\mathbf{D}|\mathbf{Y}) + \mathbb{P}(\mathbf{E}|\mathbf{Y}) + \mathbb{P}(\mathbf{F}|\mathbf{Y}) \text{ being substantially} > 1\,. \tag{3}$$

As an example of Equations 2 and 3 consider the following experiment of Fox & Birke (2002):

(Jones Versus Clinton) 200 practicing attorneys were recruited (median reported experience: 17 years) at a national meeting of the American Bar Association (in November 1997). Of this group, 98% reported that they knew at least "a little" about the sexual harassment allegation made by Paula Jones against President Clinton. At the time of that the survey, the case could have been disposed of by either A, which was an outcome other than a judicial verdict, or B, which was a judicial verdict. Furthermore, outcomes other than a judicial verdict can partition A into

(A1) settlement;

(A2) dismissal as a result of judicial action;

(A3) legislative grant of immunity to Clinton; and

(A4) withdrawal of the claim by Jones.

Each attorney was randomly assigned to judge the probability of one of these six events. The results are given in Table 1

(A) other than a Judicial verdict	.75
(B) judicial verdict	.20
Binary partition total	.95
(B) judicial verdict	.20
(A1) settlement	.85
(A2) dismissal	.25
(A3) immunity	.0
(A4) withdrawal	.19
Five fold partition total	1.49

Table 1: Median Judged Probabilities for All Events in Study

The Jones versus Clinton example illustrates the core idea of the much investigated empirically based theory of probability judgments called *Support Theory,* which is due to Tversky & Koehler (1994) and modified by Rottenstrich & Tversky (1997). Chapter 10 of Narens (2007) employs algebras of topological events to provide a foundation for Support Theory and to model its empirical results.

This foundation is based on topological algebra of events that includes considerations about boundary points. The basic premise is that in making a judgment of probability, participants use cognitive heuristics like those proposed in various seminal articles of Kahneman and Tversky (e.g., Tversky & Kahneman, 1974), and that these heuristics can be modeled through topological considerations.

For probability judgments the *availability heuristic* is particularly important. In this heuristic, the participant judges the probability of an event E in terms of the evidence for the occurrence on E and evidence for the non-occurrence of E. Namely, the judgments are based on the number and ease that instances of E are brought to mind by the event's description **E** as compared to the number and ease that non-instances of E are brought to mind by **not E**. In addition to availability, Chapter 10 of Narens (2007) models the "representativeness heuristic" by reducing it to the availability of properties of instances of an event. While it is beyond the scope of this current chapter to present a thorough discussion of Narens' foundation for Support Theory this brief description is intended to indicate the important kinds of cognitive instances of an event and their role in judgments based on the availability heuristic.

In general, the most important kind of cognitive instance of E is a *clear instance* base on **E**. These "instances" are the ones that come to mind; they are the ones that a participant views as definitely belonging to E when provided with the description **E**. The set of clear instances (to be denoted by CI) based on **E** is modeled as an open set $\mathsf{CI}_{\mathbf{E}}(\mathrm{E})$ in a topology $\mathcal{T}$. The simpler notation "$\mathsf{CI}(\mathrm{E})$" is employed when it is obvious that the set of clear instances are based on **E**. A similar convention holds for the notation "$\mathsf{CC}(\mathrm{E})$" that is presented next.

The *cognitive complement of the set of instances of **E**,* $\mathsf{CC}(E)$, consists of *all instances that come to mind that are viewed by the participant as clearly not being clear instances of **E**.* As with $\mathsf{CI}(E)$, it is assumed that $\mathsf{CC}(E)$ is an element of $\mathcal{T}$. It is also assumed that the structure of $\mathcal{T}$ is such that the pseudo complement of $\mathsf{CI}(E)$ with respect to $\mathcal{T}$, $\neg\ \mathsf{CI}(E)$, is the set of all instances i (in the domain X under consideration) such that *if i were presented to the participant as an instance described by **E**, then she would consider it to be clearly not an instance of **E**.* (Note the counterfactual nature of the definition of $\neg$ $\mathsf{CI}(E)$.) It is assumed that

$$\mathsf{CC}(E) \subseteq \neg\mathsf{CI}(\mathsf{E})\,.$$

This inclusion is a natural consequence of the meaning of "clear instance". Of interest, which is explored next, is that this expression need not be an equality.

