Entropy of capacities on lattices

Aoi Honda\textsuperscript{a,}\textsuperscript{*} Michel Grabisch\textsuperscript{b}

\textsuperscript{a}Kyushu Institute of Technology
\textsuperscript{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, 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\textsuperscript{4} or fuzzy measures\textsuperscript{17}. Roughly speaking, capacities are probability measures where the axiom of additivity has been replaced by a weaker one, monotonicity with respect to inclusion.

The first idea, due to Yager\textsuperscript{18}, was to define the entropy of a capacity by the Shannon entropy of the Shapley value of this capacity. To make the discussion more precise, let us consider a finite universal set \( N \), and a capacity \( \nu \) defined on it. The Shapley value\textsuperscript{16} is a notion coming from cooperative game theory, and can be seen as a probability distribution \( \phi \) over \( N \) which

\textsuperscript{*} corresponding author

Email address: aoi@ces.kyutech.ac.jp (Aoi Honda).
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 [12], 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 [15]. In particular, it is strictly increasing towards the capacity which maximizes entropy. An important result, due to Dukhovny [6], 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 et al. [2], and multichoice games [10].

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.

Throughout this paper, we consider as universal set a finite set \( N = \{1, \ldots, n\} \), and \( 2^N \) is the set of all its subsets.

## 2 Entropy of classical capacity

A game is a set function \( v : 2^N \to \mathbb{R} \) which satisfies \( v(\emptyset) = 0 \). A normalized monotone game is called a capacity.

**Definition 1 (classical capacity)** A set function \( v : 2^N \to [0,1] \) is a ca-
capacity if it satisfies that \( v(\emptyset) = 0, v(N) = 1 \), and \( v(A) \preceq v(B) \) whenever \( A \subseteq B \subseteq N \).

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

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

and

\[
\phi_i(v) := \sum_{A \subseteq X \setminus \{i\}} \gamma_{n-k}^n [v(A \cup \{i\}) - v(A)],
\]  

(2.1)

where

\[
\gamma_{k}^n := \frac{(n-k-1)!k!}{n!}.
\]  

(2.2)

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

**Definition 3 (Shannon Entropy[15])** 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 logarithms and by convention \( \log 0 := 0 \).

**Definition 4 (Marichal’s entropy[12])** Let \( v \) be a capacity on \( (N, 2^N) \). Marichal’s entropy of a capacity \( v \) is defined by

\[
H_M(v) := \sum_{i=1}^n \sum_{A \subseteq X \setminus \{i\}} \gamma_{n-k}^n h[v(A \cup \{i\}) - v(A)],
\]  

(2.3)

where \( \gamma_{k}^n \) is defined by (2.2).

Remark that equations (2.1) and (2.3) are similar. Dukhovny gave a representation of Marichal’s entropy using maximal chains of \( 2^N \) [6]. Let \( \mathcal{N} \subseteq 2^N \), then \( (N, \mathcal{N}) \) is called a set system.
Definition 5 (maximal chain of set system) Let \((N, \mathcal{N})\) be a set system, with \(\emptyset, N \in \mathcal{N}\). If \(C = (c_0, c_1, \ldots, c_m)\) satisfies that \(\emptyset \subseteq c_0 \subseteq c_1 \subseteq \cdots \subseteq c_m = N, c_i \in \mathcal{N}\) and there is no element \(c \in \mathcal{N}\) such that \(c_{i-1} \subseteq c \subseteq c_i\) for any \(i \in \{1, \ldots, m\}\) then we call \(C\) a maximal chain of \(\mathcal{N}\).

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

\[
p^{C, v} := (p_1^{v, C}, p_2^{v, C}, \ldots, p_m^{v, C}) = (v(c_1) - v(c_0), v(c_2) - v(c_1), \ldots, v(c_m) - v(c_{m-1})),
\]

(2.4)

where \(C = (c_0, c_1, \ldots, c_m) \in \mathcal{C}(\mathcal{N})\). Note that \(p^{v, C}\) is a probability distribution, i.e. \(p_i^{v, C} \geq 0, i = 1, \ldots, m\) 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

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 \(x \lor y\) and the greatest lower bound \(x \land y\) of \(x\) and \(y\) exist.

