An Abstract Axiomatization of the Notion of Entropy

Ivo G. Rosenberg
Université de Montréal
Dépt. de Mathématiques
et de Statistique, C. P. 6128
succ. "Centre-Ville", Montréal,
Québec, Canada H3C 3J7

rosenb@dms.umontreal.ca

Dan A. Simovici
Univ. of Massachusetts Boston,
Dept. of Comp. Science,
100 Morrissey Blvd.
Boston, Massachusetts
02125 USA

dsim@cs.umb.edu

Abstract

We present an abstract axiomatization of generalized entropy using the notion of ordinal number and the new concept of systemic set of equivalence relations. The aximoatization applies to arbitrary sets and extends previous results obtained for the finite case.

1 Introduction

Axiomatizations for entropy and entropy-like characteristics of probability distributions represent a problem with a rich history in information theory. Previous relevant work includes the results of A.I. Khinchin [Khi57], D.K. Faddeev [Fad56], R.S. Ingarden and K. Urbanik [IU62], and A. Rényi [Ŕ59], who investigated various axiomatizations of entropy, and Z. Daróczy who presented in [Dar70] an unified treatment of entropy-like characteristics of probability distributions using the notion of information function.

In our previous work (see [SR93, JS99, SJ02, SJ03a]) we introduced an axiomatization for the notion of functional entropy. This numerical characteristic of functions is related to the complexity of circuits that realize functions (see [CA90]) and serves as an estimate for power dissipation of a circuit realizing a function (see [HW97]) and it is linked to the notion of entropy for partitions, since every function $f : A \longrightarrow B$ between the finite sets A, B defines a partition on its definition domain A whose blocks are $\{f^{-1}(b) \mid b \in \mathrm{Ran}(f)\}$, where $\mathrm{Ran}(f)$ is the range of the function f.

Information measures, especially conditional entropy of a logic function and its variables, have been used for minimization of logic functions (for further references see [LRGLRR93] and [CSSY98]).

In a different direction, starting from the notion of im-

purity of a set relative to a partition, we found a common generalization of Shannon entropy and of Gini index and we used this generalization in clustering of non-categorial data (see [SJ03b]). P. A. Devijer used the Gini index in pattern recognition in [Dev74].

Our main result is an an abstract axiomatization of generalized entropy using the notion of ordinal number and the new concept of systemic set of equivalence relations. We believe that this approach clarifies the essence of previous axiomatizations.

The set of reals and the set of nonnegative reals equipped with the natural sum, product, and order are denoted by $\mathbb{R}$ and $\mathbb{R}_{\geq 0}$, respectively.

Let ρ be a binary relation on a set A. As usual, ρ is *symmetric* if $(a, b) \in \rho$ implies $(b, a) \in \rho$, *transitive* if $(a, b), (b, c) \in \rho$ implies $(a, c) \in \rho$ and *reflexive* if $(a, a) \in \rho$ for all $a, b, c \in A$. A reflexive, symmetric and transitive relation on A is an *equivalence relation*. We refer to A as the *domain of ρ* and write $A = \mathsf{D}_\rho$.

If $(A, \leq)$ and $(B, \leq')$ are partially ordered sets, a map $f : A \longrightarrow B$ is *monotone* (also called isotone, order preserving, or an order homomorphism), if for all $a_1, a_2 \in A$, $a_1 \leq a_2$ implies $f(a_1) \leq' f(a_2)$. The map is *anti-monotone* if all $a_1, a_2 \in A$, $a_1 \leq a_2$ implies $f(a_1) \geq' f(a_2)$. An *order isomorphism* between two partially ordered sets $(A, \leq)$ and $(B, \leq')$ is a bijection $f : A \longrightarrow B$ such that $a_1 \leq a_2$ if and only if $f(a_1) \leq' f(a_2)$.

A *block* (also called an equivalence class) of ρ is a non-void set B, $B \subseteq A$ such that $(b, b') \in \rho$ for all $b, b' \in B$ and $(b, c) \notin \rho$ for all $b \in B$ and $c \in A - B$.

Two equivalence relations are *disjoint* (denoted by $\rho \perp \sigma$) if $\rho \cap \sigma = \emptyset$, i.e., if $\mathsf{D}_\rho \cap \mathsf{D}_\sigma = \emptyset$. If $\rho \perp \sigma$, then $\rho \cup \sigma$ is an equivalence relation (with domain $\mathsf{D}_{\rho \cup \sigma} = \mathsf{D}_\rho \cup \mathsf{D}_\sigma$).

