Werk

Label: Article

Jahr: 1976

PURL: https://resolver.sub.uni-goettingen.de/purl?316342866_0017|log74

Kontakt/Contact

Digizeitschriften e.V.
SUB Göttingen
Platz der Göttinger Sieben 1
37073 Göttingen

info@digizeitschriften.de

COMMENTATIONES MATHEMATICAE UNIVERSITATIS CAROLINAE

17,4 (1976)

QUASI-ENTROPY OF FINITE WEIGHTED METRIC SPACES

Miroslav KATĚTOV, Praha

Abstract: The note contains a proof of the possibility to extend the notion of the entropy (in the classical sense) to finite sets endowed, besides a probability distribution, with a semimetric.

Key words: Quasi-entropy, semimetric, projective subentropy.

AMS: 94A15 Ref. Ž.: 8.721

In various questions it is useful to possess a function which is defined for an appropriate class of probability distributions on semimetric spaces and has some basic properties of the entropy in the classical sense. In this note, a construction (in a broad sense of the word) of a function of this kind for finite spaces is given. Possibly, the method, particularly suitable generalization of c.d.e. (see 3.2), may also work in a more general situation.

It seems that some concepts introduced below, though rather natural and virtually known, have not been examined in the setting presented here. Hence, definitions are given starting from the most elementary ones. On the other hand, details of proofs are omitted.

Concerning the entropy, we presuppose the elementary facts only; of the semimetric spaces as good as nothing is assumed. Therefore, no references are given.

1.1. Besides a few deviations, we use the standard terminology and notation. Parentheses are omitted whenever possible: e.g. fx stands for $f(x)$. Symbols like $\{x_a : a \in A\}$ stand for sets, $(x_a : a \in A)$ for indexed sets. The following letters, possibly with subscripts, etc., designate objects of a specified kind: D,K,Q,T stand for finite non-void sets, P,S for spaces (see 1.3), ϱ, σ for semimetrics (1.2), μ, ν for weights (1.2). The cardinality of T is denoted by $|T|$. A function is a mapping into R (the real line). If x is a segment of a string y, we write $x \prec y$. Conventions: log means $\log_2$; $0/0 = 0$; $0 \log 0 = 0$.

1.2. By definition, a semimetric on T is a function ϱ on $T \times TR$ such that $\varrho(x,y) = 0$, $\varrho(x,y) = \varrho(y,x) \geqq 0$, a weight function (or simply a weight) μ on T is a measure on T. (Observe that $\varrho(x,y)$ may be interpreted e.g. as difficulty or as importance of distinguishing x, y, or else as the "cost" of finding out, given $t \in \{x,y\}$, whether $t = x$ or $t = y$.) Symbols such as $\varrho + \sigma$, $\mu \geqq \nu$ have the usual meaning. If $\varrho(x,y) = 1$ whenever $x \neq y$, then ϱ is denoted by 1.

1.3. Let ϱ and μ be, respectively, a semimetric and a weight on Q. Then $P = \langle Q, \varrho, \mu \rangle$ will be called a finite weighted semimetric space or a FWM-space or simply a space. A space $\langle Q, \varrho, 0 \rangle$ will be denoted by O. We put

$wP = \mu Q$, $dP = \max \varphi(x,y)$, $\ell P = \log |\{q \in Q: \mu q > 0\}|$ if $\mu > 0$, $\ell 0 = 0$. If $B \subset Q$, then μ_B designates the weight function $\mu_B(X) = \mu(X \cap B)$ and P_B or $P \upharpoonright B$ stands for $\langle Q, \varphi, \mu_B \rangle$.

