Commentationes Mathematicae Universitatis Carolinae

Miroslav Katětov

Extensions of the Shannon entropy to semimetrized measure spaces

Commentationes Mathematicae Universitatis Carolinae, Vol. 21 (1980), No. 1, 171--192

Persistent URL: http://dml.cz/dmlcz/105986

Terms of use:

© Charles University in Prague, Faculty of Mathematics and Physics, 1980

Institute of Mathematics of the Academy of Sciences of the Czech Republic provides access to digitized documents strictly for personal use. Each copy of any part of this document must contain these *Terms of use*.

This paper has been digitized, optimized for electronic delivery and stamped with digital signature within the project *DML-CZ: The Czech Digital Mathematics Library* http://project.dml.cz

COMMENTATIONES MATHEMATICAE UNIVERSITATIS CAROLINAE
21, 1 (1980)

EXTENSIONS OF THE SHANNON ENTROPY TO SEMIMETRIZED MEASURE SPACES

Miroslav KATĚTOV

Abstract: Extensions of the Shannon entropy to measure spaces endowed with a semimetric are examined.

Key words: Entropy, measure space, semimetric, dyadic expansion, lower limit with respect to a filter.

Classification: 94A17

We investigate the possibility of extending the Shannon entropy to the class of all semimetrized measure spaces, i.e. sets endowed with a finite measure and a measurable semimetric. It turns out that various extensions are possible. We examine two of these, denoted by C and C^*, which satisfy certain natural conditions.

The extension C^* can be characterized, among all those functions φ on the class of semimetrized measure spaces which fulfil certain general requirements, as the maximal one satisfying the following conditions: (1) $\varphi(P) \leq \varphi(P_1) + \varphi(P_2) + H(\mu_1, \mu_2) r(P_1, P_2)$ where (P_1, P_2) is a partition of P into two measurable sets, μ_i, i = 1,2, is the measure of P_i, and $r(P_1, P_2)$ is the average distance between points of P_1 and P_2, (2) every $\gamma < \varphi(P)$ is majorized by every $\varphi(S)$ where S is the quotient space of a sufficiently fine partition of P. The extension C^* can be characterized as the lower limit of $C^*(S)$, where S is the quotient space of a "partition" of P corresponding to an arbitrary partition of unity into finitely many measurable functions.

The functions C and C^* possess various convenient properties. Some of these are stated below (3.10, 3.12 - 3.15); however, the proofs of the corresponding propositions do not appear in this note and are intended for publication elsewhere.

We also sketch the definitions of some other extensions of the Shannon entropy as well as certain analogies, different in kind from the well known ones, between dimension and entropy.

It may be asked why one should try to extend the Shannon entropy. One reason lies in the fact that the entropy of finite probability spaces, the ε-entropy of metric spaces and the differential entropy exhibit many common properties whereas no concept seems to have been introduced so far from which all of these could be obtained, at least up to minor points, in a natural way. Other reasons come from applications. Thus, the intuitive concept of the "information content" of a metrized set of possible stimuli appearing with a certain probability seems not to have been as yet expressed mathematically in a satisfactory way, in spite of its importance in problems of human information processing (for an elementary exposition of pertinent topics see e.g. Chapter 10 of the textbook [1], which also contains further references).

The problem of extending the Shannon entropy has been investigated for the case when the underlying set is finite in the author's note Quasi-entropy of finite weighted metric spaces, Comment. Math. Univ. Carolinae 17(1976), 797-806. This note, of which the present one is a free continuation, will be referred to as QE. Most definitions appearing in QE will be restated or extended in the present article.

It is to be noted that a special case of the problem of introducing an entropy for probability spaces equipped with a metric has been examined in [3]. However, the approach applied in [3] is quite different from that in QE and the present note, and only the case of a real random variable with a finite range is considered.

We prove only some of the statements in full. Many proofs are only sketched, and simple or straightforward ones are often omitted. From information theory and theory of measure we presuppose only the basic notions. From set theory and topology we need, in fact, only the concept of a filter.

1.

1.1. We recall some terminological and notational conventions, most of which have been already used in QE. - A) A function f is, by definition, a mapping of a class into $\overline{R}$, the extended set of reals. If all f(x) are real, then f is called a real function. - B) We put $a.\infty = \infty$ if $a > 0$, $0.\infty = 0$, $0/0 = 0$. We write log instead of $\log_2$. We put $0.\log 0 = 0$. For $x \geq 0$, we put $L(x) = -x \log x$. - C) Let μ be a finite (the word "finite" is often omitted) measure on a set Q. Then dom μ denotes the domain of μ (observe that we always assume that dom μ is a σ-field) and μ denotes the

Lebesgue extension of μ. Instead of $\tilde{\mu}(X)$, we often write $\mu(X)$. - D) If μ is a measure on Q and $B \in \operatorname{dom} \tilde{\mu}$, then μ_B denotes the measure defined by $\mu_B(X) = \tilde{\mu}(X \cap B)$ for $X \in \operatorname{dom} \mu$. - E) If μ, ν are measures on Q, then $\mu \geqq \nu$ means that $\operatorname{dom} \mu = \operatorname{dom} \nu$, $\mu(X) \geqq \nu(X)$ for all $X \in \operatorname{dom} \mu$. - F) Let μ_i, $i=1,2$, be a measure on Q_i. The product $\mu_1 \times \mu_2$ is defined in the usual way ($\operatorname{dom}(\mu_1 \times \mu_2)$ is the least σ-field containing all $X_1 \times X_2$, where $X_i \in \operatorname{dom} \mu_i$). - G) A semimetric on a set Q is, by definition, a non-negative real function ρ on $Q \times Q$ such that $\rho(x,x) = 0$, $\rho(x,y) = \rho(y,x)$. A semimetric ρ on a set Q will be denoted by 1 if $\rho(x,y) = 1$ whenever $x,y \in Q$, $x \neq y$. - H) For any countable indexed set $\mu = (\mu_k : k \in K)$ of non-negative reals such that $\sum \mu_k < \infty$, we put $H(\mu) = H(\mu_k : :k \in K) = \sum (L(\mu_k) : k \in K) - L(\sum(\mu_k : k \in K))$. Thus, if μ is a probability measure on a finite set Q and all $\{x\}$, $x \in Q$, are in $\operatorname{dom} \mu$, then $H(\mu(\{q\}) : q \in Q)$ is the Shannon entropy of μ. To cover the case when some $\{q\}$ is not in $\operatorname{dom} \mu$, we introduce the following trivial extension of the Shannon entropy: the entropy of a finite probability space $\langle Q, \mu \rangle$ is equal to $\sum(L(\mu(A)) : A \in \mathcal{A})$ where $\mathcal{A}$ is the set of atoms of $\operatorname{dom} \mu$.

