An Abstract Axiomatization of the Notion of Entropy

Ivo Rosenberg and Dan Simovici

rosenb@dms.umontreal.ca dsim@cs.umb.edu

Université de Montrèal
Dépt. de Mathématiques
Montreal, Quebec,
Canada

University of Massachusetts Boston,
Department of Computer Science,
Boston, Massachusetts,
USA

Axiomatics of Entropy:

- Khinchin (1957)
- Fadeev (1956)
- Ingarden and Urbanik (1962)
- Rényi (1959)
- Daróczy (1970)
- Devijer (1974)

Previous Results

Simovici and Jaroszewicz (2002, Transactions of Information Theory): axiomatics of entropy based on the notion of partition:

- makes use of the properties of the semimodular lattice of the partitions of finite sets;
- captures a broad variety of entropies (Shannon's entropy and Gini index among them);
- connects entropy (via partitions) to relational databases, and
- allows development of several information-theory-based algorithms in data mining.

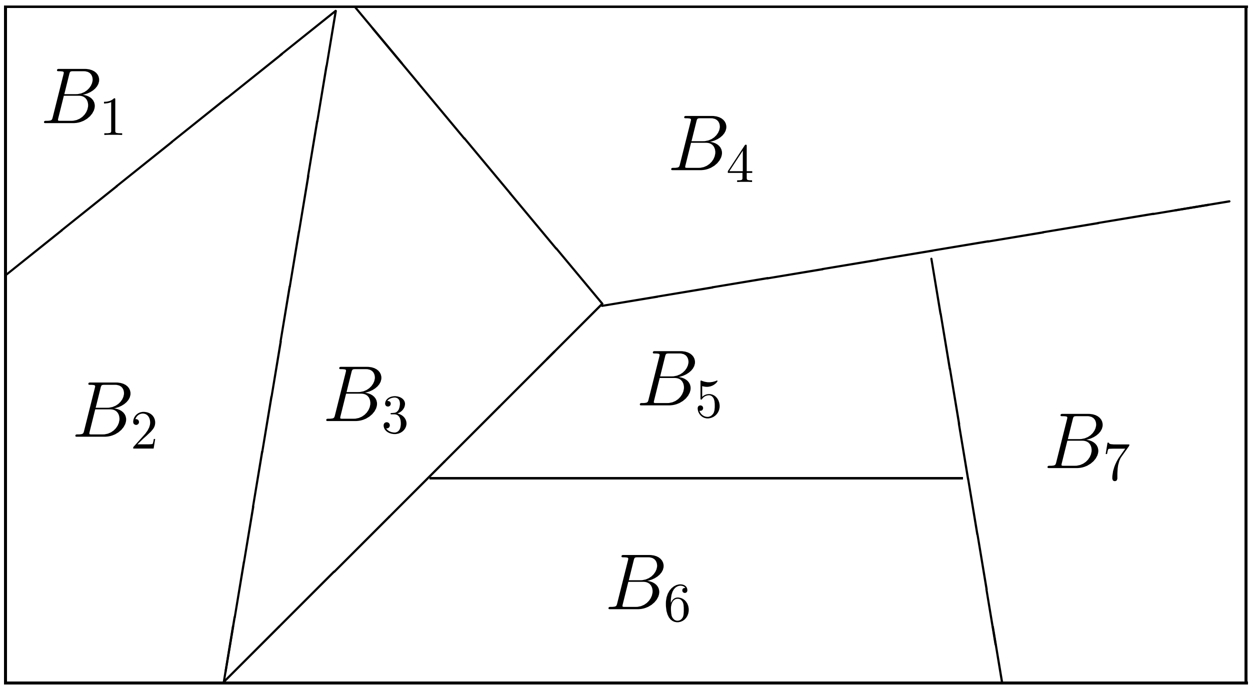
Entropy and MVL

It was shown that the entropy associated with the kernel partition of a finite function gives an evaluation of the minimum energy dissipated by a circuit implementing that function (within a certain technology).

