



HAL
open science

Entropy of capacities on lattices and set systems

Aoi Honda, Michel Grabisch

► **To cite this version:**

Aoi Honda, Michel Grabisch. Entropy of capacities on lattices and set systems. *Information Sciences*, 2006, 176 (23), pp.3472-3489. 10.1016/j.ins.2006.02.011 . hal-00179852

HAL Id: hal-00179852

<https://hal.science/hal-00179852v1>

Submitted on 13 Nov 2007

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Entropy of capacities on lattices and set systems

Aoi Honda ^{a,b,*} Michel Grabisch ^b

^a*Kyushu Institute of Technology*

^b*University Paris I-Panthéon-Sorbonne*

Abstract

We propose a definition for the entropy of capacities defined on lattices. Classical capacities are monotone set functions and can be seen as a generalization of probability measures. Capacities on lattices address the general case where the family of subsets is not necessarily the Boolean lattice of all subsets. Our definition encompasses the classical definition of Shannon for probability measures, as well as the entropy of Marichal defined for classical capacities. Some properties and examples are given.

Key words: entropy, capacity, lattice, regular set system, convex geometry, antimatroid

1 Introduction

The classical definition of Shannon for probability measures is at the core of information theory. Therefore, many attempts for defining an entropy for set functions more general than classical probability measures have been done, in particular for the so-called *capacities* [4] or *fuzzy measures* [21]. Roughly speaking, capacities are probability measures where the axiom of additivity has been replaced by a weaker one, monotonicity with respect to inclusion.

First definitions of an entropy for a capacity were proposed independently and approximately at the same time by Yager [22,23] and Marichal and Roubens [14,15,16]. The idea of Yager was to compute the Shannon entropy of the Shapley value of a capacity. To make the discussion more precise, let us consider

* corresponding author

Email address: aoi@ces.kyutech.ac.jp (Aoi Honda).

a finite universal set N , and a capacity v defined on it. The Shapley value [20] is a notion coming from cooperative game theory, and can be seen as a probability distribution ϕ over N which represents the average contribution of each element $i \in N$ in the value of v , that is, $v(S \cup i) - v(S)$, for all subsets S of $N \setminus i$. A slightly different proposition was done by Marichal and Roubens [16], just by changing the place of the function $h(x) := -x \log x$. It turned out that this definition seemed to be the right one, with properties close to the classical Shannon entropy [19]. In particular, it is strictly increasing towards the capacity which maximizes entropy. An important result, due to Dukhovny [6], and also independently found by Kojadinovic et al. [12], showed that the definition of Marichal and Roubens could be written as the average of classical entropy along maximal chains of the Boolean lattice of subsets of N .

In this paper, we consider yet more general functions than capacities, in the sense that the underlying system of sets may be not the whole collection of subsets of N , but only a part of it, provided that this collection forms a lattice. This is motivated partly by cooperative game theory, where N is the set of players, subsets are called coalitions, and the fact that all subsets may not belong to the set systems corresponds to the situation where some coalitions may be forbidden. This is considered for example by Faigle and Kern [7] (games with precedence constraints). Our approach will follow Dukhovny, in the sense that our basic material will be the maximal chains over the considered lattice, and we will try to make the least possible assumptions on the lattice in order that our construction works. This permits to consider our definition in a more abstract way, forgetting about the corresponding set system, and working only on the lattice. In this way, it is possible to consider as particular cases bi-cooperative games of Bilbao [2, Section 1.6], and multichoice games [11].

Section 2 recalls classical facts on Shannon's entropy and the definition of Marichal and Roubens, Section 3 gives the necessary material for lattices and convex geometries, while Section 4 introduces the notion of capacity on a lattice, viewed as a set system. Section 5 gives the definition of entropy for such capacities on lattices, and studies its properties. Section 6 gives examples of different lattices, so as to recover well known cases.

2 Entropy of classical capacity

Throughout this paper, we consider a finite universal set $N = \{1, 2, \dots, n\}$, and 2^N denotes the power set of N . Let us consider \mathfrak{S} a subcollection of 2^N . Then we call (N, \mathfrak{S}) (or simply \mathfrak{S} if no ambiguity occurs) a *set system*. In the following, (N, \mathfrak{S}) or simply \mathfrak{S} will always denote a set system.

Definition 1 (capacity) *Let (N, \mathfrak{S}) be a set system, with $\emptyset, N \in \mathfrak{S}$. (i) A*

game is a set function $v : \mathfrak{S} \rightarrow \mathbb{R}$ which satisfies $v(\emptyset) = 0$. (ii) A set function $v : \mathfrak{S} \rightarrow [0, 1]$ is a capacity if it satisfies that $v(\emptyset) = 0$, $v(N) = 1$, and $v(A) \leq v(B)$ whenever $A \subseteq B$.

Usually classical games and capacities are defined on $(N, 2^N)$.

Definition 2 (Shapley value) *The Shapley value of a capacity v is defined by*

$$\phi(v) := (\phi_1(v), \dots, \phi_n(v)) \in [0, 1]^n$$

and

$$\phi_i(v) := \sum_{A \subseteq N \setminus \{i\}} \gamma_{|A|}^n [v(A \cup \{i\}) - v(A)], \quad (2.1)$$

where

$$\gamma_k^n := \frac{(n-k-1)!k!}{n!}. \quad (2.2)$$

Remark that $\sum_{i=1}^n \phi_i(v) = 1$ holds.

Definition 3 (Shannon Entropy[19]) *Let p, q be probability measures on $(N, 2^N)$. The Shannon entropy of p and the relative entropy of p to q are defined by*

$$H_S(p) := \sum_{i=1}^n h[p_i],$$

$$H_S(p; q) := \sum_{i=1}^n h[p_i; q_i],$$

where $p_i := p(\{i\})$, $q_i := q(\{i\})$, $h(x) := -x \log x$, and $h(x; y) := x \log \frac{x}{y}$.

Here \log denote the base 2 logarithm and by convention $\log 0 := 0$.

Definition 4 (Marichal's entropy[16]) *Let v be a capacity on $(N, 2^N)$. Marichal's entropy of a capacity v is defined by*

$$H_M(v) := \sum_{i=1}^n \sum_{A \subseteq N \setminus \{i\}} \gamma_{|A|}^n h[v(A \cup \{i\}) - v(A)], \quad (2.3)$$

where γ_k^n is defined by (2.2).

