CAMBRIDGE

Cambridge University Press
0521416744 - Knowledge, Belief, and Strategic Interaction
Edited by Cristina Bicchieri Maria Luisa Dalla Chiara
Excerpt
More information

1

Feasibility

ISAAC LEVI

1. ACT, STATE, AND CONSEQUENCE

According to the procedure Savage (1954) and others have adopted as canonical for representing decision problems, three notions are deployed in the representation: the notion of an act, a state, and a consequence.[1] Many philosophers have followed the lead of Jeffrey (1965) in complaining about a wrong-headed ontology that insists on trinitarianism where monotheism should do. Jeffrey suggests that acts, states, and consequences are all events or propositions.

I do not want to quarrel with Jeffrey's suggestion. To me, something like it should turn out right. But I do not see why it should be supposed, as Jeffrey intimates, that Savage (or, for that matter, Ramsey) would disagree. Perhaps Ramsey and Savage may be convicted of what now seems like loose talk; but it is loose talk easily repaired without damage to the substance of their views. Instead of speaking of acts, states, and consequences, Savage could have spoken of act descriptions, state descriptions, and consequence descriptions, or of act propositions, state propositions, and consequence propositions.

There are, to be sure, important differences in a Savage framework between the attitudes the decision maker has toward act descriptions, state descriptions, and consequence descriptions. State descriptions are objects of personal or credal probability judgments; consequence descriptions are objects of utility judgment and act descriptions of *expected* utility judgment.

There is nothing in the Savage system to prevent assigning utilities to state descriptions or probabilities to consequence descriptions. Indeed, it seems clear that Savage intended the state descriptions to be evaluated with respect to utility in a certain way although, as is well known, his

Thanks are due to Teddy Seidenfeld for helpful suggestions.

© Cambridge University Press www.cambridge.org

CAMBRIDGE

Cambridge University Press
0521416744 - Knowledge, Belief, and Strategic Interaction
Edited by Cristina Bicchieri Maria Luisa Dalla Chiara
Excerpt
More information

axioms do not quite capture his intent. State descriptions cannot be assigned *unconditional* utility but may be assigned utility conditional on the act chosen. That is to say, given that act a_i is chosen and state s_j is true, the utility assigned to s_j conditional on a_i is equal to the unconditional utility assigned the consequence c_{ij}. Unless the consequences of all available options in a given state bear equal utility, the only way to derive an unconditional utility for the state is to compute the expectation of the conditional utilities of the state using unconditional probabilities for acts.[2]

Similarly credal probabilities are assignable to consequences, but these are conditional on the option chosen. Unless consequences are identical for all available options in a given state, the only way to compute unconditional probabilities for consequences is with the aid of unconditional probabilities for those options that yield them in some state or other.

Thus, unconditional utilities for states and unconditional probabilities for consequences are obtainable only if we can assign unconditional probabilities to acts. In the Savage formalism, however, unconditional credal probabilities are not assigned to acts – that is, to the agent's options. According to the Savage approach, one begins with a preference or value ranking of hypothetically available acts or options that satisfy the axioms proposed (and in which the preference ranking over the set of available options is embedded). The axioms are intended to ensure that the ranking evaluates the hypothetically available options with respect to expected utility. With this understood, one may derive a unique unconditional credal probability distribution over the states and an unconditional utility function over the consequences unique up to a positive affine transformation, where the utility of a consequence is independent of which state is true. Using this information, it is possible to derive a probability distribution over consequences conditional on acts and a utility function for states conditional on acts. However, the Savage theory fails to determine an unconditional credal probability over acts and, as a consequence, an unconditional utility function for states and an unconditional probability distribution over consequences.

Thus, the Savage trichotomy does not deny that acts, states, and consequences are propositions, sentences, or other truth-value–bearing entities. What it denies is that credal probability judgments and utility judgments are defined over the same set of propositions.

