Discussion Paper No. 835

THE VALUE OF INFORMATION--AN AXIOMATIC APPROACH*

by

Itzhak Gilboa**

and

Ehud Lehrer**

June 1989

---

*We are grateful to Ehud Kalai, Morton Kamien, Roger Myerson, Dov Samet, and especially Dov Monderer for helpful discussions. The first author gratefully acknowledges NSF Grant No. IRI-8814672, and the use of office facilities at the Cowles Foundation, Yale University.

**Department of Managerial Economics and Decision Sciences, J.L. Kellogg Graduate School of Management, Northwestern University, Evanston, Illinois 60208

Abstract

The following question is addressed: Given a real-valued function on the partitions of a measure space, what are necessary and sufficient conditions in order that it be the value of information function for a Bayesian decision maker?

A characterization is provided, and the analysis reveals a close relationship to cooperative game theory. The tools developed for characterization can also be used to derive somewhat surprising results regarding the properties of information functions.

1. Introduction

A Bayesian decision maker is assumed to maximize expected utility à la Savage (1954). Considering situations in which some information may be known to him/her prior to the actual decision, the decision maker is allowed to condition the choice of an action on this information. If we follow Aumann (1974) and model information by partitions, each partition is associated with the maximal expected payoff that can be achieved given that partition. The difference of this function between two partitions, one being a refinement of the other, is the value of the additional information to the decision maker (already having the coarser partition). (See, for instance, Hirshleifer and Riley (1979).) Thus, if we identify utility with money, for instance, this difference would be the maximal price the decision maker will be willing to pay for the additional information.

The main goal of this paper is to characterize those real-valued functions on the set of partitions that may be the information function of some Bayesian decision maker.

A fundamental tool in the analysis are expressions of the type

$$(*) \qquad [f(P) - f(P \wedge Q)] - [f(P \vee Q) - f(Q)]$$

where P and Q are partitions, $P \wedge Q$ and $P \vee Q$ are their finest coarsening and their coarsest refinement, respectively, and f is the partition function under consideration. The first brackets represent the value of P for a decision maker knowing $P \wedge Q$, while the second is the value of P if the decision maker already knows Q. (By "knowing" a partition we simply mean that this partition represents what would be the decision maker's knowledge at each state of the world.)

Thus the sign of the expression above would determine whether an information trader--a consultant, for instance--should prefer selling his/her knowledge P before or after the decision maker purchases Q, while we only consider decision makers who know $P \wedge Q$.

It turns out to be the case that f is additively separable, namely, that it is representable as

$$f(P) = \Sigma_{A \in P} \, v(A)$$

for some set function v if and only if the expression (*) vanishes whenever P and Q satisfy a simple requirement that we call "non-intersection": P and Q are non-intersecting if there is an event A measurable with respect to both P and Q, such that on A, P is finer than Q, and on A^c, Q is finer than P.

While this condition may be viewed as a strictly mathematical requirement, it can also be motivated as a reasonable axiom on partition functions, and may be interpreted as a partition-version of Savage's Sure-Thing principle.

However, this condition, which is obviously necessary for f to be an information function, is not sufficient, not even in conjunction with monotonicity, i.e., not even if f is nondecreasing with respect to refinements. Omitting some details, the characterization may be described as follows: f is an information function if and only if it is additively separable and there is a v corresponding to it, every restriction of which (to a measurable subset) has a nonempty anticore.

The analysis shows, for instance, that if the prior is nonatomic and (*) always vanishes, i.e., if every two information traders are indifferent to the order of the transactions--then f has to be constant, which means that the value of all information partitions is zero.

This result, which we found somewhat surprising, shows that the tools developed here for the characterization problem may have other applications as well and, more specifically, that the close relationship between information functions and cooperative game theory (with the states of the world playing the roles of players) may be worth further exploration.

The remainder of the paper is organized as follows: Section 2 describes the framework and definitions. In Section 3 we characterize additive separability, while the characterization of information functions is given in Section 4. Section 5 contains the result mentioned above regarding the uselessness of commutative consultants, and some comments on possible extensions are to be found in Section 6.

2. Notations and Definitions

Let there be given a measure space $(\Omega, \mathcal{B}, \mu)$ where Ω is the set of states of the world, $\mathcal{B}$ is the σ-algebra of events and μ is the decision maker's (DM) prior probability measure. Following Aumann (1974), the DM's information is a priori modeled by a partition P of Ω, i.e., a finite set of pairwise disjoint elements of $\mathcal{B}$ the union of which is Ω. For $\omega \in \Omega$, $P(\omega)$ denotes the element of P containing ω, and we interpret P as follows: should ω obtain, the minimal event known to the DM would be $P(\omega)$.

Let A be a set of actions. A strategy is an element of $S = \{s: \Omega \to A\}$. Assuming the DM has the information partition P, we would require that his/her strategy would be measurable w.r.t. (with respect to) P, while A is endowed with the σ-algebra 2^A. Namely, the DM is not allowed to condition his/her choice of action on knowledge he/she may not possess. Given a partition P, the P-measurable strategies will be called P-strategies and denoted as a set by S_P.

