An axiomatic approach to the definition of the entropy of a discrete Choquet capacity

Ivan Kojadinovic

Iremia, Université de la Réunion

15 avenue René Cassin - BP 7151, 97715 Saint-Denis messag cedex 9

Ile de la Réunion, France.

Jean-Luc Marichal

Department of Mathematics, Brigham Young University

TMCB, Provo UT 84602, U.S.A.

Marc Roubens

Institute of Mathematics, University of Liège

Grande Traverse 12 (B37), B-4000 Liège, Belgium.

Abstract

An axiomatization of the concept of entropy of a discrete Choquet capacity is given. It is based on three axioms: the symmetry property, a boundary condition for which the entropy reduces to the classical Shannon entropy, and a generalized version of the well-known recursivity property. This entropy, recently introduced to extend the Shannon entropy to non-additive measures, fulfills several properties considered as requisites for defining an entropy-like measure of uncertainty. An interpretation of it in the framework of aggregation by the discrete Choquet integral is given as well.

Keywords: entropy, Choquet capacity, Choquet integral, information theory.

1 Introduction

In 1948, Shannon introduced a measure of uncertainty of a discrete stochastic system known as *entropy* [19]. For a probability distribution p defined on a finite set $[n] = \{1, \ldots, n\}$, the Shannon entropy of p is defined by

$$H_S(p) := -\sum_{i=1}^{n} p(i) \ln p(i)$$

with the convention that $0 \ln 0 := 0$.

Although several other measures of uncertainty have been proposed as generalizations of the Shannon entropy (see e.g. [17] for an overview), the most widely used uncertainty remains that of Shannon mainly because of its attractive properties, its connections with the *Kullback-Leibler divergence* [12] and its role in the *maximum entropy principle* [9].

Several axiomatic characterizations of the Shannon entropy have been proposed in the literature (see e.g. [1, 3, 4, 10, 11]), amongst which the most famous is probably *Shannon's theorem* [19].

By replacing the additivity property of probability measures by that of monotonicity, one obtains *Choquet capacities* [2] also known as *fuzzy measures* [21] which are able to model other types of possibly uncertain information. More formally, a discrete Choquet capacity μ on $[n]$ is a monotone set function defined on the power set of $[n]$ that is zero at the empty set. Such a concept can be used to model the *importance* of a coalition of elements of $[n]$. The label set $[n]$ could correspond to criteria in a multicriteria decision problem [5, 6, 15], to players in a cooperative game [6, 8, 20], to attributes in a classification problem [7], to variables in a regression problem, voters in an opinion pooling problem, etc. In all these cases, for any subset $S \subseteq [n]$, $\mu(S)$ can be interpreted as the *degree of importance* or the *strength* of the coalition S of elements for the particular problem under consideration.

A discrete Choquet capacity being clearly a generalization of a discrete probability distribution, the following natural question arises : How could one appraise the "uncertainty" associated with a Choquet capacity in the spirit of the Shannon entropy? Recently, Marichal [13, 14, 16] proposed the notion of *entropy of a discrete Choquet capacity* as a generalization of the Shannon entropy and showed that it satisfies many properties that one would intuitively require from such a measure. This generalized entropy is defined as

$$H_M(\mu) := \sum_{i=1}^{n} \sum_{S \subseteq [n] \setminus i} \gamma_s(n)\, h[\mu(S \cup i) - \mu(S)] \quad (1)$$

where the functions

$$\gamma_s(n) := \frac{(n-s-1)!\,s!}{n!} \qquad (s = 0, 1, \ldots, n-1),$$

and

$$h(x) := \begin{cases} -x \ln x, & \text{if } x > 0, \\ 0, & \text{if } x = 0, \end{cases}$$

will be used throughout.