For equivalence relations ρ and σ we write $\rho \preceq \sigma$ if every block of ρ is included in a block of σ.

Let $\mathcal{E}$ be a set of equivalence relations. Clearly, $\preceq$ is a

partial order on $\mathcal{E}$, i.e. a reflexive, transitive and antisymmetric relation (i.e. $\rho \preceq \sigma \preceq \rho$ implies $\rho = \sigma$).

For a set X set $\iota_X = \{(x,x) \mid x \in X\}$ and $\omega_X = \{(x,y) \mid x,y \in X\}$. Clearly, for $X \neq \emptyset$ both ι_X and ω_X are equivalence relations with domain X and $\iota_X \preceq \omega_X$.

Our presentation makes intensive use of ordinal numbers. The notion of ordinal can be constructed starting from the notion of class and of membership (see, e.g. [Mon69]). The membership relation is denoted usually by $\in$. A set is a class α that is a member of another class β. A class α is $\in$-transitive if $\beta \in \gamma \in A$ implies $\beta \in A$. A class α is an *ordinal* if it is $\in$-transitive and each member of α is $\in$-transitive. For a class α we denote by $\alpha + 1$ the class $\alpha \cup \{\alpha\}$.

Let ORD be the class of all ordinals. For and $\alpha \in \mathsf{ORD}$ the relation $\{(\beta,\gamma) \mid \beta,\gamma \in \alpha \text{ and } \beta = \gamma \text{ or } \beta \in \gamma\}$ is a well-ordering. Distinct ordinals generate non-isomorphic partial orders. Furthermore, every well-ordering is isomorphic to some well-ordering derived in a standard way from an ordinal.

Let π be an equivalence relation. We assume the Axiom of Choice and consequently the blocks of π can be indexed by the ordinals smaller than an ordinal α, i.e. $\{C_\xi \mid \xi < \alpha\}$ is a list of the blocks of π. For example, if $\alpha = \omega + 1$, the list is $\{C_0, C_1, \ldots, C_\omega\}$. We refer to α as an *indexing ordinal for* π.

2 Systemic Collections of Equivalence Relations

The following notation will be needed. For an equivalence relation π having an indexing ordinal α, a list $\{C_\xi \mid \xi < \alpha\}$ of its blocks, and for an ordinal $\xi \leq \alpha$, denote by π_ξ the equivalence relation having a singleton block $\{x\}$ for every $x \in C_\tau$ whenever $\tau < \xi$ and the block C_ζ for every ζ such that $\xi \leq \zeta < \alpha$. Note that $\pi_0 = \pi$, while $\pi_\alpha = \iota_D$, where $D = \mathsf{D}_\pi$ is the domain of π.

The main notion of this section is introduced next.

Definition 2.1 A non-void set $\mathcal{E}$ of equivalence relations is *systemic* if for all $\pi, \sigma \in \mathcal{E}$ we have:

1. $\pi \perp \sigma$ implies $\pi \cup \sigma \in \mathcal{E}$;

2. $\omega_{\mathsf{D}_\pi} \in \mathcal{E}$;

3. the blocks of π can be listed as $\{C_\xi \mid \xi < \alpha\}$ so that

   (a) $\pi_\xi \in \mathcal{E}$ for all $\xi < \alpha$;

   (b) for every $\xi < \alpha$ the equivalence relation with blocks $\{C_\tau \mid \tau < \alpha \text{ and } \tau \neq \xi\}$ belongs to $\mathcal{E}$, and

   (c) $\omega_{C_\xi} \in \mathcal{E}$ for every $\xi < \alpha$.

The *support* of the systemic collection $\mathcal{E}$ is the family of sets

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

$\square$

Theorem 2.2 *Let $\mathcal{E}$ be a systemic collection of equivalence relations. Every block of an equivalence relation π of $\mathcal{E}$ belongs to* $\mathrm{supp}\ \mathcal{E}$.