Partitions

PART(S): set of partitions of set S

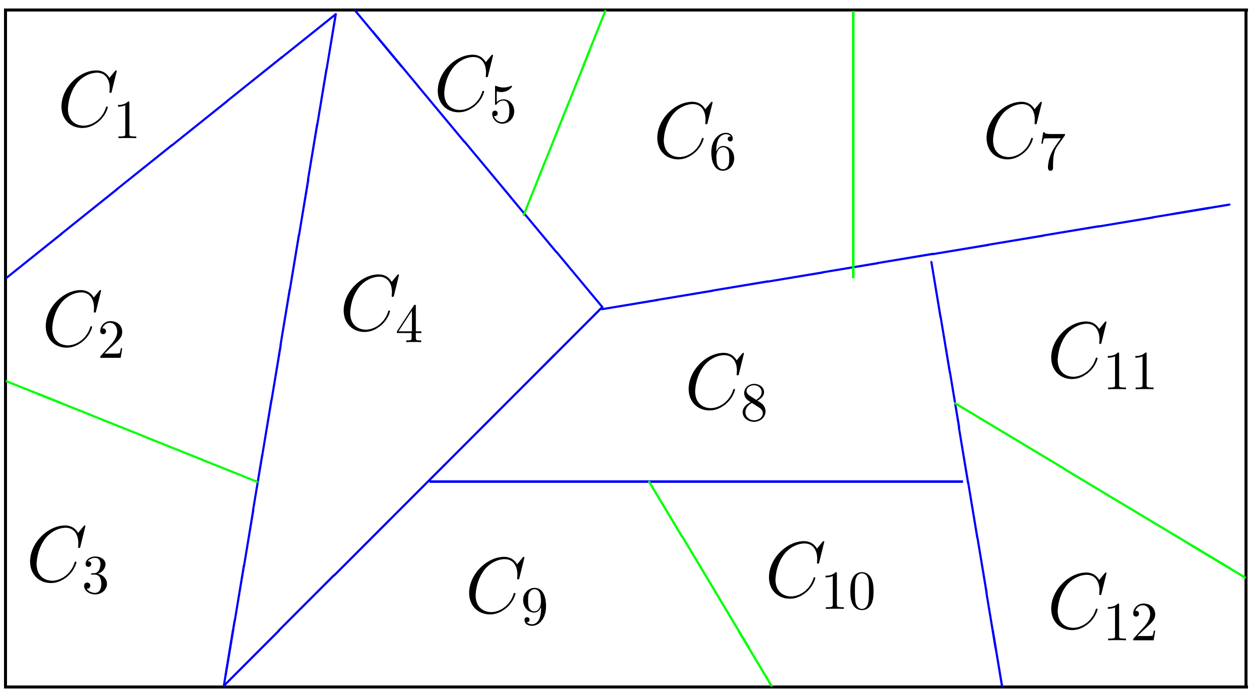
Partition $\pi = \{B_1, \ldots, B_7\}$



Partitions Partial Order

$\sigma \leq \pi$ if each block C of σ is included in a block of π.

$$\text{Partition } \sigma = \{C_1, \ldots, C_{12}\} \leq \pi$$



Operations involving partitions

- If S, T are two disjoint sets, $\pi \in \mathsf{PART}(S), \sigma \in \mathsf{PART}(T)$, $\pi = \{B_i \,|\, i \in I\}$, and $\sigma = \{C_j \,|\, j \in J\}$, then

$$\pi + \sigma = \{B_i \,|\, i \in I\} \cup \{C_j \,|\, j \in J\}$$

  is a partition on $S \cup T$.