The Savage axioms, even when construed prescriptively as norms of rationality, are designed so that one might elicit an agent's probabilities and utilities from information about that agent's preferences among acts. However, information about preferences among acts unsupplemented by other data fails to yield conclusions about unconditional probabilities of acts.

© Cambridge University Press www.cambridge.org

CAMBRIDGE

Cambridge University Press
0521416744 - Knowledge, Belief, and Strategic Interaction
Edited by Cristina Bicchieri Maria Luisa Dalla Chiara
Excerpt
More information

Some might argue that this consideration justifies the conclusion that acts are not assigned credal probabilities. This conclusion may, perhaps, be too hasty. In any case, acts, states, and consequences could all be represented as propositions. The same sentence or proposition could qualify as a state description relative to one decision problem, and as a consequence description (or, indeed, even as an act description) relative to another. Given the propositional construal of acts, states, and consequences, this is not unexpected. It does not follow, however, that we must insist that probabilities and utilities be defined over all propositions without restriction in the setting of every specific decision problem. We may therefore endorse Jeffrey's insistence on regarding acts, states, and consequences propositionally without accepting his contention that the decision maker should assign "desirabilities" (utilities or expected utilities) and probabilities to propositions of these three kinds. The point under dispute is not an ontological one; rather, the concern is whether we should restrict the applicability of unconditional utility and probability judgments along Savage's lines or should follow Jeffrey in rejecting such restrictions.

In my opinion, Spohn (1977; 1978, pp. 72–5) is right in resisting Jeffrey on this issue. Decision makers should not assign credal probabilities to the acts available to them. If this is so, we should also reject the idea developed by Jeffrey of taking as fundamental a notion of preference among propositions for the purpose of deriving an agent's credal probability and utility judgments.

Spohn's chief argument for his view is that one need not assign unconditional probabilities to acts in order to determine expected utilities of options for the purpose of identifying optimal or admissible options. Spohn's point is well taken, but it does not entail a prohibition against the deliberating agent assigning credal probabilities to hypotheses predicting his decision. At best, it serves notice to those who insist that such probability assignments may be made (and, perhaps, ought to be made) that they should identify some function such probability assignments can serve other than that of guiding choice among the options under consideration. Spohn reminds us that assigning such probabilities is not crucial in applications of the prescription to maximize expected utility, or, for that matter, of my preferred recommendation that choice be restricted to E-admissible options (Levi 1974, 1980).

I think, however, that there is an additional argument that suggests good reasons why the deliberating agent ought to avoid assigning credal probabilities to predictions of what will be chosen from among the available options in the decision problem currently faced.[3] These considerations derive from reflection on the conditions that a proposition should satisfy, from the point of view of the deliberating agent deploying principles of

© Cambridge University Press www.cambridge.org

CAMBRIDGE

Cambridge University Press
0521416744 - Knowledge, Belief, and Strategic Interaction
Edited by Cristina Bicchieri Maria Luisa Dalla Chiara
Excerpt
More information

rational choice in deciding what to do, in order for it to be an act description representing a feasible option. The remainder of this chapter proposes an answer to this question and indicates some of its ramifications.

2. Possibility for and possibility that

To say that driving from New York to Boston in four hours on Monday is a feasible option for Sam is to say something about Sam's abilities. Sam is able to drive from New York to Boston in four hours on Monday as he chooses: It is *possible for* Sam to drive from New York to Boston in four hours on Monday. Such a claim is not to be confused with the claim that it is *possible that* Sam will drive from New York to Boston in four hours on Monday. The first modal claim predicates of Sam a certain ability. It is true or false of Sam depending on whether Sam satisfies the conditions for the ability attribution. The second modal claim differs in this respect. The agent X who makes the judgment (X may be Sam or some other party) is expressing a certain propositional attitude – an attitude indicating that nothing in what X takes for granted is inconsistent with Sam's making the trip on Monday in four hours. In making this judgment, X need not assume that Sam has the ability to make the trip. There may be some doubt as to whether Sam has that ability, yet X may still judge that it is possible that Sam will take the trip. Conversely, X may take for granted that Sam has the ability and yet judge it impossible that he takes the trip – because (say) X is certain that Sam will not exercise the ability.