Proof. Let C_ξ be a block of $\pi \in \mathcal{E}$. By Property 3_c we have $\omega_{C_\xi} \in \mathcal{E}$. eh indexing ordinal of $\omega_{C_\xi} \in \mathcal{E}$ is clearly 0. By Property 3_a of systemic collections $(\omega_{C_\xi})_0 = \iota_{C_\xi} \in \mathcal{E}$, which means that $C_\xi \in \mathrm{supp}\ \mathcal{E}$. ∎

Let π, σ be two equivalence relations. We write $\pi \sqsubseteq \sigma$ if every block of π is a block of σ. If $\sigma - \pi \neq \emptyset$, then we write $\pi \sqsubset \sigma$.

Example 2.3 Let $\{A_0, \ldots, A_{n-1}\}$ be a collection on n pairwise disjoint finite sets and let $\pi^{(0)}, \ldots, \pi^{(n-1)}$ be n equivalence relations on each of these sets, respectively.[1]

Let $\mathcal{E}$ be the family of equivalence relations given by the following inductive definition:

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 $\omega_{\mathsf{D}_\pi} \in \mathcal{E}$.

We claim that $\mathcal{E}$ is a systemic family of equivalence relations. Indeed, the first property of systemic families is clearly satisfied. If C is a block of an equivalence relation $\pi \in \mathcal{E}$, then $\omega_C \sqsubseteq \pi$, so $\omega_C \in \mathcal{E}$. Thus propery $3(c)$ of systemic families is also satisfied.

Let $\pi \in \mathcal{E}$ be an equivalence relation whose blocks are $C_0, \ldots, C_{\ell-1}$. Note that $\pi_k = \iota_{C_0} \cup \cdots \iota_{C_{k-1}} \cup \omega_{C_k} \cup \cdots \cup \omega_{C_{\ell-1}}$. The equivalence relations $\iota_{C_0}, \ldots, \iota_{C_{k-1}}, \ldots, \omega_{C_k}, \ldots, \omega_{C_{\ell-1}}$ are pairwise disjoint and belong to $\mathcal{E}$, which implies property $3(a)$. Finally, property $3(b)$ follows from the fact that the equivalence relation π' that has the blocks $C_0, \ldots, C_{i-1}, C_{i+1}, \ldots, C_{\ell-1}$ can be expressed as $\pi' = \omega_{C_0} \cup \cdots \cup \omega_{C_{i-1}} \cup \omega_{C_{i+1}} \cup \cdots \cup \omega_{C_{\ell-1}}$.

$\square$

---

[1] Note that we use the parenthesized superscript notation for a family of equivalences to differentiate this notation from the notation introduced at the beginning of this section.

3 The Notion of Weak (β, λ, μ)-Entropy

Our notion of abstract entropy is based on the notion of measure on a family of sets introduced next.

Definition 3.1 A *measure* on supp $\mathcal{E}$ is a monotone map from (supp $\mathcal{E}, \leq$) into $(\mathbb{R}_{\geq 0}, \leq)$ such that for all $X, Y \in$ 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 $Z \in$ supp $\mathcal{E}$ such that $X \cap Z = \emptyset$ and $\mu(X) = \mu(Z)$.

$\square$

Usually, the first condition of Definition 3.1 has the form $\mu(X) = 0$ if and only if $X = \emptyset$. However, in our case $\emptyset \notin$ supp $\mathcal{E}$ (as we have not introduced $\iota_{\emptyset}$), and so we postulate that $\mu(X)$ is positive.

Let $\beta \in \mathbb{R}_{\geq 0}$, $\beta > 0$ and let ν be a monotonic (i.e., nondecreasing) self-mapping of $(\mathbb{R}_{\geq 0}, \leq)$. Denote by λ_{ν} the mapping defined by $\lambda_{\nu}(x) = \nu(\mu(x))$ for all $x \in \mathbb{R}_{\geq 0}$. If ν is fixed, or understood from context, then we simply write λ instead of λ_{ν}.

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

$$c_{XY} = \left(\frac{\mu(X)}{\mu(X \cup Y)} \right)^{\beta}. \tag{1}$$

Definition 3.2 Let $\mathcal{E}$ be a systemic set of equivalence relations. A *weak (β, λ, μ)-entropy* is a map $\kappa : \mathcal{E} \longrightarrow \mathbb{R}_{\geq 0}$ satisfying the following conditions:

(P_1) $\kappa(\iota_X) = \lambda(X)$ for all $X \in$ 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:

$$\kappa(\rho \cup \sigma) = c_{XY}\kappa(\rho) + c_{YX}\kappa(\sigma) + \kappa(\omega_X \cup \omega_Y). \tag{2}$$

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

$$\mathsf{d}_{\zeta}^{\pi} = \left(\frac{\mu(C_{\zeta})}{\mu(\mathsf{D}_{\pi})} \right)^{\beta}. \tag{3}$$

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

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

If $\tau \leq \alpha$ is a limit ordinal and (4_{ξ}) holds for all $\xi < \tau$, then (4_{τ}) holds.

