

States and Eventualities: How to Understand Savage without Anyone Being Hanged

Itzhak Gilboa*, Stefania Minardi†, Larry Samuelson‡ and David Schmeidler§

January 7, 2018

Abstract

We discuss the notion of a state of the world in axiomatic decision theory, and argue that it should be viewed as an “eventuality” that is implicitly assumed to be independent of the process by which preferences are observed. The distinction between states, which are assumed to resolve all uncertainty, and eventualities suggests certain limitations on the axiomatic approach for defining and measuring mental concepts (such as belief) by observed choices.

Contents

1	Introduction	1
2	What’s in a state?	1
2.1	Resolving All Uncertainty	1
2.2	The Power of Resolution	2
3	Forming and Measuring Beliefs	8
3.1	Presumably-Observable Beliefs	8
3.2	A Conflict of Principles	9
4	Eventualities	13
4.1	Observability	14
4.2	Observable Preferences	14
4.3	Eventualities in Savage’s Model	16
5	Conclusion	20
6	Bibliography	20

*HEC, Paris-Saclay, and Tel-Aviv University. tzachigilboa@gmail.com

†HEC, Paris-Saclay. minardi@hec.fr

‡Yale University. larry.samuelson@yale.edu. Financial support from the National Science Foundation (SES-1459158) is gratefully acknowledged.

§Tel-Aviv University, davidschmeidler@gmail.com.

1 Introduction

Decision theory rests on two foundations: states are taken to resolve all relevant uncertainty, and beliefs can (at least in principle) be inferred from a decision maker’s willingness to bet on the realization of various states. Section 2 argues that much of the power of decision theory comes from the ability to specify states in sufficient detail as to indeed capture all relevant uncertainty. However, Section 3 argues that in the process of specifying such states, one often scuttles the ability to elicit beliefs, even in principle. Section 4 suggests a resolution, involving the notion of an eventuality.

2 What’s in a state?

2.1 Resolving All Uncertainty

The notion of a “state of the nature” or “state of the world” has roots dating back to Leibniz’s “possible worlds” and to the notion of an outcome of an experiment. States are standardly assumed to have three fundamental properties: they are mutually exclusive, exhaustive, and resolve all relevant uncertainty. Hence, whatever transpires, there is precisely one realized state, and no two occurrences would be mapped to the same state unless they are indeed equivalent for all intents and purposes. These properties allow us to think of probability as a measure over states, as suggested by Borel (1924b) and Kolmogorov (1956). The state space is the object over which subjective probability is defined and calibrated by behavior, as in the works of de Finetti (1931, 1937) and Savage (1954).¹

Our focus here is on the requirement that states resolve all uncertainty. The Bayesian approach reinforces the call for states that are sufficiently informative to describe anything that might ever be of interest. A Bayesian adopts a prior probability over the state space and updates it according to Bayes’s rule. But Bayesian updating is a rather mechanical form of learning, which allows the generation of no new states and the formulation of no new

¹Anscombe and Aumann (1963) depart from this tradition by allowing for uncertainty that is not captured by the states: their objects of choice involve lotteries given states, and the outcome of the lottery is not described in the state. Indeed, they viewed this feature as a merit of their model, because they believe that, from a cognitive perspective, subjective probabilities are more natural when compared to objective ones than in the abstract (Aumann, personal communication on two occasions, a few decades apart). However, most researchers who employ the Anscombe-Aumann framework do so for mathematical convenience, while conceptually preferring a model with only one type of uncertainty, that can be described in terms of which state obtains.

theories. Zero probability events will never be assigned positive probabilities, and there is no mechanism that could come up with new states that were not included in the original model. A Bayesian reasoner had thus better employ all her powers of imagination when formulating the state space, since it is otherwise too late.²

Bayesian statistics assumes that the “state” describes both the unknown distribution and the values of the random variables observed. Similarly, the notion of states used in economic theory includes everything that can be relevant to the model analyzed. The language of a prior probability over this rather large state space allows one to describe knowledge, beliefs, facts and intuition. Indeed, the main criticism of this approach concerns not its lack of richness when describing knowledge, but rather its weakness in reflecting ignorance.

2.2 The Power of Resolution

A notion of a state complete enough to resolve all relevant uncertainty is not simply a matter of technical convenience, but brings tangible gains in terms of the usefulness of the model. Ensuring that states resolve all relevant uncertainty may lead to a cumbersome model, but a sufficiently detailed state space yields a more powerful language that in turn expands the scope of the theory and resolves confusion. The following examples illustrate.

1. Static, one-shot decision problems are analytically convenient, but uncomfortably limiting. Many decision problems involve sequences of decisions, interspersed with choices by Nature, to which static models seemingly do not apply.

However, we can think of dynamic decision problems as games played between the decision maker and Nature. A state is a strategy for Nature, and an act is a strategy for the decision maker. This allows us to collapse the decision problem into a single-stage problem. The construction is identical to the “normal form” representation of an “extensive form”, and consists of defining strategies as functions from

²Luce and Raiffa (1989, pp. 301–302) emphasize the importance of formulating one’s model: “In practice, however, one’s choices for a series of problems—no matter how simple—usually are not consistent. ... Once confronted with such inconsistencies, one should, so the argument goes, modify one’s initial decisions in such a manner as to be consistent. Let us assume that this jockeying—making snap judgments, checking on their consistency, modifying them, again checking on consistency, etc.—leads ultimately to a bone fide a priori distribution. ...then we are committed to a criterion which selects as optimal only acts that are best against this a priori distribution.”

all decision nodes at which players might find themselves into the available moves at these nodes. The equivalence between the extensive and the normal forms of a decision tree (or a game) requires a variety of rationality assumptions: the decision maker is assumed to know the decision tree; to correctly predict her emotional reactions; to be dynamically consistent; and to be unfazed by the size of the decision matrix (which will typically be exponential in the size of the decision tree). However, if we put various rationality limitations aside, this construction of the state space simplifies the decision problem conceptually, and leaves us the canonical decision matrix.³ For example, this representation combines with the minmax theorem to show that finite versions of chess are formally trivial (though still complicated enough to provide ample recreation), obviating the need for more involved backward induction arguments.⁴

