Choice axioms for a positive value of information

Jean-Marc Tallon[*], Jean-Christophe Vergnaud[†]

September 2002

1 Introduction

Debates around the assumption of rationality in economic theory abound and are too numerous to be listed here. Traditionally, rationality of a decision maker is captured by a set of (testable) axioms, whose meaning and strength can be discussed. These rationality axioms however do not all have the same status in the sense that they bear on different things. One could actually distinguish three categories of axioms or rationality principles. The first one bears on preferences: a decision maker's preferences are rational if they satisfy a set of axioms (transitivity and so forth). The set of axioms that is imposed might depend on the context. For instance, in the case of decision making under uncertainty, rationality axioms often include the sure thing principle (Savage (1954)). The second category of axioms is the one that bears on choices. In this view the decision maker's choices or behavior is judged to be rational or not whether it satisfies axioms such as the weak axiom of revealed preferences. The methodology of revealed preferences, which assesses that choice behavior reveals the underlying preferences ties in these two notions of rationality (Sen (1971)), making it possible to relate axioms on choice behavior to axioms on preferences. The third category is the one that assesses the rationality of the consequences of the decision maker's choices. We will name it "economic" rationality. It is based on a few principles such as, for instance, the fact that the decision maker should not be susceptible to be exploited through a money pump (that is, a dynamic process through which the decision maker loses money for sure, after a series of desirable exchanges). Another principle along this line, which will be analyzed in this paper is the fact that a decision maker should never refuse information before making a choice. Indeed, it can be shown that if this were the case, a decision maker would be willing to pay a positive amount to avoid getting informed. This behavior could then be exploited, implying that the decision maker could be made worse off in in similar fashion as in the case of a money pump.

In this paper, we look at the relationship between these three notions of rationality (namely, rationality of preferences, rationality of choices, and economic rationality) in

[*]CNRS–EUREQua, jmtallon@univ-paris1.fr
[†]CNRS–EUREQua, vergnaud@univ-paris1.fr

a dynamic setting. More precisely, we focus our attention on the link between axioms belonging to the three categories identified above: First, axioms (on choices) introduced by Hammond (1988), namely *consequentialism* and *separability*; second, the sure thing principle, which is an axiom on preferences that is at the foundation of Bayesian decision theory; third, the positive value of information, which, as we just saw, is a principle on the economic rationality of the decision maker.

Hammond (1988) showed that consequentialism and separability implied the sure thing principle. His analysis has been interpreted as giving a strong support to the Bayesian decision theory, relating the sure thing principle to axioms on dynamic choices. It is easy to see[1] that Hammond's axioms also imply that a decision maker will never refuse getting (free) information before taking a decision. Thus, Hammond's axiom on choices have the property to imply both rationality axioms on preferences and principles of economic rationality.

In this paper, we relax consequentialism in a way that, arguably, maintains rationality of dynamic choices and show that this leads to a weakening of the sure thing principle while still implying that information has a positive value for the decision maker. The setting adopted to treat these issues is the natural approach to model dynamic choices *via* decision trees. We simplify the analysis by considering only decisions that are of a very simple kind, namely, we consider only bets.

Our result can be interpreted in the light of recent advances in decision theory under uncertainty, motivated by Ellsberg (1961)'s experiments that a majority of people do not behave according to the expected utility model when the situation they face does not admit a straightforward probabilistic representation. Ellsberg identified the axiom of expected utility theory that people were most likely to violate, namely, the sure thing principle. This observation has led to a host of decision models under uncertainty in which the sure thing principle is weakened in one way or another (e.g., Yaari (1987); Quiggin (1982); Schmeidler (1989); Gilboa and Schmeidler (1989)). Descriptively, these models seem to fair better than the traditional expected utility model. However, it was also acknowledged that the non bayesian decision models were subject to difficulties in dynamic settings, leading for instance to a negative value for information. The latter is quite problematic concerning the normative aspect of these models. Thus, our result establishes that there is some room for models that are based on rationality of dynamic choices (although in a slightly weaker form than the one implied by Hammond's axioms), that satisfy the economic rationality principle of a positive value of information, while relaxing the sure thing principle. Put differently, Hammond's contribution has been seen as the justification *via* dynamic choice axioms of Bayesian decision theory under uncertainty, namely the expected utility model. It is our claim in this paper that such a foundation for the expected utility model is not warranted and that Hammond's axioms can be relaxed in a sensible way so as to be compatible with other models of decision under uncertainty.