1.2. <u>Convention</u>. In formulas, we omit parentheses whenever possible without a danger of misunderstanding, and write e.g. μX instead of $\mu(X)$, μq instead of $\mu(\{q\})$, etc.

1.3. <u>Definition</u>. A triple $P = \langle Q, \rho, \mu \rangle$, where Q is a non-void set, μ is a finite measure on Q, ρ is a $(\mu \times \mu)$-measurable semimetric on Q, will be called a semimetrized measure space or a WM-space (WM stands for "weighted semimetric", the expression used in QE). The class of all WM-spaces will be denoted by $\{WM\}$.

1.4. Remark. Clearly, FWM-spaces (finite weighted semimetric spaces) considered in QE coincide with WM-spaces $\langle Q,\varrho,\mu\rangle$ such that Q is finite and every $\{q\}$, $q\in Q$, is in dom μ. (Observe that the condition $\{q\}\in$ dom μ, not stated explicitly in QE, is tacitly understood there - see e.g. QE 1.3 and QE 1.4.)

1.5. We shall call a WM-space $\langle Q,\varrho,\mu\rangle$ finite if the underlying set Q is finite, an FWM-space if, in addition, all $\{q\}$, $q\in Q$, are in dom μ. The class of all FWM-spaces will be denoted by $\{FWM\}$. For a non-void Q, $\{WM(Q)\}$ will denote the set of all WM-spaces $\langle Q,\varrho,\mu\rangle$. If, in addition, Q is finite, then $\{FWM(Q)\}$ will denote the set of all FWM-spaces $\langle Q,\varrho,\mu\rangle$.

1.6. Examples. A) If μ is a finite measure on $Q\neq\emptyset$, then $\langle Q,1,\mu\rangle$ is a WM-space provided the diagonal of $Q\times Q$ is $(\mu\times\mu)$-measurable. - B) If $\langle Q,\tau\rangle$ is a (symmetric) graph, i.e. τ is a symmetric binary relation on Q, and μ is a finite measure on Q, then $\langle Q,\varrho_\tau,\mu\rangle$, where $\varrho_\tau(x,y)=0$ iff $x=y$ or $\langle x,y\rangle\in\tau$, $\varrho_\tau(x,y)=1$ otherwise, is a WM-space provided $\tau\cup\{\langle x,x\rangle : x\in Q\}$ is $(\mu\times\mu)$-measurable. On the other hand, every WM-space $\langle Q,\varrho,\mu\rangle$ with a $\{0,1\}$-valued ϱ is obtained from the graph $\langle Q,\{\langle x,y\rangle : \varrho(x,y)=0\}\rangle$. - C) If x is a real-valued random variable on a probability space $\langle Q,\mu\rangle$, then we associate with x the space $\langle R,\varrho,\nu\rangle$, where ϱ is the usual metric on R, $\nu Y=\mu\{q : x(q)\in Y\}$ for every Borel set $Y\subset R$.

1.7. Convention (see QE 1.4). If Q is finite non-void, then $\{FWM(Q)\}$ (see 1.5) will also denote the set $\{FWM(Q)\}$ endowed with the topology defined by $\langle Q,\varrho_n,\mu_n\rangle\longrightarrow\langle Q,\varrho,\mu\rangle$ iff $\varrho_n(x,y)\longrightarrow\varrho(x,y)$ for all $x,y\in Q$, $\mu_n x\longrightarrow\mu x$ for all $x\in Q$.

1.8. <u>Definition</u> (cf. QE 1.5). Let $P = \langle Q, \rho, \mu \rangle$ be a WM-space. If $\nu \leqq \mu$, we shall call $S = \langle Q, \rho, \nu \rangle$ a subspace of P and we shall write $S \leqq P$. If, in addition, $\nu = \mu_B$ for some $B \in \operatorname{dom} \tilde{\mu}$, then S is called pure subspace.

1.9. <u>Remark</u>. If $\langle Q, \rho, \nu \rangle$ is a subspace of $\langle Q, \rho, \mu \rangle$, then, by the Radon-Nikodým theorem, $\nu = f \cdot \mu$ (i.e. $\nu X = \int_X f(x) d\mu(x)$ for every $X \in \operatorname{dom} \mu$), where f is μ-measurable and $0 \leqq f \leqq 1$. Hence, if μ is given, we shall sometimes write $\langle Q, \rho, f \rangle$ instead of $\langle Q, \rho, \nu \rangle$, and even f instead of $\langle Q, \rho, \mu \rangle$ provided ρ is given as well.

1.10. <u>Definition</u> (cf. QE 1.5). If $P_k = \langle Q, \rho, \mu_k \rangle$, $k \in K$, are WM-spaces, the domains dom μ_k coincide, and K is finite, then the WM-space $\langle Q, \rho, \Sigma(\mu_k : k \in K) \rangle$ will be denoted by $\Sigma(P_k : k \in K)$. If $(P_k : k \in K)$ is finite and $\Sigma P_k = P$, then $(P_k : k \in K)$ will be called a partition (or a decomposition, see QE) of P. If card $K = 2$, then $(P_k : k \in K)$ will be called binary. A partition $(P_k : k \in K)$ of a space $P = \langle Q, \rho, \mu \rangle$ will be also denoted by $(f_k : k \in K)$, where f_k are functions such that $\langle Q, \rho, f_k \cdot \mu \rangle = P_k$. - A partition (P_k) of P will be called pure if all P_k are pure subspaces of P. A pure partition $(P_k : k \in K)$ of $\langle Q, \rho, \mu \rangle$ will be also denoted by $(Q(k) : k \in K)$ where Q(k) are such that $P_k = \langle Q, \rho, \mu_{Q(k)} \rangle$.