2. One of the most important examples of appropriately specifying the state space is Harsanyi's (1967, 1968) modeling of incomplete information games. Before his contribution, game theory was considered to be limited to cases in which players' utilities were known. Harsanyi's innovation can be summarized by saying, "Well, if utilities and/or beliefs are not known, let's put them in the description of the state." In a mental exercise that mirrors his own treatment of social choice problems (the exercise that Rawls (1971) later referred to as going "behind the veil of ignorance"), Harsanyi expanded the state space to describe anything that *any* player might not know, including their own attributes. At the risk of being overly dramatic, one can say that this re-definition of the state space opened the door to the conquest of economic theory by strategic game theory.
3. The application of decision theory to game theory, and, in particular, to interactive epistemology, required that a state of nature—often referred to as "a state of the world" in this context—be rich enough to describe the uncertainty as viewed from the perspective of *each* of the players involved. Because each player considers other players' moves as a source of uncertainty, a state needs to describe the choices made by all players. Moreover, to reason about these, a state needs to describe what a player knows, and therefore also what all players know.

³Savage (1954, pp. 15–17) explains this construction, describing the choice of an act as a single choice of a strategy made, as it were, before the decision maker is born.

⁴See Ewerhart (2002) for a discussion of game-theoretic analyses of chess.

Thus, states were upgraded to describe actions, information, information about others' information, and so on. The encapsulation of these hierarchies of information, or beliefs, in a single state space, pioneered by Aumann (1976, 1987), gave rise to powerful results. Among these are the demonstration that Bayesian rationality implies correlated equilibrium if players share a common prior (Aumann, 1987), and that Bayesian rationality gives rise to rationalizable play if their priors differ (Brandenburger and Dekel, 1987). It also raised fundamental problems, in the form of results that were (and are) considered counterintuitive, such as the Impossibility of Agreeing to Disagree, or unrealistic, such as no trade and no betting theorems. (See Aumann, 1976, Milgrom and Stokey, 1982, Sebenius and Geanakoplos 1983.)

4. Newcomb's paradox (see Nozick, 1969) challenges the rationality of choosing dominant strategies. While the original paradox involves the abstruse, not to say metaphysical notion of an "omniscient predictor", the paradox has a rather mundane version: the decision maker is asked to choose either only the contents of an opaque box, or the contents of both the opaque and a transparent box. The decision maker can see that the transparent box contains \$1,000, and knows that the opaque may or may not contain \$1,000,000.

The seemingly obvious representation of this problem has two states, intuitively described as "the opaque box is empty" and "the opaque box contains \$1,000,000". It then seems obvious that taking both boxes is a dominant strategy. However, the paradox is then created by supposing that many people have been faced with this choice in the past. All those who were greedy (chose both boxes) found that the opaque box was empty, whereas all the modest ones (who chose only the opaque) ended up millionaires. We are thus faced with the discomforting picture of a seemingly rational decision maker who is sorely tempted to choose a dominated strategy.

The simplest way out of this paradox is to reformulate the states more carefully (see Gibbard and Harper, 1978). The formulation of a decision matrix implicitly assumes that one's choice of act has no influence over the states of the world, or over their probabilities. This assumption often obviously fails, as in moral hazard problems, and one can well imagine the decision maker facing Newcomb's boxes coming to believe that her choice of act somehow affects whether the opaque box is bristling with cash. We cannot simply inject this effect into the exist-

ing two-state model, since a “state” whose occurrence partly depends on the decision maker’s act is not a state. However, we can think of an expanded notion of a state as a function from acts to outcomes. Thus, in Newcomb’s problem it is misleading to use only two states: such a formulation implicitly assumes that the decision maker’s choice does not affect the content of the opaque box. While a reasonable assumption a priori, evidence about the choices made by others suggests that this assumption may not hold. Hence, one has to re-phrase Newcomb’s problem with (at least) four states, describing the content of the opaque box *as a function of* the decision maker’s choice.⁵ In such a formulation, the dominance argument breaks down, as there exists a state in which the decision maker is rewarded for modesty and penalized for greediness. Moreover, the evidence cited suggests that this state of the world should have non-negligible probability.

Defining the states of the world to be all the functions from acts to outcomes (see Karni and Schmeidler, 1991, Karni, 2017) is akin to the definition of a state space as the semantic of a propositional logic model: in the latter, a state is a truth function, assigning “true” or “false” values, in a consistent way, to each proposition. In the context of decision making, the propositions of interest are “if I choose a the outcome is r ”. An assignment of truth values to all these, which is consistent with the principle that the outcomes are mutually exclusive and exhaustive, is equivalent to a function that assigns to each act a unique outcome.

5. Monty Hall’s three-door riddle is, by now, well known and well understood (Gilboa (2009, p. 120-122), Rosenhouse (2009)). Yet, at the time it posed a puzzle, and even many mathematicians tended to give the wrong answer. Clearly, the riddle is not one of mathematics, but of modeling.

People often argue that there are three states, corresponding to which door contains the coveted prize, with a prior probability $(\frac{1}{3}, \frac{1}{3}, \frac{1}{3})$ that makes each door equally likely. Once, say, the middle door has been opened to reveal that it contains no prize, the prior probability is updated to $(\frac{1}{2}, 0, \frac{1}{2})$. This is a perfectly correct mathematical calculation, only it is conducted in the wrong model. In particular, the decision

⁵Formally, there are four possible outcome and two acts, so that there should be 16 states in the problem. But to convey the main point one may retain the assumption that the decision maker knows the *possible* contents of each box, allowing a reduction to four states.

maker needs to model not only the location of the prize, but how she learns that one door does *not* contain the prize. To capture this, we can reformulate the notion of a state to identify which door contains the prize, and which door Monty opens. The probability that the prize lies behind each door remains $\frac{1}{3}$, but the states are not equally likely—Monty never opens the door chosen by the decision maker, never opens a door that contains the prize, and is equally likely to open either door not chosen by the decision maker if neither contains the prize. Once these regularities are built into the prior, the updating trivially leads to the conclusion that one’s odds of winning the prize is $\frac{1}{3}$ if sticking with one’s original door, but $\frac{2}{3}$ if switching. The key here is that the state of the world should include the underlying uncertainty, the information obtained, and also the process through which it was obtained.