1.4. Notation. $WM(T)$ is the set of all spaces $\langle T, \varphi, \mu \rangle$ endowed with the following topology: $\langle T, \varphi_k, \mu_k \rangle \longrightarrow \langle T, \varphi, \mu \rangle$ iff $\varphi_k(x,y) \longrightarrow \varphi(x,y)$, $\mu_k(x) \longrightarrow \mu(x)$ for all $x,y \in T$; (WM) is the class of all FWM-spaces; $\sim$ is the least equivalence relation on (WM) such that $\langle Q, \varphi, \mu \rangle \sim \langle T, \sigma, \nu \rangle$ if, for some $f: Q \longrightarrow T$, $\nu t = \mu(f^{-1}t)$, $\sigma(fq, fq') = \varphi(q,q')$ for all $q,q' \in Q$, $t \in T$; $Lx = -x \log x$ for $x \geqq 0$; if $\xi = (x_k)$ is finite, then $H\xi$ is the entropy, $\sum Lx_k - L(\sum x_k)$, of ξ; $V(x,y) = H(x,y)/xy$ for $x > 0$, $y > 0$.

1.5. Let $P_k = \langle Q, \varphi, \mu_k \rangle$, $k \in K$, be spaces. We put $\hat{\varphi}(P_1,P_2) = \sum (\mu_1 q \cdot \mu_2 q' \cdot \varphi(q,q') : q,q' \in Q)$, $\varphi(P_1,P_2) = \hat{\varphi}(P_1,P_2)/\mu_1 Q \cdot \mu_2 Q$. If $\mu_1 \leqq \mu_2$, we write $P_1 \leqq P_2$. If $a_k \in R$ for $k \in K$, we put $\sum (a_k P_k : k \in K) = \langle Q, \varphi, \sum a_k \mu_k \rangle$ provided $\sum a_k \mu_k(q) \geqq 0$ for all $q \in Q$. If $P = \sum (P_k : k \in K)$, we call (P_k) a decomposition of P; $[P_k]$ will designate the space $\langle K, \sigma, \nu \rangle$ where $\sigma(k,k') = \varphi(P_k, P_{k'})$, $\nu(k) = \mu_k Q$. – We put $\Gamma(P_1,P_2) = H(wP_1, wP_2)\, \varphi(P_1,P_2)$.

1.6. <u>Definition</u>. A non-negative function φ on (WM) will be called (I) a subentropy if (1) $\varphi\langle Q, a\varphi, b\mu \rangle = ab\, \varphi\langle Q, \varphi, \mu \rangle$ if $a \geqq 0$, $b \geqq 0$, (2) $\varphi P_1 = \varphi P_2$ if $P_1 \sim P_2$, (3) $\varphi\langle Q, \varphi, \mu \rangle \geqq \varphi\langle Q, \sigma, \mu \rangle$ whenever $\varphi \geqq \sigma$,

(4) if $P = \langle \{a,b\}, \varphi, \mu \rangle$, then $\varphi P \leqq H(\mu)\, \varphi\,(a,b)$;
(II) continuous if (5) φ is continuous on every $WM(T)$;
(III) a quasi-entropy if (1) - (5) hold and (6)
$\varphi \langle Q, 1, \mu \rangle = H(\mu)$; (IV) projective if, for any decomposition (P_k) of a space P, $\varphi P \leqq \sum \varphi P_k + \varphi [P_k]$.

1.7. Clearly, there exist projective subentropies, and the l.u.b. of all subentropies is a subentropy. We are going to prove

Theorem. The least upper bound of all projective subentropies is a projective quasi-entropy.

2. We shall need some well-known and/or easily proved facts concerning the entropy H.

2.1. $H(x_k) + H(y_k) \leqq H(x_k + y_k)$.

2.2. If $x_i \geqq y_i > 0$, then $V(x_1,x_2) \leqq V(y_1,y_2)$.

2.3. If $b \geqq a > 0$, then $H(1,a)/H(1,b) \geqq 1 - \log ba^{-1}/\log b$.

Proof. By the mean value theorem, $H(1,a)/H)1,b) =$
$= f(a + x)/f(b + x)$, where $f(u) = \log u + \log e$, $0 \leqq x \leqq 1$.
Hence $H(1,a)/H(1,b) \geqq 1 - \log (b + x)(a + x)^{-1}/\log (b + x)$.