1.11. <u>Notation</u>. If $P = \langle Q, \rho, \mu \rangle$ is a WM-space, we put $wP = \mu Q$, $\overline{d}(P) = \operatorname{ess\,sup}(\rho(x,y) : \langle x,y \rangle \in Q \times Q)$. If $P_i = \langle Q, \rho, \mu_i \rangle$, $i = 1,2$, are WM-spaces, dom μ_1 = dom μ_2, we put $\hat{r}(P_1, P_2) = \int_{Q \times Q} \rho \, d(\mu_1 \times \mu_2)$, $r(P_1, P_2) = \hat{r}(P_1, P_2) / wP_1 \cdot wP_2$. Clearly, if P_i are pure subspaces of a space $\langle Q, \rho, \mu \rangle$, $\mu_i = \mu_{B(i)}$, then $r(P_1, P_2)$ is the "average distance" of points $x \in B(1)$ and $y \in B(2)$. - Observe that the notation just introduced differs from that in QE.

1.12. Definition (cf. QE 1.5). Let $\mathcal{U} = (U_k : k \in K)$, where $U_k = \langle Q, \varrho, \mu_k \rangle$, be a partition of $P = \langle Q, \varrho, \mu \rangle$. If all $\hat{r}(U_i, U_j)$, $i, j \in K$, $i \neq j$, are finite, then the WM-space $\langle K, \sigma, \nu \rangle$, where $\sigma(i,j) = r(U_i, U_j)$ for $i \neq j$, and $\sigma\{k\} = \mu_k Q$ will be denoted by $[\mathcal{U}]$ or $[U_k : k \in K]$ and called the quotient space of the partition $\mathcal{U} = (U_k : k \in K)$. -

Convention. If φ is a function on a class $\mathcal{X} \subset \{WM\}$ and $\mathcal{U} = (U_k : k \in K)$ is a partition of $P \in \mathcal{X}$, then we put, by definition, (1) $\varphi(\mathcal{U}) = \varphi[\mathcal{U}]$ if $r(U_i, U_j) < \infty$ whenever $i, j \in K$, $i \neq j$, and (2) $\varphi(\mathcal{U}) = \infty$ if $r(U_i, U_j) = \infty$ for some $i, j \in K$, $i \neq j$.

1.13. Definition. Let $P = \langle Q, \varrho, \mu \rangle$, $S = \langle T, \sigma, \nu \rangle$ be WM-spaces. Let $f \subset Q \times T$. The triple $\langle f, P, S \rangle$ will be called a conservative morphism of P onto S and will be denoted by $f : P \longrightarrow S$ if the following conditions are satisfied: (1) if $X \subset Q$ is μ-measurable, then $f(X) \subset T$ is ν-measurable, $\nu f(X) = \mu X$; (1') if $Y \subset T$ is ν-measurable, then $f^{-1}(Y)$ is μ-measurable, $\mu f^{-1}(Y) = \nu Y$; (2) if $X \subset Q \times Q$ is $(\mu \times \mu)$-measurable, then $\int_Y \sigma \, d(\nu \times \nu) = \int_X \varrho \, d(\mu \times \mu)$, where $Y = (f \times f)(X)$; (2') if $Y \subset T \times T$ is $(\nu \times \nu)$-measurable, then $\int_X \varrho \, d(\mu \times \mu) = \int_Y \sigma \, d(\nu \times \nu)$, where $X = (f^{-1} \times f^{-1})(Y)$. - Remark. If $P = \langle Q, 1, \mu \rangle$, $S = \langle T, 1, \nu \rangle$ are WM-spaces, then conservative morphisms $f : P \longrightarrow S$ such that f is a one-to-one relation coincide with the "isomorphisms mod 0" for measure spaces.

1.14. Proposition. If $f_1 : P_1 \longrightarrow P_2$, $f_2 : P_2 \longrightarrow P_3$ are conservative morphisms, then so is $f_2 \circ f_1 : P_1 \longrightarrow P_3$. If $f : P \longrightarrow S$ is a conservative morphism, then so is $f^{-1} : S \longrightarrow P$.

1.15. Definition. If P, S are WM-spaces and there exists a conservative morphism $P \longrightarrow S$, we shall say that P is equivalent to S and write $P \sim S$.

1.16. It follows at once from 1.14 that the relation $\sim$ is an equivalence relation on $\{WM\}$. - It is easy to prove that the relation $\sim$ coincides on $\{FWM\}$ with the equivalence relation, also denoted $\sim$, introduced in QE 1.4.

2.

2.1. Definition. A non-negative function $\varphi : \mathcal{X} \longrightarrow \widehat{R}$ where $\mathcal{X} \subset \{WM\}$ will be called a semi-subentropy on $\mathcal{X}$ if (1) $a \geqq 0$, $b \geqq 0$, $\langle Q, \wp, \mu \rangle \in \mathcal{X}$, $\langle Q, a\wp, b\mu \rangle \in \mathcal{X}$, then $\varphi \langle Q, a\wp, b\mu \rangle = ab\, \varphi \langle Q, \wp, \mu \rangle$, (2) if $\langle Q, \wp_1, \mu \rangle \in \mathcal{X}$, $\langle Q, \wp_2, \mu \rangle \in \mathcal{X}$, $\wp_1 \leqq \wp_2$, then $\varphi \langle Q, \wp_1, \mu \rangle \leqq \varphi \langle Q, \wp_2, \mu \rangle$, (3) if $P = \langle \{q_1, q_2\}, \wp, \mu \rangle \in \{FWM\}$, then $P \in \mathcal{X}$, $\varphi P \leqq H(\mu q_1, \mu q_2)\, \wp(q_1, q_2)$; (4) if $P = \langle Q, \wp, \mu \rangle \in \mathcal{X}$, Q is finite, then $\varphi P < \infty$. If, in addition, $\varphi P_1 = \varphi P_2$ whenever $P_1 \in \mathcal{X}$, $P_2 \in \mathcal{X}$, $P_1 \sim P_2$, then φ is called a subentropy on $\mathcal{X}$. - If "on $\mathcal{X}$" is omitted, then it is understood that $\mathcal{X} = \{WM\}$.

2.2. It is easy to see that a subentropy on $\{FWM\}$ is precisely a subentropy in the sense of QE 1.6.

2.3. Definition. A semi-subentropy φ on $\mathcal{X}$ will be called projective (semiprojective) if, for any binary partition (any binary pure partition) $(P_k : k \in K)$ of a WM-space $P \in \mathcal{X}$ such that all P_j are in $\mathcal{X}$ and $\hat{r}(P_i, P_j) < \infty$ for $i, j \in K$, $i \neq j$, the quotient space $[P_k : k \in K]$ is in $\mathcal{X}$ and $\varphi P \leqq \Sigma(\varphi P_k : k \in K) + \varphi[P_k : k \in K]$.