Let \(L\) be a lattice. If \(\lor S\) and \(\land S\) exist for all \(S \subseteq L\), then \(L\) is called a complete lattice. \(\lor L\) and \(\land L\) are called the top element and the bottom element of \(L\) and written \(\top\) and \(\bot\), respectively. We denote a complete lattice by \((L, \leq, \lor, \land, \top, \bot)\). If \(L\) is a finite set, then \(L\) is a complete lattice.

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

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

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

The mapping $\eta$ 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.

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

**Definition 8 (maximal chain of lattice)** $C = (c_0, c_1, \ldots, c_m)$ is a maximal chain of $(L, \leq)$ if $c_i \in L, i = 0, \ldots, m$, and $\bot = c_0 < c_1 < \cdots < c_m = \top$.

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

**Lemma 10** If $(L, \leq)$ satisfies the following property:

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

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 11** If $L$ is \lor\text{-}minimal regular, then for every maximal chains $C = (c_0, c_1, \ldots, c_n)$, 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 $i = 1, \ldots, n$, there will be not enough \lor\text{-}irreducible elements to complete the chain.

**Definition 12 (convex geometry and antimatroid)** Let $(N, \mathcal{N})$ be a set system. $\mathcal{N}$ is called a convex geometry of $N$ if
(i) \( \emptyset, N \in \mathcal{N} \),
(ii) for any \( A, B \in \mathcal{N} \), \( A \cap B \in \mathcal{N} \),
(iii) for any \( A \in \mathcal{N} \setminus \{N\} \), there exists \( i \in N \setminus A \) such that \( A \cup \{i\} \in \mathcal{N} \).

The dual set system of the convex geometry is called antimatroid.

**Lemma 13** If \( (N, \mathcal{N}) \) is a convex geometry or an antimatroid, then \( \mathcal{N} \) is a regular set system.

This result can be found in [13][14]. We give a proof for the sake of completeness.

**PROOF.** Let \( \mathcal{N} \) be a convex geometry. Suppose that there exists \( C = (c_0, c_1, \ldots, c_k) \in \mathcal{C}(\mathcal{N}) \) such that \( |C| < n + 1 \). Then we can take \( c_i \in C \) which satisfies \( |c_i \setminus c_{i-1}| > 1 \). We have \( c_{i-1} \subseteq c_i \subseteq N \), and by (iii) of Definition 12, we can take \( j_1, \ldots, j_t \in N \) such that \( c_i \subseteq c_{i-1} \cup \{j_1\}, c_{i-1} \cup \{j_1, j_2\}, \ldots, c_{i-1} \cup \{j_1, \ldots, j_t\} \in \mathcal{N} \) and \( c_{i-1} \cup \{j_1, \ldots, j_t\} = N \), so that in these elements there exists an element \( c \) such that \( |c \cap c_i| = |c| - 1 \). By (ii), \( c \cap c_i \in \mathcal{N} \) and \( c_{i-1} \subseteq (c \cap c_i) \subseteq c_i \), which contradicts the fact that \( C \) is maximal. Hence \( |C| \geq n + 1 \). On the other hand, obviously, for any \( C \in \mathcal{C}(\mathcal{N}) \), \( |C| \leq n + 1 \), hence \( |C| = n + 1 \). And by the duality principle, the antimatroid is also a regular set system.

The convex geometry and the antimatroid are complete lattices \( (\mathcal{N}, \subseteq, \cup, \wedge, N, \emptyset) \) and \( (\mathcal{N}, \subseteq, \vee, \cap, N, \emptyset) \), where \( x \vee y := \bigcap\{z \in \mathcal{N} \setminus x \cup y \subseteq z \} \) and \( x \wedge y := \bigcup\{z \in \mathcal{N} \setminus x \cap y \subseteq z \} \), respectively.