The fact that, beyond the information one obtains, one may learn from who provided this information, what else they could have said, and so forth, is essential to information economics.⁶ It is also relevant to the understanding of language, as illustrated by Grice’s Principle (Grice, 1957).⁷ Further, there are several apparent deviations from classical decision theory that can be explained by reformulating states to including information channels. A few examples are:

- (a) It has been suggested (Morris (1997), Mukerji (1997)) that a possible resolution of Ellsberg’s “paradoxes”, that is, a way to reconcile ambiguity non-neutrality with subjective expected utility theory, is to allow for the possibility that decision makers who participate in experiments are suspicious of the experimenters. In particular, the former might trust that the latter do not lie, but could conjecture that the composition of the “unknown urn” might be determined by the experimenters after the subjects make their choices, in such a way that minimizes the subjects’ payoffs. Thus, the claim is that people who may seem to be susceptible to ambiguity in a model with “naive” states, may not be so in a more refined model, where the states describe not only the bet but also the way that the odds respond to their own choices in the experiment.

⁶This point was made by Roger Myerson, (personal communication, 1991) when discussing the importance of the Monty Hall problem for probability classes.

⁷Grice suggested that, in normal conversations, people tend to say the simplest statement that conveys the message they wish to get across, and if a simpler message has not been sent, the listener can infer that it is probably untrue.

- (b) Some phenomena that are considered to be “framing effects” (Tversky and Kahneman, 1981) may be explained by learning from the way information is gathered. To consider a simple example,⁸ suppose that a patient has to make a decision about a medical procedure, which has only “success” and “failure” as possible outcomes. Assume that the patient tends to make a different decision if his doctor tells him that the chance of success is 97% than if she tells him the chance of failure is 3%. One may argue that the very fact that the doctor chose to highlight one of the two possibilities is informative in and of itself. In particular, getting the information on 3% probability of failure, the patient may ask himself, “Why didn’t she say that success had 97% probability? Is she uncomfortable about something here?” We may embed the dialog in a richer model, in which the doctor has some uncertainty about the cited numbers, and the patient has second-order beliefs about the doctor’s uncertainty, motives, etc., and different choices given different “frames” would be perfectly in line with standard economic theory.⁹
- (c) Along similar lines, some violations of the weak axiom of revealed preference (or the “independence of irrelevant alternatives”) can be explained if one considers the possibility that the choice set is informative. For example, the “compromise effect” (see Simonson, 1989) refers to the finding that alternative b may be chosen more often from the choice set $\{a, b, c\}$ than it is from the choice set $\{a, b\}$, *if* b can be viewed as a compromise between a and c . Wernerfelt (1995) suggested that this may be a result of rational updating, assuming that the very existence of options on the menu says something about their popularity, and therefore also about their utility to a randomly chosen consumer. Thus, whereas in the presentation of the axiom in classrooms one may think of a choice set as a tool used to observe given preferences, this explanation suggests that uncertainty is involved, and that

⁸Ehud Kalai, 1990s, personal communication.

⁹This type of explanation appears rather plausible in one of the most powerful examples of framing effects, namely the choice of a default option in 401K forms (Madrian and Shea, 2001). If we assume that a typical agent considers herself to be similar to the majority, and believes that the form was designed in such a way that the modal choice is the default, choosing the default is akin to ignoring one’s private signal and following the herd. As is well known from information cascade models, such behavior is perfectly compatible with subjective expected utility maximization.

the states should describe the choice set itself.

These five classes of problems share some similarities, as do their solutions. For example, embedding a dynamic decision problem in its one-shot, “normal” form and the resolution of Newcomb’s paradox are quite similar, as both consist of defining states as functions from possible alternatives to consequences. Nonetheless, we have at least five distinct (if not disjoint) classes of conceptual problems that have been resolved by making the state space sufficiently more informative: (i) the static, one-shot decision problem model is rather limited; (ii) there may be uncertainty about the nature of a game, players’s payoffs and beliefs; (iii) an agent’s uncertainty may be intertwined with that of others with whom she interacts; (iv) acts might seem to causally affect states; (v) the very way that information is imparted might be informative. In all five cases, problems, puzzles, and paradoxes are resolved simply and elegantly by reformulating the states. When the states describe *everything* that can be of interest—be it Nature’s strategy, the parameters of the game, the decisions of others, causal relationships, the information protocol—conundrums disappear.

3 Forming and Measuring Beliefs

This section suggests that the power of the suitably enlarged state spaces described in Section 2.2 is purchased at a cost, beyond the obvious one of making the model more cumbersome. Defining, or at least imagining a large state space is one thing; forming probabilistic beliefs about it is quite another.

3.1 Presumably-Observable Beliefs

How do decision makers form their beliefs? In some cases, we might argue that beliefs can be objectively specified. We are often comfortable asserting that the probability a fair coin produces a head is $1/2$, or that a fair die produces each of its sides with probability $1/6$. Such conventions are useful for thought experiments, but leave open the gnawing suspicion that it is not obvious how we are to ascertain that the coin or die is fair. A standard response is to take a frequentist approach, appealing to previous tosses of the coin or die, and to use the relative frequencies of these tosses (at least if one has enough of them) as an estimate of the relative probabilities. This type of reasoning expands the ability to form beliefs to i.i.d. trials in general, including many for which there is no a priori idea of what “fair” means.

However, it is still uncomfortable to restrict decision theory to cases in which one has an arbitrarily large record of past draws.