The chapter is constructed as follows. Section 2 introduces the setting in which we

[1] Although Hammond did not consider this issue.

cast our argument. Section 3 contains an exposition of Hammond's argument, while in section 4 we introduce a weakening of Hammond's consequentialism and show our main result. Section 5 illustrates the result through an example based on possibility measures (a qualitative tool for representing uncertainty).

2 Decision trees

As mentioned in the introduction, one needs to introduce a language in which dynamic choices can be modelled and interpreted. This language is that of decision trees. A decision tree is a description of all available sequences of choice together with a representation of uncertainty the decision maker faces.

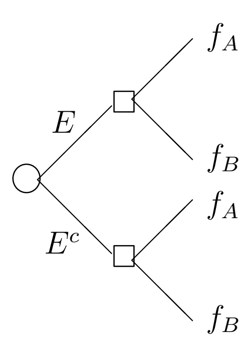
We'll cast our argument in a setting in which agents' choices are in terms of bets, which represent a particularly simple form of decision. More formally, let S be a finite set of states of nature, representing all the uncertainty there is in the model. We consider bets, which are simply variables taking either the value 0 or 1. A bet on event $A \subset S$ is denoted f_A and is given by:

$$\forall s \in S, \quad f_A(s) = \begin{cases} 1 & \text{if } s \in A, \\ 0 & \text{otherwise.} \end{cases}$$

A decision tree has two types of nodes: chance nodes (represented by circles), at which nature chooses an event; and decision nodes (represented by squares), at which the decision maker decides of an action (here a bet).

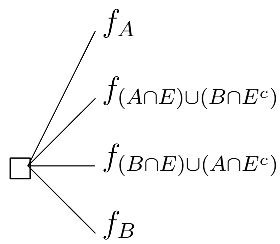
Example 1 *The following decision tree (figure 1) represents the situation in which the decision maker receives an information (whether event E occurred or not) before betting on event A or on event B):*

Figure 1: A decision tree



One can also define the equivalent strategic form, which represents the dynamic decision tree through a one shot decision tree in which the strategy space has been enlarged. For

Figure 2: Equivalent strategic form



instance, the equivalent strategic form of the decision tree of figure 1 is shown on figure 2.

The relationship between the choice of a decision maker in the decision tree and in the equivalent strategic form will be studied momentarily.

3 Positive value of information, consequentialism and the sure thing principle

Hammond (1988) based his analysis on two axioms on dynamic choices. These two axioms are:

- separability: the decision maker behavior depends only on future consequences.
- consequentialism: the decision maker's choices in two equivalent trees are identical.

For instance, if the decision maker chooses to bet on A if E occurs and on B if not in the decision tree of figure 1, then, under consequentialism he will choose $f_{(A\cap E)\cup(B\cap E^c)}$ in the strategic form of figure 2. These two axioms imply both that information always has a positive value and that the sure thing principle holds.

3.1 Positive value of information deduced

The issue of the value of information can be exemplified on the tree of figure 3. If the decision maker chooses to go up at the beginning of the tree, this means he will first learn whether E or E^c occurred and then make a choice between betting on A and betting on B. On the other hand, if he chooses to go down, this means that he will bet on A or on B without having any information on the (related) event E.

Observe that the two branches of the tree of figure 3 are equivalent to the two trees in figure 4. In the left part of the figure, the simple choice between betting on A or betting on B without any information is shown. In the right part, we represent the equivalent strategic form of a tree in which the decision maker has the possibility to get informed, and more specifically to learn whether E occurred or not.