3. We are going to define a function on (WM) which turns out to be (i) a projective quasi-entropy, (ii) the l.u.b. of all projective subentropies.

3.1. We denote by Δ the collection of all $D \subset \bigcup(\{0,1\}^n : n < \omega)$ such that (i) $x \in D$ if $x \dashv y \in D$, (ii) $x0 \in D$ iff $x1 \in D$. If $D \in \Delta$, we put $D' = \{x \in D : x0 \in D\}$, $D'' = D - D'$, $D(X) = \{y \in D : y \leqq x$ for some $x \in X\}$ where $X \subset D$.

3.2. Definition. A family $\mathcal{P} = \langle P_x : x \in D)$ will be called a dyadic expansion (d.e.) of P if (1) $D \in \Delta$, (2) $P_x \leqq P$, (3) $P_x = P_{x0} + P_{x1}$ if $x \in D'$, (4) $P_\emptyset = P$. We put $\Gamma(x) = \Gamma(\mathcal{P}, x) = \Gamma(P_{x0}, P_{x1})$, $\Gamma \mathcal{P} = \sum (\Gamma(x) : x \in D')$. If, in addition, (5) $\ell P_x = 0$ if $x \in D''$, then $\mathcal{P}$ will be called a complete dyadic expansion (c.d.e.) of P. For every P we put $cP = \inf \{ \Gamma \mathcal{P} : \mathcal{P}$ is a c.d.e. of $P \}$.

3.3. Lemma. If $\mathcal{P} = (P_x : x \in D)$ is a d.e. of P, then $cP \leqq \Gamma \mathcal{P} + \sum (cP_x : x \in D'')$.

Proof. Let $\varepsilon > 0$. Choose $\varepsilon(x) > 0$, $x \in D''$, such that $\sum \varepsilon(x) < \varepsilon$. For each $x \in D''$ choose a c.d.e. $\mathcal{P}_x = (P_{x,y} : y \in D_x)$ such that $\Gamma \mathcal{P}_x \leqq cP_x + \varepsilon(x)$. Put $D^* = D' \cup \bigcup (x.D_x : x \in D'')$. If $z \in D'$, put $P^*_z = P_z$; if $x \in D''$, $y \in D_x$, put $P^*_{x.y} = P_{x,y}$. Then $\mathcal{P}^* = (P'_z : z \in D^*)$ is c.d.e. of P, $cP \leqq \Gamma \mathcal{P}^* = \sum (\Gamma(\mathcal{P}, x) : x \in D') + \sum (\Gamma \mathcal{P}_x : x \in D'') \leqq \Gamma \mathcal{P} + \sum cP_x + \sum \varepsilon(x)$.

3.4. If $(P(t) : t \in T)$ is a decomposition of P, $S = [P(t) : t \in T]$, then $cP \leqq cS + \sum cP(t)$.

Proof. Let $\varepsilon > 0$. Choose a c.d.e. $\mathcal{S} = (S_x : x \in D)$ of $S = \langle T, \mathcal{O}, \nu \rangle$ such that $\Gamma \mathcal{S} \leqq cS + \varepsilon$. Put $S_x = \langle T, \mathcal{O}, \nu_x \rangle$, $P_x = \sum (\alpha(x,t) P(t) : t \in T)$ where $\alpha(x,t) = \nu_x(t) / \nu(t)$. Clearly, $\mathcal{P} = (P_x)$ is a d.e. of P. It is easy to see that $\Gamma \mathcal{P} = \Gamma \mathcal{S}$. Since $\ell S_x = 0$ for $x \in D''$, there are $t_x \in T$, $x \in D''$, such that $\nu_x(t) = 0$ if $t \neq t_x$. Since $P_x = \alpha(x, t_x) P(t)$, we have $\sum (cP_x : t_x = t) = cP(t)$. By 3.3, $cP \leqq \Gamma \mathcal{S} + \sum (cP_x : x \in D'') = \Gamma \mathcal{S} + \sum (cP(t) : t \in T)$.