Starting with Borel (1924a,b), Ramsey (1926a,b), and de Finetti (1931,1937), and culminating with Savage (1954) and Anscombe-Aumann (1963), the Bayesian approach relied on the idea that subjective probabilities can be defined and measured by betting behavior. The intuitive dictum, “Put your money where your mouth is” became a principle of measurement. The idea is that a decision maker, when choosing a lottery that pays 1 if state θ occurs over a lottery that pays 1 if state θ' has occurred, has just revealed that she regards state θ as more likely. This principle is backed by compelling intuition. If Sarah insists that Horse 1 is more likely to win the race, but then bets thousands of dollars on Horse 2, we might well suspect that her “true”, or at least more relevant subjective probability puts higher weight on Horse 2 winning than on Horse 1. Given the ability to assess such bets, we say that beliefs, and indeed preferences as well, are *presumably observable*. The very elegant results in the field—especially the awe-inspiring theorem by Savage (1954)—suggest that such presumably-observable preferences suffice to pin down a decision maker’s subjective probability.

Decision theory thus rests on twin foundations, consisting of properly formulated states and the ability to infer beliefs via evaluating properly chosen bets.

3.2 A Conflict of Principles

Making states sufficiently detailed as to capture all relevant uncertainty can disrupt the ability to elicit beliefs by betting on states. To begin with, as the state space becomes larger, there are more parameters to be assessed in order to describe beliefs. Worse still, often the sources of information on which one could have drawn in order to define a frequentist prior are no longer available, as they are described in the states. As noted in the preceding subsection, if a decision problem involving some state space Ω is encountered repeatedly, the relative frequencies of the states in past problems can serve as a prior probability measure for the next one. But if one wishes to allow for process that are not i.i.d—for example, temporal changes in the underlying distribution, or causal dependency between the realizations in different time periods—one needs to consider a larger state space, such as Ω^∞ , describing *sequences* of realizations in Ω . The problem of generating beliefs over this state space is two-fold: first, Ω^∞ is clearly much larger than Ω , demanding many more parameters to be estimated or guessed. Second, the history on which one wished to rely in forming beliefs is no longer there to help. While

a current replica of Ω may have been encountered in the past, the entire product Ω^∞ is encountered only once. Historical data have turned against us, so to speak: from being part of the solution they switched into part of the problem.

One might then have recourse to the presumably-observable beliefs that emerge from betting behavior. However, in each of the examples considered in Section 2.2, the specification of the state space so that it captures all relevant uncertainty requires that we form beliefs not only about more states than those we encounter in a naive state space, but also that we form beliefs about new types of information: probabilities of counterfactual statements, joint distributions over the types of players in the game, others' actions, possible causal relationships, and various conceivable regimes of information dissemination. In each case, the proposition that beliefs are at least in principle observable is called into question:

1. The construction of states of the world as Nature's strategies implies that in many problems a single state cannot be observed even in principle. For example, assume that the decision maker can choose a or b , and, as a result, obtain the outcome 0 or 1. Thus the game against Nature has two stages, and the decision tree has four leaves. In the normal form of the game, or the decision matrix, there will be four states of the world. Each state should specify what would be the outcome of act a as well as act b , *even though only one* of these acts will actually be chosen. Thus, a state can be thought of as a list of conditional statements. The antecedent of exactly one of these conditional statements materializes in any given instance, and the rest remain counterfactual.¹⁰ As a result, there are functions defined on the states that cannot be compared in this problem. The problem is attenuated if one can think of it as repeated under more or less the same conditions, as in the case of drawing balls from urns. But when the decision maker has to choose her life path, or a country has to decide whether to go to war, the counterfactual statements needed to define states are hard to assess for their truth values. And the assumption that the decision maker's beliefs can be gleaned from observed bets on states or choices between pairs of acts becomes doubtful.

2. If we take the Harsanyi-Rawls suggestion of considering the uncer-

¹⁰We can denote the states by $s_{00}, s_{01}, s_{10}, s_{11}$, where s_{ij} denotes outcome i from act a and outcome j from act b . Note that if the decision maker chooses a she will have the partition $\{\{s_{00}, s_{01}\}, \{s_{10}, s_{11}\}\}$ and if she chooses b – the partition $\{\{s_{00}, s_{10}\}, \{s_{01}, s_{11}\}\}$. She will never observe a single state.

tainty we faced before we were born seriously, we are led to discuss the observed preferences of the decision maker’s self before she came into being. But such selves do not make decisions. Clearly, there are many applications of Harsanyi’s model that need not take such a literal interpretation of prior beliefs before birth. If a worker might be of a “high” or a “low” type, or a firm can have high or low production costs, they can easily use the population statistics as the prior over their own type, and these beliefs would be reasonable beliefs for others to have over their type. These types of applications typically assume, implicitly, that the entire problem has been repeated in the past, so that empirical frequencies in the past may serve as prior beliefs for the current instantiation.¹¹ However, if a problem is not repeated in sufficiently similar ways, either because of its unique character or because the entire history is part of the definition of the problem, going “behind the veil of ignorance” is a philosophical exercise that stands at odds with the notion of revealed preference.

3. Demanding that a state resolve all uncertainty for several players simultaneously raises questions about the observability of choices among bets on such states. For example, if a game has two players, and we wish to derive the probability that a player has about the state space, we should first decompose a state into the player’s own act, and the “state” as she perceives it, that is, resolving all uncertainty that is external to her. We should then imagine that she is offered bets on such states, depending on the other player’s choices. The same exercise should be conducted for the other player. It is not obvious how one can elicit such belief from betting behavior, even in principle. (See Samet and Schmeidler, 2017.) Further, the elicited prior beliefs would be on different spaces that then need to be “patched” together.¹²
4. When we consider all possible causal relationships between acts and outcomes, we again undermine the presumably-observable basis of subjective probabilities. The problem (pointed out in Gilboa and Schmeidler, 1995)¹³ is the following: start with the set of acts that the decision maker can choose among, A , and the set of conceivable consequences, X . We are supposed to define the state space as all possible functions from the former into the latter: $S = X^A$. Then, Savage’s construction

¹¹This can also be the case when a prior is used in Bayesian statistics.

¹²See Samet (1998) and Feinberg (2000) on results that allow the different priors to be derivable from a single prior probability of an external observer.