Figure 3: Choosing to get informed

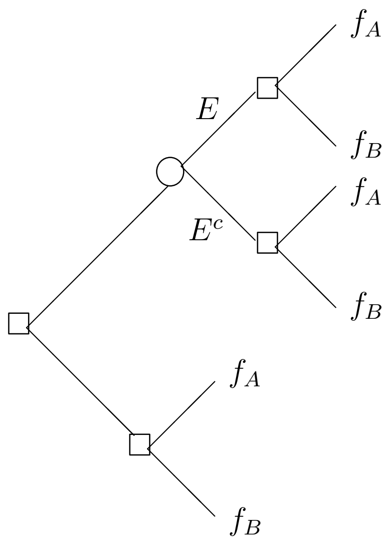
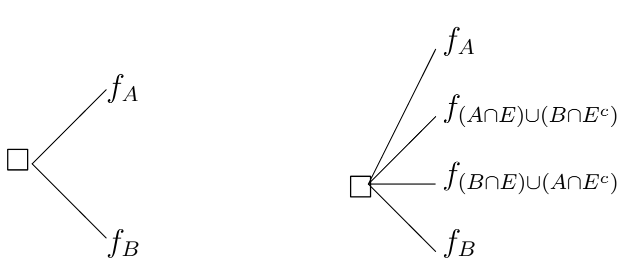


Figure 4:



The obvious advantage of getting this information is that the decision maker can condition his choice on the realization of E. In other words, the decision maker will always prefer to act in the second tree in which the available strategies include the one of the right hand side tree. Since the choices in the equivalent strategic form are the same as in the original decision tree, this means that the decision maker will always prefer to obtain some information before choosing to bet on A or B, i.e., will always choose to go up in the tree of figure 3.

Hence, by reducing the decision tree in which the decision maker chooses or not to be informed to the equivalent strategic form of figure 4, it is easy to show that a decision maker will always prefer to get the information. Thus, the principles that allowed us to indeed state that the equivalent strategic form could be used to analyze the decision maker's choice in the full blown tree entail directly that the decision maker attaches a positive value to getting informed. But these principles are precisely the axioms of separability and consequentialism. To conclude, let us summarize the nature of the argument we just made. In static choices, getting informed amounts to increase the space of available strategies. Now, consequentialism and separability ensures that the decision maker will be able to follow his intention in a dynamic decision tree whose reduced form would precisely amounts to a choice between two static trees, one having less strategy than the other. Thus, he is able to exploit the information (if he judges that it is relevant to choice to be made).

3.2 The sure thing principle deduced

Arguably, the strongest axiom that Bayesian theory imposes on preferences is the sure thing principle. This axiom is violated by preferences underlying choices as in Ellsberg's experiment, while these choices do not appear irrational[2]. Hammond's analysis however gives support to the sure thing principle by showing that it can be deduced from the two axioms on dynamic choices he introduced (consequentialism and separability)

In the simple setting we adopted, we can work with a simple version of the sure thing principle:

Axiom 1 *(sure thing principle) For all $A, B, C \subset S$ such that $(A \cup B) \cap C = \emptyset$,*

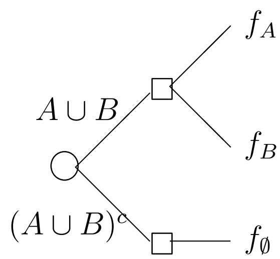
$$f_A \succeq f_B \Rightarrow f_{A \cup C} \succeq f_{B \cup C}$$

The proof that separability plus consequentialism imply the sure thing principle can be seen by representing the decision maker's choice in different ways.

Consider the choice between, betting on A (f_A) and betting on B (f_B), and suppose the decision maker prefers f_A to f_B. This choice can be represented by a simple decision tree in which there is only one decision node, the two options being f_A and f_B. Now, by consequentialism, the decision tree is equivalent to the tree in figure 5 (where $f_\emptyset$ is the bet "lose for sure", i.e., the bet yielding 0 whatever the state of nature):

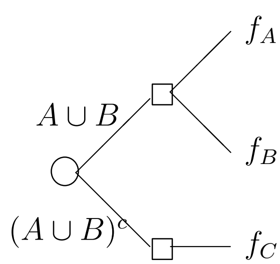
[2]Rather, they are consistent with the idea that agents are ambiguity averse.