3.5. The function c is a projective subentropy.

Proof. Consider conditions (1) - (4) from 1.6,1. Clearly, (1),(3),(4) are satisfied, (2) and the projectivity follow easily from 3.4.

3.6. If $P_i = \langle Q,1,\mu_i \rangle$, then $\Gamma(P_1,P_2) \geqq$ $\geqq H(\mu_1 + \mu_2) - H\mu_1 - H\mu_2$.

Proof. Put $[P_1,P_2] = [\{1,2\},\sigma,\nu]$. Clearly, $\sigma(1,2) = 1 - \sum(\mu_1 q.\mu_2 q: q \in Q)/m_1 m_2$ where $m_i = \mu_i Q$. By 2.2, $H(m_1,m_2)\ \mu_1 q.\mu_2 q/m_1 m_2 \leqq H(\mu_1 q,\mu_2 q)$, hence $\Gamma(P_1,P_2) \geqq H(m_1,m_2) - \sum(H(\mu_1 q,\mu_2 q): q \in Q)$, from which the assertion follows at once.

3.7. Proposition. $c\langle Q,1,\mu\rangle = H\mu$. For any P, $cP \leqq wP.dP.\ell P$.

Proof. I. Let $P = \langle Q,1,\mu\rangle$. By an easy induction proceeding on $|Q|$ it is shown that $cP \leqq H\mu$. From 3.6, it follows at once that $\Gamma\mathcal{P} \geqq H\mu$ for every c.d.e. $\mathcal{P}$ of P. II. If $P = \langle Q,\varphi,\mu\rangle$, then $cP \leqq wP.dP.c\langle Q,1,\mu/wP\rangle \leqq wP.dP.\ell P$.

3.8. If φ is a projective subentropy, then $\varphi P \leqq cP$ for every P.

Proof. Let $\mathcal{P} = (P_x: x \in D)$ be a c.d.e. of P. By 1.6, 1(4), for every $x \in D'$, $\varphi[P_{x0},P_{x1}] \leqq \Gamma(P_{x0},P_{x1})$. Since φ is projective, $\varphi P_{x0} + \varphi P_{x1} + \varphi[P_{x0},P_{x1}] \geqq \varphi P_x$ for $x \in D'$. This implies $\sum(\varphi[P_{x0},P_{x1}]: x \in D') \geqq \varphi P$, hence $\sum(\Gamma(P_{x0},P_{x1}): x \in D') \geqq \varphi P$, $\Gamma\mathcal{P} \geqq \varphi P$, which proves the assertion.

4. It remains to prove that c is continuous.

4.1. If $P' + P'' = P$, $wP \leqq 1$, $dP \leqq 1$, $wP'' \leqq a$, then $cP' \geqq cP - a.\ell P - H(a,1)$.

Proof. By 3.4, $cP' \geqq cP - cP'' - c[P',P'']$. By 3.7, $c(P'') \leqq a.\ell P$. Clearly, $c[P',P''] \leqq H(wP', wP'') \leqq H(a,1)$.

4.2. Let $p,r,s,t,u \in R$ be positive, $p < 1$, $r = p^{-1}$, $s < 1$, $t = 2^{1/s}$, $u(1 - s)(1 - p)^2 \geqq (1 + p)^2$. Let $P = \langle Q, \varphi, \mu \rangle$, $wP \leqq 1$, $dP \leqq 1$. Let $A \subset Q$, $B = Q - A$, $\varphi(x,y) = 0$ for $x \in A$, $y \in Q$. Then, for every c.d.e. $\mathcal{P} = (P_x : x \in D)$ of P, $u.\Gamma\mathcal{P} \geqq cP_B - (2 + r)a.\ell P - H(a,1) - (1 + r)a\, H(1,t)$, where $a = \mu A$.