Before considering other kinds of "instances", it is useful to describe what $\neg\neg\ \mathsf{CI}(E)$ corresponds to. It is assumed that $\mathcal{T}$ is such that $\neg\neg\ \mathsf{CI}(E)$ is the set of all instances i (in the domain X under consideration) such that *if i were presented to the participant as an instance described by **not E**, then she would consider it to be clearly not a clear instance of **not E**.* In particular, each clear instance of **not E** is not in $\mathsf{CI}(E)$, and thus

$$\mathsf{CI}(E) \subseteq \neg\neg\ \mathsf{CI}(\mathsf{E}) \text{ and } \mathsf{CC}(E) \subseteq \neg\ \mathsf{CI}(E)\,.$$

Because these expressions are of containment, but not necessarily of equality, the elements of $[\neg\neg\mathsf{CI}(E)] - \mathsf{CI}(E)]$ are of particular interest. These elements, which are called *potential clear instances of **E***, are possible clear instances of **E** that do not come to mind when judging the probability of E when presented the description **E**; mathematically, this statement means that although these elements are related to $\mathsf{CI}(\mathrm{E})$, they are not in this set. Theoretically, they are responsible for empirical observations of Equation 2 when the availability heuristic plays a primary role in probability estimations: A more specific description **F** of a subevent F of an event E is likely to bring to mind more clear instances of **F** than the subset of clear instances of **F** brought to mind when doing a probability estimation of E with a description **E** of E.

Indeed, it is the mathematical boundary structure, where even if a boundary point for a set is not in the set it still shares aspects of the set's structure, that provides an appropriate framework to describe two additional and important kinds of "cognitive instance" — ambiguity and vagueness. Element i is said to be a *weakly ambiguous instance of* ***E*** if and only if when making a probability judgment of E with **E**, i is an element of the boundary of $\mathsf{CI}(E)$ and $\mathsf{CC}(E)$. Notice, i is not in either $\mathsf{CI}(E)$ or $\mathsf{CC}(E)$.

Similarly, i is said to be an *vague instance of* ***E*** if and only if when making a probability judgment of E with **E**, i is an element of the boundary of $\mathsf{CI}(E)$ but it is *not* an element of the boundary of $\mathsf{CC}(E)$.

A weakly ambiguous instance comes to mind in the judging of both **E** and **not E** and *the participant is aware of this.* Because of this awareness, it is neither a clear instance of **E** nor a clear instance of **not E**. A distinction is made between weakly ambiguous instances and another kind of ambiguous instance called "strongly ambiguous". Consider a situation where a participant judges $\mathbb{P}(\mathbf{E}\,|\,\mathbf{E}\ \mathbf{or}\ \mathbf{F})$ and later judges $\mathbb{P}(\mathbf{F}\,|\,\mathbf{E}\ \mathbf{or}\ \mathbf{F})$, where the conjunction **E and F** describe the empty event. Then i is said to be *strongly ambiguous instance* of these judgments if and only if it is a clear instance of **E** when $(\mathbf{E}\,|\,\mathbf{E}\ \mathbf{or}\ \mathbf{F})$ is judged and it is a clear instance of **F** when $(\mathbf{F}\,|\,\mathbf{E}\ \mathbf{or}\ \mathbf{F})$ is judged.

Our reading of Tversky & Koehler (1994) suggests that Support Theory has participants ignoring weakly ambiguous and vague instances in their calculations of probability estimates. However, their calculations take into account strongly ambiguous instances, causing the sum,

$$\mathbb{P}(\mathbf{E}\,|\,\mathbf{E}\ \mathbf{or}\ \mathbf{F}) + \mathbb{P}(\mathbf{F}\,|\,\mathbf{E}\ \mathbf{or}\ \mathbf{F}) \,,$$

to be an increase over what one would expect from standard probability theory, because of the strongly ambiguous instances that happen for both **E** and **F**.