¹³See also Gilboa (2009), Gilboa, Postlewaite, and Schmeidler (2012).

would require a complete preference relation (satisfying P1-P7) on *all* “conceivable acts”, which are all the functions from S to X . That is, while we can only observe choices between pairs in A , we are asked to imagine choices between all pairs in $F = X^S = X^{(X^A)}$, a set that is larger than A by two orders of magnitude.

We might respond by either enriching the set of acts available to the decision maker, or reducing the set of states. Both approaches are problematic. The obvious expanded set of acts is F . But if we take F as our starting point, with no means of limiting the set of states, then we simply put this construction mechanism in motion again, giving us the state space X^F and the set of conceivable acts $X^{(X^F)}$, giving an even greater mismatch. We might instead argue that there are some states we can a priori eliminate from consideration. When discussing Newcomb’s paradox in Section 2.2, for example, we took it for granted that only four of the sixteen possible states need be considered. But we are now assuming that we have a priori knowledge of the very beliefs we hope to infer from behavior. This may sometimes be appropriate, but appears to be a rather fragile foundation for decision theory.¹⁴

5. The fact that information might be part of the state space suggests that the measurement of beliefs might also be intertwined with the beliefs measured. Whatever is our measurement device, such as a lab experiment, it is possible that the decision makers involved suspect that there are some causal relationships between their behavior in the experiment and the outcome they would experience in the future. As suggested by the explanation of ambiguity aversion in Ellsberg’s experiment in Section 2.2, it is possible that, when participating in an experiment, people think of it as a game involving the experimenters. Indeed, rationality seems to suggest that they should. To push this idea to a ridiculous extreme, one may explain any observed behavior as maximization of subjective expected utility: all one needs to do is to include a state at which, should the decision maker behave differently than she actually has, she would be kidnapped by malicious aliens. Clearly, the alien threat does not only allow to explain away any anomaly; it also renders prediction impossible, as behavior that

¹⁴Schipper (2016) similarly notes that there is a tension between incorporating information in states and the ability to elicit beliefs about these states from choice behavior. He considers restrictions on the set of states and on the set of acts as possible responses. He recognizes the weakness of either approach, in that it undermines the ability to elicit beliefs from choice behavior.

was observed under the threat would not be indicative of behavior when the aliens are no longer there. But even when limiting attention to more reasonable conjectures about the outcome of the experiment, the elicitation of beliefs by betting behavior is problematic when the elicitation procedure might be causally intertwined with the beliefs to be elicited.

To conclude, larger state spaces increase the descriptive power of the model, and they can resolve a variety of unpleasant paradoxes. But this enhanced theoretical power comes at a cost. From a psychological point of view, the task of forming beliefs over the state space becomes more demanding, often attempting to estimate more parameters based on less information. From a philosophical viewpoint, the presumably-observed behavioral foundations of subjective probability become shaky, at times to such a degree that the exercise does not seem to be coherent even “in principle”.

We close this section with a casual observation on the sociology of the field. Decision and game theory are more “conceptual frameworks” than specific theories. Even when we couple the framework with a solution concept (such as expected utility maximization, or Nash equilibrium), there is a considerable degree of freedom in defining “players”, “states”, “acts”, and so forth. This freedom is very useful, and it is responsible for many interesting and insightful analogies. But it is also responsible for conceptual pitfalls. For example, we may explain Savage’s model using balls drawn from urns, or perhaps an upcoming election. The states are quite naturally defined in these problems, and we derive the notion of a prior belief. Then we move on to incomplete information games, and assume that the unborn players have a prior over their subsequent identities, “as in Savage”. Similarly, facing a causal paradox such as Newcomb’s, we may re-define the states to deal with it, but, facing the next problem, go back to a simpler definition of the state space. Our claim above is not that Savage’s derivation of prior probabilities is logically flawed. The claim is that a construction that makes sense in the classroom examples of the state space may be quite shaky in many of the applications in which economic theory resorts to Savage’s model.

4 Eventualities

It follows from the discussion above that one may define rich state spaces at will, but that one should be careful when defining beliefs over them, and take extra care when quoting results that pin down prior beliefs over the state space. In the following we attempt to be more precise about the

way such results should be interpreted. We first present a general notion of observability. We apply the concept to decision theory in general, and then to Savage’s (1954) model in particular.

4.1 Observability

When saying that an event or magnitude is observable, it is typically assumed that the event or magnitude can be measured, and, furthermore, that the measurement can be repeated. It is further assumed, at least implicitly, that the measurement could be taken by different scientists, who would reach similar conclusions. For example, if we think of a physical quantity $\mu \in \mathbb{R}$, a measurement thereof would be a real-valued random variable X_t , ideally with $E(X_t) = \mu$, and it would be natural to assume that different measurements generate a sequence of $(X_t)_t$ that are i.i.d. given μ . As μ is not known, a Bayesian approach would suggest that the measurements $(X_t)_t$ be exchangeable. In any event—taking a Bayesian or a Classical approach—one can test some implications of the conditional-i.i.d. assumption. For instance, statistics that are computed based on long sequences should have similar distributions, and if we can compute their averages, the resulting values should be similar.

Notice that this notion of observability implicitly assumes that the phenomenon to be measured has a certain degree of *stability* and *causal independence*: it has to be sufficiently stable so that it would not change from one measurement to another, or from the measurement of one attribute to another, and it has to be causally independent of the process of measurement.

4.2 Observable Preferences

When axiomatic derivations of decision rules are interpreted descriptively, it is common to assume that they deal with observable phenomena, such as choices.¹⁵ If we couple this assumption with the definition of observability above, the content of the assumption is that for any choice there exist repeated measures of that choice that are conditionally i.i.d. (or exchangeable).

¹⁵In some cases, such as the notion of “qualitative probability”, the observable ranking might be interpreted as a cognitive construct, “at least as likely as”. Our discussion applies to these as well as to choice data. Gul and Pesendorfer (2008) argue that choices are the only conceivable economic data.