Thus, (P_3) allows the extension of (4_{ξ}), $\xi < \tau$ to (4_{τ}).

We start with the following lemma.

Lemma 3.3 *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 $X \in$ supp $\mathcal{E}$.*

Proof. Let $X \in$ supp $\mathcal{E}$. By the third part of the definition of measure there exists $Z \in$ supp $\mathcal{E}$, disjoint from X, such that $\mu(X) = \mu(Z)$. By Part (2) of Definition 2.1, we get $\omega_X, \omega_Z \in \mathcal{E}$. From the second part of the definition of measure (Definition 3.1) we obtain $\mu(X \cup Z) = \mu(X) + \mu(Z) = 2\mu(X)$. Since $c_{XZ} = c_{ZX} = 2^{-\beta}$, Equality (2) of Property (P_2) yields:

$$\kappa(\omega_X \cup \omega_Z) = 2^{-\beta}(\kappa(\omega_X) + \kappa(\omega_Z)) + \kappa(\omega_X \cup \omega_Z),$$

hence $\kappa(\omega_X) + \kappa(\omega_Z) = 0$. Here $\kappa(\omega_X) \geq 0$ and $\kappa(\omega_Z) \geq 0$ and, therefore, $\kappa(\omega_X) = 0$. $\blacksquare$

Lemma 3.4 *Let $\mathcal{E}$ be a systemic family of equivalence relations, μ a measure on supp $\mathcal{E}$, $\sigma \in \mathcal{E}$, and let $Y \in$ supp $\mathcal{E}$ be a set disjoint from $X = \mathsf{D}_{\sigma}$. Then, for the (β, λ, μ)-entropy κ we have:*

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

Proof. Apply Equality (2) twice, Lemma 3.3 and Property (P_1) to obtain:

$$\begin{aligned}
&\kappa(\sigma \cup \iota_Y) - \kappa(\sigma \cup \omega_Y) \\
&= c_{XY}\kappa(\sigma) + c_{YX}\kappa(\iota_Y) + \kappa(\omega_X \cup \omega_Y) \\
&\quad - c_{XY}\kappa(\sigma) - c_{YX}\kappa(\omega_Y) - \kappa(\omega_X \cup \omega_Y) \\
&= c_{YX}\kappa(\iota_Y) = c_{YX}\lambda(Y).
\end{aligned}$$

Theorem 3.5 *Let $\mathcal{E}$ be a systemic set, κ be a (β, λ, μ)-entropy satisfying (P_1)-(P_3), and let $\pi \in \mathcal{E}$ be an equivalence relation whose blocks are listed as $\{C_{\xi} \mid \xi < \alpha\}$, where α is an indexing ordinal of π, $\alpha > 0$. Then, if κ is a (β, λ, μ)-entropy we have:*

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

Proof. We start with two claims.

`Claim 1:` If ξ is an ordinal and $\xi + 1 < \alpha$, then

$$\kappa(\pi_{\xi+1}) = \kappa(\pi_{\xi}) + \mathsf{d}_{\xi}^{\pi} \lambda(C_{\xi}).$$

$$(4_{\xi})$$

`Proof of Claim 1:` Set $Z = \bigcup_{\zeta < \xi} C_{\zeta}$,

$$\sigma = \iota_Z \cup \bigcup_{\xi < \tau < \alpha} \omega_{\tau},$$

and $Y = C_\xi$. Notice that

$$\begin{aligned}
X = \mathsf{D}_\sigma &= \mathsf{D}_\pi - C_\xi, \\
\sigma \cup \omega_Y &= \pi_\xi, \\
\sigma \cup \iota_Y &= \pi_{\xi+1}.
\end{aligned}$$

From Lemma 3.4 we get

$$\kappa(\pi_{\xi+1}) - \kappa(\pi_\xi) = c_{C_\xi X}\lambda(C_\xi).$$

Recall that

$$c_{C_\xi X} = \left(\frac{\mu(C_\xi)}{\mu(\mathsf{D}_\pi)}\right)^\beta = \mathsf{d}_\xi^\pi,$$

which concludes the justification of the claim.