**Lemma 14** If \( (N, \mathcal{N}) \) is a convex geometry, then \( |\mathcal{J}(\mathcal{N})| = n \). Similarly, if \( (N, \mathcal{N}) \) is an antimatroid, then \( |\mathcal{M}(\mathcal{N})| = n \).

**PROOF.**

Suppose that \( \mathcal{N} \) is a convex geometry. By Lemma 13, for any \( a \in \mathcal{N} \), we have \( a \setminus a \in \mathcal{N} \), where \( a \prec a \). And for any \( b, c \in \mathcal{J}(\mathcal{N}) \) such that \( b \neq c \), we have \( b \setminus b \neq c \setminus c \), because when \( b = c \), we have \( b \setminus b \neq c \setminus c \) obviously, and when \( b \neq c \), \( b \setminus b = c \setminus c \) means \( b \cap c \supseteq b \setminus b \) and \( b = (b \cap c) \cup b \), which contradicts that \( b \) is a \( \vee \)-irreducible element, so that \( |\mathcal{J}(\mathcal{N})| \leq n \). On the other hand, for any chain \( C = (c_0, \ldots, c_n) \in \mathcal{C}(\mathcal{N}) \), \( |\eta(c_0)| \geq |\eta(c_{n-1})| \) so that \( n \leq |\eta(\top)| = |\mathcal{J}(\mathcal{N})| \). Therefore \( |\mathcal{J}(\mathcal{N})| = n \). By the duality principle, the same is true for antimatroids.
For example, \( \mathcal{N}_1 \) in Fig. 1 is an antimatroid and a regular set system of \( N = \{1, 2, 3\} \). \( |\mathcal{F}(\mathcal{N}_1)| = |\{1, 3, 12, 23\}| = 4 \) and \( |\mathcal{M}(\mathcal{N}_1)| = |\{12, 13, 23\}| = 3 \).

![Diagram of an antimatroid](image1)

Fig. 1. Antimatroid

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

![Diagram of a regular set system](image2)

Fig. 2. Regular set system

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

Definition 15 (capacity on lattice) A mapping \( v : L \rightarrow [0,1] \) is a capacity on \( L \) if it satisfies \( v(\bot) = 0, v(\top) = 1 \) and for any \( x, y \in L \), \( v(x) \leq v(y) \) whenever \( x \leq y \).

Definition 16 (uniform capacity) A capacity on \((N, \mathcal{N})\) is uniform if \( v(A) \) depends only on \( |A| \) for any \( A \in \mathcal{N} \).

Definition 17 (additive uniform capacity) The additive uniform capacity on \((N, \mathcal{N})\) is defined by

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

for any \( A \in \mathcal{N} \).

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 18 (Bilbao and Edelman’s Shapley value) Let \( v \) be a game on a convex geometry or an antimatroid \((N, \mathcal{N})\). For \( i \in N \), the Shapley value of \( v \) is defined by

\[
\hat{\phi}_i(v) := \frac{1}{|\mathcal{C}(\mathcal{N})|} \sum_{\substack{A \in \mathcal{C}(\mathcal{N}) \setminus \{\{i\}\} \in \mathcal{C} \setminus \{\{i\}\}}} (v(A \cup \{i\}) - v(A)). \tag{4.1}
\]

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

By Lemma 10, 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 19 (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

\[
\phi_x(v) := \frac{1}{|\mathcal{C}(L)|} \sum_{\substack{C \in \mathcal{C}(L) \setminus \{\emptyset\} \cap \mathcal{C}(L) \setminus \{\{x\}\} \setminus \{\{x\}\}}} (v(C) - v(C \setminus \{x\})).
\]
\[
\frac{1}{|\mathcal{C}(L)|} \sum_{C \in \mathcal{C}(L) \atop \eta(c_3) \neq \eta(c_2) = \eta(c_1) = x} p^B_C,
\]