Using Savage (1954) as a conceptual basis, we assume that there is a utility $u: A \times \Omega \to \mathbb{R}$ such that the DM's behavior is equivalent to maximization of $\int_\Omega u(a)d\mu$ over $a \in A$ (where $u(a): \Omega \to \mathbb{R}$ is defined by $u(a)(\omega) = u(a,\omega)$.) (To be precise, we should start out with a measurable space $(\Omega,\mathcal{B})$ and introduce μ only at this point. The current formulation will, however, facilitate presentation. See also Remark 6.2 in the sequel.) Savage's axioms also imply that u is bounded (see Fishburn (1970) for a proof and a bibliographical note) and w.l.o.g. we shall assume that $0 \le u \le M$. Although Savage's theorem states that μ is nonatomic, we will adopt all definitions to other cases as well when the need arises, and most notably to the case of a finite Ω. We also deviate from Savage's framework by allowing a general σ-algebra $\mathcal{B}$ (rather than the special case of $\mathcal{B} = 2^\Omega$) and requiring μ to be σ-additive (as opposed to finitely additive). (See Chateauneuf (1985).) Finally, note that our framework is general enough to deal with state-dependent utility.

For expository reasons we would like to impose a restriction that would allow us to deal with maxima--instead of suprema--as a first stage. Later on we will relax this condition, and obtain basically the same results. However, we were not very successful at finding general conditions (on A and u) that would guarantee this property and could still be considered an exposition simplification. We therefore prefer to explicitly assume that

(A,u) satisfy the following condition w.r.t. $(\Omega, \mathcal{B}, \mu)$: for every $B \in \mathcal{B}$ there exists $\text{Max}_{a \in A} \int_B u(a) d\mu$. This condition will be referred to as the <u>maximality property</u>.

Given a partition P we can compute the DM's expected utility:

$$e_{(A,u)}(P) = \max\{\int_\Omega u(s) d\mu | s \in S_p\},$$

which is well defined since (A,u) satisfies the maximality property and P is finite.

The set of all partitions is partially ordered by the "finer than" relation: P is <u>finer than</u> Q, denoted $P \geq Q$, if for every $A \in P$ there is $B \in Q$ such that $A \subseteq B$. Note that this relation is antisymmetric, i.e., $P \leq Q$ and $Q \leq P$ imply $P = Q$. The notation "$P \geq Q$" will be equivalent to "$Q \leq P$," to saying that "Q is coarser than P," that "P is a refinement of Q," or that "Q is a coarsening of P."

The set of partitions is a lattice: for every two partitions P and Q there exists a unique finest partition that is coarser than both, called their <u>join</u> and denoted $P \wedge Q$, and a unique coarsest partition that is finer than both, called their <u>meet</u>, and denoted by $P \vee Q$.

Two partitions, P and Q, are said to be <u>non-intersecting</u>, denoted $P \circ Q$, if the following holds: for every $A \in P$, either (i) there is $B \in Q$ such that $A \subseteq B$; or (ii) there are $\{B_i\}_{i=1}^n \subseteq Q$ such that $A = \cup_{i=1}^n B_i$. Note that o is a symmetric relation.

All the definitions regarding partitions will be extended in the natural way to elements of $\mathcal{B}$ other than Ω. Furthermore, if P^A is a partition of $A \in \mathcal{B}$, P^B is a partition of $B \in \mathcal{B}$ and $A \cap B = \emptyset$, we will use

$P^A \cup P^B$ as the obvious partition of $A \cup B$.

We note without proof that:

Observation 2.1: Two partition of Ω, P and Q, are non-intersecting iff there is an event $A \in \mathcal{B}$ such that $P = P^A \cup P^{A^c}$, $Q = Q^A \cup Q^{A^c}$ (where $P^{A(A^c)}, Q^{A(A^c)}$ are partitions of $A(A^c)$) with $P^A \le Q^A$ and $P^{A^c} \ge Q^{A^c}$.

We will denote by $\mathcal{P}^A$ the set of all partitions of $A \in \mathcal{B}$, where an omission of the superscript A will generally mean $A = \Omega$ (or an oversight). A partition function $f: \mathcal{P} \to \mathbb{R}$ is called an information function if there are (A,u) satisfying the maximality property w.r.t. $(\Omega,\mathcal{B},\mu)$ such that $f = e_{(A,u)}$. We can finally formulate the main question this paper deals with: Given a space $(\Omega,\mathcal{B},\mu)$ and a partition function f, what are necessary and sufficient conditions on f to be an information function?

3. Additive Separability

In this section we will do some first steps towards the characterization of information functions. First, we note that an information function $e_{(A,u)}(\cdot)$ is monotonic w.r.t. the "finer than" relation, hence monotonicity is surely a necessary condition on a partition function f to be an information function.

However, monotonicity is not a sufficient condition as is shown by:

Example 3.1: Let $\Omega = \{1,2,3,4\}$ and

$P_1 = \{\{1\},\{2\},\{3,4\}\}$ $\qquad$ $P_2 = \{\{1,2\},\{3\},\{4\}\}$

$P_3 = \{\{1\},\{2\},\{3\},\{4\}\}$ $\qquad$ $P_4 = \{\{1,2\},\{3,4\}\}$.