Note that this formulation is very close to that of the *Shapley value* [20] of a Choquet capacity μ on $[n]$, which is a fundamental concept in game theory. For an element $i \in [n]$, it is defined by

$$\phi_i(\mu) := \sum_{S \subseteq [n] \setminus i} \gamma_s(n)\,[\mu(S \cup i) - \mu(S)]$$

and can be interpreted as the average marginal contribution of i to a coalition not containing it. It is worth noting that a basic property of the Shapley value is

$$\sum_{i=1}^{n} \phi_i(\mu) = \mu([n]). \qquad (2)$$

In this paper, after defining the notion of *uncertainty* in the setting of discrete Choquet capacities (Section 2), we propose a characterization of the generalized entropy (1) by means of three axioms (Section 3). We also list some of its properties (Section 4) and present an interpretation of it in the framework of aggregation by the discrete Choquet integral (Section 5).

2 Uncertainty contained in a discrete Choquet capacity

Although discrete Choquet capacities can be seen as generalizations of discrete probability distributions, it is not clear what "uncertainty" should mean for such non-additive measures. After introducing the notation, we propose an intuitive definition of the notion of *uncertainty* based on the lattice representation of a discrete Choquet capacity.

2.1 Notation and first definitions

Throughout this paper, $[n] = \{1, \ldots, n\}$ is a finite label set, the power set of which is denoted $\mathcal{P}([n])$.

A *discrete Choquet capacity* [2] or *discrete fuzzy measure* [21] on $[n]$ is a set function $\mu : \mathcal{P}([n]) \to \mathbb{R}^+$ satisfying the following conditions :

1. $\mu(\emptyset) = 0$,
2. for all $S, T \subseteq [n]$, $S \subseteq T \Rightarrow \mu(S) \leq \mu(T)$.

The Choquet capacity is said to be *normalized* if $\mu([n]) = 1$. In the sequel, we restrict ourselves to the study of the notion of *uncertainty* only for normalized Choquet capacities. For any integer $n \geq 1$, we denote by $\mathcal{F}_{[n]}$ the set of all normalized Choquet capacity on $[n]$.

In order to avoid a heavy notation, we adopt that used in [15]. Thus, we will omit braces for singletons, e.g., by writing $\mu(i)$, $[n] \setminus i$ instead of $\mu(\{i\})$, $[n] \setminus \{i\}$. Furthermore, cardinalities of subsets $S, T, \ldots,$ will be denoted by the corresponding lower case letters $s, t, \ldots$

A Choquet capacity $\mu \in \mathcal{F}_{[n]}$ is said to be:

- *additive* if $\mu(S \cup T) = \mu(S) + \mu(T)$ for all disjoint subsets $S, T \subseteq [n]$,
- *cardinality-based* if, for all $T \subseteq [n]$, $\mu(T)$ depends only on the cardinal of T. In this case, there exist $\mu_1, \ldots, \mu_n \in [0, 1]$ such that $\mu(T) = \mu_t$ for all $T \subseteq N$.

There is only one Choquet capacity on $[n]$ that is both additive and cardinality-based. We shall call it *the uniform Choquet capacity* on $[n]$ and denote it by μ^*. It is easy to check that μ^* is given by

$$\mu^*(T) = t/n \qquad \forall T \subseteq [n].$$

2.2 Choquet capacities and maximal chains

Consider the lattice $\mathcal{L}_{[n]}$ related to the power set of $[n]$ under the inclusion relation. The lattice $\mathcal{L}_{[n]}$ can be represented by a graph $\mathcal{H}_{[n]}$ called Hasse Diagram whose nodes correspond to subsets $S \subseteq [n]$ and whose edges represent adding an element to the bottom subset to get the top subset. Figure 1 shows an example of such a graph for $[n] = \{1, 2, 3, 4\}$.

A *maximal chain* m of $\mathcal{H}_{[n]}$ is an ordered collection of $n+1$ nested distinct subsets denoted

$$m = (\emptyset \subsetneq \{i_1\} \subsetneq \{i_1, i_2\} \subsetneq \cdots \subsetneq \{i_1, \ldots, i_n\} = [n]).$$

For instance, in Figure 1, the maximal chain $(\emptyset \subsetneq \{1\} \subsetneq \{1, 4\} \subsetneq \{1, 2, 4\} \subsetneq \{1, 2, 3, 4\})$ is emphasized.

We denote by $\mathcal{C}_{[n]}$ the set of maximal chains of $\mathcal{H}_{[n]}$. The cardinality of $\mathcal{C}_{[n]}$ is clearly $n!$.