(4.2)

where \( C = (c_0, c_1, \ldots, c_n) \) and \( n = |J(L)|. \)

By Lemma 11, 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 \( a < b. \) 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

\[
\phi_x(v) := \frac{1}{|\mathcal{C}(L)|} \sum_{C \in \mathcal{C}(L) \atop \eta(c_3) = \eta(c_2) = \eta(c_1) = x} (v(c_i) - v(c_{i-1}))
\]

\[
= \frac{1}{|\mathcal{C}(L)|} \sum_{C \in \mathcal{C}(L) \atop \eta(c_3) = \eta(c_2) = \eta(c_1) = x} p^B_C,
\]

where \( C = (c_0, c_1, \ldots, c_n), n = |\mathcal{M}(L)| \) 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 lattice

Let \( (N, \mathcal{N}) \) be a regular set system, and let \( v \) and \( u \) be capacities on \( \mathcal{N} \).

Definition 20 (entropy) Let \( v \) be a capacity on \( \mathcal{N}. \) The entropy of \( v \) is defined by

\[
H(v) := \frac{1}{|\mathcal{C}(\mathcal{N})|} \sum_{C \in \mathcal{C}(\mathcal{N})} H_S(p^B_C),
\]

(5.1)

where \( C = (c_0, c_1, \ldots, c_n) \) and \( h(x) := -x \log x. \)
Definition 21 (relative entropy) Let $v$ and $u$ be capacities on $\mathcal{N}$. The relative entropy of $v$ to $u$ is defined by

$$H(v; u) := \frac{1}{|\mathcal{C}(\mathcal{N})|} \sum_{C \in \mathcal{C}(\mathcal{N})} H_S(p^{vC}; p^{uC}).$$

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

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

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

We can consider that $H(v)$ is an average of Shannon entropies, and also that $H(u; v)$ is an average of Shannon relative entropies. Therefore they satisfy several properties which are required for entropies (cf. [11]).

Proposition 22 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 $C \in \mathcal{C}(\mathcal{N})$, $p^{vC}$ takes value only 0 or 1, which means that for all $A \in \mathcal{N}$, $v(A)$ 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 $a \in \mathcal{N}$, $v(A) = |A|/n$, which completes the proof.

Proposition 23 For any uniform capacity $v$ on $\mathcal{N}$, we have

$$H(v) = H_S(p^{vC})$$

for any $C \in \mathcal{C}(\mathcal{N})$.\[\]

PROOF. In this case, for all $C \in \mathcal{C}(\mathcal{N})$, $p^{vC}$ is the same probability distribution, hence we have
\[
H(v) = \frac{1}{|\mathcal{C}(\mathcal{N})|} \sum_{C \in \mathcal{C}(\mathcal{N})} H_S(p^v_C)
= \frac{1}{|\mathcal{C}(\mathcal{N})|} \left|\mathcal{C}(\mathcal{N})\right| H_S(p^v_C) = H_S(p^v_C).
\]

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

**Proof.** We show that \( \frac{dH_S(p^{v_\lambda}_C)}{d\lambda} > 0 \) for any \( C \in \mathcal{C}(\mathcal{N}) \) such that \( p^{v_\lambda}_C \neq p^v \).

\[
H_S(p^{v_\lambda}_C) = \sum_{i=1}^n h[v_\lambda(c_i) - v_\lambda(c_{i-1})]
= \sum_{i=1}^n h\left[ p_i^{v_\lambda}_C + \lambda \left( \frac{1}{n} - p_i^{v_\lambda}_C \right) \right],
\]

where \( C = (c_0, \ldots, c_n) \) and \( p_i^{v_\lambda}_C := v(c_i) - v(c_{i-1}) \).