In Support Theory, a participant's estimation $\mathbb{P}(A|A \cup B)$ of the conditional probability $A|A \cup B$, where $A \cap B = \varnothing$, is computed by the formula,

$$\mathbb{P}(A|A \cup B) = \frac{S(A)}{S(A) + S(B)} \,, \qquad (4)$$

where S is a function with nonnegative real values. Tversky & Koehler (1994) calls S a *support function.* Equation 4 have been used by Luce, (1959) and others to model choice situations where $\mathbb{P}$ is an observed probability function instead of a subjective estimation. Below, it is generalized slightly to model situations where a subject's probability estimations violate finite additivity.

The availability heuristic assumes that $S(A)$ and $S(B)$ are determined, respectively, by ease and number of instances of A and B come to mind when presented with appropriate instructions to the participant using descriptions **A** and **B**. Note that such instructions are asymmetric with respect to **A** and **B**: During this phase of the experiment, the participant is instructed to estimate the conditional probability of A given $A \cup B$, while no instruction (during this phase of the experiment) is given to estimate the conditional probability of B given $A \cup B$. The form of these instructions allows for asymmetric approaches for calculating $S(A)$ and $S(B)$.

In terms of the foundational concepts presented here, this asymmetry becomes more apparent. The reason is that the above foundation replaces Equation 4 with

$$\mathbb{P}(A|A \cup B) = \frac{S[\mathsf{CI}(A)]}{S[\mathsf{CI}(A)] + S[\mathsf{CC}(A)]}. \tag{5}$$

An important difference is that Equation 5 includes the possibility that the structure of $\mathcal{T}$ is such that $\mathsf{CC}(A) \neq \mathsf{CI}(B)$. This provides the possibility for structurally asymmetric cognitive processing of A and B in the estimation of $\mathbb{P}(A|A \cup B)$, e.g., the clear instances A can be processed without consideration of the clear instances of B, but the processing of $\mathsf{CC}(A)$ requires also processing a relationship between A and B describing which clear instances come to mind that are instances B but not instances of A. It is assumed that subjects employ the processing described by Equation 5. For some situations, this results in different predictions than the formula

$$\mathbb{P}(A|A \cup B) = \frac{S[\mathsf{CI}(A)]}{S[\mathsf{CI}(A)] + S[\mathsf{CI}(B)]}.$$

Topological modeling of events is a promising alternative to boolean modeling for describing subjective probability estimations. The reason is that its internal "logic" matches better with the forms of cognitive processing entering into the estimations. This assertion becomes apparent when memory is involved. In particular, one of the more robust empirical findings in memory research is that, for the vast majority of times, recognition is easier than recall.

In fact, one of the simplest models of recall memory, which is called the *generation-recognition theory of recall,* relies on this fact. The model assumes that in response to a recall task of the "Name the wild African animals" type, the participant generates a set of animal names (the generation phase) and selects those that she believes are names of wild African animals (the recognition phase).

In contrast, the recognition task presents a list of animal names and asks to participant to select those that name wild African animals. This approach eliminates the need to generate possible names, which makes recognition generally an easier and more accurate task (in terms of percent correct) than recall.

Narens (2009) shows that the logical relationship of recognition and recall can be nicely modeled in a topological algebra of events by the operation of pseudo complementation, $\neg$: To see how this is done, in the topological algebra $\mathcal{T}$, let E be a set of items that is recalled from a category **E** or a set of items that is recognizable as belonging to a category **E**. For example, the universe of $\mathcal{T}$ can be the set of animals, **E** a description of the category of African animals, and E the items recalled with description **E**. By definition, $\neg\, E$ is the set of items of **not E** that is recognized. As **E** is a description of the category of African animals, it follows that $\neg\, E$ is the set of animals that is recognized as being non-African. In turn, $\neg\neg\, E$ becomes the set of items of items of **not not E** that is recognized, which coincides the set of items of **E** that is recognized.

By properties of pseudo complementation,

$$E \subseteq (\neg\neg\, E)\,.$$

When E is a set of recalled items of $\mathbf{E}$, it is a subset of recognized items of $\mathbf{E}$.