Remark that equations (2.1) and (2.3) are similar. Dukhovny gives a representation of Marichal's entropy using maximal chains of 2^N [6].

Definition 5 (maximal chain of set system) Let (N, \mathfrak{S}) be a set system, with $\emptyset, N \in \mathfrak{S}$. If $C = (c_0, c_1, \dots, c_m)$ satisfies that $\emptyset = c_0 \subsetneq c_1 \subsetneq \dots \subsetneq c_m = N, c_i \in \mathfrak{S}$ and there is no element $c \in \mathfrak{S}$ such that $c_{i-1} \subsetneq c \subsetneq c_i$ for any $i \in \{1, \dots, m\}$ then we call C a maximal chain of \mathfrak{S} .

We denote the set of all maximal chains of \mathfrak{S} by $\mathcal{C}(\mathfrak{S})$. Let v be a capacity on \mathfrak{S} . Define $p^{v,C}$ by

$$\begin{aligned} p^{v,C} &:= (p_1^{v,C}, p_2^{v,C}, \dots, p_m^{v,C}) \\ &= (v(c_1) - v(c_0), v(c_2) - v(c_1), \dots, v(c_m) - v(c_{m-1})), \end{aligned} \quad (2.4)$$

where $C = (c_0, c_1, \dots, c_m) \in \mathcal{C}(\mathfrak{S})$. Note that $p^{v,C}$ is a probability distribution, i.e. $p_i^{v,C} \geq 0, i = 1, \dots, m$ and $\sum_{i=1}^m p_i^{v,C} = 1$. Dukhovny showed that Marichal's entropy can be represented as an average of Shannon entropies of all probabilities $p^{v,C}$ such that $C \in \mathcal{C}(2^N)$:

$$H_M(v) = \frac{1}{n!} \sum_{C \in \mathcal{C}(2^N)} H_S(p^{v,C}).$$

Remark that $|\mathcal{C}(2^N)| = n!$.

3 Lattices and related ordered structures

In this section, we investigate the relations between lattices and set systems. In particular we introduce a general class of sets systems called regular set systems, and also consider known classes of set systems called convex geometries and antimatroids.

Definition 6 (lattice) Let (L, \leq) be a partially ordered set, i.e. \leq is a binary relation on L being reflexive, antisymmetric and transitive. (L, \leq) is called a lattice if for all $x, y \in L$, the least upper bound $x \vee y$ and the greatest lower bound $x \wedge y$ of x and y exist.

Let L be a lattice. If $\bigvee S$ and $\bigwedge S$ exist for all $S \subseteq L$, then L is called a complete lattice. $\bigvee L$ and $\bigwedge L$ are called the top element and the bottom element of L and written \top and \perp , respectively. We denote a complete lattice by $(L, \leq, \vee, \wedge, \perp, \top)$. If L is a finite set, then L is a complete lattice.

The dual of a statement about lattices phrased in terms of \vee and \wedge is obtained by interchanging \vee and \wedge . If a statement about lattice is true, then the dual statement is also true. This fact is called the duality principle.

Definition 7 (\vee -irreducible element) An element $x \in (L, \leq)$ is \vee -irreducible if for all $a, b \in L$, $x \neq \perp$ and $x = a \vee b$ implies $x = a$ or $x = b$.

The dual of a \vee -irreducible element is called a \wedge -irreducible element, which satisfies that if for all $a, b \in L$, $x \neq \top$ and $x = a \wedge b$ implies $x = a$ or $x = b$. We denote the set of all \vee -irreducible elements of L by $\mathcal{J}(L)$ and the set of all \wedge -irreducible elements of L by $\mathcal{M}(L)$.

The mapping η for any $a \in L$, defined by

$$\eta(a) := \{x \in \mathcal{J}(L) \mid x \leq a\}$$

is a lattice-isomorphism of L onto $\eta(L) := \{\eta(a) \mid a \in L\}$, that is, $(L, \leq) \cong (\eta(L), \subseteq)$. Obviously $(\mathcal{J}(L), \eta(L))$ is a set system (see Section 6.1).

We say a is *covered* by b , and write $a \prec b$ or $b \succ a$, if $a < b$ and $a \leq x < b$ implies $x = a$.

Definition 8 (maximal chain of lattice) $C = (c_0, c_1, \dots, c_m)$ is a maximal chain of (L, \leq) if $c_i \in L, i = 0, \dots, m$, and $\perp = c_0 \prec c_1 \prec \dots \prec c_m = \top$.

We denote the set of all maximal chains of L by $\mathcal{C}(L)$.

We introduce the regular property for set systems.

Definition 9 (regular set system) Let (N, \mathfrak{S}) be a set system. We say that \mathfrak{S} is a regular set system if for any $C \in \mathcal{C}(\mathfrak{S})$, the length of C is n , i.e. $|C| = n + 1$.

Definition 10 (\vee -minimal regular) If (L, \leq) satisfies that the length of C is $|\mathcal{J}(L)|$, i.e. $|C| = |\mathcal{J}(L)| + 1$, for any $C \in \mathcal{C}(L)$ then we say that $(\mathcal{J}(L), \eta(L))$ is \vee -minimal regular.

Lemma 11 If (L, \leq) is \vee -minimal regular then $(\mathcal{J}(L), \eta(L))$ is a regular set system.

PROOF. Since $\eta(L)$ is isomorphic to L , for any $C \in \mathcal{C}(\mathcal{J}(L))$, $|C| = |\mathcal{J}(L)| + 1$ holds.

Lemma 12 If L is \vee -minimal regular, then for every maximal chains $C = (c_0, c_1, \dots, c_n)$, where $n = |\mathcal{J}(L)|$, it holds that $\eta(c_i) = \eta(c_{i-1}) \cup \{j\}$ for some $j \in \mathcal{J}(L)$.

PROOF. It suffices to show that $|\eta(c_i)| = i$. Suppose that there exists i_0 such that $|\eta(c_{i_0})| > i_0$. Since $|C| = n$ and $|\eta(c_i) \setminus \eta(c_{i-1})| \geq 1$ for any $i = 1, \dots, n$,

there will be not enough \vee -irreducible elements to complete the chain.