For example, consider the revealed-preference foundation for utility maximization over a finite set of consumption bundles, X . One assumes that a binary relation over X can be observed, and proves that this relation can be represented by maximization of a numerical function if and only if it is complete and transitive. The assumption that the preference relation between $x, y \in X$ is observable means that there are random variables $xy_t \in \{0, 1\}$, for $t \geq 1$, where $xy_t = 1$ is interpreted as “at measurement t , x was found to be at least as desirable as y ” (typically, because x and y were both feasible, but x was chosen), and that these variables are conditionally i.i.d. Note that this assumption implicitly requires that the preference one presumably measures would not change from one t to another—neither exogenously, say, because of the passage of time, nor endogenously, say, because of habit formation.¹⁶

The observability assumption is an empirical assertion about the data that are available. To see this more clearly, it is useful to think of axiomatizations as second-order scientific theories, along the following lines. Let there be an external reality, consisting of people, households, or organizations making decisions. Let there be a working economist, WE, who models these decisions. We can think of the economist as examining the data and describing it by a model. For example, WE may consider a household choosing between pairs of bundles, and attempt to estimate a utility function that would explain the household’s past choices and help predict future ones. There is, however, also a “second-order-” or “meta-” theorist, Th, who models WE’s work. Th observes WE, and asks, when can he (WE) assume that a utility function exists? Suppose that Th proves the theorem mentioned above, stating that any complete and transitive binary relation can be represented by utility maximization. The theorem is a mathematical result, showing the identity of two sets of binary relations. Binary relations and sets thereof are, obviously, mathematical constructs. In the present context they are part of a model that is designed by Th to describe reality. The reality that Th considers, however, isn’t the household’s choices, but the scientific work that WE performs in trying to describe these choices. That is, Th considers WE, who collects data and writes models. She (Th) then writes down a formal model, describing WE’s work, by defining two classes of binary relations. One binary relation refers to the choice data that WE gathers, and the other refers to the preferences arising out of the

¹⁶ A more elaborate model would allow the study of changes in preferences due to considerations such as habit formation. An observation would then specify the preferences before and after the change in preference, while presuming that such an observation could be repeated in a conditional-iid (or exchangeable) manner.

utility maximization WE uses in his paper. The theorem that Th proves allows her to say when WE’s attempt to describe the household’s behavior via utility maximization might be successful. For WE, there is one formal model, that of utility maximization; for Th there are two—WE’s model of utility maximization, as well as another model to describe WE’s data (by a binary relation).¹⁷

Clearly, Th’s work as a social scientist should be judged not only for its mathematical beauty or depth, but also for its relevance. For example, assume that another theorist, Th’, comes along and says to Th, “Your theorem is really nice, and even correct. But I have to tell you, as a model of reality it’s rather poor. I have talked to many empirical economists, and none of them seems to think that a binary relation can be observed in reality. Maybe it can be in experiments. But with empirical data—not really. What’s observable is, at best, a choice of a bundle out of a budget set. That is, given prices p , WE can observe a chosen bundle x . His resulting database is, therefore, of the type $\{(p_i, x_i)\}_{i \leq n}$ —but no choice among finite subsets of bundles can be assumed observable.” Thus, Th’ might suggest that Th abandon her analysis of WE’s work in favor of an alternative model based on Afriat’s Theorem (Afriat, 1967), which characterizes databases $\{(p_i, x_i)\}_{i \leq n}$ that can be rationalized by utility maximization.

From the perspective of Th’, Th was using a model that was too highly idealized. However, just as working economists may make idealized assumptions that can be misleading—say, ignoring transaction costs in a particular environment—so might theorists make excessively idealized assumptions about observability, allowing Th to defend her model as usefully sacrificing some realism in order to gain tractability. Moreover, just as a mathematical model of a working economist might be a reasonable approximation of reality in some contexts but not in others, so can the assumptions of observability in the case of the theorist’s work. For example, binary relations might well be observable in experiments, but not in empirical work. However every model of decision theory that is committed to the principle of presumably-observable beliefs makes some assumptions about observability.

4.3 Eventualities in Savage’s Model

We now consider the interpretation of Savage’s model as a description of the way working economists model decision making under uncertainty. These

¹⁷If we liken WE to an artist, who represents reality by a painting, Th is akin to a painter who represents another painter in the act of painting. Th’s canvas might have a representation of the canvas that WE draws on, as well as of the reality WE observes.

economists write down models involving probabilities and utilities, which are supposed to describe choices among alternative acts. Savage's theorem shows the equivalence between two formal entities: (i) the formal models working economists write, with utility functions and probability measures; and (ii) the data that these economists wish to explain, modeled as a binary relation over acts. The appropriateness of this model depends on the assumption that this binary relation is at least in principle observable. In particular, because completeness is a key axiom in Savage's derivation, the assumption of being presumably observable applies to all pairs of pairs of acts, f, g . And it states that (at least in principle), for each such pair preferences can be observed, where we interpret this assumption as above, with measurements of the observation of $f \succsim g$ being conditionally i.i.d. (or exchangeable) random variables.¹⁸

In order for this observability condition to be a reasonable assumption (about the reality of WE's work), we cannot blindly construct the set of states S in Savage's model by taking each state to resolve all uncertainty. Instead, the set S must consist of entities that allow repeated measures of preferences. For example, for any pair $s, s' \in S$, we should be able to imagine the agent choosing between a lottery that pays if s occurs, and a lottery that pays if s' occurs. Moreover, we must be able to make multiple measures of the same preference question, in a way that does not change with time, nor with the measurement procedure itself.

We call such entities eventualities. More precisely, a set S is a set of *eventualities* for a decision maker relative to a set of outcomes X if there exists a complete binary relation \succsim such that, for any $f, g \in X^S$, one can in principle observe a sequence of random variables $f g_t \in \{0, 1\}$ that are conditionally i.i.d.