`Claim 2:` We have

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

`Proof of Claim 2:` The argument is by transfinite induction on $\xi < \alpha$.

Let $\xi = 0$. Notice that $\pi_0 = \pi$ and (7_0) reduces to the obvious $\kappa(\pi_0) = \kappa(\pi)$.

Suppose that (7_ξ) holds for all $\xi < \tau$ for some $0 < \tau < \alpha$.

The inductive steps has two cases.

I. Let τ be an isolated ordinal. From $(6_{\tau-1})$ and $(7_{\tau-1})$ we have

$$\begin{aligned}
\kappa(\pi_\tau) &= \kappa(\pi_{\tau-1}) + \mathsf{d}_{\tau-1}^\pi \lambda(C_{\tau-1}) \\
&= \kappa(\pi) + \sum_{\zeta < \tau-1} \mathsf{d}_\zeta^\pi \lambda(C_\zeta) + \mathsf{d}_{\tau-1}^\pi \lambda(C_{\tau-1}) \\
&= \kappa(\pi) + \sum_{\zeta < \tau} \mathsf{d}_\zeta^\pi \lambda(C_\zeta),
\end{aligned}$$

proving (7_τ).

II. Let τ be a limit ordinal. From the hypothesis and (P_3) we get (4_τ) and thus (7_τ). This concludes the induction step and the claim follows.

Now we can prove the theorem. Let α be isolated. Then $\pi_\alpha = \iota_{\mathsf{D}_\pi}$ by the definition of π_α, hence $\kappa(\pi_\alpha) = \lambda(\mathsf{D}_\pi)$ and (7_α) becomes (5). Thus, let α be a limit ordinal. Now (7_ξ) holds for all $\xi < \alpha$ and from (P_3) we get the validity of (4_α). Using again $\pi_\alpha = \iota_{\mathsf{D}_\pi}$, clearly (4_α) converts into (5). ∎

We extend Equality (2) to any number of equivalence relations. For this we need a lemma and an additional property.

Lemma 3.6 *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:*

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

Proof. The statement follows from Equality (2) and $\kappa(\omega_Y) = 0$ by Lemma 3.3. ∎

We need the following property:

> (P_4) Let $\alpha > 0$ be a limit ordinal and let $\{\pi^{(\xi)} \in \mathcal{E} \mid \xi < \alpha\}$ 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
>
> $$\kappa(\sigma^{(\alpha)}) = \inf\{\kappa(\pi^{(\xi)}) \mid \xi < \alpha\}$$

Here $\inf$ is the standard infimum in $(\mathbb{R}_{\geq 0}, \leq)$ which exists since the set $\{\kappa(\pi^{(\xi)}) \mid \xi < \alpha\}$ is a set of reals bounded from below by 0.

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

Theorem 3.7 *Let $\mathcal{E}$ be a systemic set and let κ be a (β, λ, μ)-entropy satisfying (P_1)-(P_4). Let α be an ordinal, $\alpha \geq 3$, and 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 $0 < \tau \leq \alpha$. 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 $X = \mathsf{D}_{\sigma^{(\alpha)}}$.

Proof. The argument is by transfinite induction on $\alpha \geq 3$.

Let $\alpha = 3$. Then $\sigma^{(1)} = \rho^{(0)} \in \mathcal{E}$ and also $\sigma^{(2)} = \rho^{(0)} \cup \rho^{(1)} \in \mathcal{E}$ by the first part of Definition 2.1. Thus, Part (1) of the theorem holds. Next, (11_3) coincides with (2) and so Part (2) holds.

Suppose now that $\alpha > 3$ and that 1) and 2) hold for all $\tau < \alpha$. Let $\rho^{(\xi)} \in \mathcal{E}$ for all $\xi < \alpha$ satisfy Parts 1 and 2 of the hypothesis of the theorem.

`Claim 1:` We have $\sigma^{(\xi)} \in \mathcal{E}$ for all $\xi < \alpha$.

`Proof of Claim 1:` We need to consider the following two cases:

I. Let α be an isolated ordinal. Then $\sigma^{(\alpha)} = \sigma^{(\alpha-1)} \cup \rho^{(\alpha-1)}$, where $\sigma^{(\alpha-1)}$ and $\rho^{(\alpha-1)}$ belong to $\mathcal{E}$ and $\sigma^{(\alpha-1)} \perp \rho^{(\alpha-1)}$ by hypothesis. Then, by the first part of Definition 2.1 we have $\sigma^{(\alpha)} \in \mathcal{E}$.