Ability attributions are relative to what may be called initial or test conditions or experiments. For example, a person may be able to play the piano with training but lack the ability to play the piano by choice. The ability to play is relative to the test condition describable as undergoing training in piano playing. The inability is relative to the test condition describable as undertaking a deliberation.

The relativity, moreover, is not restricted to the abilities of agents to do things or have things done to them; it applies also to other objects or systems. A given coin may have the ability to land heads on a toss and also the ability to land tails on a toss. But it may lack the ability to land heads on a toss by Morgenbesser (although, if this were true, Morgenbesser himself would have some remarkable abilities). Less fancifully, the coin will lack the ability to land heads on a toss situating it in a mechanical state which, according to the laws of physics, destines it to land tails. Abilities are like dispositions in this respect; this is not surprising, because abilities are duals of dispositions. The ability to respond in manner R on a trial of kind T applies to x if and only if it is false that x

© Cambridge University Press www.cambridge.org

CAMBRIDGE

Cambridge University Press
0521416744 - Knowledge, Belief, and Strategic Interaction
Edited by Cristina Bicchieri Maria Luisa Dalla Chiara
Excerpt
More information

has the (sure-fire) disposition to fail to respond in manner R on a trial of kind T. And since dispositions are relative to test conditions, so are abilities.

Judgments of possibility (termed judgments of "serious possibility" in Levi 1980) are also relative. But the relativity is to the assumptions the inquiring agent takes for granted and hence judges with certainty to be true. To say that it is seriously possible that h, according to agent X at time t, is to say that h is consistent with X's full or settled beliefs at t.

Of course, whether X judges it seriously possible that h or not is true or false as the case may be. But X's judgment itself is neither true nor false any more than X's preferences, utility judgments, or judgments of credal probability are truth-valued. As a consequence, it makes no sense to say that X fully believes that it is seriously possible that h because the that-clause cannot represent a truth-value–bearing proposition. (One *can* say that X fully believes that it is seriously possible that h according to X.) On the other hand, it does make sense to say that X fully believes that Sam is able to take the trip to Boston on Monday at will. The claim that Sam has the ability is truth-value–bearing.

Although possibility for and possibility that are different notions, there is an important connection between them. Sometimes an agent X can ground or justify judgments of serious possibility in knowledge of ability.

Suppose then that X is certain that a coin has the ability to land heads on a toss. Suppose further that X is certain that coin will be tossed at time t. Do these assumptions warrant X's judging it to be possible (for all X knows) that the coin will land heads? This cannot be sufficient: X might also be certain that Morgenbesser will toss the coin at time t while being fully convinced that the coin lacks the ability to land heads on a toss by Morgenbesser (i.e., has the sure-fire disposition to fail to land heads on a toss). There is no inconsistency in believing that the coin has the ability to land heads on a toss, lacks the ability to land heads on a toss by Morgenbesser, and is tossed by Morgenbesser. If X has these convictions, it is incompatible with what he takes for granted that the coin will land heads. Of course, if in doubt as to whether the coin had the ability to land heads on a toss by Morgenbesser, X could judge it possible that it will land heads. But X could do that also if in doubt as to whether the coin had the ability to land heads on a toss. The important point is that X may know that a person or system has the ability to R on a trial of kind S and that a trial of kind S is being implemented. If X knows that the trial is also of kind S', then X may consistently continue to fully believe that the trial is of kind S, that the system has the ability to R on a trial of kind S and lack the ability to R on a trial of both kind S and kind S'. Under these circumstances, X should rule out as

© Cambridge University Press www.cambridge.org

CAMBRIDGE

Cambridge University Press
0521416744 - Knowledge, Belief, and Strategic Interaction
Edited by Cristina Bicchieri Maria Luisa Dalla Chiara
Excerpt
More information

impossible the hypothesis that the system will respond in manner R, even though X is convinced that it is possible for the system to do so.[4]

These remarks have an important bearing on judgments of feasibility and how to understand act descriptions.