- For any sets S, T, $\pi \in \mathsf{PART}(S), \sigma \in \mathsf{PART}(T)$, and $\pi = \{B_i \,|\, i \in I\}$, and $\sigma = \{C_j \,|\, j \in J\}$, the partition $\pi \times \sigma \in \mathsf{PART}(S \times T)$ is

$$\pi \times \sigma = \{B_i \times C_j \,|\, i \in I \text{ and } j \in J\}.$$

Axioms for Entropy

A family of (Φ, β)-entropies is a a collection of functions $\kappa_S : \mathsf{PART}(S) \leftarrow \mathbb{R}_{\geq 0}$ such that:

ENT1: if $\pi, \pi' \in \mathsf{PART}(S)$, $\pi \leq \pi'$ implies $\kappa_S(\pi') \leq \kappa_S(\pi)$;

ENT2: if A, B are two finite sets, $|A| \leq |B|$ implies $\kappa_A(\iota_A) \leq \kappa_B(\iota_B)$;

Axioms for Entropy continued ...

ENT3: if $\pi \in \mathsf{PART}(A), \sigma \in \mathsf{PART}(B)$, and $A \cap B = \emptyset$, then:

$$\kappa_{A \cup B}(\pi + \sigma) = \left(\frac{|A|}{|A| + |B|}\right)^{\beta} \kappa_A(\pi) + \left(\frac{|B|}{|A| + |B|}\right)^{\beta} \kappa_B(\sigma) + \kappa_{A \cup B}(\{A, B\})$$

ENT4: For $\pi \in \mathsf{PART}(S), \sigma \in \mathsf{PART}(T)$, $\kappa_{S \times T}(\pi \times \sigma)$ is a function of $\kappa_A(\pi)$ and $\kappa_B(\sigma)$.

Entropy of Partitions of Finite Sets

A family of functions of the form $\kappa_A : \mathsf{PART}(A) \leftarrow \mathbb{R}_{\geq 0}$, indexed by finite sets has the form:

$$\kappa(\pi) = k(\beta) \left(1 - \sum_{i=1}^{n} \frac{|C_i|}{|A|} \right)^{\beta},$$

where $\pi = \{C_1, \ldots, C_n\} \in \mathsf{PART}(A)$ and $\lim_{\beta \to 1} k(\beta)(1 - \beta)$ is finite.
When $\beta \to 1$ we have the Shannon entropy
$\beta = 2$ gives the Gini index.

Partitions and Equivalences

A relation ρ on A is an equivalence if it is reflexive, symmetric and transitive.
A is the *domain* of ρ, $A = \mathsf{D}_\rho$.
A *block* of ρ is a non-empty set B, $B \subseteq A$ such that

- $B \times B \subseteq \rho$, and
- $B \times (A - B) \cup (A - B) \times B \subseteq A \times A - \rho$.

Ordinal Numbers

Primary notions: class and membership

- A set is a class α such that $\alpha \in \beta$ for some class β.
- A class α is $\in$-transitive if $\beta \in \gamma \in \alpha$ implies $\beta \in \alpha$.
- A class is an ordinal if it is $\in$-transitive and each member of α is $\in$-transitive.

Well-orders and Ordinals

If α is a class, then $\alpha + 1$ denotes the class $\alpha \cup \{\alpha\}$. ORD is the class of all ordinals.

- For every ordinal α the relation $\{(\beta, \gamma) \mid \beta, \gamma \in \alpha$ and $\beta = \gamma$ or $\beta \in \gamma\}$ is a well-ordering on α.
- Every well-ordering is isomorphic to some well-ordering derived as above from an ordinal.

Ordinals and the AC

Let π be an equivalence.
AC $\Rightarrow$: The blocks of π can be indexed by an ordinal α,

$$\pi = \{C_\xi | \xi < \alpha\}$$

α is the indexing ordinal for π.
Example: If $\alpha = \omega + 1$, then $\pi = \{C_0, C_1, \ldots, C_\omega\}$.