2.4. <u>Remark</u>. If "binary" is omitted in 2.3, we get two further properties (with respect to which we have stated the definition in a somewhat involved form). I do not know under what conditions these properties of φ are equivalent, respectively, to those defined in 2.3.

2.5. <u>Example</u>. Put $\varphi P = 2\hat{r}(P,P)/wP$. It is easy to show that φ is a projective subentropy (and possesses also the stronger property mentioned in 2.4).

2.6. It is to be noted that there is an error in QE, 1.6(IV): "decomposition" should be replaced by "binary decomposition" and the definition should run as follows: "(IV) projective, if, for any binary decomposition (P_k) of a space P, $\varphi P \leqq \sum \varphi P_k + \varphi[P_k]$". In QE 3.4 "decomposition" should be replaced by "binary decomposition", too. After these corrections and a further correction mentioned below (2.8), all propositions in QE remain valid.

2.7. <u>Notation</u> (see QE 3.1). We denote by Δ the collection of all finite non-void sets $D \subset \bigcup(\{0,1\}^n : n < \omega)$ (i.e. of all finite non-void sets of strings, including the void string, of elements 0 and 1) such that (1) if $x \in D$, then every segment of x is in D, (2) if $x \in D$, then either $\{x0, x1\} \subset D$ or $\{x0, x1\} \cap D = \emptyset$. If $D \in \Delta$, we put $D' = \{x \in D : x0 \in D\}$, $D'' = D \setminus D'$.

2.8. In QE 3.1, the words "finite non-void" (sets of strings) have been omitted by mistake. However, in the subsequent text of QE it is tacitly assumed that D is finite non-void.

2.9. <u>Definition</u> (cf. QE 3.2). Let $P = \langle Q, \varphi, \mu \rangle$ be a

WM-space. A family (indexed set) $(P_z : z \in D)$ will be called a dyadic expansion (abbreviated "d.expansion") of P if (1) $D \in \Delta$; (2) $P_x = P_{x0} + P_{x1}$ if $x \in D'$, (3) $P_\emptyset = P$. Instead of $(P_z : z \in D)$ we sometimes write $(f_z : z \in D)$ where f_z are such that $\langle Q, \wp, f_z \rangle = P_z$. If $\mathcal{P} = (P_z : z \in D)$ is a d.expansion, then $(P_z : z \in D'')$ will be denoted by $\mathcal{P}''$. - A d.expansion $(P_z : z \in D)$ will be called pure if all P_z are pure subspaces.

2.10. Remarks. 1) Dyadic expansions are closely related to "questionnaires" introduced by C.F. Picard [4]. - 2) The name "expansion" has been chosen with a view to the following more general concept which also embraces e.g. expansions into series. Let A be a non-void set. We shall call a "polydromic structure" on A every set of strings on A (thus, in fact, a polydromic structure on A is a language with the alphabet A; however, this terminology is not convenient for the present purpose). A polydromic structure D on A will be called (for the purpose of the present remark) regular if it is non-void and (1) if $x \in D$, then every segment of x is in D; (2) if $x \in D$, then either $xa \in D$ for every $a \in A$ or xa non$\in$ D for every $a \in A$; (3) there exists no infinite sequence (a_i), $a_i \in A$, such that all $(a_i : i < n)$ are in D. If D is regular, we define D' and D'' in a way completely analogous to that in 2.7 (observe that Δ consists exactly of all regular polydromic structures on $\{0,1\}$). If D is a regular polydromic structure on N, then a family $(f_z : z \in D)$ of functions on a given set T will be called an N-polydromic expansion of $f: T \longrightarrow R$ into the family $(f_z : z \in D'')$ if (1) for every $x \in D'$, $f_x = \Sigma (f_{xi} : i \in N)$, (2) $f_\emptyset = f$. Clearly, expansion into series is a special case (with D consisting of $\emptyset$ and all (i), $i \in N$). It is not difficult to prove that a function f on a topological space is a Baire function iff it possesses an N-polydromic expansion into continuous functions.

2.11. Notation (see QE 1.5, 3.2). If $P_i = \langle Q, \varrho, \mu_i \rangle$, $i = 1,2$, are WM-spaces, $\mathrm{dom}\,\mu_1 = \mathrm{dom}\,\mu_2$, we put $\Gamma(P_1,P_2) = H(wP_1,wP_2)r(P_1,P_2)$. If $\mathcal{P} = (P_z : z \in D)$ is a dyadic expansion of a WM-space, we put $\Gamma(\mathcal{P}) = \Sigma(\Gamma(P_{z0},P_{z1}) : z \in D')$.

2.12. Definition. Let $\mathcal{U} = (U_k : k \in K)$, $\mathcal{V} = (V_m : m \in M)$ be partitions of a WM-space P. If there exists a partition, in the usual sense, $(M(k) : k \in K)$ of the set M such that, for each $k \in K$, $U_k = \Sigma(V_m : m \in M(k))$, then we shall say that $\mathcal{V}$ refines $\mathcal{U}$. If $\mathcal{U}$ refines $\mathcal{V}$ and is refined by $\mathcal{V}$, then $\mathcal{U}$ and $\mathcal{V}$ are said to be equivalent.

2.13. Notation. For any $P \in \{WM\}$, we shall consider four filters described below (2.16). Let $De^a(P)$, $De^p(P)$, $Pt^a(P)$, and $Pt^p(P)$ denote, respectively, the set of all d.expansions, all pure d.expansions, all partitions, and all pure partitions of P. If $\mathcal{U} \in Pt^a(P)$, we put $\Phi^a_{de}(\mathcal{U}) = \{\mathcal{P} \in De^a(P) : \mathcal{P}'' \text{ refines } \mathcal{U}\}$, $\Phi^a_{pt}(\mathcal{U}) = \{\mathcal{V} \in Pt^a(P) : \mathcal{V} \text{ refines } \mathcal{U}\}$, $\Phi^p_{de}(\mathcal{U}) = \{\mathcal{P} \in De^p(P) : \mathcal{P}'' \text{ refines } \mathcal{U}\}$, $\Phi^p_{pt}(\mathcal{U}) = \{\mathcal{V} \in Pt^p(P) : \mathcal{V} \text{ refines } \mathcal{U}\}$. - Observe that in Φ^a_{de}, Φ^p_{de} etc., a stands for "all", p for "pure".