As these examples demonstrate, because various concepts derivable in topological algebras of events have structural correspondences with notions coming from cognitive psychology, a topological algebra of events can be an attractive alternative to a boolean algebras of events. Although Kolmogorov probability theory can be avoided for measuring uncertainty on a boolean algebra of events by using systems of weights on events instead of probabilities, such weightings do not have sufficient logical structure to provide a foundation for a subjective probability theory with a rich mathematical calculus for manipulating and calculating measurements of uncertainty. It is precisely having such a calculus that makes Kolmogorov probability theory so useful in applications. Narens (2007, 2014b) show that topological algebras of events have rich probabilistic calculi.

4 Rationality

It is claimed by many that rational decision making under uncertainty requires that a particular model of decision making, the *Subjective Expected Utility Model*—or *SEU* for short—must hold. This model assumes that the decision maker has a a utility function u over outcomes and lotteries and a fintely additive Kolmogorov probability function P over events such that for all lotteries

$$L = (a_1, A_1; \ldots; a_i, A_i; \ldots; a_n, A_n)\,,$$

where a_i is a pure outcome, A_i is an event, and "a_i, A_i" stands for receiving a_i if A_i occurs, and

$$u(L) = \sum_{i=1}^{n} \frac{\mathsf{P}(A_i)}{\mathsf{P}(A_1) + \cdots + \mathsf{P}(A_n)} \cdot u(a_i)\,. \tag{6}$$

Equation 6 is called the *SEU Model* for L.

Note that in Equation 6,

$$\frac{\mathsf{P}(A_i)}{\mathsf{P}(A_1) + \cdots + \mathsf{P}(A_n)}$$

is the subjective conditional probability of A_i occurring given that $\bigcup_{i=1}^{n} A_i$ has occurred.

The basis for claiming the rationality of SEU rests on axiomatizations for which the individual axioms are argued to be rational, for example, the famous axiomatization of Savage (1954), or the axiomatization of a conditional form of SEU by Luce & Krantz (1971).

Humans tend to violate SEU in systematic ways. While economics generally consider these examples to be violations of rationality, some have argued that for human decision making, SEU is an inappropriate model of rationality. Instead, it is proposed that rationality should be evaluated in terms of a form of optimality that takes into account various constraints the decision maker encounters while making decisions. These include cognitive constraints like limitations of memory and the ability to make complicated mathematical calculations as well as inherent biological constraints such as the effects of emotion generated by the decision task on the final decision. Forms of rationality that take into account constraints like these are called *bounded rationality* (Simon, 1957).

This section focuses on situations where the decision maker experiences different states while making decisions about lotteries, and it develops a notion of "rationality" for these situations. This form of rationality, which is called *cognitive rationality,* is illustrated in Figure 3. It is distinguished from the rationality inherent in the SEU model called *objective rationality* and illustrated in Figure 2.

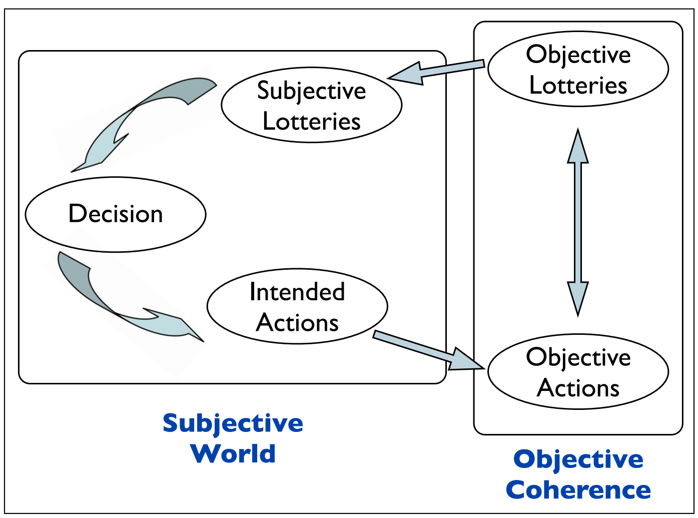

Figure 2: Objective Rationality

Both Figures 2 and 3 are concerned with a situation where lotteries from a set of lotteries $\mathcal{L}$ are presented to a participant. The elements of $\mathcal{L}$ are called the *objective lotteries*; they can be considered to be part of the everyday world.

For purposes of evaluating utilities, the participant needs to interpret them subjectively. From a mathematical perspective, objective rationality assumes there is an isomorphism between each objective lottery and a particular subjective representation that is used for calculating utilities. Namely, each item of an event in an objective lottery has a corresponding isomorphic item in the subjective representation.