II. Let α be a limit ordinal. Then $\sigma^{(\alpha)} \in \mathcal{E}$ by the second part of the hypothesis. This concludes the inductive step and proves the claim.

`Claim 2:` The equation (11_α) holds.

`Proof of Claim 2:` Again, we need to consider two cases.

I. Let α be an isolated ordinal. Then $\sigma^{(\alpha)} = \sigma^{(\alpha-1)} \cup \rho^{(\alpha-1)}$. Denote by X and X_ξ domains of $\sigma^{(\alpha-1)}$ and $\rho^{(\xi)}$, respectively, where $\xi < \alpha$, and set $Z = X \cup X_\xi$. From equation (2) we get:

$$\kappa(\sigma^{(\alpha)}) = \mu(Z)^{-\beta}\left[\mu(X)^\beta \kappa(\sigma^{(\alpha-1)}) + \mu^\beta(X_{\alpha-1})\kappa(\rho^{(\xi-1)})\right] + \kappa(\omega_X \cup \omega_{X_{\xi-1}})$$

By applying equality $(10_{\alpha-1})$ we obtain:

$$\kappa(\sigma_\alpha)$$
$$= \mu(Z)^{-\beta}\left[\mu(X)^\beta\mu(X)^{-\beta}\left\{\sum_{\tau<\alpha-1}\mu(X_\tau)^\beta\kappa(\rho^{(\tau)}) + \kappa\left(\bigcup_{\tau<\alpha-1}\omega_{X_\tau}\right)\right\}\right.$$
$$\left. + \mu(X_{\alpha-1})^\beta\kappa(\rho^{(\alpha-1)})\right] + \kappa(\omega_X \cup \omega_{X_{\alpha-1}})$$
$$= \mu(Z)^{-\beta}\sum_{\tau<\alpha}\mu(X_\tau)^\beta\kappa(\rho^{(\tau)}) + \mu(Z)^{-\beta}\kappa\left(\bigcup_{\tau<\alpha-1}\omega_{X_\tau}\right) + \kappa(\omega_X \cup \omega_{X_{\alpha-1}})$$

By (8) (for $\rho = \bigcup_{\tau<\alpha-1}\omega_{X_\tau}$ and $Y = \omega_{X_{\xi-1}}$) this simplifies to (11_α).

II. Let now α be a limit ordinal. Then (11_ξ) holds for every $\xi < \alpha$. Notice that $\sigma^{(\xi)}$ (for $\xi < \alpha$) satisfy $\sigma^{(\tau)} \sqsubset \sigma^{(\xi)}$ whenever $0 < \xi < \tau < \alpha$. From (9_α) we obtain:

$$\kappa(\sigma^{(\alpha)}) = \inf\{\kappa(\rho^{(\xi)}) \mid \xi < \alpha\}.$$

We also have that

$$\bigcup_{\lambda<\tau}\omega_{X_\lambda} \sqsubset \bigcup_{\lambda<\xi}\omega_{X_\lambda}$$

whenever $0 < \tau < \xi < \alpha$. By the same token

$$\kappa\left(\bigcup_{\xi<\alpha}\omega_{X_\xi}\right) = \inf\left\{\kappa\left(\bigcup_{\tau<\xi}\omega_{X_\tau}\right) \mid \tau < \xi\right\}$$

We claim that in $\mathbb{R}_{\geq 0}$ the validity of (12), (13) and (10_ξ) for all $\xi < \alpha$ implies (10_α). Indeed, the reals $\kappa(\sigma^{(\xi)})$ ($\xi < \alpha$) form a (decreasing) sequence "converging" to $\kappa(\sigma^{(\alpha)})$ and, similarly, the reals $\kappa\left(\bigcup_{\tau<\xi}\omega_{X_\tau}\right)$ ($\xi < \alpha$) form a "decreasing" sequence "converging" to $\kappa\left(\bigcup_{\tau<\alpha}\omega_{X_\tau}\right)$. It follows that the middle terms in (10_ξ) ($\xi < \alpha$) form an (increasing) sequence "converging" to