Figure 5:



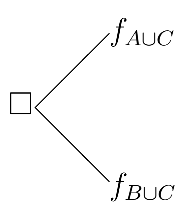
Consider now an event C disjoint from $(A \cup B)$. The choice between f_A and f_B in the tree in figure 5 is the same, by separability, to the one in the tree in figure 6, since what happens outside $(A \cup B)$ is immaterial for the decision maker's choice between f_A and f_B.

Figure 6:



The last step consists in writing the equivalent strategic form of the tree of figure 6, which is represented on figure 7.

Figure 7:



Since the decision maker was assumed to choose f_A in the first tree, it comes, following the chain of equivalent trees described, that he will choose $f_{A \cup C}$ in the last of these trees. This is nothing but the expression, in the language of decision trees, of the sure thing principle.

Hence, Hammond's argument amounts to deduce the sure thing principle, on the basis of axioms on dynamic choices, i.e., having to do with the way a decision maker should analyze a decision tree.

4 A weaker axiom on dynamic choices for a positive value of information

We now show that the full strength of consequentialism is not necessary to obtain a positive value of information. More precisely, we replace Hammond's consequentialism by the weaker principle, that the revealed choices in a decision tree are subset of the optimal choices in the equivalent strategic form. Call this principle *selection of optimal strategies*, which can be illustrated as follows: if the revealed optimal choice in the decision tree represented in figure 1 is f_A if E and f_B is E^c, then, it must be the case that the strategy $f_{(A\cap E)\cup(B\cap E^c)}$ is an optimal strategy in the equivalent strategic form represented on figure 2. But whereas *consequentialism* imposes that it is the only optimal strategy (assuming that the described optimal strategy in the decision tree is unique), *selection of optimal strategies* allows that there is some other optimal (and hence equivalent) strategy in the strategic form, such as, for instance f_A. Thus, the axiom we introduce can arguably be seen as a rationality axiom, since it states that the decision maker will always implement an optimal strategy.

Now, it can be shown that *separability* together with *selection of optimal strategies* imply that the decision maker will never assess a negative value for information (see Vergnaud (2002)). The intuitive reason is similar to the one we gave in the previous section: the optimal choice in the full decision tree is always an optimal choice in the equivalent strategic form; furthermore when comparing strategic forms the decision maker always prefers to get information (since it enlarges his set of strategies); hence, it must be the case that the decision maker would rather be in the decision tree in which he gets the information than in the decision tree without any possibility to get informed.

We now look at the implication on preferences of relaxing *consequentialism.* More specifically, we establish that *separability* together with *selection of optimal strategies* imply a weaker axiom than the sure thing principle, that we dub the *weak sure thing principle* and that goes as follows:

Axiom 2 *(weak sure thing principle) For all* $A, B, C \subset S$ *such that* $(A \cup B) \cap C = \emptyset$*,*

$$f_{A\cup C} \succ f_{B\cup C} \Rightarrow [\forall D \text{ s.th. } (A \cup B) \cap D = \emptyset,\ f_{A\cup D} \succeq f_{B\cup D}]$$

The proof goes as follows.

The two trees represented on figure 8 are equivalent. Assume that $f_{A\cup C} \succ f_{B\cup C}$ in the first tree. Then, by *selection of optimal strategies*, $f_{A\cup C}$ is chosen over $f_{B\cup C}$ in the upper part of the second tree.

Now, by separability, the second tree of figure 8 is equivalent to the first tree represented on figure 9. Hence, what happens outside of $(A\cup B)$ is immaterial for the choice in the upper branch, and, conditionally on $(A\cup B)$, $f_{A\cup D} = f_A$ is thus chosen over $f_{B\cup D} = f_B$.