Proof. 1. Put $P_x = \langle Q, \varphi, \mu_x \rangle$. For $x \in D$, put $\mu x = \mu_x Q$, $\alpha x = \mu_x A$, $\beta x = \mu_x B$. Put $Z = \{x \in D : \alpha x > p.\beta x\}$. Choose an antichain $X \subset Z$ with $Z \subset D(X)$. Then (1) $\sum(P_x : x \in X) \leqq P$, (2) $wP_x \leqq (1 + r).\alpha x$ if $x \in Z$. Put $Y = \{x \in X : \alpha x \geqq \beta x\}$, $P'_x = P_x - \sum(P_y : y \in D(X) \cap Y)$ for $x \in D - D(Y)$, $P'_x = 0$ for $x \in D(Y)$, $P' = P'_\emptyset$, $P'_x = \langle Q, \varphi, \mu'_x \rangle$, $\beta' x = \mu'_x B$, $P'_{B,x} = P'_x \upharpoonright B$. Then $(P'_x : x \in D)$ is a c.d.e. of P'. Clearly, (3) $\beta x - \beta' x \leqq \alpha x$ if $x \in D - D(Y)$, (4) $P'_x = P_x$ if $x \in X - Y$. It is easy to see that (5) $(1 + p).\beta' x \geqq (1 - p).\mu x$ if $x \notin Z \cup D(Y)$.

II. Let $k = 0,1$; put $\bar{k} = 1 - k$. Put $E_1 = D' \cap D(Y)$, $E' = D' - E_1$, $E_2 = \{x \in E' : xk \in Y \text{ for some } k\}$, $E_3 = \{x \in E' : \text{for some } k,\ xk \in X - Y,\ t\,\beta'(xk) \geqq \beta'(x\bar{k})\}$, $E_4 = \{x \in E' : xk \in X - Y \text{ for some } k\} - E_3$, $E_5 = \{x \in E' : xk \notin Z,\ k = 0,1\}$. It is easy to see that E_i are disjoint, $\bigcup E_i = D'$. Put $\Gamma(x) = \Gamma(P_{x0}, P_{x1})$, $\Gamma'(x) = \Gamma(P'_{B,x0}, P'_{B,x1})$, $g_i = \sum(\Gamma(x) : x \in E_i)$, $g'_i = \sum(\Gamma'(x) : x \in E_i)$,

$g_1^* = \sum(cP'_{B,x} : x \in E_1)$. Lemma 3.3 implies (6) $g_1^* + \sum(g_i : i = 2,\dots,5) \geq cP'_B$. By 3.7, $cP'_{B,x} \leq \beta' x.\ell P$, hence (7) $g_1^* \leq (1 + r)\, a.\ell P$, by (1),(2). Clearly, $g'_2 = 0$.

III. If $x \in E_3$, $xk \in X - Y$, then $\Gamma'(x) \leq H(\beta'(xk), \beta'(x\bar{k})) \leq \beta'(xk)\, H(1,t)$, hence, by (1),(2), $g'_3 \leq (1 + r)\, a\, H(1,t)$. - If $x \in E_4$, $xk \in X - Y$, put $m = \mu(xk)$, $\bar{m} = \mu(x\bar{k})$, $b = \beta'(xk)$, $\bar{b} = \beta'(x\bar{k})$. Clearly, (8) $tb < \bar{b}$. Since $x\bar{k} \notin Z$, (5) implies (9) $\bar{b}\bar{m} \geq (1 - p)(1 + p)^{-1}$. Since $xk \in X - Y$, we have, by (4), $\alpha(xk) < b$, hence (10) $2b > m$. By (8)-(10) and 2.3, $V(m,\bar{m})/V)b,\bar{b}) \geq (1 - s)(1 - p)(1 + p)^{-1}$ hence $ug_4 \geq g'_4$. - If $x \in E_5$, then, by (5), $\beta'(xk) \leq (1 - p)(1 + p)^{-1} \mu(xk)$, $k = 0,1$, hence, by 2.2, $ug_5 \geq g'_5$.