We shall need the following simple facts.

2.14. If $\mathcal{U}$ and $\mathcal{V}$ are partitions (pure partitions) of a WM-space P, then there is a partition (pure partition) $\mathcal{W}$ refining both $\mathcal{U}$ and $\mathcal{V}$.

Proof. Write $\mathcal{U}$, $\mathcal{V}$ in the form $\mathcal{U} = (\langle Q, \varphi, f_k \rangle : k \in K)$,

$\mathcal{V} = (\langle Q,\varphi,g_m\rangle : m \in M)$ and put $\mathcal{W} = (\langle Q,\varphi,h_{k,m} : \langle k,m\rangle \in K\times M)$, where $h_{k,m}$ is the product of functions f_k, g_m.

2.15. If $\mathcal{U}$ is a partition of a WM-space P, then there exists a dyadic expansion $\mathcal{P}$ of P such that $\mathcal{P}''$ is equivalent to $\mathcal{U}$.

2.16. It follows from 2.14 and 2.15 that each of the following sets is a base of a filter: (1) $\{\Phi^{a}_{de}(\mathcal{U}): \mathcal{U} \in \in Pt^{a}(P)\}$, (2) $\{\Phi^{a}_{pt}(\mathcal{U}): \mathcal{U} \in Pt^{a}(P)\}$, (3) $\{\Phi^{P}_{de}(\mathcal{U}): \mathcal{U} \in \in Pt^{P}(P)\}$, (4) $\{\Phi^{P}_{pt}(\mathcal{U}): \mathcal{U} \in Pt^{P}(P)\}$. The corresponding filters will be called, respectively, the projective filter on d.expansions of P, the projective filter on partitions of P, the semiprojective filter on pure d.expansions of P, the semiprojective filter on pure partitions of P. If there is no danger of confusion, the symbol $\mathcal{F}$ or $\mathcal{F}(P)$ will be used for any of these filters.

2.17. We recall that if $\mathcal{G}$ is a filter on a set A and h is a function on A, then the lower limit of h with respect to $\mathcal{G}$ is defined as follows: $\mathcal{G}$-$\underline{\lim}$ h = sup (inf h(G):$G \in \mathcal{G}$).

2.18. If $\mathcal{F}$ is the filter generated by the collection (1) or (2) (respectively, (3) or (4)) described in 2.16, then $\mathcal{F}$-$\underline{\lim}$ h will be called the lower projective (respectively, lower semiprojective) limit of h.

3.

3.1. <u>Definition</u>. If P is a WM-space, then, by definition, $C^{*}(P)$ is equal to the lower semiprojective limit of $\Gamma(\mathcal{P})$, where $\mathcal{P}$ is a pure dyadic expansion, and C(P) is equal to the lower projective limit of $\Gamma(\mathcal{P})$, where $\mathcal{P}$ is a dyadic expansion. The functions $P \mapsto C(P)$ and $P \mapsto C^*(P)$, defined on $\{WM\}$, will be denoted by C and C^*, respectively.

3.2. Theorem. The function C^* is a semiprojective semi-subentropy. For every WM-space P, $C^*(P)$ is equal to the lower semiprojective limit of $C^*(\mathcal{U})$, where $\mathcal{U}$ is a pure partition. If φ is a semiprojective semi-subentropy such that, for every WM-space P, φP does not exceed the lower semiprojective limit of $\varphi(\mathcal{U})$, where $\mathcal{U}$ is a pure partition of P, then $\varphi \leqq C^*$. The function C^* is continuous on every $\{FWM(Q)\}$. If $\langle Q, \mu \rangle$ is a finite probability field, then $C^* \langle Q, 1, \mu \rangle$ coincides with the Shannon entropy of $\langle Q, \mu \rangle$.

3.3. We shall need the following simple facts. - (A) For any d.expansion $\mathcal{P} = (P_z : z \in D)$ of a WM-space P, $\Gamma(\mathcal{P}) \geqq C^*(\mathcal{P}'')$. - (B) For any partition (pure partition) $\mathcal{V}$ of a WM-space P there exists a d.expansion (pure d.expansion) $\mathcal{P}$ such that $\mathcal{P}''$ is equivalent to $\mathcal{V}$ and $\Gamma(\mathcal{P}) = C^*(\mathcal{V})$. - (C) If a semi-subentropy φ is semiprojective, then for any pure d.expansion $\mathcal{P} = (P_z : z \in D)$ of P we have $\varphi P \leqq \Sigma(\varphi[P_{z0}, P_{z1}] : z \in D') + \Sigma(\varphi(P_z) : z \in D'')$. - (D) If P is a finite WM-space, then $C^*(P)$ is equal to the least $\Gamma(\mathcal{P})$, where $\mathcal{P} = (P_z : z \in D)$ is a pure dyadic expansion of P such that no P_z, $z \in D''$, has a nontrivial pure partition.

3.4. We now give an outline of the proof of 3.2. - The fact that $C^*(P)$ is equal to the lower limit of $C^*(\mathcal{U})$ follows from 3.3 A, B. - The assertion concerning φ is proved as follows: 3.3 D implies that, for a finite WM-space P, there exists a dyadic expansion $\mathcal{P} = (P_z : z \in D)$ such that $C^*(P) = \Gamma(\mathcal{P})$ and $\varphi P_z = 0$ for $z \in D''$; 3.3 C implies that

$\varphi P \leqq \sum(\varphi[P_{zo},P_{z1}]:z \in D') \leqq \sum(\Gamma(P_{zo},P_{z1}):z \in D') = \Gamma(\mathcal{P})$. Since, for any $P \in \{WM\}$, φP does not exceed the lower limit of $\varphi(\mathcal{U})$, the assertion is proved. - To prove that C^* is continuous on $\{FWM(Q)\}$, Q finite, let $\mathcal{T}$ be the set of all "proper dyadic expansions" of the set Q, i.e. of families $T = (T(z):z \in D)$ such that $D \in \Delta$, $T(\emptyset) = Q$, and if $z \in D'$, then $T(z0) \cup T(z1) = T(z)$, $T(zo) \cap T(z1) = \emptyset$, $T(zo) \neq \emptyset$, $T(z1) \neq \emptyset$. For any $T = (T(z):z \in D) \in \mathcal{T}$ and $P = \langle Q,\varphi,\mu\rangle \in \{FWM\}$ put $f(P,T) = \Gamma(\mathcal{P})$, where $\mathcal{P} = (P_z:z \in D)$, $P_z = \langle Q,\varphi,\mu_{T(z)}\rangle$. Clearly, for any fixed T, $P \mapsto f(T,P)$ is continuous on $\{FWM(Q)\}$. This proves the continuity of C^* since $\mathcal{T}$ is finite and $C^*(P) = \min(f(P,T):T \in \mathcal{T})$.

