Commentationes Mathematicae Universitatis Carolinae
13,4 (1972)

ON INFORMATION IN CATEGORIES

Miroslav KATĚTOV, Praha

In this note we consider real-valued functions defined on morphisms of a given category and satisfying certain natural conditions. It is shown that if the category in question is that of all finite non-void sets, then every such a function is of the form well-known from the information theory.

Terminology and notation. For basic concepts concerning categories we refer to [3]. The classes of objects and morphisms of a category $\mathcal{C}$ will be denoted by $\mathcal{O}bj\,\mathcal{C}$ and $\mathcal{M}orph\,\mathcal{C}$, respectively. Letters f, g, h, possibly with subscripts, will designate morphisms of $\mathcal{C}$. The domain of a morphism (in particular, of a mapping) f will be denoted by Df. A sum (product) of f_i, $i = 1, \dots, n$, will be denoted by $f_1 + \dots + f_n$ (by $f_1 \times \dots \times f_n$). Sometimes we will write $\sum f_i$ instead of $f_1 + \dots + f_n$ and nf instead of $f + \dots + f$ (n times). If f is isomorphic to g (in the sense that there are isomorphisms h_1, h_2 such that $f = h_1 g h_2$), we write $f \sim g$.

AMS, Primary: 94A15, 18B99 Ref. Ž. 8.721, 2.726.1

The cardinality of a set X will be denoted by $|X|$. If X, Y are non-void sets, $|Y| = 1$, then the (unique) mapping $f: X \to Y$ will be denoted by $i(X, Y)$ or by $i(X)$.

The set of all real numbers will be denoted by R, that of non-negative ones by R^+. For an $x > 0$, $\log x$ is the dyadic logarithm of x; we put $0 \log 0 = 0$.

Definition. Let $\mathcal{C}$ be a category. A function $\varphi : \mathrm{Morph}\, \mathcal{C} \to R^+$ will be called an ID-function (ID stands for "information decrement") for $\mathcal{C}$ if the following conditions hold:

(1) $f \approx g$ implies $\varphi(f) = \varphi(g)$;

(2) $\varphi(fg) \geq \varphi(g)$ provided fg is defined;

(3) if $f = f_1 + \dots + f_n$ and all Df_i are mutually isomorphic, then $\varphi(f) = \frac{1}{n} \sum \varphi(f_i)$;

(4) if h is a product of f and g, then

$$\varphi(h) = \varphi(f) + \varphi(g).$$

Conventions. If $\mathcal{C}$ is the category of finite non-void sets and $\varphi : \mathrm{Morph}\, \mathcal{C} \to R^+$ satisfies (1), we will put: (i) for any $X \in \mathrm{Obj}\, \mathcal{C}$, $\varphi(X) = \varphi(i(X))$; (ii) for any $n = 1, 2, \dots$, $\varphi(n) = \varphi(X)$, where $|X| = n$.

Theorem. Let $\mathcal{C}$ be the category of all finite non-void sets (with mappings as morphisms). A function $\varphi : \mathrm{Morph}\, \mathcal{C} \to R^+$ is an ID-function if and only if there is a number $c \geq 0$ such that, for every morphism $f: A \to B$ we have

$$\varphi(f) = \frac{c}{|A|} \sum_{b \in B} |f^{-1} b| \log |f^{-1} b| .$$

Proof. It is easy to see that every φ of the form described above is an ID-function. To show the converse, we need some lemmas. In what follows, $\mathcal{C}$ is the category of finite non-void sets.

Lemma 1. Assume that $\varphi : \mathcal{M}orph\, \mathcal{C} \longrightarrow R^+$ satisfies conditions (1),(3) from the definition of an ID-function. If $f : A \longrightarrow B$ is surjective, then

$$\varphi(f) = \frac{1}{|A|} \sum_{b \in B} |f^{-1} b|\, \varphi(f^{-1} b) \ .$$

Proof. If $b \in B$, put $m_b = |f^{-1} b|$. Put $m = \sum m_b$, $s = \prod m_b$, $s_b = s m_b^{-1}$. For every $b \in B$, put $q_b = s_b\, i(m_b)$. Clearly, for every $b \in B$, $\varphi(q_b) = \varphi(i(m_b)) = \varphi(f^{-1} b)$, $|Dq_b| = s$. Put $f' = \sum_{a \in A} q_{fa}$, $f'' = sf$. It is easy to see that $f' \approx f''$. Since $\varphi(f') = \frac{1}{m} \sum m_b\, \varphi(q_b)$, $\varphi(f'') = \varphi(f)$, we obtain

$$\varphi(f) = \frac{1}{m} \sum_{b \in B} m_b\, \varphi(q_b) \ .$$

This proves the assertion.