Define $f(P_1) = f(P_2) = f(P_3) = 1$, $f(P_4) = 0$, and extend f to all $\mathcal{P}$ in a monotonic way. Suppose $f = e_{(A,u)}$ for some (A,u) satisfying the maximality property, and define for $C \in \mathcal{B}$

$$v(C) = \mathrm{Max}_{a \in A} \int_C u(a)d\mu$$

so that $f(P) = \Sigma_{C \in P} v(C)$ for all $P \in \mathcal{P}$.

It is easy to check that

$$\begin{aligned} f(P_1) + f(P_2) &= v(\{1\}) + v(\{2\}) + v(\{3,4\}) \\ &\quad + v(\{1,2\}) + v(\{3\}) + v(\{4\}) \\ &= f(P_3) + f(P_4) \end{aligned}$$

where, in fact, $f(P_1) + f(P_2) = 2 \neq 1 = f(P_3) + f(P_4)$. //

The above example suggests the following definition: a partition function is <u>additively separable</u> if there exists a set function $v: \mathcal{B} \to \mathbb{R}$ such that $f(P) = \Sigma_{C \in P} v(C)$. In this case we say that v <u>corresponds</u> to f. In the sequel we will use the term "set function" for $v: \mathcal{B} \to \mathbb{R}$ with $v(\emptyset) = 0$. It is obvious that any information function $e_{(A,u)}$ is additively separable (with v defined as in 3.1).

However, the example also suggests another condition that can be formulated in terms of partitions: note that $P_1 \vee P_2 = P_3$, $P_1 \wedge P_2 = P_4$ and that P_1 and P_2 are non-intersecting. We therefore introduce a new definition: a partition function f is <u>partially commutative</u> if, whenever P and Q are non-intersecting, $f(P) + f(Q) = f(P \wedge Q) + f(P \vee Q)$.

Before stating the main result of this section, let us briefly motivate the choice of this term. Rearranging the terms in the equation above, we obtain

$$f(P \vee Q) - f(P) = f(Q) - f(P \wedge Q).$$

On the left side we have the value of the information partition Q for a DM who already has P. On the right side we find the value of Q for a DM who only knows $P \wedge Q$ (what is common knowledge between an agent with the partition P and one with Q). Hence, if f satisfies this equation for all $P,Q \in \mathcal{P}$, it is commutative in the following sense: the value of the information Q is the same, regardless of whether Q is acquired before or after P was acquired. Such an f will be called <u>commutative</u>. Since we require this property only for non-intersecting P and Q, we choose the name "partially commutative." As we will see in the sequel, (full) commutativity is much too strong a property that means, under fairly general conditions, that all the partitions are worthless.

We can finally state:

<u>Theorem 3.2</u>: A partition function f is additively separable iff it is partially commutative.

<u>Proof</u>: The "only if" part is quite straightforward. Assume $f(P) = \sum_{C \in P} v(C)$ for all $P \in \mathcal{P}$, and let P and Q be non-intersecting. There are events $\{A_{ij} | 1 \le i \le n,\ 1 \le j \le k_i\}$ and a set $I \subseteq \{1,\ldots,n\}$ such that:

(i) $P \vee Q = \{A_{ij} | 1 \le i \le n,\ 1 \le j \le k_i\}$

(ii) $P \wedge Q = \{\cup_j A_{ij} | 1 \le i \le n\}$

(iii) $P = \{A_{ij} | i \in I,\ 1 \le j \le k_i\} \cup \{\cup_j A_{ij} | i \notin I\}$

(iv) $Q = \{A_{ij} | i \notin I,\ 1 \le j \le k_i\} \cup \{\cup_j A_{ij} | i \in I\}$.

Hence, we obtain

$$f(P) + f(Q) = \Sigma_{i,j}\ v(A_{ij}) + \Sigma_i\ v(\cup_j A_{ij}) = f(P \wedge Q) + f(P \vee Q).$$

To prove the converse, assume that f is partially commutative. We will show that it is additively separable in two steps: first, by assuming $\mathcal{B}$ is finite, and then in the general case.

<u>Step 1</u>: A finite $\mathcal{B}$.

We first introduce the following notations: for $A \in \mathcal{B}$, two partitions of it, $P_1, P_2 \in \mathcal{P}^A$ and a partition of its complement, $Q \in \mathcal{P}^{A^c}$, such that $P_1 \le P_2$, let $d(P_1,P_2,Q) = f(P_1 \cup Q) - f(P_2 \cup Q)$.

For $B \in \mathcal{B}$ let P_c^B be the coarsest partition of B (which is always well defined) and let P_f^B be its finest partition (which is well defined, $\mathcal{B}$ being finite).

<u>Claim</u>: The function d is independent of its third argument.

<u>Proof</u>: Given $Q_1, Q_2 \in \mathcal{P}^{A^c}$ we wish to show that $d(P_1,P_2,Q_1) = d(P_1,P_2,Q_2)$.