3. Passing judgment and deciding

If X supposes that the proposition that Sam will travel to Boston in four hours on Monday represents an act or option for Sam, then X presupposes that Sam has the ability to choose that he (Sam) will take the trip as the outcome of deliberation and that this choice will be efficacious. Moreover, X also assumes that Sam has or will engage in a process of deliberation by the appropriate time.

In general, if X takes the sentence "Sam behaves in manner R" to be an act description vis-à-vis a decision problem faced by Sam, then X is in a state of full belief that has the following contents:

(i) Sam has the ability to choose that Sam will R on a trial of kind S, where the trial of kind S is a process of deliberation eventuating in choice.[5] Let us call this the *ability* condition.
(ii) Sam is subject to a trial of kind S at time t; that is, Sam is deliberating at t. This is the *deliberation* condition.

Sam's state of full belief should also satisfy the following requirement:

(iii) Adding the claim that Sam chooses that he will R to X's current body of full beliefs entails that Sam will R – that is, travel to Boston. This is the *efficaciousness* condition.[6]

I shall say that X is committed to judging that Sam's doing R is a feasible option for Sam, for the purposes of passing judgment on the rationality of Sam's choice, if and only if X's state of full belief satisfies conditions (i)–(iii).

Suppose then that X judges that Sam is able to travel to Boston from New York on Monday in under four hours by his own choice. Suppose further that X is convinced that Sam will not choose to take the trip. X may be certain that (1) Sam will not even engage in deliberation resulting in choice or that (2) Sam will engage in such deliberation but that the deliberation will result in a decision not to make the trip.

In the first case, the "test condition" relative to which Sam is supposed to have the ability will not be realized (so X judges). If X also rules out Sam's being transported to Boston against his will, then X will judge it impossible that Sam travels to Boston on Monday. In this case, it should

© Cambridge University Press www.cambridge.org

CAMBRIDGE

Cambridge University Press
0521416744 - Knowledge, Belief, and Strategic Interaction
Edited by Cristina Bicchieri Maria Luisa Dalla Chiara
Excerpt
More information

be clear that traveling to Boston is not optional for Sam, so far as X is concerned, in any decision problem Sam is facing.

In the second case, X is convinced that Sam is facing a decision problem. X is also certain that the choice made will be not to travel to Boston. Again, if X also rules out Sam's taking the trip against his will, it is not a serious possibility (according to X) that Sam will take the trip. All of this holds even though X is convinced that conditions (i)–(iii) obtain.

For the purposes of passing judgment on the rationality of Sam's choice, counting Sam's taking the trip as an act description may seem quite acceptable. Even if X supposes that Sam will choose irrationally – perhaps, in an akratic manner – this seems to make good sense. But as we shall see shortly, appearances can be deceiving.

However, suppose X is not concerned to pass judgment on the rationality of Sam's choice but is rather engaged in advising Sam as to what his choice should be. In such a context, criteria of rational choice are being used not to pass judgment on what Sam did, does, or will do, but instead to decide what Sam should do. This use of rational choice principles is especially relevant when X is Sam himself (although it is not necessary that X be identical with Sam).

In this context, the fact that X's state of full belief satisfies (i)–(iii) is necessary but not sufficient for X's judging it feasible that Sam will R. To see that it is not sufficient, consider the situation where X is certain that Sam will choose not to R and that Sam's choice will be efficacious. From X's point of view, it is no longer a serious possibility that Sam will R. Although X might deplore this as a mark of Sam's irrationality, this is relevant only insofar as X is passing judgment on the rationality of Sam's decisions. If X is merely giving advice, it is pointless to advise Sam to do something X is sure Sam will not do. In this setting, advice concerning what Sam should rationally do should take into account only those propositions judged to represent feasible options in a stricter sense than the one captured by conditions (i)–(iii); a fourth condition must be satisfied by X's state of full belief:

(iv) Nothing in X's state of full belief is incompatible with Sam's choosing to travel to Boston on Monday. This is the *serious possibility* condition.