Given a Choquet capacity μ on $[n]$, with each maximal chain $m \in \mathcal{C}_{[n]}$ can be associated a discrete probability distribution p_m^μ on $[n]$ defined by

$$p_m^\mu(i) := \mu(m_i) - \mu(m_{i-1}) \qquad \forall i \in [n]$$

where m_i denotes the subset of cardinal i of m.

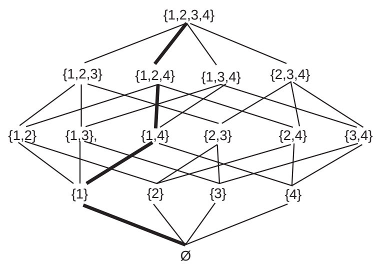

Figure 1: Hasse diagram $\mathcal{H}_{[n]}$ corresponding to the lattice of subsets of $[n] = \{1, 2, 3, 4\}$.

Denote by $\Pi_{[n]}$ the set of permutations on $[n]$. With each permutation $\pi \in \Pi_{[n]}$ is associated a unique maximal chain $m^\pi \in \mathcal{C}_{[n]}$ defined by

$$m^\pi = (\emptyset \subsetneq \{\pi(n)\} \subsetneq \{\pi(n-1), \pi(n)\} \subsetneq \cdots \subsetneq \{\pi(1), \dots, \pi(n)\} = [n]).$$

We then write

$$p^\mu_\pi(i) := p^\mu_{m^\pi}(i) = \mu(\{\pi(i), \dots, \pi(n)\}) - \mu(\{\pi(i+1), \dots, \pi(n)\}) \qquad \forall i \in [n].$$

The set $\{p^\mu_\pi\}_{\pi\in\Pi_{[n]}} = \{p^\mu_m\}_{m\in\mathcal{C}_{[n]}}$ of $n!$ probability distributions obtained from μ will be denoted P^μ as we continue.

Notice that, if μ is cardinality-based, then there exists a unique probability distribution p^μ on $[n]$ such that $p^\mu_\pi = p^\mu$ for all $\pi \in \Pi_{[n]}$. If μ is additive then we simply have $p^\mu_\pi(i) = \mu(\pi(i))$ for all $i \in [n]$.

2.3 Uncertainty associated with a discrete Choquet capacity

For a discrete probability distribution, the notion of *uncertainty* has an intuitive meaning which is directly linked with that of *uniformity*. Indeed, the more "uniform" a discrete probability distribution, the higher the uncertainty of the underlying discrete stochastic system (for a discussion on the notion of uncertainty, see e.g. [17]).

Although discrete Choquet capacities can be clearly seen as generalizations of discrete probability distributions, it is not clear what "uncertainty" should mean for such non-additive measures. As we have seen in the previous subsection, a Choquet capacity on $[n]$ can be represented by the set $P^\mu = \{p^\mu_m\}_{m\in\mathcal{C}_{[n]}}$ of $n!$ probability distributions on $[n]$. We therefore propose to define the intuitive notion of *uncertainty associated with a discrete Choquet capacity* as a kind of *average* of the uncertainties contained in the probability distributions $\{p^\mu_m\}_{m\in\mathcal{C}_{[n]}}$. Hence, the more uniform on average the probability distributions p^μ_m, $m \in \mathcal{C}_{[n]}$, the higher the *uncertainty* contained in the discrete Choquet capacity μ. As no maximal chain should be privileged, the average uncertainty should be defined by means of a symmetric function over all the $n!$ maximal chains m of $H_S(p^\mu_m)$. It is worth mentioning that, in terms of maximal chains, the entropy H_M can be rewritten as

$$H_M(\mu) = \frac{1}{n!} \sum_{m\in\mathcal{C}_{[n]}} H_S(p^\mu_m) \tag{3}$$

or equivalently,

$$H_M(\mu) = \frac{1}{n!} \sum_{\pi\in\Pi_{[n]}} H_S(p^\mu_\pi). \tag{4}$$

This result immediately follows from the next proposition.