It would suffice to prove this for $Q_2 = P_c^{A^c}$ (and an arbitrary Q_1). Note that $P_1 \cup Q_2$ and $P_2 \cup Q_1$ are non-intersecting partitions of Ω. Since f is partially commutative,

$$f(P_1 \cup Q_2) + f(P_2 \cup Q_1) = f((P_1 \cup Q_2) \vee (P_2 \cup Q_1)) + f((P_1 \cup Q_2) \wedge (P_2 \cup Q_1)).$$

But $(P_1 \cup Q_2) \vee (P_2 \cup Q_1) = P_1 \cup Q_1$ and $(P_1 \cup Q_2) \wedge (P_2 \cup Q_1) = P_2 \cup Q_2$, so that

$$f(P_1 \cup Q_2) + (P_2 \cup Q_1) = f(P_1 \cup Q_1) + f(P_2 \cup Q_2)$$

or

$$f(P_1 \cup Q_2) - f(P_2 \cup Q_2) = f(P_1 \cup Q_1) - f(P_2 \cup Q_1),$$

namely, $d(P_1,P_2,Q_2) = d(P_1,P_2,Q_1)$. //

In view of this claim, we define $d(P_1,P_2)$ to be $d(P_1,P_2,Q)$ for some $Q \in \mathcal{P}^{A^c}$.

We now turn to define the set function v. Given $A \in \mathcal{B}$, $A \neq \emptyset$, let

$$v(A) = |P_f^A|w - d(P_f^A,P_c^A),$$

where

$$w = f(P_f^\Omega)/|P_f^\Omega|.$$

We wish to show that for $P \in \mathcal{P}$

$$f(P) = \Sigma_{A \in P}\ v(A).$$

Assume, then, that $P = \{A_i\}_{i=1}^n$ is given. The right side is equal to

$$\Sigma_{i=1}^{n} v(A_i) = \Sigma_{i=1}^{n} [|P_f^{A_i}|w - d(P_f^{A_i}, P_c^{A_i})]$$

$$= w\, \Sigma_{i=1}^{n} |P_f^{A_i}| - \Sigma_{i=1}^{n} d(P_f^{A_i}, P_c^{A_i})$$

$$= f(P_f^{\Omega}) - \Sigma_{i=1}^{n} d(P_f^{A_i}, P_c^{A_i}).$$

To see that this expression indeed equals f(P), define $P_0, P_1, \ldots, P_n \in \mathcal{P}$ by

$$P_i = (\cup_{1 \le j \le i} P_f^{A_j}) \cup (\cup_{i < j \le n} P_c^{A_j}), \qquad 0 \le i \le n$$

so that $P_0 = P$ and $P_n = P_f^{\Omega}$. Note that for $0 \le i \le n - 1$, P_{i+1} and P_i differ only on A_{i+1} and

$$f(P_{i+1}) - f(P_i) = d(P_f^{A_i}, P_c^{A_i}).$$

Finally, consider the expression

$$f(P_f^{\Omega}) - f(P) = f(P_n) - f(P_0) = \Sigma_{i=0}^{n-1} [f(P_{i+1}) - f(P_i)]$$

$$= \Sigma_{i=1}^{n} d(P_f^{A_i}, P_c^{A_i}).$$

Hence,

$$f(P_f^{\Omega}) - \Sigma_{i=1}^{n} d(P_f^{A_i}, P_c^{A_i}) = f(P)$$

and this completes the proof of Step 1. //

Step 2: An arbitrary $\mathcal{B}$.

Consider the set

$$H = \{(h,a) \mid h: \mathcal{B} \to \mathbf{R},\ |\{A \in \mathcal{B} \mid h(A) \neq 0\}| < \infty, \text{ and } a \in \mathbf{R}\}.$$

An element $(h,a) \in H$ is associated with the linear equation $\Sigma_{A\in\mathcal{B}}\, h(A)v(A) = a$.

A subset E of H will be called a system of equations and it is said to have a solution if there is a set function v satisfying all the equations associated with the elements of E.

We wish to prove that the system

$$E_0 = \{(1_P,\ f(P)) \mid P \in \mathcal{P}\}$$

has a solution. (Note that the equation associated with $(1_P, f(P)) \in H$ is $\Sigma_{A\in P}\, v(A) = f(P)$.)

We know, however, that every finite subsystem of E_0 has a solution: given the subsystem corresponding to $P_1, P_2, \ldots, P_n \in \mathcal{P}$, choose the set function v provided by Step 1 for the (finite) algebra $\bar{\mathcal{B}}$ induced by their common refinement $P_1 \vee P_2 \vee \ldots \vee P_n$.

This fact suggests the following definition:

$$\mathcal{E} = \{E \subseteq H \mid E_0 \subseteq E \text{ and every finite subsystem of E has a solution}\}.$$

Obviously, $E_0 \in \mathcal{E}$ so that $\mathcal{E} \neq \emptyset$. $\mathcal{E}$ is partially ordered by inclusion, and it is easy to see that every chain has an upper bound in $\mathcal{E}$, namely, the union of all the chain elements. Hence, we may apply Zorn's lemma and conclude that there is a maximal element $\bar{E}$ in $\mathcal{E}$.