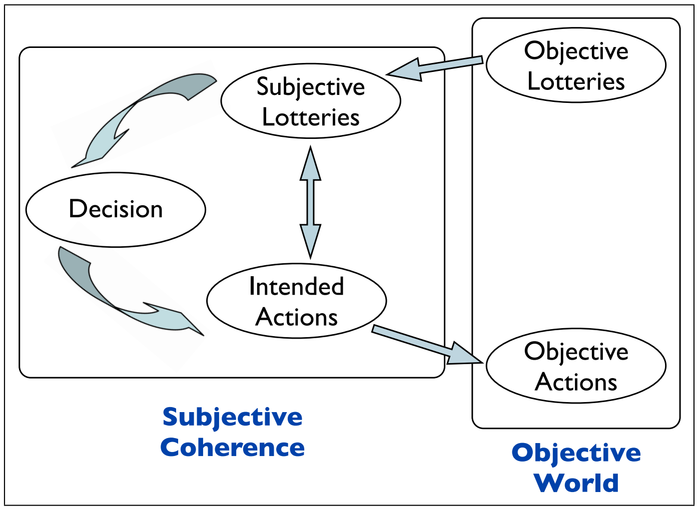

Figure 3: Subjective Rationality

While subjective rationality also assumes the existence of subjective representations of objective lotteries, the representations are not required to be isomorphisms of objective lotteries. They must, however, have the logical form of lotteries.

Both objective rationality and subjective rationality assume that each subjective lottery is an input to a decision process. The outcome of the decision process has two steps. The first yields intended actions, which then yield objective actions that take place in the everyday world. For this discussion, the intended actions can be assumed to produce preference orderings on the subjective lotteries, $\precsim_{\text{obj}}$ for objective rationality, and $\precsim_{\text{sub}}$ for subjective rationality. The intended actions are carried out in the everyday world producing preference orderings on objective lotteries, $\precsim'_{\text{obj}}$ for objective rationality and $\precsim'_{\text{sub}}$ for subjective rationality. Objective rationality assumes that its subjective lotteries with the ordering $\precsim_{\text{obj}}$ is isomorphic to objective Lotteries with the ordering $\precsim'_{\text{obj}}$. Subjective rationality does not make this assumption.

Notice how the principal difference between objective and subjective rationality is the kind of coherence that relates lotteries with preference orderings. Objective rationality assumes that $\precsim'_{\text{obj}}$ is *objectively coherent* in that it demonstrates the following consistency with SEU: There is a utility function u_{obj} on the set of outcomes occurring in objective lotteries and a probability function P_{obj} on the set of events occurring in objective lotteries such that

(i) each objective lottery satisfies the SEU Model (Equation 6) with u_{obj} and P_{obj}, and

(ii) for all objective lotteries K and L,

$$K \precsim'_{\text{obj}} L \text{ iff } u_{\text{obj}}(K) \leq u_{\text{obj}}(L)\,.$$

Subjective reality assumes a similar kind of consistency for $\precsim_{\text{sub}}$. Specifically, $\precsim_{\text{sub}}$ is *subjectively consisent* if and only if there is a utility function u_{sub} on the set of outcomes occurring in objective lotteries and a probability function P_{sub} on the set of events occurring in objective lotteries such that

(i) each objective lottery satisfies the SEU Model (Equation 6) with u_{sub} and P_{sub} , and

(ii) for all objective lotteries K and L,

$$K \precsim_{\text{sub}} L \text{ iff } u_{\text{sub}}(K) \leq u_{\text{sub}}(L)\,.$$

The participant is assumed to enter into various states. Let S be the set of such states. It is important to understand how changes of state affect her subjective representations of lotteries and her subjective preference ordering. The following notation is useful for this. For each s in S and each objective lottery $L = (a_1, A_1; \ldots; a_i, A_i; \ldots; a_n, A_n)$, let

$$L^s = (a_1^s, A_1^s; \ldots; a_i^s, A_i^s; \ldots; a_n^s, A_n^s)$$

denote the participant's subjective representation of L when she is in state s. Various theories of subjective rationality can be formulated by postulating relationships among a_i, A_i, a_i^s, A_i^s, and a_i^t, A_i^t for states s and t. The following are the relationships postulated by a theory of Narens (2014a) called *descriptive subjective expected utility* or *DSEU* for short, where

$$\mathcal{L}^S = \{L^s \mid s \in S\} = \text{ the set of objective lotteries}\,.$$