Proposition 2.1 *Let μ be any Choquet capacity on $[n]$ (normalized or not) and let f be any function. Then, we have*

$$\frac{1}{n!} \sum_{m\in\mathcal{C}_{[n]}} \sum_{j=1}^{n} f[\mu(m_j) - \mu(m_{j-1})] = \sum_{i=1}^{n} \sum_{S\subseteq[n]\setminus i} \gamma_s(n)\, f[\mu(S\cup i) - \mu(S)].$$

Proof. For all $i \in [n]$, for all $S \subseteq [n]\setminus i$, let us denote by $\mathcal{C}^{S,S\cup i}_{[n]}$ the subset of $\mathcal{C}_{[n]}$ composed of maximal chains whose subsets of cardinal s and $s+1$ are equal to S and $S \cup i$ respectively. It is easy to check that $|\mathcal{C}^{S,S\cup i}_{[n]}| = s!(n-s-1)!$.

It follows therefore that, for a given $i \in [n]$ and for a subset $S \subseteq [n] \setminus i$, when summing the term $\sum_{j=1}^{n} f[\mu(m_j) - \mu(m_{j-1})]$ over the set of maximal chains, the term $f[\mu(S\cup i) - \mu(S)]$ will appear $s!(n-s-1)!$ times. □

If $\Pi_{[n]}$ is considered as a probability space, a straightforward probabilistic interpretation of H_M directly follows from Eq. (4): for any $\mu \in \mathcal{F}_{[n]}$, $H_M(\mu)$ is the mathematical expectation of $H_S(p^\mu_\pi)$ with respect to the uniform distribution on $\Pi_{[n]}$.

3 Axiomatization of the entropy H_M

Before stating the axioms that a measure of *uncertainty* or *entropy* of a discrete Choquet capacity should satisfy, we define some additional concepts that will be needed in the sequel.

3.1 Additional definitions

Let $\mu \in \mathcal{F}_{[n]}$ and let S and T be two disjoint subsets of $[n]$. The *Choquet capacity on S in the presence of T* [8] is denoted $\mu^S_{\cup T}$ and is defined by

$$\mu^S_{\cup T}(K) := \mu(K \cup T) - \mu(T), \qquad \forall K \subseteq S.$$

Clearly, the Choquet capacity $\mu^S_{\cup T}$ is not normalized. The normalized version of the Choquet capacity $\mu^S_{\cup T}$ is denoted $\bar{\mu}^S_{\cup T}$ and is defined by

$$\bar{\mu}^S_{\cup T}(K) := \begin{cases} \dfrac{\mu^S_{\cup T}(K)}{\mu^S_{\cup T}(S)} & \text{for all } K \subseteq S \text{ if } \mu^S_{\cup T}(S) \neq 0, \\ 0 & \text{for all } K \subsetneq S \text{ if } \mu^S_{\cup T}(S) = 0, \\ 1 & \text{if } K = S \text{ and } \mu^S_{\cup T}(S) = 0. \end{cases}$$

Let μ be a Choquet capacity on $[n]$ and let $A_1, \ldots, A_k$ form a partition of $[n]$. The *reduced Choquet capacity with respect to* $A_1, \ldots, A_k$ [8] is a Choquet capacity denoted $\mu^{[A_1]\ldots[A_k]}$ defined on a set of k elements noted as $[A_1]\ldots[A_k]$, where, for all $i \in [k]$, $[A_i]$ stands for an hypothetical element which is the union (or the representative) of the elements in A_i, that is,

$$\mu^{[A_1]\ldots[A_k]}\Big(\bigcup_{i\in S}[A_i]\Big) = \mu\Big(\bigcup_{i\in S} A_i\Big) \qquad \forall S \subseteq [k]. \quad (5)$$

For any Choquet capacity $\mu \in \mathcal{F}_{[n]}$ and any permutation $\pi \in \Pi_{[n]}$, we denote by $\pi\mu$ the Choquet capacity on $[n]$ defined by

$$\pi\mu(\pi(S)) = \mu(S) \qquad \forall S \subseteq [n],$$

where $\pi(S) := \{\pi(i) \mid i \in S\}$.