Equivalences

- $\rho \perp \sigma$ if ρ and σ are disjoint, that is $\mathsf{D}_\rho \cap \mathsf{D}_\sigma = \emptyset$.
- $\rho \sqsubseteq \sigma$ if every block of ρ is included in a block of σ.
- $\iota_X = \{(x, x) \mid x \in X\}$, $\omega_X = \{(x, y) \mid x, y \in X\}$, $\iota_X \sqsubseteq \omega_X$.

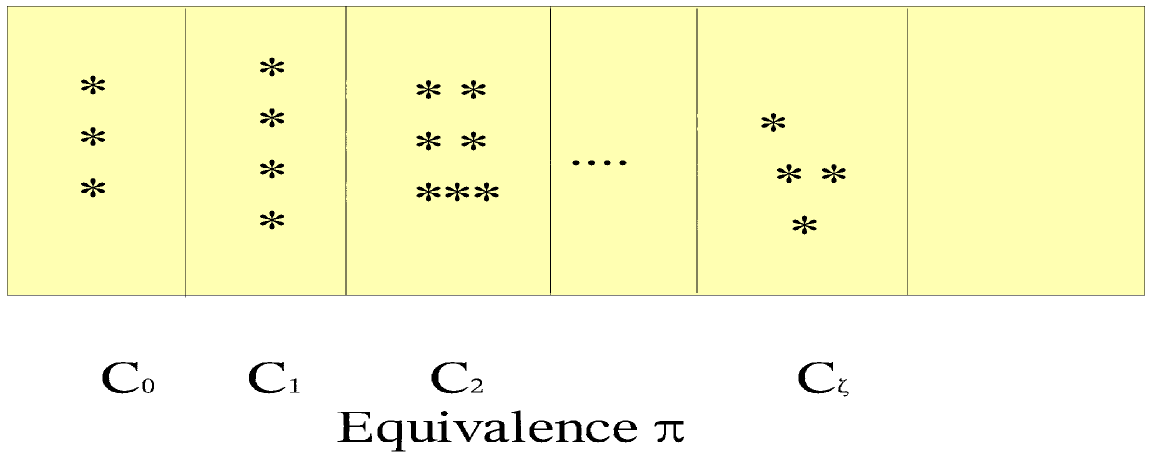
Systemic Collections of Equivalences

Let

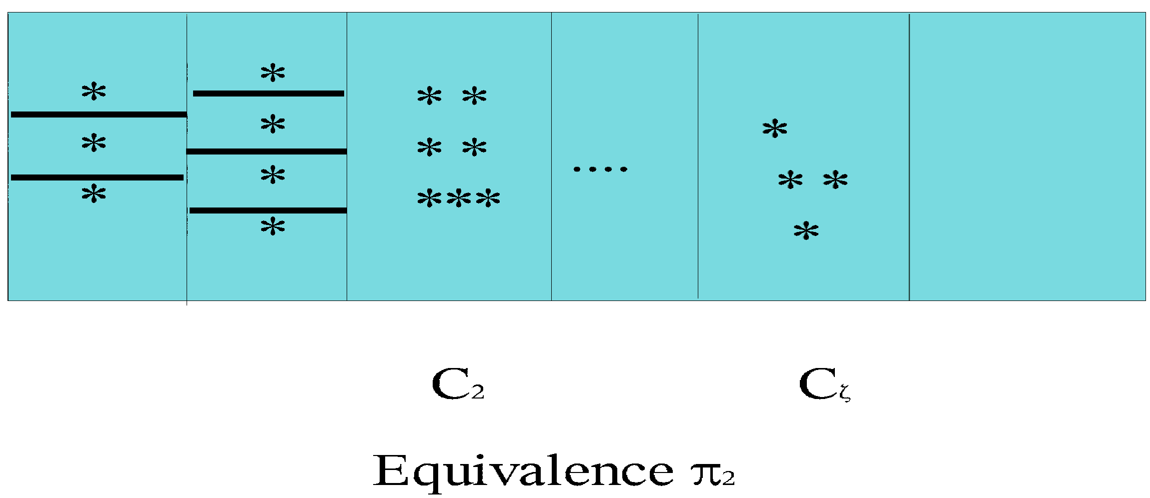
- π be an equivalence,
- α, an indexing ordinal for π,
- ξ an ordinal such that $\xi \leq \alpha$.