Definition 13 (convex geometry and antimatroid) *Let (N, \mathfrak{S}) be a set system. \mathfrak{S} is called a convex geometry of N if*

- (i) $\emptyset, N \in \mathfrak{S}$,
- (ii) for any $A, B \in \mathfrak{S}$, $A \cap B \in \mathfrak{S}$,
- (iii) for any $A \in \mathfrak{S} \setminus \{N\}$, there exists $i \in N \setminus A$ such that $A \cup \{i\} \in \mathfrak{S}$.

Let (N, \mathfrak{S}) be a convex geometry. The dual system of (N, \mathfrak{S}) defined by $\mathfrak{A} = \{N \setminus A \mid A \in \mathfrak{S}\}$, (N, \mathfrak{A}) is called antimatroid.

Following result can be found in [17][18]. We give a proof for the sake of completeness.

Lemma 14 *If (N, \mathfrak{S}) is a convex geometry or an antimatroid, then \mathfrak{S} is a regular set system.*

PROOF. Let \mathfrak{S} be a convex geometry. Suppose that there exists $C = (c_0, c_1, \dots, c_k) \in \mathcal{C}(\mathfrak{S})$ such that $|C| < n + 1$. Then we can take $c_i \in C$ which satisfies $|c_i \setminus c_{i-1}| > 1$. We have $c_{i-1} \subsetneq c_i \subseteq N$, and by (iii) of Definition 13, we can take $j_1, \dots, j_t \in N$ such that $c_{i-1} \cup \{j_1\}, c_{i-1} \cup \{j_1, j_2\}, \dots, c_{i-1} \cup \{j_1, \dots, j_t\} \in \mathfrak{S}$ and $c_{i-1} \cup \{j_1, \dots, j_t\} = N$, so that in these elements there exists an element c such that $|c \cap c_i| = |c_i| - 1$. By (ii), $c \cap c_i \in \mathfrak{S}$ and $c_{i-1} \subsetneq (c \cap c_i) \subsetneq c_i$, which contradicts the fact that C is maximal. Hence $|C| \geq n + 1$. On the other hand, obviously, for any $C \in \mathcal{C}(\mathfrak{S})$, $|C| \leq n + 1$, hence $|C| = n + 1$. And by the duality principle, the antimatroid is also a regular set system.

Convex geometries and antimatroids are complete lattices $(\mathfrak{S}, \subseteq, \vee, \cap, N, \emptyset)$ and $(\mathfrak{S}, \subseteq, \cup, \wedge, N, \emptyset)$, respectively, where $x \vee y := \cap\{z \in \mathfrak{S} \mid x \cup y \subseteq z\}$ and $x \wedge y := \cup\{z \in \mathfrak{S} \mid x \cap y \subseteq z\}$.

Lemma 15 *If (N, \mathfrak{S}) is a convex geometry, then $|\mathcal{J}(\mathfrak{S})| = n$. Similarly, if (N, \mathfrak{S}) is an antimatroid, then $|\mathcal{M}(\mathfrak{S})| = n$.*

PROOF.

Suppose that \mathfrak{S} is a convex geometry. By Lemma 14, for any $a \in \mathfrak{S}$, we have $a \setminus \underline{a} \in N$, where $\underline{a} \prec a$. And for any $b, c \in \mathcal{J}(\mathfrak{S})$ such that $b \neq c$, we have $b \setminus \underline{b} \neq c \setminus \underline{c}$, because when $\underline{b} = \underline{c}$, we have $b \setminus \underline{b} \neq c \setminus \underline{c}$ obviously, and when $\underline{b} \neq \underline{c}$, $b \setminus \underline{b} = c \setminus \underline{c}$ means $b \cap c \supseteq b \setminus \underline{b}$ and $b = (b \cap c) \cup \underline{b}$, which contradicts that b is a \vee -irreducible element, so that $|\mathcal{J}(\mathfrak{S})| \leq n$. On the other hand, for any chain $C = (c_0, \dots, c_n) \in \mathcal{C}(\mathfrak{S})$, $|\eta(c_i)| > |\eta(c_{i-1})|$ so that

$n \leq |\eta(\top)| = |\mathcal{J}(\mathfrak{S})|$. Therefore $|\mathcal{J}(\mathfrak{S})| = n$. By the duality principle, the same is true for antimatroids.

For example, \mathfrak{S}_1 in Fig. 1 is an antimatroid and a regular set system of $N = \{1, 2, 3\}$. $|\mathcal{J}(\mathfrak{S}_1)| = |\{1, 3, 12, 23\}| = 4$ and $|\mathcal{M}(\mathfrak{S}_1)| = |\{12, 13, 23\}| = 3$.

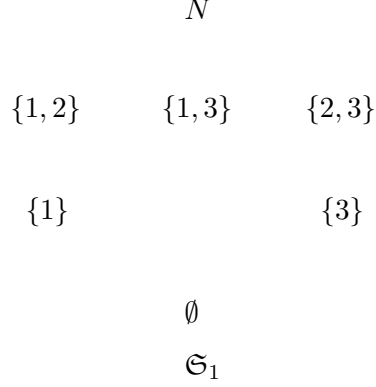


Fig. 1. Antimatroid

If \mathfrak{S} is a regular set system, it does not necessarily hold that $|\mathcal{J}(\mathfrak{S})| = n$ nor $|\mathcal{M}(\mathfrak{S})| = n$. Consider the lattice \mathfrak{S}_2 in Fig. 2. \mathfrak{S}_2 is a regular set system of $\{1, 2, 3\}$, but $\mathcal{J}(\mathfrak{S}) = \mathcal{M}(\mathfrak{S}) = \{1, 3, 12, 23\}$.