3.2 Axioms

Consider a sequence of functions $H^{(n)} : \mathcal{F}_{[n]} \to \mathbb{R}$ ($n = 2, 3, \ldots$), where, for any $\mu \in \mathcal{F}_{[n]}$, $H^{(n)}(\mu)$ is a measure of entropy of μ. When $n = 1$ we simply set $H^{(1)}(\mu) = 0$.

Given $\mu \in \mathcal{F}_{[n]}$, it is easy to check that a permutation on $[n]$ leaves unchanged the set P^μ of probability distributions. Hence, naturally, the first axiom is:

- *Symmetry axiom (S)*: For any $n \geq 2$, any $\mu \in \mathcal{F}_{[n]}$, and any $\pi \in \Pi_{[n]}$, we have $H^{(n)}(\pi\mu) = H^{(n)}(\mu)$.

Recall from the previous section that, for a cardinality-based Choquet capacity $\mu \in \mathcal{F}_{[n]}$, there exits a probability distribution p^μ such that all elements of the set P^μ are equal to p^μ. This suggests measuring the uncertainty of a cardinality-based Choquet capacity μ as that of the probability distribution p^μ. The choice of the Shannon entropy being natural as a measure of uncertainty contained in a probability distribution, our second axiom is:

- *Shannon entropy axiom (SE)*: For any $n \geq 2$ and any $\mu \in \mathcal{F}_{[n]}$, if μ is cardinality-based, then $H^{(n)}(\mu) = H^{(n)}_S(p^\mu)$.

Note that the previous axiom implies that among all the cardinality-based capacities on $[n]$, μ^* is the one that has maximum uncertainty.

The Shannon entropy is known to satisfy the so-called *recursivity property* [1, 3, 4, 17, 18, 19], which basically states that the entropy of a discrete stochastic system can be calculated either directly or by dividing the system into subsystems[1]. Let p be a probability distribution on $[n]$ and assume that there exists a partition $\{A_1, A_2\}$ of $[n]$ such that $p(A_1) \neq 0$ and $p(A_2) \neq 0$. Then, we have

$$H^{(n)}_S(p) = H^{(2)}_S(p^{[A_1][A_2]}) + p(A_1)\, H^{(a_1)}_S(\bar{p}^{A_1}) + p(A_2)\, H^{(a_2)}_S(\bar{p}^{A_2}), \quad (6)$$

where $p^{[A_1][A_2]}$ is a probability distribution on the set $\{[A_1], [A_2]\}$ defined by $p^{[A_1][A_2]}([A_i]) = p(A_i)$, $i = 1, 2$, and where $\bar{p}^{A_1}$ and $\bar{p}^{A_2}$ are the normalized probability distributions on A_1 and A_2 respectively obtained from p, that is,

$$\bar{p}^{A_i}(j) = \frac{p(j)}{p(A_i)} \qquad (j \in A_i;\, i = 1, 2).$$

As a generalization of the Shannon entropy, we require that our measure of uncertainty $H^{(n)}(\mu)$ satisfy a similar property that would thus reflect the possibility to decompose *in an additive way* the calculation of the uncertainty of the discrete Choquet capacity $\mu \in \mathcal{F}_{[n]}$. Under certain conditions and when μ is cardinality-based, such a decomposition already exists if axiom SE is satisfied. To demonstrate this, let us consider a cardinality-based Choquet capacity

[1] Note that the recursivity property is important because it implies amongst other things the additivity of the Shannon entropy, i.e.,

$$H^{(n^2)}_S(p * q) = H^{(n)}_S(p) + H^{(n)}_S(q),$$

for all probability distributions p, q on $[n]$ where $p * q$ denotes the distribution

$$(p_1q_1, \ldots, p_1q_n, \ldots, p_nq_1, \ldots, p_nq_n).$$

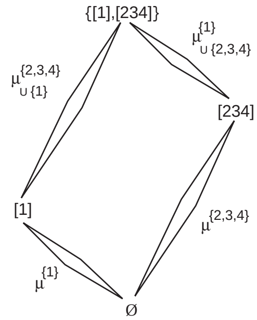

Figure 2: Decomposition resulting from a partition of $[n] = \{1, 2, 3, 4\}$ into the subsets $\{1\}$ and $\{2, 3, 4\}$.