For every $A \in \mathcal{B}$ consider

$$H_A = \{(1_{\{A\}},a) | a \in \mathbb{R}\} \subseteq H,$$

which are associated with equations of the type $v(A) = a$ for some $a \in \mathbb{R}$. Obviously, $|H_A \cap \bar{E}| < 2$, since the subsystem $\{(1_{\{A\}},a),(1_{\{A\}},b)\}$ will not have a solution for $a \neq b$. If $|H_A \cap \bar{E}| = 1$ for all $A \in \mathcal{B}$, one may define $v(A)$ to be such that $(1_{\{A\}},v(A)) \in \bar{E}$ and this v will be a solution of $\bar{E}$, perforce of E_0. We will now show that this must be the case, that is to say, that $H_A \cap \bar{E} \neq \emptyset$ for every $A \in \mathcal{B}$.

Assume, then, that for some $A \in \mathcal{B}$, $H_A \cap \bar{E} = \emptyset$. Consider $h = (1_{\{A\}},0)$, i.e., the equation $v(A) = 0$. By the maximality of $\bar{E}$, $\bar{E} \cup \{h\} \notin \mathcal{E}$, which means that it has a finite subsystem that does not have a solution. Hence, there is a finite $E_f \subseteq \bar{E}$ such that E_f has a solution but $E_f \cup \{h\}$ does not. Consider the linear space of all solutions to E_f. The set $V = \{v(A) | v$ is a solution to $E_f\}$, which is known to be nonempty, cannot equal $\mathbb{R}$ since this would imply that $E_f \cup \{h\}$ does have a solution. Hence $V = \{a\}$ for some $a \in \mathbb{R}$. It is obvious that $E \cup \{(1_A,a)\}$ will also be in $\mathcal{E}$, which contradicts the maximality of $\bar{E}$.

This completes the proof of Theorem 3.2. //

Given a set-function v, we have a uniquely defined partially commutative partition function f. However, given such an f, a set function v corresponding to it need not be unique. In fact, it will only be unique if $\mathcal{B}$ is the trivial algebra $\{\emptyset,\Omega\}$, as shown in the following:

<u>Proposition 3.3</u>: Suppose f is an additively separable partition function and v_1 corresponds to it. Then v_2 also corresponds to f iff $\nu = v_1 - v_2$ is a finitely additive set function with $\nu(\Omega) = 0$.

<u>Proof</u>: First assume that v_1 corresponds to f and that ν is finitely additive with $\nu(\Omega) = 0$. Then

$$\Sigma_{A\in P} (v_1 + \nu)(A) = \Sigma_{A\in P} v_1(A) + \Sigma_{A\in P} \nu(A) = f(P) + \nu(\Omega).$$

Hence, $v_2 = v_1 + \nu$ also corresponds to f.

As for the converse, assume v_1 and v_2 correspond to f, define $\nu = v_1 - v_2$, and for $A,B \in \mathcal{B}$ with $A \cap B = \emptyset$, let

$$P_1 = \{A,B,(A \cup B)^c\},\ P_2 = \{A \cup B,(A \cup B)^c\}.$$

Then

$$\begin{aligned} f(P_2) - f(P_1) &= v_1(A \cup B) - v_1(A) - v_1(B) \\ &= v_2(A \cup B) - v_2(A) - v_2(B) \end{aligned}$$

whence

$$\nu(A \cup B) = v_1(A \cup B) - v_2(A \cup B)$$
$$= v_1(A) - v_2(A) + v_1(B) - v_2(B) = \nu(A) + \nu(B).$$

Finally, $v_1(\Omega) = v_2(\Omega) = f(P_c^{\Omega})$, whence $\nu(\Omega) = 0$. //

Before concluding this section we would like to suggest a slightly different interpretation of partial commutativity. Our approach is to take Savage's model as a foundation and look for mathematical conditions on partition functions that characterize the information functions of that model. However, one could start out with partition functions as a primitive of the decision making model, and look for behavioral axioms in this framework. A reasonable condition in such a set-up could have been the following.

A partition function f is said to satisfy the <u>sure thing principle</u> if the following holds: for every $A \in \mathcal{B}$, $P_1, P_2 \in \mathcal{P}^A$, $Q_1, Q_2 \in \mathcal{P}^{A^c}$, with $P_1 \geq P_2$,

$$f(P_1 \cup Q_1) - f(P_2 \cup Q_1) = f(P_1 \cup Q_2) - f(P_2 \cup Q_2).$$

The interpretation of this condition should be as follows: in all four partitions the DM would know whether A has occurred or not. Hence, by the sure thing principle, the DM should not care about what he/she will know should A not occur in order to evaluate information <u>given</u> A. Thus, the left side, which is the marginal value of P_1 to a DM having P_2 (in case he/she has Q_1 for A^c), should be the same as in the case Q_2 is the DM's information on A^c.

We note that:

<u>Observation 3.4</u>: An information function f satisfies the sure thing principle iff it is partially commutative, i.e., iff it is additively separable.

<u>Proof</u>: The fact that partial commutativity implies the sure thing principle was shown in the proof of 3.2. The converse is trivial in view of 2.1. //

4. <u>Characterization of Information Functions</u>

In the previous section we have formulated two necessary conditions on f to be an information function: monotonicity and partial commutativity. It seems natural to ask at this point whether these conditions are also sufficient.