3.5. Theorem. The function C is a projective subentropy. For any WM-space P, C(P) is equal to the lower projective limit of $C^*(\mathcal{U})$, where $\mathcal{U}$ is a partition of P. The function C is continuous on every $\{FWM(Q)\}$, Q finite. If $\langle Q,\mu\rangle$ is a finite probability space, then $C\langle Q,1,\mu\rangle$ is equal to the Shannon entropy of $\langle Q,\mu\rangle$.

Proof. The first assertion of the theorem follows at once from the definition. The second is a consequence of 3.3 A, B. It is clear that if $P \in \{FWM\}$, then C(P) coincides with cP introduced in QE. Hence, the last two assertions of the theorem follow from QE 4.4 and QE 2.7 (the case of a finite P not in $\{FWM\}$ is easily reduced to that of a space in $\{FWM\}$). - We take the opportunity to correct some distracting technical errors in the proof of QE 4.2, on which QE 4.4 is based. In part I of the proof, line 5, replace D(X) by D(x); in part II, line 1, replace D(Y) by D(X); in part II, line 8 (first line on p. 804) replace $x \in X_1$ by $x \in X$; in part II, line 9 (line 2 on p. 804) replace g_i by g_i'; in part III, line 5, replace $\bar{b}\bar{m}$ by $\bar{b}/\bar{m}$; in part III, lines 8 and 9, reverse the inequality involving $\beta'(xk)$.

3.6. We shall call C(P) and $C^*(P)$ the projective (Shannon) entropy (or C-entropy) and the semiprojective (Shannon) semientropy (or C^*-semientropy) of P, respectively. The functions C and C^* will be called the projective entropy and the semiprojective semientropy.

3.7. Remark. C^* is not invariant with respect to conservative morphisms and is distinct from C. This can be shown by simple examples (e.g. a $P = \langle Q,\varphi,\mu\rangle$ with card $Q = 3$ and suitable φ, μ).

3.8. The following propositions (3.9, 3.10, 3.12, 3.13) show that C and C^* are non-trivial in the sense that $C(P) > 0$, $C^*(P) > 0$ unless $P \sim 0$ and C(P), $C^*(P)$ are finite for a fairly wide class of spaces.

3.9. Proposition. For any WM-space P, $C(P) \geqq 2\hat{r}(P,P)/wP$, $C^*(P) \geqq 2\hat{r}(P,P)/wP$.

Proof. Clearly, for any d.expansion $\mathcal{P} = (P_z : z \in D)$ of P we have $\Gamma(\mathcal{P}) \geqq \sum(4\hat{r}(P_{z0},P_{z1})/wP_z : z \in D') \geqq \sum(4\hat{r}(P_{z0},P_{z1}) : z \in D')/wP = 2\hat{r}(P,P)/wP - 2\sum(\hat{r}(P_z,P_z) : z \in D'')$. It is easy to see that, for sufficiently fine $(P_z : z \in D'')$, $\sum(\hat{r}(P_z,P_z) : z \in D'')$ is arbitrarily small.

3.10. Proposition. For every non-trivial WM-space $P = \langle Q,\varphi,\mu\rangle$, $C^*(P) > 0$, $C(P) > 0$. More exactly, the following properties of P are equivalent: (1) $\hat{r}(P,P) = 0$, (2) $P \sim \langle Q,0,\mu\rangle$, (3) there are no disjoint sets $X \subset Q$, $Y \subset Q$

with $\mu X > 0$, $\mu Y > 0$, $\hat{r}(X,Y) > 0$, (4) $C^*(P) = 0$, (5) $C(P) = 0$. - This follows easily from 3.9.

3.11. Proposition. Let $P \in \{WM\}$ and let $\mathcal{U}_n = (U_{nk} : k \in K_n)$ be partitions of P such that $\max(\bar{d}(U_{nk}) : k \in K_n) \to 0$ for $n \to \infty$. Then (1) $C(P) \leqq \underline{\lim}\, C^*(\mathcal{U}_n)$, (2) if $\mathcal{U}_n$ are pure, then $C^*(P) \leqq \underline{\lim}\, C^*(\mathcal{U}_n)$.

Proof. We may assume $\underline{\lim}\, C^*(\mathcal{U}_n) < \infty$. Choose $\alpha > \underline{\lim}\, C^*(\mathcal{U}_n)$. Let $\mathcal{V} = (g_m : m \in M)$ be a partition of P. Choose n such that $C^*(\mathcal{U}_n) + \bar{m}\, \varepsilon\, wP < \alpha$, where $\bar{m}$ = card M, $\varepsilon = \max(\bar{d}(U_{nk}) : k \in K_n)$. By 3.3 A, B, there is a d.expansion $\mathcal{P} = (f_x : x \in D_0)$ of P such that $\Gamma(\mathcal{P}) = C^*[\mathcal{U}_n]$ and $\mathcal{P}''$ is equivalent to $\mathcal{U}_n$. Let $\mathcal{S} = (g_z : z \in D_1)$ be a d.expansion of P such that $\mathcal{S}$ is equivalent to $\mathcal{V}$. Let D_2 consist of all $x \in D_0$ and all $y = x.z$, where $x \in D_0''$, $z \in D_1$. Put $h_x = f_x$ if $x \in D_0$, $h_y = f_x.h_z$ if $y = x.z$, $x \in D_0''$, $z \in D_1$. Then $\mathcal{T} = (h_y : y \in D_2)$ is a d.expansion of P, $\mathcal{T}''$ refines $\mathcal{V}$. It is easy to see that $\Gamma(\mathcal{T}) \leqq \Gamma(\mathcal{P}) + \Sigma(\bar{m}\, \varepsilon \int f_x d\mu : x \in D_0'') = C^*[\mathcal{U}_n] + \bar{m}\, \varepsilon\, wP < \alpha$. - The proof of (2) is analogous.