Figure 8:

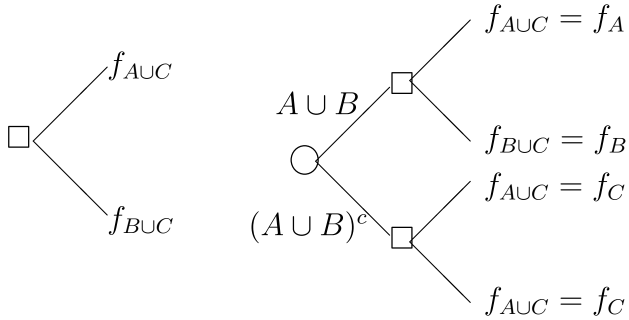
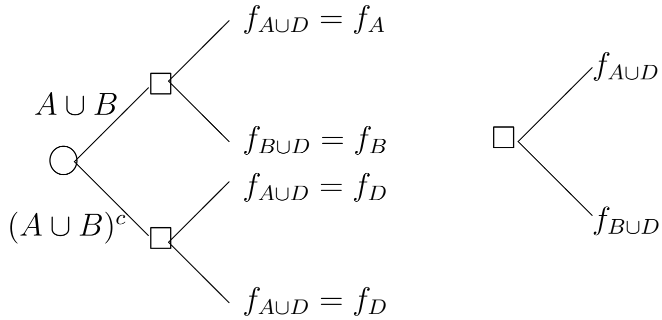


Figure 9:



Since $f_{A\cup D}$ is an optimal choice in the first tree of figure 9, it must be the case, by *selection of optimal strategies* that $f_{A\cup D}$ is an optimal choice (not necessarily the only one) in the second tree, proving the weak sure thing principle.

5 Positive value of information without probabilistic beliefs

We illustrate in this section the claim that non probabilistic beliefs can be compatible with a positive value of information. Observe that given that we work with bets, preferences directly reveal the decision maker's beliefs. We can therefore content ourselves to exhibit an example in which only beliefs are modelled explicitly.

The example is based on the concept of possibility measures. This concept has been introduced in artificial intelligence by Zadeh (1978) (see also Dubois and Prade (1985)). It is in particular useful to represent situations in which the information is qualitative as opposed to probabilities that are by nature quantitative and can be seen as the formal representation of ideas introduced in economics by Shackle (1952).

Definition 1 *A possibility measure is a set function* $\Pi : 2^{\Omega} \to [0,1]$*, such that* $\Pi(\emptyset) = 0$*,* $\Pi(\Omega) = 1$ *and*

$$\forall A, B \subseteq \Omega, \quad \Pi(A \cup B) = Max(\Pi(A), \Pi(B))$$

This definition reads as follows: the union of two events is "as possible as" the most possible of the two events. Thus, possibility measures do not have the additivity property of probability measures.

Now, suppose there exists a possibility measure Π such that the preference relation on bets is such that

$$\forall A, B \subseteq \Omega, f_A \succeq f_B \Leftrightarrow \Pi(A) \geq \Pi(B)$$

That is, a bet is preferred to another bet if the winning event is more possible than the winning event of the second bet. One can check that these preferences satisfy the weak sure thing principle (axiom 2):

Let $A, B, C, D \subseteq \Omega$ such that $(A \cup B) \cap C = (A \cup B) \cap D = \emptyset$. Suppose $f_{A\cup C} \succ f_{B\cup C}$. Thus $\Pi(A\cup C) > \Pi(B\cup C)$ and therefore $Max(\Pi(A), \Pi(C)) > Max(\Pi(B), \Pi(C))$. Then necessarily $\Pi(A) > \Pi(B)$. Hence $Max(\Pi(A), \Pi(D)) \geq Max(\Pi(B), \Pi(D))$ which proves that $f_{A\cup D} \succeq f_{B\cup D}$.

We establish now on an example that such a preference relation on bets built on a possibility measure can satisfy the *selection of optimal strategies* principle.