$\mu \in \mathcal{F}_{[n]}$ and a partition of $[n]$ into two subsets A_1 and A_2. The Choquet capacities that appear from such a decomposition are:

- the reduced Choquet capacity $\mu^{[A_1][A_2]}$ on $\{[A_1], [A_2]\}$, which is not necessarily cardinality-based,
- the Choquet capacities on A_1: $\bar{\mu}^{A_1}$ and $\bar{\mu}^{A_1}_{\cup A_2}$, which are cardinality-based,
- and the Choquet capacities on A_2: $\bar{\mu}^{A_2}$ and $\bar{\mu}^{A_2}_{\cup A_1}$, which are cardinality-based as well.

For instance, when $[n] = \{1, 2, 3, 4\}$, Figure 2 shows the decomposition resulting from considering $A_1 = \{1\}$ and $A_2 = \{2, 3, 4\}$.

Assume now that n is even so that the subsets A_1 and A_2 can be chosen to have the same cardinal $n/2$. Then, we have $\bar{\mu}^{A_1} = \bar{\mu}^{A_2}$, $\bar{\mu}^{A_1}_{\cup A_2} = \bar{\mu}^{A_2}_{\cup A_1}$, and the reduced Choquet capacity $\mu^{[A_1][A_2]}$ is cardinality-based. According to axiom SE and Eq. (6), the following functional equation holds

$$H^{(n)}(\mu) = H^{(2)}(\mu^{[A_1][A_2]}) + \mu^{A_1}(A_1)\, H^{(a_1)}(\bar{\mu}^{A_1}) + \mu^{A_2}_{\cup A_1}(A_2)\, H^{(a_2)}(\bar{\mu}^{A_2}_{\cup A_1}). \tag{7}$$

The following question then arises: How could we generalize the previous functional equation to situations when the subsets A_1 and A_2 do not have the same cardinal anymore? Indeed, for a general choice of n and of the subsets A_1 and A_2, the Choquet capacities $\bar{\mu}^{A_1}$ and $\bar{\mu}^{A_1}_{\cup A_2}$ as well as the capacities $\bar{\mu}^{A_2}$ and $\bar{\mu}^{A_2}_{\cup A_1}$ are not necessarily equal. Furthermore, the reduced Choquet capacity $\mu^{[A_1][A_2]}$ is not cardinality-based anymore. As an extension of the previous case, we then require that the following functional equation, which is a generalization of (7) and is still in accordance with our intuitive additivity requirement, holds for any cardinality-based Choquet capacity $\mu \in \mathcal{F}_{[n]}$:

$$\begin{aligned} &H^{(n)}(\mu) = H^{(2)}(\mu^{[A_1][A_2]}) \\ &+ \alpha^1\, \mu^{A_1}(A_1)\, H^{(a_1)}(\bar{\mu}^{A_1}) + \alpha^1_2\, \mu^{A_1}_{\cup A_2}(A_1)\, H^{(a_1)}(\bar{\mu}^{A_1}_{\cup A_2}) \\ &+ \alpha^2\, \mu^{A_2}(A_2)\, H^{(a_2)}(\bar{\mu}^{A_2}) + \alpha^2_1\, \mu^{A_2}_{\cup A_1}(A_2)\, H^{(a_2)}(\bar{\mu}^{A_2}_{\cup A_1}), \end{aligned}$$

where α^1, α^1_2, α^2, and α^2_1 are real numbers left undetermined thus far.

By generalizing the previous functional equation to any partition $\{A_1, \ldots, A_k\}$ of $[n]$ into k subsets, we obtain our third axiom:

- *Recursivity axiom (R)*: For any integers $n \geq k \geq 2$, there exists a family of real coefficients $\{\alpha^i_S(n, k) \mid i \in [k],\ S \subseteq [k] \setminus i\}$ such that, for any partition $\{A_1, \ldots, A_k\}$ of $[n]$ and any cardinality-based capacity $\mu \in \mathcal{F}_{[n]}$,

$$\begin{aligned} &H^{(n)}(\mu) = H^{(k)}(\mu^{[A_1]\cdots[A_k]}) \\ &+ \sum_{i=1}^{k} \sum_{S \subseteq [k] \setminus i} \alpha^i_S(n, k)\, \mu^{A_i}_{\cup_{j \in S} A_j}(A_i)\, H^{(a_i)}(\bar{\mu}^{A_i}_{\cup_{j \in S} A_j}). \end{aligned} \tag{8}$$

3.3 Characterization

We can now state our main result.