Suppose s and t are arbitrary states in S, $\mathcal{T}$ is a topological algebra of events, P is a probability function on $\mathcal{T}$, and u is a real valued function on the set of outcomes of objective lotteries. Then the following hold:

- Subjective rationality holds for $\mathcal{L}^S$.
- *Invariance of lotteries:* Each subjective lottery is a lottery with pure outcomes. (This holds automatically, because it is a consequence of subjective rationality. It is stated here to emphasize that for subjective rationality, the concept of "being a lottery" remains invariant under changes of state.)
- *Invariant utilities of outcomes:* $u(a^s) = u(a)$ for each outcome a of each objective lottery.
- *Invariance of disjointness of events across states:* For each event A of each objective lottery, A is in the topology $\mathcal{T}$ and

$$A^s \subseteq \neg\neg A \text{ and } A^t \subseteq \neg\neg A\,. \tag{7}$$

Note that in the principle of invariance of disjointness of events across states, the topology $\mathcal{T}$ and the pseudo complementation operator $\neg$ depends on the subject. (Also note that it implies that if C and D are distinct events occurring in some objective lottery, then $C^s \cap D^t = \varnothing$. This is part of the reason that this assumption is called "invariance of disjointness of events across states". It provides a much stronger constraint than invariance of lotteries. Also note that it provides a strong—but in applications a workable—constraint on the subjective representations of an objective event: They are related by Equation 7.)

- *Subjective SEU with invariant probability and invariant utility across states:* For all objective lotteries $L = (a_1, A_1; \ldots; a_i, A_i; \ldots; a_n, A_n)$,

$$u(L^s) = \sum_{i=1}^{n} \frac{\mathsf{P}(A_i^s)}{\mathsf{P}(A_1^s) + \cdots + \mathsf{P}(A_n^s)} \cdot u(a_i)\,. \tag{8}$$

(Note that P and u do not depend on the state s.)

The idea behind the DSEU is to produce a model that satisfies much of the experimental literature designed to violate SEU while retaining much of the rationality expressed by SEU. Another approach in the economic literature for generalizing SEU replaces SEU's utility function with a family of utility functions, where the utility of an outcome can vary with state. This approach is reasonable for some situations, and DSEU can be easily modified to incorporate it as an additional feature. However, there are many situations where it is unreasonable to think that the driving force for the failure of SEU is due to changes in utilities of outcomes. This appears to be likely for most situations involving emotional states, where changes in subjective probabilities appear to be a more reasonable choice.

5 Connections

An interesting feature of the above discussion is how the described method permits positive weights to be attached to boundary elements. Namely, part of the strength of this approach comes from the ability to assign added weight to important events that might otherwise be ignored.

A similar concern partly motivated the work of Chichilniski (2009), where she examined decision analysis in settings that include rare but catastrophic events. As she accurately points out, a weakness of standard expected utility approaches is that the small likelihood (the measure) of an horrific event could cause it to drop out of the decision analysis.

To see how this can happen, suppose an event has the extremely large negative utility of $-M$ where M has an arbitrary large value. But if the likelihood of this event is very rare, say M^{-10}, then in expected utility considerations the event becomes the unnoticeable $-M\frac{1}{M^{10}} = -\frac{1}{M^9}$. Stated in other words, with standard expected utility considerations, rare but crucially important events (such as earthquakes, attacks such as 9/11) might not receive sufficient consideration when it is part of a standard policy/decision analysis.