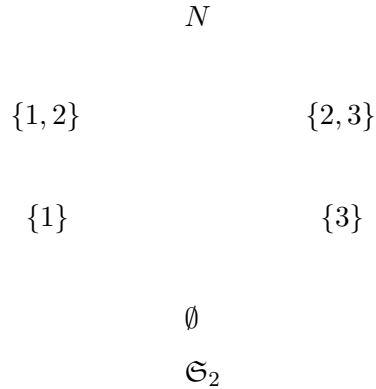


Fig. 2. Regular set system

Remark A segment $[a, b]$ of L , for $a, b \in L$, is the set of all elements x which satisfy $a \leq x \leq b$. If \mathfrak{S} is a regular set system then \mathfrak{S} satisfies the *Jordan-Dedekind chain condition*, that is, all maximal chains in any segments of \mathfrak{S} have the same length. The converse does not hold. For instance, $(N, \mathfrak{S}) = (\{1, 2, 3\}, \{\emptyset, \{1, 2\}, \{3\}, N\})$ satisfies the Jordan-Dedekind chain condition but is not a regular set system. Incidentally, \mathfrak{S} is \vee -minimal regular. Similarly, If L is a convex geometry or an antimatroid, then L satisfies the Jordan-Dedekind chain condition, but the converse does not hold. For instance, consider the lattice $(\mathfrak{S}, \subseteq) = (\{1, 2, 3, 4\}, \{\emptyset, \{1\}, \{1, 2\}, \{3\}, \{3, 4\}, N\})$

4 Capacity on lattice

Definition 16 (capacity on lattice) *A mapping $v : L \rightarrow [0, 1]$ is a capacity on L if it satisfies $v(\perp) = 0, v(\top) = 1$ and for any $x, y \in L$, $v(x) \leq v(y)$ whenever $x \leq y$.*

Definition 17 (cardinality-based capacity) *A capacity on (N, \mathfrak{S}) is cardinality-based if $v(A)$ depends only on $|A|$ for any $A \in \mathfrak{S}$.*

Definition 18 (additive uniform capacity) *The additive uniform capacity on (N, \mathfrak{S}) is defined by*

$$v^*(A) := \frac{|A|}{n}$$

for any $A \in \mathfrak{S}$.

Uniform capacities and the additive uniform capacity can be defined on any lattice L by putting $|x| := |\eta(x)|$ for any $x \in L$.

Faigle and Kern generalized the Shapley value to that of a game on a lattice [7], and Bilbao defined it for games on convex geometries [3] and on antimatroids [1].

Definition 19 (Bilbao and Edelman's Shapley value) *Let v be a game on a convex geometry or an antimatroid (N, \mathfrak{S}) . For $i \in N$, the Shapley value of v is defined by*

$$\phi_i(v) := \frac{1}{|\mathcal{C}(\mathfrak{S})|} \sum_{\substack{C \in \mathcal{C}(\mathfrak{S}) \\ A, A \cup \{i\} \in C}} (v(A \cup \{i\}) - v(A)). \quad (4.1)$$

When \mathfrak{S} is a regular set system of N , we can also define the Shapley value of games on \mathfrak{S} by (4.1).

By Lemma 11, regarding the lattice as a set system of $\mathcal{J}(L)$, we can also calculate the Shapley value of capacities on the regular lattice as follows.

Definition 20 (Shapley value on L (cf. [7])) *Suppose that (L, \leq) is \vee -minimal regular and let v be a capacity on L . For $x \in \mathcal{J}(L)$, the Shapley value of v on L is defined by*

$$\phi_x(v) := \frac{1}{|\mathcal{C}(L)|} \sum_{\substack{C \in \mathcal{C}(L) \\ \eta(c_i) \setminus \eta(c_{i-1}) = x}} (v(c_i) - v(c_{i-1}))$$

$$= \frac{1}{|\mathcal{C}(L)|} \sum_{\substack{C \in \mathcal{C}(L) \\ \eta(c_i) \setminus \eta(c_{i-1}) = x}} p_i^{v,C}, \quad (4.2)$$

where $C = (c_0, c_1, \dots, c_n)$ and $n = |\mathcal{J}(L)|$.

By Lemma 12, if L is \vee -minimal regular, for any $C \in \mathcal{C}(L)$, $\eta(a) \setminus \eta(b) \in \mathcal{J}(L)$ for any $a, b \in C$ such that $a \prec b$. Hence formulas (4.2) are well-defined. Similarly, if L satisfies the following property:

(\wedge -minimal regular) any $C \in \mathcal{C}(L)$, the length of C is $|\mathcal{M}(L)|$, i.e. $|C| = |\mathcal{M}(L)| + 1$,

for $x \in \mathcal{M}(L)$, we can calculate the Shapley value of capacities on L in a similar manner as follows. For $x \in \mathcal{M}(L)$, the Shapley value of v on L is defined by

$$\begin{aligned} \phi_x(v) &:= \frac{1}{|\mathcal{C}(L)|} \sum_{\substack{C \in \mathcal{C}(L) \\ \eta^d(c_{i-1}) \setminus \eta^d(c_i) = x}} (v(c_i) - v(c_{i-1})) \\ &= \frac{1}{|\mathcal{C}(L)|} \sum_{\substack{C \in \mathcal{C}(L) \\ \eta^d(c_{i-1}) \setminus \eta^d(c_i) = x}} p_i^{v,C}, \end{aligned}$$

where $C = (c_0, c_1, \dots, c_n)$, $n = |\mathcal{M}(L)|$ and $\eta^d(a) := \{x \in \mathcal{M}(L) \mid x \geq a\}$. We have $(L, \leq) \cong (\mathcal{M}(L), \eta^d(L))$.

If L is both \vee and \wedge -minimal regular, then we can use both $\mathcal{J}(L)$ and $\mathcal{M}(L)$ for calculating the Shapley value. However $\mathcal{J}(L)$ is better, because elements of $\mathcal{J}(L)$ are in general easier to interpret (cf. Section 6.5).

5 Entropy of capacities on lattices and set systems

In this section, we suppose that (N, \mathfrak{S}) is a regular set system and let v and u be capacities on (N, \mathfrak{S}) .

Definition 21 (entropy) *Let v be a capacity on \mathfrak{S} . The entropy of v is defined by*

$$H(v) := \frac{1}{|\mathcal{C}(\mathfrak{S})|} \sum_{C \in \mathcal{C}(\mathfrak{S})} H_S(p^{v,C}), \quad (5.1)$$

where $C = (c_0, c_1, \dots, c_n)$.