Theorem 3.1 *The sequence $H^{(n)} : \mathcal{F}_{[n]} \to \mathbb{R}$ $(n \geq 2)$ fulfills axioms S, SE, and R, if and only if*

$$H^{(n)} = H^{(n)}_M \qquad (n \geq 2).$$

The proof of this theorem can be found in an article downloadable on http://www.univ-reunion.fr/~ikojadin/entropy.html.

References

[1] J. Aczél and Z. Daróczy. *On measures of information and their characterizations.* Academic Press, New York–San Francisco–London, 1975.

[2] G. Choquet. Theory of capacities. *Annales de l'Institut Fourier*, 5:131–295, 1953.

[3] B. Ebanks, P. Sahoo, and W. Sander. *Characterizations of information measures.* World Scientific, Singapore, 1997.

[4] D.K. Fadeev. Zum begriff der entropie einer endlichen wahrscheinlichkeitschemes. *Arbeit zur Informationstheorie, Deutscher Verlag der Wissenschaften*, 1, 1957.

[5] M. Grabisch. Alternative representations of discrete fuzzy measures for decison making. *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems*, 5(5):587–607, 1997.

[6] M. Grabisch, J.-L. Marichal, and M. Roubens. Equivalent representations of set functions. *Mathematics of Operations Research*, 25(2):157–178, 2000.

[7] M. Grabisch and J-M. Nicolas. Classification by fuzzy integrals: Performance and test. *Fuzzy Sets and Systems*, 65:255–271, 1994.

[8] M. Grabisch and M. Roubens. An axiomatic approach to the concept of interaction among players in cooperative games. *Int. J. Game Theory*, 124:547–565, 1999.

[9] E. T. Jaynes. Information theory and statistical mechanics. *Phys. Rev.*, 106:620–630, 1957.

[10] E. T. Jaynes. *Probability Theory: The Logic of Science.* Fragmentary Edition of March 1996, 1996.

[11] A.I. Khinchin. *Mathematical foundations of information theory.* Dover, 1957.

[12] S. Kullback and R. A. Leibler. On information and sufficiency. *Ann. Math. Stat.*, 22:79–86, 1951.

[13] J.-L. Marichal. Entropy of discrete choquet capacities. *European Journal of Operational Research.* to appear.

[14] J.-L. Marichal. *Aggregation operators for multicriteria decison aid.* PhD thesis, University of Liège, Liège, Belgium, 1998.

[15] J.-L. Marichal. An axiomatic approach of the discrete choquet integral as a tool to aggregate interacting criteria. *IEEE Transactions on Fuzzy Systems*, 8(6):800–807, 2000.

[16] J.-L. Marichal and M. Roubens. Entropy of discrete fuzzy measures. *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems*, 8(6):625–640, 2000.

[17] D. Morales, L. Pardo, and I. Vajda. Uncertainty of discrete stochastic systems: general theory and statistical theory. *IEEE Trans. on System, Man and Cybernetics*, 26(11):1–17, 1996.

[18] A. Rényi. On the foundations of information theory. *Review of the International Statistical Institute*, 33(1):1–14, 1965.

[19] C. E. Shannon. A mathematical theory of communication. *Bell Systems Technical Journal*, 27:379–423,623–656, 1948.

[20] L.S. Shapley. A value for n-person game. In H.W. Kuhn and A.W. Tucker, editors, *Contributions to the Theory of Games*, volume 2, pages 307–317. Princeton University Press, 1953.

[21] M. Sugeno. *Theory of fuzzy integrals and its applications.* PhD thesis, Tokyo Institute of Technology, Tokyo, Japan, 1974.