This notion of an eventuality then gives us a tool for evaluating models of decision making. Savage (1954) notes that any exercise in decision making begins by molding states so as to capture relevant uncertainty and exclude irrelevant uncertainty. If the resulting states in the model can be reasonably viewed as eventualities, then Savage's analysis provides an elegant model

¹⁸The two-level model of economic modeling can also apply to normative interpretations of decision theory. For example, a working economist might be advising a government agency on a policy issue, suggesting that the agency maximizes expected utility, where the (meta-)theoretical economist might provide a result that justifies this particular model of choice based on properties of choices that necessitate it. In this interpretation actual observability of choices is of lesser importance. However, if some arguments resort to choices that can never be observed even in principle (as discussed above), their validity might be questionable.

of the way economists model decision making. In particular, actions in such a model map eventualities into outcomes in such a way that we can imagine designing decision problems that will allow us to observe an agent's preferences over the resulting states. If this is not the case, then preferences are fundamentally unobservable, and we are wary of appealing to Savage as a foundation for decision making. Therefore, from our "second-order" perspective, Savage's model is appraised for its plausibility—that is, for its potential to be empirically accurate, rather than for its effective empirical accuracy.

In the simplest examples of Savage's model, the naturally defined "states" obviously correspond to eventualities satisfying our desiderata. For example, if we draw balls from urns, with monetary payoffs depending on the color drawn, then the colors of the ball can serve as eventualities for all decision makers. Indeed, an experimenter who wishes to measure the decision maker's subjective probability for the different colors can attach to each eventuality any monetary outcome, and observe the decision maker's preferences. In assuming that these preferences are observable, we assume that they can be repeatedly measured, perhaps also by different experimenters, generating a conditionally i.i.d. (or exchangeable) sequence of random variables. In addition, we assume that sufficiently long sub-sequences of such repeated measures yield approximately the same frequencies of results. While these are empirical assumptions that, in principle, can and should be statistically tested, they appear to be reasonable enough to use as a working hypothesis even when not formally tested. Similarly, for most voters, the possible results of an upcoming election would constitute a set of eventualities with respect to monetary outcomes, because one can imagine financial assets that yield various amounts of money if, say, a Democrat or a Republican wins.

In the famous exchange between Aumann and Savage (1970, 1971, see Drèze (1987)) regarding axioms P3 and P4, Aumann attacked the implicit assumption of state-independent utility, providing as examples the outcomes "swimsuit" and "umbrella", whose utilities change as a function of the weather (described by the state of the world). Savage replied that these goods aren't outcomes, and that an outcome should specify all aspects relevant to the individual's well-being, say, "lying on the beach on a sunny day" or "running for shelter half naked in the rain". He was fully aware of the fact that his model required one to imagine acts that assign an outcome such as "lying on the beach on a sunny day" to the state "Rain". He viewed his own model as asking the reader for some indulgence in imagining that the decision maker had preferences over all such acts and that these preferences

satisfied P1-P7.

We would re-phrase Savage's answer by using the term "eventuality". We can view rain and sun as eventualities, and elicit beliefs about these events. The exercise is relatively simple if the outcomes, X , are only monetary. It becomes more challenging when we require the decision maker to imagine sunbathing in the rain, though still possible (perhaps with some imagination). In particular, we understand Savage's reply as saying that weather conditions can be viewed as eventualities, satisfying observability conditions along the lines described above. More generally, we view Savage as arguing that many decision problems of interest are such that the states, again perhaps with some imagination, can reasonably be viewed as eventualities. Note that we take a permissive view of the assumption of observability here, requiring that preferences are at least *in principle* observable.

The definition of eventualities is more problematic when the states of the world involve uncertainty about one's life or health (physical or psychological). Savage was aware of these limitations. Ridiculing his own mental exercise, he famously wrote, "... I should not mind being hung so long as it be done without damage to my health or reputation" (see Drèze (1987), p. 78). Being hanged is clearly not an eventuality. The assumption that a working economist can gather data about the decision maker's preferences over acts attaching any imaginable outcome to the event in which he is hung is plainly not true. One need not run sophisticated statistical tests, nor interview empirical economists to realize that such databases of preferences do not exist. Indeed, such preferences cannot seriously be assumed observable even in principle. At this point, we cannot view states as eventualities, and instead we view the prospect of eliciting beliefs from observed choices as breaking down.

Note that what is considered to be an acceptable idealization of eventualities may depend on the decision maker in question. For example, assume that we wish to elicit the decision maker's beliefs about a financial crisis. If we are dealing with a small investor, we can imagine offering her a collection of financial assets that would match practically any financial outcome to any eventuality. In particular, we can offer her a bet that would make her very rich if the DJIA loses 50% of its value. By contrast, if we are considering a very rich decision maker, who owns large shares in many DJIA companies, such an exercise would become harder to execute for us in a lab, and even hard for the decision maker to imagine. Along similar lines, we can ask a resident of Country A to consider various bets on a military coup in Country B, but it may not make sense to assume that the choices of Country A's leader between such bets are observable. In other words, we can think of

Savage being hanged as an eventuality for us, but probably not for him.

5 Conclusion

If we insist on interpreting beliefs as being presumably observable and also take states to resolve all uncertainty, including all possible causal relationships between acts and states, all possible causal relationships between the measurement procedure and the measured phenomenon, and also the history that preceded the decision maker's birth, then we are bound to run into conceptual difficulties. We see little hope for arguing that we can observe beliefs over such states in a meaningful way. In response, we propose limiting applications of axiomatic decision theories to cases in which the states in the resulting model can be viewed as eventualities. Actions then attach outcomes to eventualities in combinations that allow us to at least imagine constructing decision problems that will reveal preferences over these actions, allowing us to attach meaningful interpretations to beliefs and preferences. By contrast, in problems that necessitate state-outcome pairs that are implausible to appear in any real decision problem, one should be careful in citing Savage as a foundation for Bayesianism. We may still enjoy the beauty and simplicity of this approach, but we cannot refer to it as the coherent way of decision making.

6 Bibliography

Afriat, S. N. (1967), "The Construction of a Utility Function from Expenditure Data," *International Economic Review*, **8**: 67-77.

Anscombe, F. J., and R. J. Aumann (1963), "A Definition of Subjective Probability", *The Annals of Mathematics and Statistics*, **34**: 199-205.

Aumann, Robert J. (1974), "Subjectivity and Correlation in Randomized Strategies", *Journal of Mathematical Economics*, **1**:67-96.

Aumann, Robert J. (1976), "Agreeing to Disagree", *Annals of Statistics*, **4**:1236-1239.