When and only when X's state of full belief satisfies conditions (i)–(iv) does X judge that doing R is feasible for Sam for the purpose of advising Sam concerning what to do.

I contend that the primary prescriptive function of principles of rational choice is to furnish criteria for determining what a rational agent

© Cambridge University Press www.cambridge.org

CAMBRIDGE

Cambridge University Press
0521416744 - Knowledge, Belief, and Strategic Interaction
Edited by Cristina Bicchieri Maria Luisa Dalla Chiara
Excerpt
More information

ought to do, given the agent's goals and values. In particular, such criteria ought to be applicable when the deliberating agent is concerned with identifying optimal or admissible options from a given set of options judged to be feasible relative to the agent's state of full belief and values. (An admissible option is one that the agent is not forbidden to choose according to the principles of choice.)

From what has been said thus far, it appears that judgments of feasibility differ in the context of passing judgment on the rationality of choice versus the context of deciding what is to be done. But the difference can be exaggerated.

When passing judgment on the rationality of Sam's choice, there are two cases to consider. First, there is the question of whether Sam chose rationally relative to the information available to him. In this case, X should take the feasible options to be those that are judged to be feasible relative to Sam's state of full belief, when such judgments satisfy conditions (i)–(iv) applied to *Sam's* state of full belief. Even if X (who passes the judgment) is distinct from Sam, X should make the assessment on the assumption that Sam's state of full belief met these requirements prior to reaching a decision. X might actually disagree with Sam on some issue of feasibility; in particular, X can have views of his own as to what Sam will do. But when assessing the rationality of Sam's choice relative to the information available to him, feasibility should be as judged by Sam in the sense of meeting (iv) as well as the other requirements. The reason is that X is passing judgment on the rationality of Sam's judgment concerning what to do in the "thin" sense of rationality, a sense that covers minimal conditions of coherence and consistency of Sam's attitudes regardless of whether Sam's beliefs and values are sensible or accurate by X's lights or from any other point of view.

There is, to be sure, another approach to passing judgment on Sam's decision: X may evaluate the rationality of Sam's choice not relative to Sam's point of view prior to reaching a decision but relative to another body of information that (so X thinks) should have been available to Sam. Observe, however, that if the rationality of Sam's choice is at issue, then the body of information being used should meet the requirements for feasibility relevant to deciding what to do, so that – relative to that body of information – condition (iv) as well as conditions (i)–(iii) should be met.

To be sure, X might be in a state of belief of already knowing how Sam has chosen, so that condition (iv) cannot be satisfied for X's current judgments of feasibility. That is why I suggested that, in passing judgment on the rationality of Sam's choice, we may require only conditions (i)–(iii).

However, the importance of waiving condition (iv) can be exaggerated even in this context. If X judges that Sam's doing R represents a feasible

© Cambridge University Press www.cambridge.org

CAMBRIDGE

Cambridge University Press
0521416744 - Knowledge, Belief, and Strategic Interaction
Edited by Cristina Bicchieri Maria Luisa Dalla Chiara
Excerpt
More information

option for Sam in a given context where X is certain that Sam did not choose to do R, then X may still ask whether Sam should have chosen to do R relative to a system of beliefs like X's minus the information regarding what Sam did choose. By contracting X's state of full belief concerning Sam's decisions in a certain manner, the set of propositions judged to be feasible options relative to X's current state (and hence in the sense satisfying conditions (i)–(iii)) is converted into a set of propositions judged to be feasible relative to a contraction of X's current state meeting conditions (i)–(iv).