We will shortly present the negative answer. It will prove useful to introduce the following definition first: a set function v is <u>subadditive</u> if $v(A) + v(B) \geq v(A \cup B)$ whenever $A \cap B = \emptyset$. It is easy to verify the following:

<u>Observation 4.1</u>: A partially commutative partition function f is monotonic iff any (by 3.3-all) of its corresponding set functions are subadditive.

The insufficiency is proven by:

<u>Example 4.2</u>: Let $\Omega = \{1,2,3\}$, $\mathcal{B} = 2^{\Omega}$, $\mu(\{i\}) = 1/3$ for $i \in \Omega$, and define

$v(\Omega) = 2$, $v(A) = 1$, for $A \neq \Omega, \emptyset$. Finally, define f via v. This f is certainly partially commutative and it is monotonic since v is subadditive. However, we claim it cannot be an information function. To prove this, suppose to the contrary that $f = e_{(A,u)}$ where (A,u) satisfy the maximality property with respect to $(\Omega, \mathcal{B}, \mu)$. For $B \in \mathcal{B}$ let $a_B \in A$ be an action maximizing the DM's expected utility given B, and denote

$$w(B) = \int_B u(a_B)d\mu.$$

Since w corresponds to f, there are (ν_1, ν_2, ν_3) such that $w(B) = v(B) + \Sigma_{i \in B}\, \nu_i$ and $\nu_1 + \nu_2 + \nu_3 = 0$.

Next, consider a_Ω and denote $u(a_\Omega, i)$ by u_i $(1 \leq i \leq 3)$, for short. Then we obtain

$$w(\Omega) = v(\Omega) = 2 = (1/3)(u_1 + u_2 + u_3).$$

For any $B \in \mathcal{B}$,

$$\int_B u(a_\Omega)d\mu \leq w(B)$$

whence, for $B = \{1,2\}$,

$$(1/3)(u_1 + u_2) \leq w(B) = 1 + \nu_1 + \nu_2.$$

Similarly, we obtain for {2,3} and {1,3}

$$(1/3)(u_2 + u_3) \le w(\{2,3\}) = 1 + \nu_2 + \nu_3$$
$$(1/3)(u_1 + u_3) \le w(\{1,2\}) = 1 + \nu_1 + \nu_3,$$

and summation yields

$$(2/3)(u_1 + u_2 + u_3) \le 3 + 2(\nu_1 + \nu_2 + \nu_3)$$
$$(1/3)(u_1 + u_2 + u_3) \le 3/2 < 2,$$

a contradiction. //

At this point the analysis seems close enough to cooperative game theory to suggest the following definition: for a set function v (which is, in fact, a game) and $B \in \mathcal{B}$, the <u>B-anticore</u> of v, denoted $AC_B(v)$, is the set of all σ-additive measures λ on $\mathcal{B}$ satisfying: (i) λ is absolutely continuous w.r.t. μ; (ii) for $A \in \mathcal{B}$, $\lambda(A) \le v(A)$; and (iii) $\lambda(B) = v(B)$.

We are about to require that v would have a nonempty B-anticore for all $B \in \mathcal{B}$. Let us first note some of the properties implied by this constraint.

<u>Observation 4.3</u>: If $AC_B(v) \neq \emptyset$ for all $B \in \mathcal{B}$, then v is monotone (namely, $A \subseteq B$ implies $v(A) \le v(B)$) and subadditive.

<u>Proof</u>: For monotonicity, let $A \subseteq B$ and let $\lambda \in AC_A(v)$. then $v(A) = \lambda(A) \le \lambda(B) \le v(B)$. As for subadditivity, let $A \cap B = \emptyset$ and choose $\lambda \in AC_{A \cup B}(v)$. Then $v(A) + v(B) \ge \lambda(A) + \lambda(B) = \lambda(A \cup B) = v(A \cup B)$. //

Note, however, that while subadditivity is invariant w.r.t. "shifting" v by an additive ν, monotonicity is not.

We can finally provide a characterization of information functions:

Theorem 4.4: Given $(\Omega,\mathcal{B},\mu)$ and a partition function f, the following are equivalent:

(i) There are (A,u) satisfying the maximality property w.r.t. $(\Omega,\mathcal{B},\mu)$, with $0 \le u \le M$, such that $f = e_{(A,u)}$.

(ii) f is additively separable and there is a set function v corresponding to it with $AC_B(v) \neq \emptyset$ and $v(B) \le \mu(B)M$ for all $B \in \mathcal{B}$.

Proof: First assume (i). Define $v(B) = \max_{a\in A} \int_B u(a)d\mu$. f is obviously additively separable and v corresponds to it. $v(B) \le \mu(B)M$ is trivial, so it is only left to show that $AC_B(v) \neq \emptyset$. Given $B \in \mathcal{B}$, let a_B be an optimal action, namely, $\int_B u(a_B)d\mu = v(B)$. Define

$$\lambda_B(A) = \int_A u(a_B)d\mu.$$

λ_B thus defined is a σ-additive measure that is absolutely continuous w.r.t. μ. Moreover, $\lambda_B(B) = v(B)$ by the choice of a_B, and $\lambda_B(A) \le v(A)$ by optimality. Hence $\lambda_B \in AC_B(v)$.