3.12. Proposition. Let ρ be the metric on R^n defined by $\rho((x_i),(y_i)) = \Sigma |x_i - y_i|$. Let λ be the Lebesgue measure on R^n. If $Q \subset E$ is bounded λ-measurable, then $C\langle Q,\rho,\lambda\rangle$ and $C^*\langle Q,\rho,\lambda\rangle$ are finite.

3.13. Let f be a Lebesgue measurable non-negative function on R with the following property: there are numbers $k > 0$, $\gamma > 0$ such that, if $|x|$ is large, then $f(x) \leqq k\, |x|^{-2-\gamma}$ and $f(x)$ is non-increasing for $x > 0$, non-decreasing for $x < 0$. Let ρ denote the usual metric on R and let λ denote the Lebesgue measure. Then $C\langle R,\rho,f.\lambda\rangle < \infty$.

3.14. Theorem. If $a, b \in R$, $a < b$, then $C\langle [a,b], \varphi, \lambda\rangle = C^*\langle [a,b], \varphi, \lambda\rangle = |b - a|^2$.

3.15. Theorem. Let $P = \langle Q, 1, \mu\rangle$ be a WM-space. Let A consist of $q \in Q$ such that $\mu q > 0$. Then $C(P) = C^*(P) = \infty$ if $\mu(Q \setminus A) > 0$, and $C(P) = C^*(P) = H(\mu q : q \in A)$ if $\mu(Q \setminus A) = 0$.

Remark. In fact, the theorem asserts that, for any probability space $\langle Q, \mu\rangle$ such that the diagonal is measurable, the values $C\langle Q, 1, \mu\rangle$, $C^*\langle Q, 1, \mu\rangle$ coincide with the entropy of $\langle Q, \mu\rangle$ in the well known sense (see e.g. [5], where a special class of probability spaces is considered).

As stated in the introduction, we omit the proofs of 3.12 - 3.15.

3.16. Remark. We are going to point out certain analogies between the definition of C and C^* by means of dyadic expansions and a possible approach (which seems not to have been applied explicitly as yet) to the dimension of topological space and also of closure spaces, etc. The facts stated below are not used in the present note, and therefore we restrict ourselves to a few definitions and propositions. For concepts not defined here we refer to [2]. - (1) If P is a closure space, $X \subset P$, $Y \subset P$, then we define $\nabla(X,Y) = \nabla(X,Y;P)$ as follows: $\nabla(X,Y) = 1$ if $c\ell X \cap c\ell Y \neq \emptyset$, and $\nabla(X,Y) = 0$ otherwise. - (2) A dyadic expansion of a space P is, by definition, a family $\mathcal{M} = (M_z : z \in D)$, where $D \in \Delta$ and $M = M_{z0} \cup M_{z1}$, $M_{z0} \cap M_{z1} = \emptyset$ for every $z \in D'$. If $\mathcal{M} = (M_z : z \in D)$ is a dyadic expansion of a closure space P, we put $\nabla(\mathcal{M}) = \nabla(\mathcal{M};P) = \Sigma(\nabla(M_{z0}, M_{z1}) : z \in D')$. - (3) If P is a closure space, then $\mathcal{U} = \mathcal{U}(P)$ denotes the collection of all finite families $\xi = (X_k : k \in K)$ such that $\bigcup \operatorname{Int} X_k = P$, where we put $\operatorname{Int} X = P \setminus c\ell\,(P \setminus X)$, and $\mathcal{U}^* = \mathcal{U}^*(P)$ denotes the collection of all finite $\xi = (X_k : k \in K)$ such that $\bigcup X_k = P$ and, for some $\eta = (Y_k : k \in K) \in \mathcal{U}$ and some finite F_k, we have $X_k = Y_k \setminus F_k$. Clearly, $\mathcal{U} \subset \mathcal{U}^*$, and if every finite $F \subset P$ is closed, then $\mathcal{U}(P) = \mathcal{U}^*(P)$. - (4) For any $\xi \in \mathcal{U}^*(P)$ let $\Phi(\xi)$ denote the set of all dyadic expansions $\mathcal{P}$ of P such that $\mathcal{P}''$ refines ξ, let $\mathcal{F} = \mathcal{F}(P)$ be the filter generated by all $\Phi(\xi)$, $\xi \in \mathcal{U}$, and let $\mathcal{F}^* = \mathcal{F}^*(P)$ the filter generated by all $\Phi(\xi)$, $\xi \in \mathcal{U}^*$. - (5) For a closure space P, let $\nabla(P)$ ($\nabla^*(P)$) denote the lower limit of $\nabla(\mathcal{P})$ with respect to $\mathcal{F}$ ($\mathcal{F}^*$). - (6) It is not difficult to prove that, for any normal topological space P, $\nabla(P) = \nabla^*(P) = \dim P$ ($\dim P$ is the dimension defined by means of finite open coverings). - (7) Let every symmetric graph $G = \langle G, \tau \rangle$ be considered as a closure space (tolerance space): $y \in c\ell X$ iff $y \in X$ or $\langle y, x \rangle \in \tau$ for some $x \in X$. Then $\nabla(G) = 0$, $\nabla^*(G) = \chi(G) - 1$, where $\chi(G)$ is the chromatic number of G. - (8) If $\nabla(X,Y)$ is introduced in a different way, we can get different kinds of dimension.

4.

4.1. This section contains a remark concerning extensions of the Shannon entropy distinct from C and C*; definitions of "entropies" for probability spaces and semimetric spaces; definitions of concepts (closely connected with that of the ε-entropy of a metric space) of "proximal" and "graded" modifications of a given subentropy and a proposition concerning the relation of "graded" entropies to the differential entropy. - Proofs of statements contained in the present section as well as of some further related results are intended for publication in another note.

4.2. Besides C and C^*, there are various other extensions of the Shannon entropy satisfying some natural conditions. In particular, if $\Gamma(P_1,P_2)$ is replaced by another suitable function defined for pairs of subspaces, then various kinds of "projective" entropies can be obtained. Thus, if we put $E(P_1,P_2) = H(wP_1,wP_2)\overline{d}(P_1 + P_2)$, $E(P_z : z \in D) = \Sigma(E(P_{z0},P_{z1}) : z \in D')$, $E(D) = \underline{\lim}\, E(\mathcal{P})$, where $\mathcal{P}$ denotes dyadic expansions of P, we get a function E, which seems to possess various convenient properties, though in some respects it is less natural than C.