Definition 22 (relative entropy) *Let v and u be capacities on \mathfrak{S} . The relative entropy of v to u is defined by*

$$H(v; u) := \frac{1}{|\mathcal{C}(\mathfrak{S})|} \sum_{C \in \mathcal{C}(\mathfrak{S})} H_S(p^{v,C}; p^{u,C}). \quad (5.2)$$

Let v and u be capacities on L . If L is \vee or \wedge -minimal regular, then regarding L as a set system $(\mathcal{J}(L), \eta(L))$ or $(\mathcal{M}(L), \eta^d(L))$, we can also define the entropy $H(v)$ and the relative entropy $H(v; u)$ as follows.

$$H(v) := \frac{1}{|\mathcal{C}(L)|} \sum_{C \in \mathcal{C}(L)} H_S(p^{v,C}), \quad (5.3)$$

$$H(v; u) := \frac{1}{|\mathcal{C}(L)|} \sum_{C \in \mathcal{C}(L)} H_S(p^{v,C}; p^{u,C}). \quad (5.4)$$

We can consider that $H(v)$ is an average of Shannon entropies, and also that $H(v; u)$ is an average of Shannon relative entropies. Therefore they satisfy several properties which are required for entropies (cf. [12]).

Proposition 23 *For any v , $H(v)$ is a continuous function, and $0 \leq H(v) \leq \log n$, with equality on left side if and only if v is $\{0, 1\}$ -valued capacity, and with equality on right side if and only if v is the additive uniform capacity v^* .*

PROOF. The continuity is obvious. For any probability p , $H_S(p) \geq 0$, so that $H(v) \geq 0$ holds. $H_S(p) = 0$ if and only if p is deterministic, i.e. there exists i such that $p_i = 1$ and otherwise $p_j = 0$. Hence for all $C \in \mathcal{C}(\mathfrak{S})$, $p_i^{v,C}$ takes value only 0 or 1, which means that for all $A \in \mathfrak{S}$, $v(A)$ takes value only 0 or 1. Similarly, $H_S(p) \leq \log n$, so that an average of $H_S(p)$ is dominated by $\log n$. $H_S(p) = \log n$ if and only if for all i , $p_i = 1/n$, hence for all $a \in \mathfrak{S}$, $v(A) = |A|/n$, which completes the proof.

Proposition 24 *For any uniform capacity v on \mathfrak{S} , we have*

$$H(v) = H_S(p^{v,C})$$

for any $C \in \mathcal{C}(\mathfrak{S})$.

PROOF. In this case, for all $C \in \mathcal{C}(\mathfrak{S})$, $p^{v,C}$ is the same probability distribution, hence we have

$$\begin{aligned}
H(v) &= \frac{1}{|\mathcal{C}(\mathfrak{S})|} \sum_{C \in \mathcal{C}(\mathfrak{S})} H_S(p^{v,C}) \\
&= \frac{1}{|\mathcal{C}(\mathfrak{S})|} |\mathcal{C}(\mathfrak{S})| H_S(p^{v,C}) = H_S(p^{v,C}).
\end{aligned}$$

Define $v_\lambda := (1 - \lambda)v + \lambda v^*$ for $0 < \lambda < 1$. Then for any $v (\neq v^*)$, $H(v_\lambda)$ is strictly increasing toward the additive uniform capacity v^* .

Proposition 25 *For any $v (\neq v^*)$, $H(v_\lambda)$ is a strictly increasing function of λ .*

PROOF. We show that $\frac{dH_S(p^{v_\lambda,C})}{d\lambda} > 0$ for any $C \in \mathcal{C}(\mathfrak{S})$ such that $p^{v,C} \neq p^*$.

$$\begin{aligned}
H_S(p^{v_\lambda,C}) &= \sum_{i=1}^n h[v_\lambda(c_i) - v_\lambda(c_{i-1})] \\
&= \sum_{i=1}^n h \left[p_i^{v,C} + \lambda \left(\frac{1}{n} - p_i^{v,C} \right) \right],
\end{aligned}$$

where $C = (c_0, \dots, c_n)$ and $p_i^{v,C} := v(c_i) - v(c_{i-1})$.

$$\begin{aligned}
\frac{dH_S(p_i^{v_\lambda,C})}{d\lambda} &= \sum_{i=1}^n \left(p_i^{v,C} - \frac{1}{n} \right) \left(1 + \log \left(p_i^{v,C} + \lambda \left(\frac{1}{n} - p_i^{v,C} \right) \right) \right) \\
&= \sum_{i=1}^n \left(p_i^{v,C} - \frac{1}{n} \right) \left(\log \left(p_i^{v,C} + \lambda \left(\frac{1}{n} - p_i^{v,C} \right) \right) \right)
\end{aligned}$$

If $1/n \geq p_i^{v,C}$, then

$$p_i^{v,C} + \lambda \left(\frac{1}{n} - p_i^{v,C} \right) \in \left[p_i^{v,C}, \frac{1}{n} \right)$$

and otherwise, that is, $1/n < p_i^{v,C}$, we have

$$p_i^{v,C} + \lambda \left(\frac{1}{n} - p_i^{v,C} \right) \in \left(\frac{1}{n}, p_i^{v,C} \right),$$

so that we have

$$\frac{dH_S(p_i^{v_\lambda,C})}{d\lambda} > \sum_{i:1/n \geq p_i^{v,C}} \left(\frac{1}{n} - p_i^{v,C} \right) \log \frac{1}{n} + \sum_{i:1/n < p_i^{v,C}} \left(\frac{1}{n} - p_i^{v,C} \right) \log \frac{1}{n}$$

$$= \log \frac{1}{n} \sum_{i=1}^n \left(\frac{1}{n} - p_i^{v,C} \right) = \log \frac{1}{n} \left(1 - \sum_{i=1}^n p_i^{v,C} \right) = 0$$

Since $v \not\equiv v^*$, there exist at least an $C \in \mathcal{C}(\mathfrak{S})$ such that $p^{v,C} \not\equiv p^{v^*,C}$, therefore

$$H(v) = \frac{1}{|\mathcal{C}(\mathfrak{S})|} \sum_{C \in \mathcal{C}(\mathfrak{S})} H_S(p_i^{v\lambda,C})$$

is a strictly increasing function of λ .

Proposition 26 $H(v; u) \geq 0$ and that $H(v; u) = 0$ if and only if $v \equiv u$.

PROOF. Non-negativity is obvious by $H_S(p; q) \geq 0$. And $H(v; u) = 0$ if and only if $H_S(p^{v,C}; p^{u,C}) = 0$ for all $C \in \mathcal{C}(\mathfrak{S})$, which is true if and only if $p^{v,C} \equiv p^{u,C}$ for all $C \in \mathcal{C}(\mathfrak{S})$, which means $v \equiv u$.