To show the converse, assume f is additively separable with v satisfying the boundedness and anticores' nonemptiness conditions. Define the set of actions as follows: for each $B \in \mathcal{B}$ choose $\lambda_B \in AC_B(v)$. Let a_B be the Radon-Nikodym derivative of λ_B w.r.t. μ, ($a_B(\omega) \in [0,M]$ since $0 \le \lambda_B(C) \le M\mu(C)$ for all $C \in \mathcal{B}$.) and let $A = \{a_B | B \in \mathcal{B}\}$. Finally, define $u(a_B,\omega) = a_B(\omega)$. Then for all $C \in \mathcal{B}$:

$$\int_C u(a_B)d\mu = \lambda_B(C) \le v(C)$$

with equality for C = B. Hence, $f = e_{(A,u)}$, (A,u) satisfy the maximality property w.r.t $(\Omega,\mathcal{B},\mu)$ and u is bounded as required. //

We conclude this section by characterizing the information functions $e_{(A,u)}$ in the more general case where (A,u) need not satisfy the maximality property. It turns out to be the case that very minor modifications are required. First we redefine $e_{(A,u)}$ by replacing "max" by "sup." Next we define, for $B \in \mathcal{B}$ and $\varepsilon > 0$, the <u>ε-B-anticore</u> of v, denoted $\varepsilon\text{-}AC_B(v)$ to be the set of σ-additive measures λ that satisfy: (i) λ is absolutely continuous w.r.t. μ; (ii) $\lambda(A) \le v(A)$ for all $A \in \mathcal{B}$; and (iii) $\lambda(B) \ge v(B) - \varepsilon$.

We may now formulate:

<u>Theorem 4.5</u>: Given $(\Omega,\mathcal{B},\mu)$ and a partition function f, the following are equivalent:

(i) There are (A,u) with $0 \le u \le M$ such that $f = e_{(A,u)}$;

(ii) f is additively separable and there is a set function v corresponding to it such that $\varepsilon\text{-}AC_B(v) \ne \emptyset$ for all $B \in \mathcal{B}$ and $\varepsilon > 0$, and $v(B) \le \mu(B)M$ for all $B \in \mathcal{B}$.

The proof is a straightforward adaptation of that of Theorem 4.4.

5. <u>Applications</u>

The results presented in this section are not applications of the characterization theorems themselves; rather, they are by-products of the analysis described above. Throughout this section we will assume that f is an information function and that μ is nonatomic. (The results would have natural asymptotic counterparts for a finite but large Ω.)

We are interested in expressions of the form

$$[f(P) - f(P \wedge Q)] - [f(P \vee Q) - f(Q)]$$

where P and Q are two partitions of Ω. If they are non-intersecting, this expression has to equal zero. This fact, which is a simple result of the above analysis, may be interpreted as follows: suppose P represents the expertise of consultant I and Q represents that of consultant II. The two consultants are selling their a priori knowledge to a third party (say, a firm), which is already guaranteed to know events that will be common knowledge between the two consultants. That is, the firm has the partition $P \wedge Q$.

Should consultant I be the first to sell his/her partition, the firm should be willing to pay him/her $f(P) - f(P \wedge Q)$. Should he/she be the second, the payoff will be $f(P \vee Q) - f(Q)$. Hence, the expression above measures the marginal utility derived from information; if it is, say, always nonnegative, we would say that <u>f exhibits a decreasing marginal utility</u> (from information). Similarly, if it is always nonpositive (zero), we would say that <u>f exhibits an increasing (constant) marginal utility</u> (from information).

We here quote some traditional game theoretic definitions. A game v is <u>concave</u> if for every $A, B \in \mathcal{B}$,

$$v(A) + v(B) \geq v(A \cup B) + v(A \cap B).$$

It is convex if the reverse inequality holds. Note that it is both concave and convex iff it is additive.

The relationship between those game properties and the information function properties is given by:

Proposition 5.1: Let f be an information function on a nonatomic measure space $(\Omega, \mathcal{B}, \mu)$ for (A,u) and let v be a corresponding set function. If f exhibits a decreasing (increasing) marginal utility from information, then v is concave (convex).

Proof: Let $A, B \in \mathcal{B}$ and we want to show that $v(A) + v(B) \geq v(A \cup B) + v(A \cap B)$. (The proof for convexity is symmetric.)

First assume that $A \cap B \neq \emptyset$, and let $P = \{A, B\backslash A, (A \cup B)^c\}$ and $Q = \{A\backslash B, B, (A \cup B)^c\}$. Then $P \vee Q = \{A\backslash B, A \cap B, B\backslash A, (A \cup B)^c\}$ and $P \wedge Q = \{A \cup B, (A \cup B)^c\}$, and the concavity inequality follows.

Next, if $A \cap B = \emptyset$, distinguish between two cases: if $\mu(A) = 0$ the inequality surely holds; otherwise, choose $A_\varepsilon \subseteq A$ such that $\mu(A_\varepsilon) = \varepsilon$ for an arbitrary $\varepsilon > 0$, and define $B_\varepsilon = B \cup A_\varepsilon$. Then $B_\varepsilon \cap A \neq \emptyset$ and

$$v(A) + v(B_\varepsilon) \geq v(A \cup B) + v(A_\varepsilon) \geq v(A \cup B),$$

which implies $v(A) + v(B) \geq v(A \cup B)$. //

<u>Corollary 5.2</u>: If the marginal utility from information is constant (i.e., both decreasing and increasing), then f is constant.