Lemma 2. Assume that $\varphi : \mathcal{M}orph\, \mathcal{C} \longrightarrow R^+$ satisfies conditions (1),(2),(3) and that $\varphi(1) = 0$. Then, for $m = 1, 2, \ldots$, we have

$$m\, \varphi(m) \leqq (m+1)\, \varphi(m+1) \ .$$

Proof. Let A, B, C be sets, $|A| = m+1$, $|B| = 2$, $|C| = 1$. Choose $g : A \rightarrow B$, $g = i(m) + i(1)$, $f : B \rightarrow C$. Clearly, $\varphi(fg) = \varphi(m+1)$, and, by condition (2), we have $\varphi(fg) \geqq \varphi(g)$. By Lemma 1, $\varphi(g) = \frac{m}{m+1}\, \varphi(m)$. This proves the assertion.

Lemma 3. Let ψ be a non-negative real-valued function on the set of positive integers. Assume that $m.\psi(m) \leq (m+1)\psi(m+1)$ for $m = 1, 2, \ldots$ and that $\psi(p^n) = n.\psi(p)$ for $p, n = 1, 2, \ldots$. Then, for every $m = 1, 2, \ldots$ we have

$$\psi(m) = \psi(2) . \log m \quad .$$

The proof is standard and may be omitted.

We are now going to prove the theorem. Let $\varphi : Morph\, \mathcal{C} \longrightarrow R^+$ satisfy (1) - (4). By Lemma 2, we have $m\, \varphi(m) \leq (m+1)\varphi(m+1)$ for $m = 1, 2, \ldots$. Since (4) is fulfilled, we have $\varphi(p^n) = n\, \varphi(p)$ for $p, n = 1, 2, \ldots$. Hence, by Lemma 3, $\varphi(m) = c \log m$, where $c = \varphi(2)$. Lemma 1 now implies that, for any surjective $f : A \longrightarrow B$, we have

$$(f) = \frac{c}{|A|} \sum_{b \in B} |f^{-1} b| \log f^{-1} b| \quad .$$

If $f : A \longrightarrow B$ is an arbitrary morphism of $\mathcal{C}$, let $j : f(A) \longrightarrow B$ be the embedding and let $\kappa : B \longrightarrow f(A)$ be such that $\kappa(x) = x$ for all $x \in f(A)$. Then $g = \kappa f$ is surjective, $f = jg$. By condition (2), we have $\varphi(f) = \varphi(g)$, which proves the theorem.

Remarks. 1) Clearly, there exist categories for which there is no ID-function (except 0). An example: the category $\mathcal{L}$ of finite-dimensinal linear spaces (over some fixed field). However, for this category there exist functions $Morph\, \mathcal{L} \longrightarrow R^+$ satisfying (1), (2) and (4). - 2) It may be of some interest to investigate those categories for which there exist non-trivial ID-functions. - 3) Since the cartesian product in the category of sets plays two distinct roles, that of categorical product and that of tensor product (see e.g. [2],[1]), it might be interesting to investigate, in closed categories (see e.g. [2],[1]), another concept of an ID-function with (4) replaced by an analogous condition on tensor product.

R e f e r e n c e s

[1] S. EILENBERG, G.M. KELLY: Closed categories, Proc. Conf.Categorical Algebra (La Jolla,Calif., 1965),pp.421-562.Springer,New York,1966.

[2] G.M. KELLY, S.MacLANE: Coherence in closed categories, J.Pure Appl.Algebra 1(1971), 97-140.

[3] B. MITCHELL: Theory of categories, Academic Press, 1965.

Matematický ústav ČSAV

Žitná 25

Praha 1

(Oblatum 23.11.1972)