Proposition 27 Let $v \not\equiv u$ and $v_\lambda^u := \lambda u + (1 - \lambda)v$. Then $H(v_\lambda^u; u)$ is a strictly decreasing function of λ .

PROOF. We show that

$$\frac{dH_S(p^{v_\lambda^u,C}; p^{u,C})}{d\lambda} < 0$$

for any $C \in \mathcal{C}(\mathfrak{S})$ such that $p^{v_\lambda^u,C} \not\equiv p^{u,C}$.

$$\begin{aligned} H(p^{v_\lambda^u,C}; p^{u,C}) &= \sum_{i=1}^n p_i^{v_\lambda^u,C} \log \frac{p_i^{v_\lambda^u,C}}{p_i^{u,C}} \\ &= \sum_{i=1}^n h \left[\lambda(p_i^{u,C} - p_i^{v,C}) + p_i^{v,C}; p_i^{u,C} \right] \end{aligned}$$

$$\frac{dH(p^{v_\lambda^u,C}; p^{u,C})}{d\lambda} = (p_i^{u,C} - p_i^{v,C}) \left(\log \frac{\lambda(p_i^{u,C} - p_i^{v,C}) + p_i^{v,C}}{p_i^{u,C}} + 1 \right).$$

If $p_i^{u,C} \geq p_i^{v,C}$, then

$$\frac{\lambda(p_i^{u,C} - p_i^{v,C}) + p_i^{v,C}}{p_i^{u,C}} \in \left[\frac{p_i^{v,C}}{p_i^{u,C}}, 1 \right),$$

and otherwise, that is, $p_i^{u,C} < p_i^{v,C}$

$$\frac{\lambda(p_i^{v,C} - p_i^{v,C}) + p_i^{v,C}}{p_i^{u,C}} \in \left(1, \frac{p_i^{v,C}}{p_i^{v,C}}\right),$$

so that we have

$$\begin{aligned} \frac{dH(p^{v,\lambda}; p^{u,C})}{d\lambda} &< \sum_{i: p_i^{u,C} \geq p_i^{v,C}} (p_i^{u,C} - p_i^{v,C}) + \sum_{i: p_i^{u,C} < p_i^{v,C}} (p_i^{u,C} - p_i^{v,C}) \\ &= \sum_{i=1}^n p_i^{u,C} - \sum_{i=1}^n p_i^{v,C} = 0 \end{aligned}$$

Since $v \neq u$, there exists at least un $C \in \mathcal{C}(\mathfrak{S})$ such that $p^{v,C} \neq p^{u,C}$, therefore

$$H(v; u) = \frac{1}{|\mathcal{C}(\mathfrak{S})|} \sum_{C \in \mathcal{C}(L)} H_S(p^{v,C}; p^{u,C})$$

is a strictly decreasing function of λ .

6 Examples

In this section, we show several examples. Most games and capacities which appear in applications are particular capacities on regular set systems.

6.1 Regular lattice

L_1 in Fig. 3 is \vee -minimal regular, and is also isomorphic to a convex geometry.

a	$\{d, e, f\}$				
b	c	$\{d, e\}$	$\{e, f\}$		
d	e	f	$\{d\}$	$\{e\}$	$\{f\}$
g	\emptyset				
L_1	$\eta(L_1)$				

Fig. 3.

In fact, $\mathcal{J}(L_1) = \{d, e, f\}$, and L_1 is also represented by $\eta(L_1)$. $\mathcal{C}(\eta(L_1)) = \{(\emptyset, d, de, def), (\emptyset, e, de, def), (\emptyset, e, ef, def), (\emptyset, f, ef, def)\}$. Let v be a capacity on L_1 . Then the Shapley values and the entropy of v on L_1 are as follows.

$$\begin{aligned}\phi_d(v) &= \frac{1}{4}(v(d) - v(g)) + \frac{1}{4}(v(b) - v(e)) + \frac{1}{2}(v(a) - v(c)) \\ \phi_e(v) &= \frac{1}{2}(v(e) - v(g)) + \frac{1}{4}(v(b) - v(d)) + \frac{1}{4}(v(c) - v(f)) \\ \phi_f(v) &= \frac{1}{4}(v(f) - v(g)) + \frac{1}{4}(v(c) - v(e)) + \frac{1}{2}(v(a) - v(b))\end{aligned}$$

and

$$\begin{aligned}H(v) &= \frac{1}{4}h[v(d) - v(g)] + \frac{1}{4}h[v(b) - v(e)] + \frac{1}{2}h[v(a) - v(c)] \\ &\quad + \frac{1}{2}h[v(e) - v(g)] + \frac{1}{4}h[v(b) - v(d)] + \frac{1}{4}h[v(c) - v(f)] \\ &\quad + \frac{1}{4}h[v(f) - v(g)] + \frac{1}{4}h[v(c) - v(e)] + \frac{1}{2}h[v(a) - v(b)].\end{aligned}$$

6.2 Distributive lattice

(L, \leq) is said to be distributive if it satisfies the distributive law, $a \wedge (b \vee c) = (a \wedge b) \vee (a \wedge c)$ for any $a, b, c \in L$. If (L, \leq) is distributive then (L, \leq) is also \vee and \wedge -minimal regular. Remark that a regular set system, even the convex geometry and the antimatroid are not necessarily distributive (cf. Fig. 2, Fig. 3).

6.3 Capacity on 2^N (classical capacity)

The classical capacity is a monotone function on the Boolean lattice 2^N . 2^N is a distributive lattice and also a complemented lattice, i. e. for any $A \in 2^N$, there exists a complement $B \in 2^N$ such that $A \wedge B = \perp = \emptyset$ and $A \vee B = \top = N$. For any capacity on 2^N , (4.1) is equals to the Shapley value (2.1), and our entropies (5.1) and (5.2) are equal to Marichal's entropy (2.3) (cf. Section 2).

6.4 Bi-capacity [8][9]

A bi-capacity is a monotone function on $\mathcal{Q}(N) := \{(A, B) \in 2^N \times 2^N \mid A \cap B = \emptyset\}$ which satisfies that $v(\emptyset, N) = -1$, $v(\emptyset, \emptyset) = 0$ and $v(N, \emptyset) = 1$.