<u>Proof</u>: In this case, f is fully commutative, i.e.,

$$f(P) + f(Q) = f(P \wedge Q) + f(P \vee Q),$$

which means that v is additive. Hence, $f(P) = v(\Omega)$ for all $P \in \mathcal{P}$. //

Note that when f is constant the marginal utility from any partition is zero. This last result may be interpreted as follows: if the consultants do not care which one sells his/her expertise first, their information must be worthless anyhow (which menas that the DM has an (ε-) dominant strategy).

We conclude this section with two remarks.

<u>Remark 5.3</u>: Result 5.2 may be slightly strengthened since it is enough to require

$$f(P \vee Q) - f(Q) \geq f(P) - f(P \wedge Q)$$

for all $P,Q \in \mathcal{P}$ to obtain convexity of v, while subadditivity is guaranteed by f's monotonicity.

<u>Remark 5.4</u>: It is natural to ask whether the converse of 5.1 holds. The following example shows it does not.

Let Ω be [0,1], $\mathcal{B}$ the Borel sets, and μ the Lebesgue measure. Define $v = g(\mu)$ where $g\colon [0,1] \to \mathbb{R}$ is defined by

$$g(x) = \begin{cases} 8x, & 0 \le x \le 1/4 \\ 1 + 4x, & 1/4 < x \le 1. \end{cases}$$

g is a concave real function; hence, v is concave. However, the partition function f induced by v does not exhibit a decreasing marginal utility from information: let $A = [0, 1/4)$, $B = [1/4, 1/2)$, $C = [1/2, 3/4)$, and $D = [3/4, 1]$. Consider $P = \{A \cup B, C \cup D\}$ and $Q = \{A \cup C, B \cup D\}$, with $P \vee Q = \{A,B,C,D\}$, $P \wedge Q = \{\Omega\}$. Then

$$f(P) + f(Q) = v(A \cup B) + v(C \cup D) + v(A \cup C) + v(B \cup D) = 12,$$

but

$$f(P \wedge Q) + f(P \vee Q) = v(A) + v(B) + v(C) + v(D) + v(\Omega) = 13. \quad //$$

6. Concluding Remarks

Remark 6.1: One may not regard the nonemptiness of the anticores as a primitive enough condition. In this case, one may adapt the known theorems regarding existence of a nonempty core (Shapley (1967), Bondareva (1963)) and for the existence of σ-additive measures in it (Schmeidler (1972)).

However, a more challenging task would be to formulate the anticore conditions in terms of the partition function f directly. We were not able to find axioms on f that would not be straightforward (and awkward) translations of those we have on v.

Remark 6.2: The analysis of the problem presented here may be carried out to a certain extent without presupposing a prior μ. A measurable space $(\Omega,\mathcal{B})$ is enough to define a partition function f and it will then be perfectly meaningful to ask what are necessary and sufficient conditions on f for it to be the information function for some (A,μ,u).

The discussion in Section 3 would still be valid and the equivalence of additive separability and partial commutativity would hold. However, given a set function v, one has to find conditions on it for the existence of a σ-additive μ w.r.t. which v is absolutely continuous. Separation methods may be invoked at this point, as, say, in Kelley (1959), but they do not seem to provide very enlightening conditions. At present, we are not aware of any nice-looking characterization of such v's.

At any rate, one should note that the specification of μ is not very restrictive from a conceptual viewpoint: given $(\Omega,\mathcal{B},\mu)$ and a partition function f, f is an information function for $(\Omega,\mathcal{B},\mu)$ iff it is for some $(\Omega,\mathcal{B},\nu)$ where ν is a probability measure that is Lipschitz-continuous w.r.t. μ. Thus, the question this paper deals with does allow a certain freedom in the choice of the DM's prior.

References

Aumann, R. J. (1974), "Subjectivity and Correlation in Randomized Strategies," J. of Math. Econ. 1, 67-96.

Bondareva, O. N. (1963), "Some Applications of Linear Programming Methods to the Theory of Cooperative Games" (in Russian), Problemy Kibernetiki 10, 119-139.

Chateauneuf, A. (1985), "On the Existence of a Probability Measure Compatible with a Total Order on a Boolean Algebra," J. of Math. Econ. 14, 43-52.

Fishburn, P. C. (1970), Utility Theory for Decision Making, Wiley, New York.

Hirshliefer, J. and J. G. Riley (1979), "The Analytics of Uncertainty and Information--An Expository Survey," J. of Economic Literature, 17, 1375-1421.

Kelley, J. L. (1959), "Measure on Boolean Algebras," Pacific Journal of Mathematics, 9, 1165-1177.

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

Schmeidler, D. (1972), "Cores of Exact Games, I," J. of Math. Analysis and Applications 40, 214-225.

Shapley, L. S. (1967), "On Balanced Sets and Cores," Naval Res. Logist. Quart. 14, 453-460.