π_ξ is the equivalence having the following quotient set:

- for every $x \in C_\tau$ with $\tau < \xi$ a singleton block $\{x\}$, and
- C_ζ for every ζ such that $\xi \leq \zeta < \alpha$.



....
C0 C1 C2 Cζ
Equivalence π



....
C2 Cζ
Equivalence π2

A non-void collection of equivalence relations $\mathcal{E}$ is systemic if for every $\pi, \sigma \in \mathcal{E}$:

- $\pi \perp \sigma$ implies $\pi \cup \sigma \in \mathcal{E}$;
- $\omega_{\mathsf{D}_\pi} \in \mathcal{E}$;
- the blocks of π can be listed as $\{C_\xi \,|\, \xi < \alpha\}$ such that:
  - $\pi_\xi \in \mathcal{E}$ for $\xi < \alpha$;
  - for $\xi < \alpha$ the equivalence with blocks $\{C_\tau \,|\, \tau \neq \xi\}$ belongs to $\mathcal{E}$;
  - $\omega_{C_\xi} \in \mathcal{E}$.

Let $\mathcal{E}$ be a systemic collection of equivalence relations. The support of $\mathcal{E}$ is the family of sets

$$\mathrm{supp}(\mathcal{E}) = \{X \mid \iota_X \in \mathcal{E}\}.$$

Every block of an equivalence relation π of $\mathcal{E}$ belongs to $\mathrm{supp}(\mathcal{E})$.
Let C_ξ be a block of $\pi \in \mathcal{E}$.

- $3_c \Rightarrow$: $\omega_{C_\xi} \in \mathcal{E}$.
- The indexing ordinal of $\omega_{C_\xi} \in \mathcal{E}$ is 0.
- $3_a \Rightarrow$: $(\omega_{C_\xi})_0 = \iota_{C_\xi} \in \mathcal{E}$, so $C_\xi \in \mathrm{supp}(\mathcal{E})$.

An Example ...

$\{A_0, \ldots, A_{n-1}\}$: n pairwise disjoint finite sets;
$\pi^{(0)}, \ldots, \pi^{(n-1)}$: n equivalence relations on each of these sets, respectively.
Let $\mathcal{E}$ be the family of equivalence relations given by:

1. every equivalence $\pi^{(i)}$ belongs to $\mathcal{E}$;
2. if C is a block of an equivalence relation of $\mathcal{E}$, then $\iota_C \in \mathcal{E}$;
3. if $\pi, \pi' \in \mathcal{E}$ and $\pi \perp \pi'$, then $\pi \cup \pi' \in \mathcal{E}$;
4. if $\sigma \in \mathcal{E}$ and $\rho \sqsubseteq \sigma$, then $\rho \in \mathcal{E}$;
5. if $\pi \in \mathcal{E}$, then $\omega_{\mathsf{D}_\pi} \in \mathcal{E}$.

$\mathcal{E}$ is a systemic family of equivalence relations.

Measure on Support Sets

A *measure* on $\mathrm{supp}(\mathcal{E})$ is a monotone map from $(\mathrm{supp}(\mathcal{E}), \leq)$ into $(\mathbb{R}_{\geq 0}, \leq)$ such that for all $X, Y \in \mathrm{supp}(\mathcal{E})$ we have:

1. $\mu(X) > 0$;
2. $X \cap Y = \emptyset$ implies $\mu(X \cup Y) = \mu(X) + \mu(Y)$, and
3. there exists $Z \in \mathrm{supp}(\mathcal{E})$ such that $X \cap Z = \emptyset$ and $\mu(X) = \mu(Z)$.