For any $(A_1, A_2), (B_1, B_2) \in \mathcal{Q}(N)$, $(A_1, A_2) \sqsubseteq (B_1, B_2)$ iff $A_1 \subseteq B_1$ and $A_2 \supseteq B_2$. $\mathcal{Q}(N) \cong 3^N$. It can be shown that $(\mathcal{Q}(N), \sqsubseteq)$ is a finite distributive lattice. Sup and inf are given by $(A_1, A_2) \vee (B_1, B_2) = (A_1 \cup B_1, A_2 \cap B_2)$ and $(A_1, A_2) \wedge (B_1, B_2) = (A_1 \cap B_1, A_2 \cup B_2)$, and we have

$$\mathcal{J}(\mathcal{Q}(N)) = \{(\emptyset, N \setminus \{i\}), i \in N\} \cup \{(\{i\}, N \setminus \{i\}), i \in N\},$$

where $i \in N$. Normalizing v by $v' : \mathcal{Q}(N) \rightarrow [0, 1]$ such that

$$v' := \frac{1}{2}v + \frac{1}{2},$$

we can regard v as a capacity on $\mathcal{Q}(N)$. Then, applying (4.2) and (5.3), we have

$$\begin{aligned} \phi_i^+(v') &:= \phi_{(\{i\}, N \setminus \{i\})}(v') \\ &= \sum_{\substack{A \subseteq N \setminus \{i\} \\ B \subseteq N \setminus (A \cup \{i\})}} \gamma_{|A|, |B|}^n (v'(A \cup \{i\}, B) - v'(A, B)), \\ \phi_i^-(v') &:= \phi_{(\emptyset, N \setminus \{i\})}(v') \\ &= \sum_{\substack{A \subseteq N \setminus \{i\} \\ B \subseteq N \setminus (A \cup \{i\})}} \gamma_{|A|, |B|}^n (v'(B, A) - v'(B, A \cup \{i\})) \end{aligned}$$

and

$$\begin{aligned} H(v') &= \sum_{i=1}^n \sum_{\substack{A \subseteq N \setminus x_i \\ B \subseteq N \setminus (A \cup \{i\})}} \gamma_{|A|, |B|}^n (h[v'(A \cup \{i\}, B) - v'(A, B)] \\ &\quad + h[v'(B, A) - v'(B, A \cup \{i\})]). \end{aligned}$$

where $\gamma_{k, \ell}^n := \frac{(n - k + \ell - 1)! (n + k - \ell)! 2^{n-k-\ell}}{(2n)!}$, and $h(x) := -x \log x$.

ϕ_i^+ and ϕ_i^- mean positive and negative degrees of i 's contribution to v , respectively, hence the contribution of i to v is given by $\phi_i(v) := \phi_i^+(v) + \phi_i^-(v)$. $\gamma_{|A|, |B|}^n$ is the rate of the number of chains which contain $(A \cup \{i\}, B)$ and (A, B) . In fact,

$$\begin{aligned} &|\{C \in \mathcal{C}(\mathcal{Q}(N)) \mid C \ni (A \cup \{i\}, B), (A, B)\}| \\ &= \frac{(n + |A| - |B|)!}{(2!)^{|A|}} \cdot \frac{(n - |A| + |B| - 1)!}{(2!)^{|B|}} \end{aligned}$$

and $|\mathcal{C}(\mathcal{Q}(N))| = (2n)!/(2!)^n$. These Shapley values are different from those in [8].

6.5 Multichoice game

Multichoice games have been proposed by Hsiao and Raghavan [11]. They have been proposed also independently in the context of capacities by Grabisch and Labreuche [10], under the name k -ary capacities.

Let $N := \{0, 1, \dots, n\}$ be a set of players, and let $L := L_1 \times \dots \times L_n$, where (L_i, \leq_i) is a totally ordered set $L_i = \{0, 1, \dots, \ell_i\}$ such that $0 \leq_i 1 \leq_i \dots \leq_i \ell_i$. Each L_i is the set of choices of player i . (L, \leq) is a regular lattice. For any $(a_1, a_2, \dots, a_n), (b_1, b_2, \dots, b_n) \in L$, $(a_1, a_2, \dots, a_n) \leq (b_1, b_2, \dots, b_n)$ iff $a_i \leq_i b_i$ for all $i = 1, \dots, n$. We have

$$\mathcal{J}(L) = \{(0, \dots, 0, a_i, 0, \dots, 0) \mid a_i \in \mathcal{J}(L_i) = L_i \setminus \{0\}\}$$

and $|\mathcal{J}(L)| = \sum_{i=1}^n \ell_i$. The lattice in Fig. 4 is an example of a product lattice, which represents a 2-players game. Players 1 and 2 can choose among 3 and

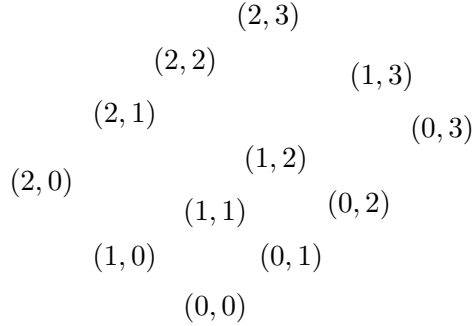


Fig. 4. 2-players game

4 choices. Let v be a capacity on L , that is, $v(0, \dots, 0) = 0$, $v(\ell_1, \dots, \ell_n) = 1$ and, for any $a, b \in L$, $v(a) \leq v(b)$ whenever $a \leq b$. In this case, applying (4.2) and (5.3), we have

$$\begin{aligned} \phi_i^j(v) &= \phi_{(0, \dots, 0, a_i=j > 0, 0, \dots, 0)}(v) \\ &= \sum_{a \in L/L_i} \xi_i^{(a,j)} (v(a, j) - v(a, j-1)) \end{aligned}$$

and

$$H(v) = \sum_{\substack{i \in N \\ j \in L_i}} \sum_{a \in L/L_i} \xi_i^{(a,j)} h[v(a, j) - v(a, j-1)]$$

where $L/L_i := L_1 \times \dots \times L_{i-1} \times L_{i+1} \times \dots \times L_n$, $(a, a_i) := (a_1, \dots, a_{i-1}, a_i, a_{i+1}, \dots, a_n) \in$

L such that $a \in L/L_i$ and $a_i \in L_i$, and

$$\xi_i^{(a,a_i)} := \left(\prod_{k=1}^n \binom{\ell_k}{a_k} \right) \cdot \left(\frac{\sum_{k=1}^n \ell_k}{\sum_{k=1}^n a_k} \right)^{-1} \cdot \frac{a_i}{\sum_{k=1}^n a_k}$$

and $h(x) := -x \log x$.