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

Thus, (10_α) holds. This concludes the induction step and proves the theorem. ∎

Example 3.8 Let $\{A_0, \ldots, A_{n-1}\}$ be a collection on n pairwise disjoint finite sets and let $\pi^{(0)}, \ldots, \pi^{(n-1)}$ be n equivalence relations on each of these sets, respectively. Denote $A = \bigcup_{i=0}^{n-1} A_i$. Consider again the systemic family of equivalences $\mathcal{E}$ introduced in Example 2.3. Choose the measure μ as $\mu(X) = |X|$ for $X \in \text{supp } \mathcal{E}$ and let $\nu(p) = p$ for $p \in \mathbb{R}_{\geq 0}$. Let κ be a (β, λ, μ)-entropy that satisfies (P_1)–(P_4). Then, Theorem 3.7 yields:

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

which is the central result (Corollary II.7) of [SJ02] that allows us to retrieve the specific formulas of generalized entropies for partitions of finite sets. ∎

4 Conclusions

We extended the axiomatization of entropies of partitions of finite sets discussed in [SJ02] to partitions of arbitrary sets by using measures, ordinal numbers and transfinite induction. In addition to its application to a wider class of entropies, this approach has the advantage of illuminating the mathematical bases of the argument made for partitions of finite sets.

In our opinion further exploration of this approach would be useful for various class of measures on systemic families of equivalence relations.

(12)

References

[CA90]   K.T. Cheng and V. Agrawal. An entropy measure of the complexity of multi-output

boolean function. In *Proc. of the 27th Design Automation Conference*, pages 302–305, 1990.

[CSSY98] V. Cheushev, V. Shmerko, D. A. Simovici, and S. Yanushkevich. Functional entropy and decision trees. In *Proceedings of IS-MVL 1998*, pages 257–262, Fukuoka, Japan, 1998.

[Dar70] Z. Daróczy. Generalized information functions. *Information and Control*, 16:36–51, 1970.

[Dev74] P. A. Devijer. Entropie quadratique et reconnaisance des formes. In J.-C. Simon, editor, *Proceedings of the NATO Advanced Study Institute: Computer Oriented Learning Processes*, pages 257–278, Château de Bonas, France, 1974.

[Fad56] D. K. Faddeev. On the notion of entropy of finite probability distributions (in Russian). *Usp. Mat. Nauk*, 11:227–231, 1956.

[HW97] C. H. Hwang and A. C. H. Wu. An entropy measure for power estimation of boolean functions. In *Proc. of Asia and South Pacific Design Automation Conference (ASP-DAC)*, pages 101–106, 1997.

[IU62] R. S. Ingarden and K. Urbanik. Information without probability. *Coll. Math.*, 1:281–304, 1962.

[JS99] S. Jaroszewicz and D. A. Simovici. On axiomatization of conditional entropy of functions between finite sets. In *Proc.of the 29th ISMVL*, pages 24–31, Freiburg, Germany, 1999.

[Khi57] A. Ia. Khinchin. *Mathematical Foundations of Information Theory*. Dover Publ., New York, 1857.

[LRGLRR93] A. Lloris-Ruiz, J. F. Gomez-Lopera, and R. Roman-Roldan. Entropic minimization of multiple-valued logic functions. In *Proceedings of the 23rd ISMVL*, pages 24–28, 1993.

[Mon69] J. D. Monk. *Introduction to Set Theory*. McGraw-Hill Book Company, New York, 1969.

[Ŕ59] A. Rényi. On the dimension and entropy of probability distributions. *Acta Math. Acad. Sci. Hung.*, 10:193–215, 1959.

[SJ02] D. A. Simovici and S. Jaroszewicz. An axiomatization of partition entropy. *Transactions on Information Theory*, 48:2138–2142, 2002.

[SJ03a] D. A. Simovici and S. Jaroszewicz. An algebraic approach to entropy. In *Beyond Two: Theory and Applications of Multiple-Valued Logic*, pages 101–115. Springer-Verlag, Heidelberg, New York, 2003.

[SJ03b] D. A. Simovici and S. Jaroszewicz. Generalized entropy and decision trees. In *EGC 2003 - Journées francophones d'Extraction et de Gestion de Connaissances*, pages 369–380, Lyon, France, 2003.

[SR93] D.A. Simovici and C. Reischer. On functional entropy. In *Proceedings of the 23rd ISMVL*, pages 100–104, 1993.