Resolutions for this kind of difficulty are immediate: The goal is to find ways to attach stronger, more commensurate attention to these concerns. This can be done through concepts involving the double negation operator $\neg$. To review how this can be done, let the standard, everyday events be represented by E. In this setting, rare, possibly catastrophic events can be treated as being contained in the boundary of E: It can be shown that in many topologies

$$\text{boundary of } E \cap \text{boundary of } (\neg\neg\ E) - E$$

contain subsets of points that are natural candidates for representing rare, possibly catastrophic events. As described in the first section of this chapter, a difficulty with boundary events is that, with standard probabilistic approaches, they tend to be lost by being assigned a probability of zero. But similar to approaches described earlier, positive values can be attached to subsets contained in boundary of $E \cap$ boundary of $(\neg\neg\ E) - E$. However, unlike to the approaches described previously, such subsets in this case are not like events considered in these earlier approaches: They are not elements of the underlying topology.

Using this approach to boundaries, it becomes a direct exercise to convert the utility approach described in the last section into one that handles these kinds of subsets of boundaries. This is because the measure of a set supporting rare but important events (that is, an event C contained in the subset of boundary points $[E \cap$ boundary of $(\neg\neg\ E) - E]$) can be assigned a weight commensurate with its actual importance, while retaining the measures of the non-rare events in E. This allows for an establishment of a coherent probability function without the use of inappropriate values coming from the mathematical structure of an adopted, but perhaps inappropriate decision method.

A new kind of interpretation needs to be given to the rare event C described just above. From the perspective of the decision method used for calculating E, C has very small but non-specifiable, non-infinitesimal chance of occurrence. Its non-specifiability puts it outside of the subsets determinable by the decision method with definite probabilities, whereas it is still described by $\mathbf{E}$, and therefore is contained in $\neg\neg\ E$. A natural place for it is as a subset of the boundary of $E \cap$ the boundary of $(\neg\neg\ E) - E$. As such, C is not an open subset of E or an open subset of $\neg\neg\ E$, that is, $C \notin \mathcal{T}$. Events of $(\neg\neg\ E) - E$ are assigned probabilities by a new method. This gives rise to two probability functions, $\mathbb{P}_1$ by the old decision method for events in $\mathcal{T}$ and $\mathbb{P}_2$ by the new method for events contained in $(\neg\neg\ E) - E$ for each E in $\mathcal{E}$. Let

$$\mathcal{C} = \{\mathcal{C} \mid C \notin \mathcal{T},\ C \text{ is an event, and for some } E \text{ in } \mathcal{T},\ [C \ \subseteq\ (\neg\neg\ E) - E]\}\,.$$

A probability function $\mathbb{P}$ is then defined on the boolean algebra generated $\mathcal{T} \cup \mathcal{C}$ having the following properties:

- On $\mathcal{T}$, $\mathbb{P} = \mathbb{P}_1$.
- On $\mathcal{C}$, $\mathbb{P} = \mathbb{P}_2$.

- On $\mathcal{T} \cup \mathcal{C}$, $\mathbb{P}$ is defined as the following weighted average: There exists $0 < \alpha < 1$ such that for all E in $\mathcal{T}$ and C in $\mathcal{C}$,

$$\mathbb{P}(E \cup C) = \alpha \mathbb{P}_1(E) + (1 - \alpha)\mathbb{P}_2(C)\,.$$

Although Chichilniski adopted a different approach, it is interesting to note some of the similarities. She noted that if the utility function u is assumed to be in L^p, $p \geq 1$, (that is, the space of functions $f(x)$ where $\int |f(x)|^p\, dx$ is bounded), then the above same effect can occur causing an important rare event to be ignored. While the above example with the negative utility of $-M$ occurs on a set of measure M^{-10} will be picked up by placing the analysis in L^{11} (because now $\int |u(x)|^{11}\, dx$ includes the computation $|-M|^{11}(M^{-10}) = |M|$ where the $|M|$ value is noticed). But the same problems would be ignored in this space if the supporting measure is M^{-20} (because now the computation $|-M|^{11}(M^{-20}) = |M|^{-9}$ where the spike is ignored). In other words, a realistic issue is that, a priori, it is not known what would be the underlying measure of a serious rare event.

On the other hand, no matter how small the supporting measure, if it is positive, then this event will be picked up for functions in L^∞ (where the norm of a function can be viewed as being given by the supremum of $|f(x)|$ over sets of positive measure). With the above choice where u can have the negative utility of $-M$, no matter how small the supporting set for this value, if it has a positive measure, then the $|-M|$ value will dominate attention.