$\phi_i^j(v)$ represents the contribution of player i playing at level j compared to level $j-1$, where $j, j-1 \in \mathcal{J}(L_i) = L_i \setminus \{0\}$, hence player i 's overall contribution is given by

$$\phi_i(v) = \sum_{j=1}^{\ell_i} \phi_i^j(v).$$

$\xi_i^{(a,a_i)}$ is the rate of the number of chains which contain (a, a_i) and $(a, a_i - 1)$. In fact,

$$|\{C \in \mathcal{C}(L) \mid C \ni (a, a_i), (a, a_i - 1)\}| = \frac{(\sum_{k=1}^n a_k - 1)!}{(\prod_{k=1}^n (a_k!))(a_i - 1)!/(a_i!)} \cdot \frac{(\sum_{k=1}^n (\ell_k - a_k))!}{\prod_{k=1}^n ((\ell_k - a_k)!)}$$

and $|\mathcal{C}(L)| = (\sum_{k=1}^n \ell_k)! / \prod_{k=1}^n (\ell_k!)$.

Regarding a bi-capacity in Section 6.4 as a special case of multichoice game such that n players and $\ell_i = 2$ for all i which is fixed a value $v'(\emptyset, \emptyset) = 1/2$, we obtain the same Shapley values and the entropy.

7 Conclusion

We have proposed a general definition of entropy for capacities defined on a large class of ordered structures we call regular set systems, which encompasses the original definition of Marichal for classical capacities. Regular set systems contain as particular important classes, distributive lattices, convex geometries and antimatroids. Hence our approach permits to define the entropy of multichoice games, also called k -ary capacities.

References

- [1] E. Algaba, J.M. Bilbao, R. van den Brink and A. Jiménez-Losada, Cooperative games on antimatroids, *Discrete Mathematics* 282 (2004), 1–15.
- [2] J.M. Bilbao, *Cooperative games on combinatorial structures*, Kluwer Academic Publishers, Boston, 2000.
- [3] J.M. Bilbao and P.H. Edelman, The Shapley value on convex geometries, *Discrete Applied Mathematics* 103 (2000), 33–40.

- [4] G. Choquet, Theory of capacities, *Annales de l'Institut Fourier* 5 (1953), 131–295.
- [5] B.A. Davey and H.A. Priestley, *Introduction to lattices and orders*, Cambridge University Press, 1900.
- [6] A. Dukhovny, General entropy of general measures, *Internat. J. Uncertain. Fuzziness Knowledge-Based Systems* 10 (2002), 213–225.
- [7] U. Faigle and W. Kern, The Shapley value for cooperative games under precedence constraints, *Int. J. of Game Theory* 21 (1992), 249–266.
- [8] M. Grabisch and Ch. Labreuche, Bi-capacities — Part I: definition, Mobius transform and interaction, *Fuzzy Sets and Systems* 151 (2005), 211–236.
- [9] M. Grabisch and Ch. Labreuche, Bi-capacities — Part II: the Choquet integral, *Fuzzy Sets and Systems* 151 (2005), 237–259.
- [10] M. Grabisch and Ch. Labreuche, Capacities on lattices and k -ary capacities, *3d Int. Conf. of the European Soc. for Fuzzy Logic and Technology (EUSFLAT 2003)* (2003), Zittau, Germany, 304–307.
- [11] C.R. Hsiao and T.E.S. Raghavan, Shapley value for multichoice cooperative games, I, *Games and Economic Behavior* 5 (1993), 240–256.
- [12] I. Kojadinovic, J.-L. Marichal and M. Roubens, An axiomatic approach to the definition of the entropy of a discrete Choquet capacity, *Information Sciences* 172 (2005), 131–153.
- [13] J.-L. Marichal, Aggregation Operations for Multicriteria Decision Aid. Ph.D. Thesis, Department of Mathematics, University of Liege, Liege, Belgium, 1998.
- [14] J.-L. Marichal, Entropy of discrete Choquet capacities, *Eur. J. of Oper. Res.*, 137 (2002), 612–624.
- [15] J.-L. Marichal and M. Roubens, Entropy of a Choquet capacity, *Proc. of 1999 Eusflat-Estylf Joint Conference, Palma de Mallorca, Spain* (1999), 383–385.
- [16] J.-L. Marichal and M. Roubens, Entropy of discrete fuzzy measure, *Internat. J. Uncertain. Fuzziness Knowledge-Based Systems* 8 (2000), 625–640.
- [17] B. Monjardet, The consequences of Dilworth's work on lattice with unique irreducible decompositions, In K. Bogart, R. Freese and J. Kung (eds.), *The Dilworth theorems, selected papers of Robert P. Dilworth*, 192–200, Birkhäuser, (1990).
- [18] B. Monjardet, The presence of lattice theory in discrete problems of mathematical social sciences. Why. *Math. Soc. Sci.* 46 (2003), 103–144.
- [19] C.E. Shannon, A mathematical theory of communication, *Bell System Tech. Journ.* 27 (1948), 374–423, 623–656.
- [20] L.S. Shapley, A value for n -person games, in: H.W. Kuhn and A.W. Tucker (eds.), *Contributions to the Theory of Games*, Vol. II, Annals of Mathematics Studies 28 (Princeton University Press, 1953), 307–317.

- [21] M. Sugeno, Theory of fuzzy integrals and its applications, *Phd. Thesis of Tokyo Institute of Technology* 8 (1974), 453–461.
- [22] R.R. Yager, A class of fuzzy measures generated from a Dempster-Shafer belief structure, *Int. J. of Intelligent Systems* 14 (12) (1999) 1239-1247.
- [23] R.R. Yager, On the entropy of fuzzy measures, *IEEE Transaction on Fuzzy Systems* 8 (2000), 453–461.