4.3. If φ is a semi-subentropy, then the corresponding functions, still denoted by φ, on the class of all measure spaces $\langle Q,\mu\rangle$, $\mu Q < \infty$, and on the class of all semimetric spaces (hence, in particular on the class of all symmetric graphs) can be defined as follows: (1) for a measure space $\langle Q,\mu\rangle$, $\varphi\langle Q,\mu\rangle$ is the least upper bound of $\varphi\langle Q,\rho,\mu\rangle$, where $\langle Q,\rho,\mu\rangle$ is a WM-space, $\rho \leqq 1$; (2) for a semimetric space $\langle Q,\rho\rangle$, $\varphi\langle Q,\rho\rangle$ is the least upper bound of $\varphi\langle Q,\rho,\mu\rangle$, where $\langle Q,\rho,\mu\rangle$ is a WM-space, and $\mu Q \leqq 1$. Due to 3.15, the values $C\langle Q,\mu\rangle$, $C^*\langle Q,\mu\rangle$ are not of much interest. On the other hand, $C\langle Q,\rho\rangle$, $C^*\langle Q,\rho\rangle$, and $E\langle Q,\rho\rangle$ may deserve a closer examination.

4.4. Let φ be a semi-subentropy. Then the following mapping φ_{pr} ("pr" stands for "proximal") can be associated with φ : for a WM-space $P = \langle Q,\rho,\mu\rangle$ and any $\varepsilon > 0$ put $P\wedge\varepsilon = \langle Q,\rho\wedge\varepsilon,\mu\rangle$, where $(\rho\wedge\varepsilon)(x,y) = \min(\rho(x,y),\varepsilon)$. Put $\varphi_{pr}(P;\varepsilon) = \varphi(P\wedge\varepsilon)$ and let $\varphi_{pr}(P)$ be the function $\varepsilon \longmapsto \varphi_{pr}(P;\varepsilon)$. The mapping φ_{pr} may be called the proximal modification of φ . It resembles the ε-entropy of totally bounded metric spaces; however, in general, it is more difficult to handle than the "graded" entropy described below.

4.5. If φ is a semi-subentropy, $P = \langle Q,\rho,\mu\rangle$ is a WM-space, and $\varepsilon > 0$, then we define $\varphi_{gr}(P;\varepsilon)$ as follows. For any semimetric ρ on a set Q, and any $\varepsilon > 0$, let $\rho_\varepsilon(x,y) = 1$ if $\rho(x,y) \geqq \varepsilon$, $\rho(x,y) = 0$ if $\rho(x,y) < \varepsilon$. Clearly, $\rho = \int_0^\infty \rho_t dt$ (thus, in a sense, we have a "spectral representation" of ρ by means of graphs). Now put $\varphi_{gr}(P;\varepsilon) = \varphi\langle Q,\rho_t,\mu\rangle$, and let $\varphi_{gr}(P)$ be the function $\varepsilon \longmapsto \varphi_{gr}(P;\varepsilon)$. The mapping φ_{gr} may be called the "graded" modification of φ . It can be shown that, for totally bounded metric spaces $\langle Q,\rho\rangle$, $\varphi_{gr}\langle Q,\rho,\mu\rangle$, where $\varphi = E$ (see 4.2), is closely connected with the ε-entropy.

4.6. There is a close connection between the graded entropy E_{gr} and the differential entropy in the usual sense (defined e.g. for a probability measure on R as follows: if $\mu = f\cdot\lambda$, f continuous, then the differential entropy of μ , i.e. of $\langle R,\rho,\mu\rangle$ is equal to $-\int_{-\infty}^{\infty} f(x)\log f(x)dx$ provided the integral exists). In fact, define the differential φ-entropy $\varphi_{diff}(P;S)$ as follows: if P, S are WM-spaces, $wP = wS = 1$, then $\varphi_{diff}(P;S)$ is equal to $\lim_{\varepsilon\to 0}(\varphi_{gr}(P;\varepsilon) - \varphi_{gr}(S;\varepsilon))$. It turns out that, under certain fairly weak conditions, $E_{diff}(\langle [a,b],\rho,f\cdot\lambda\rangle$, $\langle [a,b],\rho,\lambda/|b-a|\rangle)$ is equal to the differential entropy of the probability measure $f\cdot\lambda$ on R, and that this result can be extended in various ways. However, I do not know whether results of this kind are valid if E is replaced by C or if φ_{diff} is defined in terms of φ_{pr} instead of φ_{gr}.

4.7. Concepts of a semi-subentropy, a subentropy, etc., can be extended to more general spaces. E.g. instead of semimetrics on Q, "diameters" δ can be considered, i.e. non-negative functions assigning a value $\delta(X)\in\bar{R}$ to every $X\subset Q$. In this setting, the definition of E may remain unchanged; it is not clear, however, how to introduce functions corresponding to C, C^* , etc.

4.8. In conclusion, we mention two important and probably difficult questions not touched upon in this note: introduction of an appropriate concept (or rather concepts) of the product of WM-spaces, and of entropies of mappings (morphisms).

References

[QE] M. KATĚTOV: Quasi-entropy of finite weighted metric spaces, Comment. Math. Univ. Carolinae 17 (1976), 797-806.

[1] C.H. COOMBS, R.M. DAWES, A. TVERSKY: Mathematical Psychology. An elementary introduction, Prentice-Hall, Englewood Cliffs, N.J., U.S.A., 1970.

[2] R. ENGELKING: Dimension Theory, PWN, Warszawa, North-Holland Publ., Amsterdam-Oxford-New York, 1978.

[3] B. FORTE: Subadditive entropies for a random variable, Boll. Un. Mat. Ital. B(5)14(1977), no. 1, 118-133.

[4] C.F. PICARD: Théorie des questionnaires, Gauthier-Villars, Paris 1965.

[5] V.A. ROHLIN: Lekcii po entropiǐnoǐ teorii preobrazovaniǐ, Uspehi mat. nauk 22(1967), 5(137), 3-56.

Matematický ústav ČSAV
Žitná 25
Praha 1
Československo

(Oblatum 5.10. 1979)