Aumann, Robert J. (1987), "Correlated Equilibrium as an Expression of Bayesian Rationality", *Econometrica*, **55**:–18.

Borel, E. (1924a), "A propos d'un traité de probabilité", *Revue philosophique*, **98**: 321-336.

- Borel, E. (1924b), *Elements de la theorie des probabilités*, Paris: Librairie Scientifique Hermann.
- Brandenburger, Adam and Eddie Dekel (1987), “Rationalizability and Correlated Equilibria”, *Econometrica*, **55**:1391-1402.
- Feinberg, Yossi (2000), “Characterizing Common Priors in the Form of Posteriors,” *Journal of Economic Theory*, **91**:127-179.
- de Finetti, B. (1931), “Sul Significato Soggettivo della Probabilità”, *Fundamenta Mathematicae*, **17**: 298-329.
- (1937), “La Prevision: ses Lois Logiques, ses Sources Subjectives”, *Annales de l’Institut Henri Poincaré*, **7**: 1-68.
- Drèze, J. H. (1987), *Essays on Economic Decisions under Uncertainty*, Cambridge: Cambridge University Press.
- Ewerhart, C. (2002), “Backward Induction and Game-Theoretic Analysis of Chess”, *Games and Economic Behavior*, **39**: 206-214.
- Gibbard, A., and W. L. Harper (1978), “Counterfactuals and Two Kinds of Expected Utility”, *Foundations and Applications of Decision Theory*, **1**: 125-162.
- Gilboa, I. (2009), *Theory of Decision under Uncertainty*, Cambridge: Cambridge University Press.
- Gilboa, I., A. Postlewaite, and D. Schmeidler (2009), “Is It Always Rational to Satisfy Savage’s Axioms?”, *Economics and Philosophy*, **25**: 285-296.
- (2012), “Rationality of Belief”, *Synthese*, **187**: 11-31.
- Gilboa, I., and D. Schmeidler (1995), “Case-Based Decision Theory”, *The Quarterly Journal of Economics*, **110**: 605-639.
- Grice, H. P. (1957), “Meaning”, *The Philosophical Review*, **66**: 377-388.
- Gul, Faruk and Wolfgang Pesendorfer (2008), “The Case for Mindless Economics”, in *The Foundations of Positive and Normative Economics: A Handbook*, Andrew Caplin and Andrew Schotter (Eds.), Oxford: Oxford University Press.
- Harsanyi, J. C. (1967), “Games with Incomplete Information Played by “Bayesian” Players, I-III: Part I. The Basic Model”, *Management Science*, **14**: 159-182.

- Harsanyi, J. C. (1968), "Games with Incomplete Information Played by "Bayesian" Players, I-III. Part II. Bayesian Equilibrium Points", *Management Science*, **14**: 320-334.
- Karni, E. (2017), "States of Nature and the Nature of States", *Economics and Philosophy*, **33**: 73-90.
- Karni, E., and D. Schmeidler (1991), "Utility Theory with Uncertainty", in *Handbook of Mathematical Economics*, Vol. 4, Werner Hildenbrand and Hugo Sonnenschein (Eds.), New York: Elsevier Science.
- Kolmogorov, Andrey (1956 (originally 1933)), *Foundations of the Theory of Probability*, New York: Chelsea.
- Luce, R. D., and H. Raiffa (1989 (originally 1957)), *Games and Decisions*, New York: Dover Publications, Inc..
- Madrian, B. C., and D. F. Shea (2001), "The Power of Suggestion: Inertia in 401(k) Participation and Savings Behavior", *The Quarterly Journal of Economics*, **116**: 1149-1187.
- Milgrom, Paul R. and Nancy Stokey (1982), "Information, Trade, and Common Knowledge", *Journal of Economic Theory*, **26**:17-27.
- Morris, S. (1997), "Risk, uncertainty and hidden information", *Theory and Decision*, **42**: 235-269.
- Mukerji, S. (1997), "Understanding the Nonadditive Probability Decision Model", *Economic Theory*, **9**: 23-46.
- Nozick, R. (1969), "Newcomb's problem and two principles of choice", in *Essays in Honor of Carl G. Hempel*, Nicholas Rescher (ed.), Dordrecht Holland: Reidel, pp. 114-146.
- Ramsey, F. P. (1926a), "Truth and Probability", in R. Braithwaite (ed.), (1931), *The Foundation of Mathematics and Other Logical Essays*. London: Routledge and Kegan.
- Ramsey, F. P. (1926b), "Mathematical Logic", *Mathematical Gazette*, **13**: 185-194.
- Rawls, John (1971), *A Theory of Justice*, Cambridge: Cambridge University Press.
- Rosenhouse, Jason (2009), *The Monty Hall Problem*, Oxford: Oxford University Press.

- Samet, Dov (1998), “Common Priors and Separation of Convex Sets,” *Games and Economic Behavior*, **24**:172-174.
- Samet, Dov and David Schmeidler (2017), “Desirability”.
- Savage, L. J. (1954), *The Foundations of Statistics*. New York: John Wiley and Sons. (Second addition in 1972, Dover)
- Savage, L. J. (1971), “Letter from Leonard Savage to Robert Aumann, 27 January 1971,” in Robert Aumann (2000), *Collected Papers, Volume 1*, p. 307. Cambridge: MIT Press.
- Schipper, B. C. (2016), “A Note on States and Acts in Theories of Decision under Uncertainty”.
- Sebenius, James K. and John Geanakoplos (1983), “Don’t Bet on it: Contingent Agreements with Asymmetric Information”, *Journal of the American Statistical Association*, **78**, 424-426.
- Simonson, I. (1989), “Choice based on reasons: The case of attraction and compromise effects”. *Journal of Consumer Research*, **16**: 158-174.
- Tversky, A., and D. Kahneman (1981), “The Framing of Decisions and the Psychology of Choice”, *Science*, **211**: 453-458.
- Wernerfelt B. (1995), “A Rational Reconstruction of the Compromise Effect: Using Market Data to Infer Utilities”, *Journal of Consumer Research*, **21**: 627-633.