\[
\frac{dH_S(p^{v_\lambda}_C)}{d\lambda} = \sum_{i=1}^n \left( p_i^{v_\lambda}_C - \frac{1}{n} \right) \left( 1 + \log \left( p_i^{v_\lambda}_C + \lambda \left( \frac{1}{n} - p_i^{v_\lambda}_C \right) \right) \right) \\
= \sum_{i=1}^n \left( p_i^{v_\lambda}_C - \frac{1}{n} \right) \left( \log \left( p_i^{v_\lambda}_C + \lambda \left( \frac{1}{n} - p_i^{v_\lambda}_C \right) \right) \right)
\]

If \( 1/n \geq p_i^{v_\lambda}_C \), then
\[
p_i^{v_\lambda}_C + \lambda \left( \frac{1}{n} - p_i^{v_\lambda}_C \right) \in \left[ p_i^{v_\lambda}_C, \frac{1}{n} \right]
\]

and otherwise, that is, \( 1/n < p_i^{v_\lambda}_C \), we have
\[
p_i^{v_\lambda}_C + \lambda \left( \frac{1}{n} - p_i^{v_\lambda}_C \right) \in \left( \frac{1}{n}, p_i^{v_\lambda}_C \right),
\]

so that we have

\[
\frac{dH_S(p^{v_\lambda}_C)}{d\lambda} > \sum_{i: 1/n \geq p_i^{v_\lambda}_C} \left( \frac{1}{n} - p_i^{v_\lambda}_C \right) \log \frac{1}{n} + \sum_{i: 1/n < p_i^{v_\lambda}_C} \left( \frac{1}{n} - p_i^{v_\lambda}_C \right) \log \frac{1}{n}
\]
\[
\frac{1}{n} \sum_{i=1}^{n} \left( \frac{1}{n} - p_{i}^{v\mathcal{C}} \right) = \log \frac{1}{n} \left( 1 - \sum_{i=1}^{n} p_{i}^{v\mathcal{C}} \right) = 0
\]

Since \( v \not\equiv v^{u} \), there exist at least an \( C \in \mathcal{C}(\mathcal{N}) \) such that \( p^{v\mathcal{C}} \not\equiv p^{v\mathcal{C}} \), therefore

\[
H(v) = \frac{1}{|\mathcal{C}(\mathcal{N})|} \sum_{C \in \mathcal{C}(\mathcal{N})} H_{S}(p_{i}^{v\mathcal{C}})
\]

is a strictly increasing function of \( \lambda \).

**Proposition 25**  \( H(v; u) \geq 0 \) 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 \( H_{S}(p^{v\mathcal{C}}; p^{v\mathcal{C}}) = 0 \) for all \( C \in \mathcal{C}(\mathcal{N}) \), which is true if and only if \( p^{v\mathcal{C}} \equiv p^{v\mathcal{C}} \) for all \( C \in \mathcal{C}(\mathcal{N}) \), which means \( v \equiv u \).

**Proposition 26**  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 \( \lambda \).

**PROOF.** We show that

\[
\frac{dH_{S}(p^{v\mathcal{C}}; p^{u\mathcal{C}})}{d\lambda} < 0
\]

for any \( C \in \mathcal{C}(\mathcal{N}) \) such that \( p^{v\mathcal{C}} \not\equiv p^{v\mathcal{C}} \).

\[
H(p^{v\mathcal{C}}; p^{u\mathcal{C}}) = \sum_{i=1}^{n} p_{i}^{v\mathcal{C}} \log \frac{p_{i}^{v\mathcal{C}}}{p_{i}^{u\mathcal{C}}}
\]

\[
= \sum_{i=1}^{n} h \left( \lambda(p_{i}^{v\mathcal{C}} - p_{i}^{u\mathcal{C}}) + p_{i}^{v\mathcal{C}} ; p_{i}^{u\mathcal{C}} \right)
\]

\[
\frac{dH(p^{v\mathcal{C}}; p^{u\mathcal{C}})}{d\lambda}(p_{i}^{u\mathcal{C}} - p_{i}^{v\mathcal{C}}) \left( \log \frac{\lambda(p_{i}^{u\mathcal{C}} - p_{i}^{v\mathcal{C}}) + p_{i}^{v\mathcal{C}}}{p_{i}^{u\mathcal{C}}} + 1 \right).
\]