Notations

Let $\beta \in \mathbb{R}_{\geq 0}$, $\beta > 0$ and let ν be a monotonic (i.e., nondecreasing) self-mapping of $(\mathbb{R}_{\geq 0}, \leq)$. The mapping λ_ν is defined by $\lambda_\nu(X) = \nu(\mu(X))$ for every set X.

If ν is fixed, or understood from context, then we simply write λ instead of λ_ν.

For $X, Y \in \mathrm{supp}(\mathcal{E})$ such that $X \cap Y = \emptyset$ define

$$c_{XY} = \left(\frac{\mu(X)}{\mu(X \cup Y)} \right)^{\beta}.$$

Weak (β, λ, μ)-entropy

Let $\mathcal{E}$ be a systemic set of equivalence relations. A *weak* (β, λ, μ)*-entropy* is a mapping $\kappa : \mathcal{E} \longleftarrow \mathbb{R}_{\geq 0}$ satisfying the following conditions:

(P_1) $\kappa(\iota_X) = \lambda(X)$ for all $X \in \mathrm{supp}(\mathcal{E})$.

(P_2) For all $\rho, \sigma \in \mathcal{E}$ with $\rho \perp \sigma$ and $X = \mathsf{D}_\rho$, $Y = \mathsf{D}_\sigma$ we have:

$$\kappa(\rho \cup \sigma) = c_{XY}\kappa(\rho) + c_{YX}\kappa(\sigma) + \kappa(\omega_X \cup \omega_Y).$$

(P_3) If $\pi \in \mathcal{E}$ and π is indexed by an ordinal α, where $\{C_\xi \,|\, \xi < \alpha\}$ is a listing of the blocks of π, then for each $\zeta < \alpha$ define

$$\mathsf{d}^{\pi}_{\zeta} = \left(\frac{\mu(C_\zeta)}{\mu(\mathsf{D}_\pi)}\right)^{\beta}.$$

For an ordinal $\xi \leq \alpha$ consider the equation:

$$\kappa(\pi) = \kappa(\pi_\xi) - \sum_{\zeta<\xi} \mathsf{d}^{\pi}_{\zeta} \lambda(C_\xi). \tag{4_ξ}$$

If $\tau \leq \alpha$ is a limit ordinal and (4_ξ) holds for all $\xi < \tau$, then (4_τ) holds.

Let $\mathcal{E}$ be a systemic family of equivalence relations and let κ be a weak (β, λ, μ)-entropy.
We have $\kappa(\omega_X) = 0$ for all $X \in \mathrm{supp}(\mathcal{E})$.

Let $\mathcal{E}$ be a systemic family of equivalence relations, μ a measure on $\mathrm{supp}(\mathcal{E})$, $\sigma \in \mathcal{E}$, and let $Y \in \mathrm{supp}(\mathcal{E})$ be a set disjoint from $X = \mathsf{D}_\sigma$. Then, for the (β, λ, μ)-entropy κ we have:

$$\kappa(\sigma \cup \iota_Y) - \kappa(\sigma \cup \omega_Y) = c_{Y\,X}\lambda(Y).$$

Main Result

Let

- $\mathcal{E}$ be a systemic family,
- κ be a (β, λ, μ)-entropy satisfying (P_1)-(P_3),
- $\pi \in \mathcal{E}$ be an equivalence relation whose blocks are listed as $\{C_\xi \,|\, \xi < \alpha\}$, where α is an indexing ordinal of π, $\alpha > 0$.