The next step is to find a way to determine the underlying measure *and* to find ways to assign positive values to small events. As a review to describe what is done, the well known Riesz representation theorem (e.g., Dunford and Schwartz (1957)) states that a linear functional $\mathbb{L}(f)$ for $f \in L^p$ can be represented as $\mathbb{L}(f) = \int f(x)\, g(x)\, dx$ for a particular $g(x) \in L^q$ where $\frac{1}{p} + \frac{1}{q} = 1$. That is, the linear functional can be identified with an element in the dual space of L^p, which is L^q; the linear functional has the representation of a scalar product (which is an integral here).

While the $\frac{1}{p} + \frac{1}{q} = 1$ dual space representation holds only for finite $p, q > 1$, it suggests with Chichilniski's setting of $p = \infty$ (so $\frac{1}{p} = 0$ or $\frac{1}{q} = 1$) that the dual space (which defines the underlying measure) should involve L^1. It does; the dual space for L^∞ is the combination of L^1 with bounded, finitely additive signed measures that are absolutely continuous with the underlying measure. (See p. 296 of Dunford-Schwarts (1957).) These finitely additive measures, which normally are difficult to use, are what provide the extra structure where positive weights can be assigned to objects of small size. In this way, the decision structure confronts and incorporates the rare but significant events into the decision analysis.

There is a certain similarity in how the two approaches elevate the importance of small but highly relevant sets; both incorporate a sense of the double dual, or double negation. In a degenerate topological space where every open set is also a closed set,

$$E = (\neg\neg\, E)\,,$$

and the topology becomes a boolean algebra of events. That is, certain degenerate topological models would require a set to be equal to its second negation. A richer setting arises by adopting a modeling environment where the double negation introduces new sets through

$$E \subseteq (\neg\neg\, E)\,.$$

As shown, the identity of these new sets vary with what is being modeled; they can range from the modeling of ambiguity or vagueness to providing a way to attach appropriate attention to rare but crucial events.

A similar mathematical effect occurs with expected utility theory with the duality operation. Here, the L^p spaces are reflexive in that the dual space for L^p is L^q and the dual space for L^q returns to L^p. That is, a fairly normal modeling environment is where a space equal to it second dual. But a richer setting arises by adopting a modeling environment where the second dual contains, but does not equal, the original space. This is the effect of assuming that the utility functions are in L^∞; the dual of this space, or the second dual of L^1, introduces the new finitely additive measures that can be used to handle rare but crucial events.

References

Chichilniski, G. (2009). The topology of fear. *Journal of Mathematical Economics, 45,* 807–816.

Dunford, N., and Schwartz, J. (1957) *Linear Operators*, New York: Interscience Publishers.

Fox, C. R. & R. Birke. (2002). Forecasting trial outcomes: Lawyers assign higher probabilities to scenarios that are described in greater detail. *Law and Human Behavior, 26,* 159–173.

Luce, R. D. (1959a). *Individual Choice Behavior.* New York: Wiley.

Luce, R. D. & Krantz, D. H. (1971). Conditional expected utility. *Econometrica, 39,* 253–271.

Narens, L. (2007). *Theories of Probability: An Examination of Logical and Qualitative Foundations.* London: World Scientific.

Narens L. (2009). A foundation for support theory based on a non-Boolean event space. *Journal of Mathematical Psychology, 53,* 399–407.

Narens L. (2011). Probabilistic lattices: theory with an application to decision theory. In E. N. Dzhafarov & L. Perry (Eds), *Descriptive and Normative Approaches to Human Behavior.* New Jersey: World Scientific, 161–202.

Narens, L. (2014a). Multimode Utility Theory. Manuscript.

Narens, L. (2014b). *Probabilistic Lattices with Applications to Decision Theory.* Manuscript

Rottenstreich, Y., & Tversky, A. (1997). Unpacking, repacking, and anchoring: Advances in Support Theory. *Psychological Review, 104,* 203–231.

Savage, L. J. (1954). *The Foundations of Statistics.* New York: Wiley.

Simon, H. (1957). *Models of Man.* New York: Wiley.

Tversky, A., & Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. *Science, 185,* 1124–1131.

Tversky, A., & Koehler, D. (1994). Support Theory: A nonextensional representation of subjective probability. *Psychological Review, 101,* 547–567.