If \( p_{i}^{u\mathcal{C}} \geq p_{i}^{v\mathcal{C}} \), then

\[
\frac{\lambda(p_{i}^{u\mathcal{C}} - p_{i}^{v\mathcal{C}}) + p_{i}^{v\mathcal{C}}}{p_{i}^{u\mathcal{C}}} \in \left[ \frac{p_{i}^{u\mathcal{C}}}{p_{i}^{v\mathcal{C}}}, 1 \right],
\]

and otherwise, that is, \( p_{i}^{u\mathcal{C}} < p_{i}^{v\mathcal{C}} \)

\[
\frac{\lambda(p_{i}^{u\mathcal{C}} - p_{i}^{v\mathcal{C}}) + p_{i}^{v\mathcal{C}}}{p_{i}^{u\mathcal{C}}} \in \left( 1, \frac{p_{i}^{v\mathcal{C}}}{p_{i}^{u\mathcal{C}}} \right),
\]

12
so that we have

\[
\frac{dH(p_{\mu}^{u}; p_{\mu}^{v})}{d\lambda} < \sum_{i \in p_{\nu}^{u} \geq p_{\mu}^{u}} (p_{i}^{\mu} - p_{i}^{u}) + \sum_{i \in p_{\nu}^{u} < p_{\mu}^{u}} (p_{i}^{\mu} - p_{i}^{v}) \\
= \sum_{i=1}^{n} p_{i}^{\mu} - \sum_{i=1}^{n} p_{i}^{v} = 0
\]

Since \( v \not\equiv u \), there exists at least un \( C \in \mathcal{C}(\mathcal{N}) \) such that \( p_{\nu}^{u} \not\equiv p_{\mu}^{v} \), therefore

\[
H(v; u) = \frac{1}{|\mathcal{C}(\mathcal{N})|} \sum_{C \in \mathcal{C}(\mathcal{L})} H_{S}(p_{\nu}^{u}; p_{\mu}^{v})
\]

is a strictly decreasing function of \( \lambda \).

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.

![Diagram of a regular lattice](image)

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

13
\[ \phi_d(v) = \frac{1}{4} (v(d) - v(g)) + \frac{1}{4} (v(b) - v(e)) + \frac{1}{2} (v(a) - v(c)) \]

\[ \phi_c(v) = \frac{1}{2} (v(e) - v(g)) + \frac{1}{4} (v(b) - v(d)) + \frac{1}{4} (v(c) - v(f)) \]

\[ \phi_f(v) = \frac{1}{4} (v(f) - v(g)) + \frac{1}{4} (v(c) - v(e)) + \frac{1}{2} (v(a) - v(b)) \]

and

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

6.2 Distributive lattice

\((L, \leq)\) is said to be distributive if it satisfies the distributive law, \(a \land (b \lor c) = (a \land b) \lor (a \land c)\) for any \(a, b, c \in L\). If \((L, \leq)\) is distributive then \((L, \leq)\) is also \(\lor\) and \(\land\)-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 \(A \land B = \bot = \emptyset\) and \(A \lor B = \top = N\). For any capacity on \(2^N\), (4.1) is equal 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 \(Q(N) := \{(A, B) \in 2^N \times 2^N \mid A \cap B = \emptyset\}\) which satisfies that \(v(\emptyset, N) = -1, v(\emptyset, \emptyset) = 0\) and \(v(N, \emptyset) = -1\). For any \((A_1, A_2), (B_1, B_2) \in Q(N)\), \((A_1, A_2) \sqsubseteq (B_1, B_2)\) iff \(A_1 \subseteq B_1\) and \(A_2 \supseteq B_2\), \(Q(N) \cong 3^N\). It can be shown that \((Q(N), \sqsubseteq)\) is a finite distributive lattice. Sup and inf are given by \((A_1, A_2) \lor (B_1, B_2) = (A_1 \cup B_1, A_2 \cap B_2)\) 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 \(v' : \mathcal{Q}(N) \to [0, 1]\) 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