If κ is a (β, λ, μ)-entropy we have:

$$\kappa(\pi) = \lambda(\mathsf{D}_\pi) - \sum_{\zeta < \alpha} \mathsf{d}_\zeta^\pi \lambda(C_\zeta).$$

Extension to any number of equivalence relations

Another Preliminary Result

Let $\mathcal{E}$ be a systemic collection of equivalence relations and let $\rho \in \mathcal{E}$ be an arbitrary equivalence relation. Set $X = \mathsf{D}_\rho$ and let $\rho \perp \omega_Y$ for some $\omega_Y \in \mathcal{E}$. Then, we have:

$$\kappa(\rho \cup \omega_Y) = \mu(X \cup Y)^{-\beta} \mu(Y) \kappa(\rho) + \kappa(\omega_X \cup \omega_Y).$$

Property P_4

Let $\alpha > 0$ be a limit ordinal and let $\{\pi^{(\xi)} \in \mathcal{E} \,|\, \xi < \alpha\}$ be a family of equivalences indexed by α. If

1. $\pi^{(\zeta)} \sqsubseteq \pi^{(\xi)}$ whenever $\zeta < \xi < \alpha$, and
2. $\sigma^{(\alpha)} = \bigcup_{\xi<\alpha} \pi^{(\xi)}$ belongs to $\mathcal{E}$,

then

$$\kappa(\sigma^{(\alpha)}) = \inf\{\kappa(\pi^{(\xi)}) \,|\, \xi < \alpha\}$$

The first part of (P_4) means that $\{\pi^{(\xi)} \,|\, \xi < \alpha\}$ form a well-ordered chain of type α.

Main Result - II

Hypothesis:

- $\mathcal{E}$ is a systemic set and let κ be a (β, λ, μ)-entropy satisfying (P_1)-(P_4);
- let α be an ordinal, $\alpha \geq 3$;
- let $\rho^{(\xi)} \in \mathcal{E}$ for all $\xi < \alpha$ be such that $\rho^{(\zeta)} \perp \rho^{(\xi)}$ whenever $\zeta < \xi < \alpha$;
- for $0 < \xi < \alpha$ set $\sigma^{(\xi)} = \bigcup_{\zeta < \xi} \rho^{(\zeta)}$

Suppose that $\sigma^{(\tau)} \in \mathcal{E}$ for all limit ordinals $0 < \tau \leq \alpha$.

continuation...

Conclusion:
Then,

1. $\sigma^{(\xi)} \in \mathcal{E}$ for all $0 < \xi < \alpha$,
2. and

$$\kappa(\sigma^{(\alpha)}) = \mu(X)^{-\beta} \sum_{\xi<\alpha} \mu(X_\xi)^{\beta} \kappa(\rho^{(\xi)}) + \kappa\left(\bigcup_{\xi<\alpha} \omega_{X_\xi}\right),$$

where $X_\xi = \mathsf{D}_{\rho^{(\xi)}}$ for every $\xi < \alpha$ and $X = \mathsf{D}_{\sigma^{(\alpha)}}$.

An Example

Assumptions and Notations:

- $\{A_0, \ldots, A_{n-1}\}$ is a collection on n pairwise disjoint finite sets and $\pi^{(0)}, \ldots, \pi^{(n-1)}$ are n equivalence relations on each of these sets, respectively.
- Denote $A = \bigcup_{i=0}^{n-1} A_i$.
- Let $\mathcal{E}$ previously introduced. Choose $\mu(X) = |X|$ for $X \in \text{supp}(\mathcal{E})$ and let $\nu(p) = p$ for $p \in \mathbb{R}_{\geq 0}$.

Example continued ...

Let κ be a (β, λ, μ)-entropy that satisfies (P_1)–(P_4). Then,

$$\kappa\left(\bigcup_{i=0}^{n-1} \pi^{(i)}\right) = |A|^{-\beta} \sum_{i<n} |A_i|^{\beta} \kappa(\pi^{(i)}) + \kappa\left(\bigcup_{i<n} \omega_{A_i}\right),$$

which is the central result (Corollary II.7) of Simovici and Jaroszewicz.

Extension of Entropy Axiomatization

to partitions of arbitrary sets by using:

- measures,
- ordinal numbers, and
- transfinite induction.

This approach clarifies the mathematical bases of the argument made for partitions of finite sets.
Further exploration of this approach appears useful for various class of measures on systemic families of equivalence relations.