Consider $A, B, C \subseteq \Omega$ three events such that $\Pi(B) = .5 < \Pi(A) = .7 < \Pi(C) = .8$ and go back to figure 8. Since $f_{A\cup C} \sim f_{B\cup C}$, according to the first tree there are several optimal strategies in the second tree. Take for instance, the two following strategies are optimal strategies:

- choose $f_{A\cup C}$ whatever the information received,
- choose $f_{B\cup C}$ whatever the information received.

Observe however that this second strategy will not be implemented in the second tree since in the upper part of the second tree, the agent will strictly prefer $f_{A\cup C}$. In order to see this, we need to specify the way the decision maker revises his beliefs once he learned that $A \cup B$ occurred. A natural updating rule (see Dubois and Prade (1985)) in this context is simply Bayes' rule, applied to the possibility measure Π. Notice that $\Pi(A \cup B) = .7$ and hence $\Pi(A|A \cup B) = .7/.7 = 1$ while $\Pi(B|A \cup B) = .5/.7 < 1$. Therefore, conditionally on the fact that $A \cup B$ occurred, the decision maker strictly prefers to bet on A rather than to bet on B.

Hence, to conclude on this example, the strategies derived from the study of the decision tree are a subset of the "optimal" strategies (computed on the equivalent strategic form). Thus, the dynamic choice selects a subset of the optimal strategies. This behavior is consistent with the weak sure thing principle and hence with the basic requirement of a positive value of information and hence cannot be ruled out on the basis that is not "rational".

6 Concluding remarks

We established that if *selection of optimal strategies* and *separability* are satisfied then the decision maker always prefers to be informed, and on the other hand, that his behavior is not necessarily in accordance with the sure thing principle. It might seem that the weakening of the sure thing principle that we established does not differ that much from the standard sure thing principle. Hence, it is not clear that the type of behavior allowed by this weakening of the axiom are of great interest. However, there is an entire family of models that violate the sure thing principle while satisfying its weak form. These decision models are based on a rather qualitative description of the uncertainty, relying on the use of specific capacities named possibility measures (and their dual, necessity measures).

Finally, we should mention that the weakening of consequentialism that we proposed in this chapter entails in some sense a minimal departure from Hammond's argument. A completely different approach has been followed by Machina (1989) and McClennen (1990). These authors assume that the decision maker is able to commit to dynamic choices, a behavior they label *resolute choice*. Although Machina's and McClennen's *resolute choice* are rather different from one another, they both build in the idea that the decision maker will always be able to follow his original intentions. Hence, consequentialism is abandoned and replaced by a behavioral assumption of a different nature. In this approach, information always has a positive value as well.

References

Dubois, D., and H. Prade (1985): *Théorie des possibilités*. Masson, Paris.

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

Gilboa, I., and D. Schmeidler (1989): "Maxmin expected utility with a non-unique prior," *Journal of Mathematical Economics*, 18, 141–153.

Hammond, P. (1988): "Consequentialist foundations for expected utility," *Theory and Decision*, 25, 25–78.

Machina, M. (1989): "Dynamic consistency and non-expected utility models of choice under uncertainty," *Journal of Economic Literature*, 28, 1622–1668.

McClennen, E. (1990): *Rationality and Dynamic Choice: Foundational Explorations*. Cambridge University Press.

Quiggin, J. (1982): "A theory of anticipated utility," *Journal of Economic Behavior and Organization*, 3, 323–343.

Savage, L. (1954): *The foundations of statistics*. New-York, John Wiley.

Schmeidler, D. (1989): "Subjective probability and expected utility without additivity," *Econometrica*, 57(3), 571–587.

Sen, A. (1971): "Choice functions and revealed preferences," *Review of Economic Studies*, 38, 307–317.

Shackle, G. (1952): *Expectation in economics*. Cambridge University Press.

Vergnaud, J.-C. (2002): "Information and capacities," *Statistical Papers*, 43, 111–125.

Yaari, M. (1987): "The dual theory of choice under risk," *Econometrica*, 55(1), 95–115.

Zadeh, L. (1978): "Fuzzy sets as a basis for a theory of possibility," *Fuzzy sets and systems*, 1, 3–28.