\[
\phi_i^+(v') := \phi(\emptyset, N \setminus \{i\}) \frac{v'}{\mathcal{P}(A)}
\]

\[
= \sum_{A \subseteq N \setminus \{i\}} \sum_{B \subseteq N \setminus \{i\}} \gamma^{n}_{[A,B]} \left( v'(A \cup \{i\}, B) - v'(A, B) \right),
\]

\[
\phi_i^-(v') := \phi(\emptyset, N \setminus \{i\}) \frac{v'}{\mathcal{P}(A)}
\]

\[
= \sum_{A \subseteq N \setminus \{i\}} \sum_{B \subseteq N \setminus \{i\}} \gamma^{n}_{[A,B]} \left( v'(B, A) - v'(B, A \cup \{i\}) \right)
\]

and

\[
H(v') = \sum_{i=1}^{n} \sum_{A \subseteq N \setminus \{i\}} \sum_{B \subseteq N \setminus \{i\}} \gamma^{n}_{[A,B]} \left( h \left[ v'(A \cup \{i\}, B) - v'(A, B) \right] + h \left[ v'(B, A) - v'(B, A \cup \{i\}) \right] \right).
\]

where \(\gamma^{n}_{[A,B]} := \frac{(n-k+\ell-1)! (n+k-\ell)! 2^{n-k-\ell}}{(2n)!}\), and \(h(x) := -x \log x\).

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

\[
|\{C \in \mathcal{C}(\mathcal{Q}(N)) \mid C \ni (A \cup \{i\}, B), (A, B)\}| = \frac{(n+|A|-|B|)! (n-|A|+|B|-1)!}{(2!)^{|A|} (2!)^{|B|}}
\]

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

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

\[
\mathcal{J}(L) = \{(0, \ldots, 0, a_i, 0, \ldots, 0) \mid a_i \in \mathcal{J}(L_i) = L_i \setminus \{0\}\}
\]

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 4 choices. Let \( v \) be a capacity on \( L \), that is, \( v(0, \ldots, 0) = 0 \), \( v(\ell_1, \ldots, \ell_n) = 1 \) 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

\[
\phi_i^J(v) = \phi_{(0, \ldots, 0, a_i = j > 0, \ldots, 0)}(v)
= \sum_{a \in L/L_i} \xi_i^{(a, j)} (v(a, j) - v(a, j - 1))
\]

and

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

where \( L/L_i := L_1 \times \cdots \times L_{i-1} \times L_{i+1} \times \cdots \times L_n \), \((a, a_i) := (a_1, \ldots, a_{i-1}, a_i, a_{i+1}, \ldots, a_n) \in L \) 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 \( h(x) := -x \log x \).

16
\( \phi_i^j(v) \) 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,

\[
|\{C \in \mathcal{C}(L) | C \ni (a, a_i), (a, a_i - 1)\}| = \frac{(\sum_{k=1}^{\ell} a_k - 1)!}{(\prod_{k=1}^{\ell} (a_k!))} \frac{(\sum_{k=1}^{\ell} (\ell_k - a_k))!}{(\prod_{k=1}^{\ell} ((\ell_k - a_k)!))}
\]

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

Regarding a bi-capacity in Section 6.4 as a special case of multi-choice game such that \( n \) players and \( \ell_i = 2 \) for all \( i \) which is fixed a value \( v'(0, 0) = 1/2 \), we obtain the same Shapley values and the entropy.

References