The point I mean to belabor is that passing judgment on the rationality of Sam's choices has little merit unless it gives advice as to how one should choose in predicaments similar to Sam's in relevant respects. In such situations, the evaluation of feasible options with respect to admissibility is undertaken on the assumption that conditions (i)–(iv) obtain.

Thus, even though prescriptions for rational choice can be used to pass a judgment on the rationality of a decision maker's choice, the fundamental type of application of such prescriptions is in determining what a rational agent ought to do. In particular, it becomes important to determine what a rational agent ought to do given the information available to that agent.

4. Foreknowledge of rationality

The aim of a prescriptive theory of rational choice is to provide criteria for identifying a set of options that are optimal or, at least, "admissible" in the sense that they are not ruled out by the principles of choice, given the agent's beliefs and values. If the agent is to reach a stage in deliberation where such principles can be used to identify a set of admissible options, the agent must first identify a set A of feasible options or option descriptions from which the set $C(A)$ of admissible options is selected. No proposition can be admissible unless it is an option description – that is, represents a feasible option. And insofar as the deliberating agent is deciding what he or she is to do, and not passing judgment on what the deliberating agent or someone else has done or will do, judgments of feasibility must meet conditions (i)–(iii) as well as the serious possibility condition (iv).

In order for the agent to apply the choice criteria in determining what is to be done, the agent must not only be in a position to judge what is feasible (in the sense already explained) but must make other judgments as well. In particular, the agent must know enough about his own values (goals, preferences, utilities) and beliefs (both full beliefs and probability judgments), and have enough logical omniscience and computational

© Cambridge University Press www.cambridge.org

CAMBRIDGE

Cambridge University Press
0521416744 - Knowledge, Belief, and Strategic Interaction
Edited by Cristina Bicchieri Maria Luisa Dalla Chiara
Excerpt
More information

capacity, to use his principles of choice to determine the set $C(A)$ of optimal or admissible options. That is to say, the deliberating agent's state of belief must meet a *self-knowledge* condition and a *logical omniscience* condition. The agent does not need perfect self-knowledge or perfect logical omniscience – just enough so that, having identified a set A of feasible options, the agent's principles of choice can be used to identify the admissible set $C(A)$.

Suppose the agent meets conditions (i)–(iv), as well as the self-knowledge and logical omniscience conditions. In that case, the agent is certain what the elements of the admissible set $C(A)$ are. That is to say, the deliberating agent has identified which of the feasible options ought not to be chosen and which he is entitled to choose as a rational agent with given beliefs and goals.

Observe that nothing in the agent's beliefs need imply that the agent will choose rationally – that is, restrict the choice to an admissible option. The deliberating agent can determine what he ought rationally to do without assuming that he will actually succeed in doing it.

But suppose that the agent is confident of his or her own rationality and satisfies the *smugness* condition. The smugness condition states that the agent is certain that in the deliberation taking place at time t, X will choose an admissible option.

Assuming that logical omniscience enables the agent to identify elementary logical consequences of adding the smugness condition to (i)–(iv), to self-knowledge, and to logical omniscience, the agent must be certain of not choosing an inadmissible option. From conditions (iii) and (iv), it follows that no inadmissible option is feasible from the deliberating agent's point of view when deciding what to do; $C(A) = A$.

Though this result is not contradictory, it implies the vacuousness of principles of rational choice for the purpose of deciding what to do. To be nonvacuous, these principles should offer criteria for reducing a set of feasible options to a set of admissible options that is a proper subset of the feasible options. Sometimes the reduction fails and the two sets will coincide. If the reduction always fails, then the principles of rational choice are useless as criteria that deliberating agents might use to determine what they ought to do. If they are useless for this purpose then, by the argument of the previous section, they are useless for passing judgment on the rationality of choice as well. If principles of rational choice have (or are intended to have) a prescriptive application, we should reconsider the assumptions that lead to this result.

Notice that the argument invoked does not depend on endorsing any particular system of principles of rational choice. One can favor maximizing

© Cambridge University Press www.cambridge.org