IV. By (6),(7), and III, $u.\Gamma\mathcal{P} - cP'_B \geq - (1 + r)\, a.\ell P - (1 + r)\, a\, H(1,t)$. By 4.1, $cP'_B \geq cP_B - a.\ell P - H(a,1)$. This completes the proof.

4.3. For every T and every $\varepsilon > 0$ there exists a $\delta > 0$ such that if $P' \leq P$, $P \in WM(T)$, $wP \leq 1$, $dP \leq 1$, $w(P - P') \leq \delta$, then $|cP - cP'| \leq \varepsilon$.

This is easily deduced from 4.1, 4.2.

4.4. Let $P = \langle Q, \varphi, \mu \rangle$, $P(n) = \langle Q, \varphi_n, \mu \rangle$ be spaces, $P(n) \to P$ (in $WM(Q)$). Then $cP_n \to cP$.

Proof. I. Let $\mathcal{P} = \{P_x : x \in D\}$ be a c.d.e. of P. Let $P_x(n)$ be the space obtained from P_x by replacing φ with φ_n. It is easy to see that $\mathcal{P}_n = \{P_x(n) : x \in D\}$ is a c.d.e. of P(n) and $\Gamma\mathcal{P}_n \to \Gamma\mathcal{P}$, hence $\limsup cP_n \leq cP$. - II. There exist $t_n < 1$ such that $t_n \to 1$, and $\varphi_n(x,y) \geq t_n\, \varphi(x,y)$ for all n and all $x,y \in Q$. Hence

$cP_n \geq t_n . cP$.

4.5. Propositions 4.3 and 4.4 imply that c is continuous. Hence, by 3.5 and 3.7, c is a projective quasi-entropy. By 3.8, c is equal to the l.u.b. of all projective subentropies. The theorem 1.7 is proved.

5. Remarks. 1) I do not know whether there exist projective quasi-entropies $\varphi \neq c$. - 2) A subentropy φ may be called inductive if $\varphi P \geq \varphi P_1 + \varphi P_2$ whenever $P = P_1 + P_2$. Problem: do there exist inductive quasi-entropies? - 3) Let ρ and μ be, respectively, a semimetric and a measure on M. If $X \subset M$, $Y \subset M$, put $\hat{\rho}(X,Y) = \int_{X \times Y} \rho(x,y) d\mu(x) d\mu(y)$, $\rho(X,Y) = \hat{\rho}(X,Y) / \mu X . \mu Y$. Consider finite measurable decompositions $\mathcal{V} = (V_k : k \in K)$, $\bigcup V_k = M$. Assume that there are sufficiently many, in a sense which can be specified, decompositions (V_h) such that $\rho(V_h, V_k) < \infty$ for $h \neq k$. Define $[\mathcal{V}] = [V_k]$ in a way quite analogous to that in 1.5, and define $C \langle M, \rho, \mu \rangle$ to be the limit, provided it exists, of $c[\mathcal{V}]$ with respect to the filter of all $\mathcal{V}$ described above. Problems: (a) to find $\langle M, \rho, \mu \rangle$ for which the definition works, possibly after a suitable modification; (b) to find a characterization of C analogous to that of c in 1.7; (c) to introduce C directly by means of suitably defined c.d. expansions (cf. 3.2). - 4) Concepts just described may be useful e.g. if, in addition, a topology on M is given and ρ is continuous at every $\langle x,y \rangle$, $x \neq y$.

Matematický ústav ČSAV
Žitná 25, Praha 1
Československo

(Oblatum 18.10.1976)