MONOIDAL KLEISLI CATEGORY STRUCTURE OF CLASSES OF INFORMATION TRANSFORMERS [1]

P. V. GOLUBTSOV

Department of Physics, Moscow State Lomonosov University
(119899, Moscow, Russia, e-mail: P_V_G@mail.ru)

UDC 532

It is proposed to consider any uniform class of information transforming systems as morphisms of a certain category — a category of information transformers (ITs). Composition of ITs corresponds to their "consecutive application". The paper introduces an axiomatics for a category of ITs as a monoidal category that contains a subcategory (of deterministic ITs) with finite products and satisfies a certain set of axioms. Besides, it shows that many IT-categories can be constructed as Kleisli categories.

Introduction

Currently the growing interest is attracted to various mathematical ways of describing uncertainty, most of them being different from the probabilistic one, (e.g., based on the apparatus of fuzzy sets). For adequate theoretical study of the corresponding "nonstochastic" systems of information transforming and, in particular, for the study of important notions, such as sufficiency, informativeness, etc., we need to develop an approach general enough to describe different classes of information transforming systems in a uniform way.

It is convenient to consider different systems that take place in information acquiring and processing as particular cases of so-called *information transformers* (ITs). Besides, it is useful to work with families of ITs in which certain operations, e.g., *sequential* and *parallel compositions* are defined.

It was noticed fairly long ago [1–6] that the adequate algebraic structure for describing information transformers (initially for the study of statistical experiments) is the structure of *category* [7–10].

Analysis of general properties for the classes of linear, multivalued, and fuzzy information transformers, studied in [5, 6, 11–18, 20, 21], allowed one to extract general features shared by all these classes. Namely, each of these classes can be considered as a family of morphisms in an appropriate category, where the composition of information transformers corresponds to their "consecutive application." Each category of ITs (or IT-category) contains a subcategory (of the so-called *deterministic* ITs) that has products. Moreover, the operation of morphism product is extended in a "coherent way" to the whole category of ITs.

Works [6, 19] undertook an attempt to formulate a set of "elementary" axioms for a category of ITs, which would be sufficient for an abstract expression of the basic concepts of the theory of information transformers and for study of informativeness, decision problems, etc. This paper proposes another, significantly more compact axiomatics for a category of ITs. According to this axiomatics, a category of ITs is defined in effect as a *monoidal* category [7, 9] containing a subcategory (of *deterministic* ITs) with finite products.

Among the basic concepts connected to information transformers, there is one that plays an important role in the uniform construction of a wide spectrum of IT-categories — the concept of *distribution*. Indeed, fairly often an IT $a\colon \mathcal{A} \to \mathcal{B}$ can be represented by a mapping from $\mathcal{A}$ to the "space of distributions" on $\mathcal{B}$ (see, e.g., [6, 14,15,18,20,21]). For example, a probabilistic transition distribution (an IT in the category of stochastic ITs) can be represented by a certain measurable mapping from $\mathcal{A}$ to the space of distributions on $\mathcal{B}$. This observation suggests to construct a category of ITs as a *Kleisli* category [7, 22, 23] arising from the following components: an obvious category of deterministic ITs; a functor that takes an object $\mathcal{A}$ to the object of "distributions" on $\mathcal{A}$; and a natural transformation of functors, describing an "independent product of distributions".

It appears that a rather general axiomatic theory obtained in this way makes it possible to express, in terms of IT-categories, the basic concepts for information transformers and to derive their main properties.

[1] This article was presented at the II International Symposium "Fundamental Problems in Modern Quantum Theories and Experiments" (September 2—7 2002, Odessa, Ukraine) on the occasion of Professor Walter Thirring's 75th birthday.

Of course, the most developed theory of uncertainty is probability theory (and statistics based on probability). Certainly, mathematical statistics accumulated a rich conceptual experience. It introduced and deeply investigated such notions as joint and conditional distributions, independence, sufficiency, and others.

At the same time, it appears that all these concepts have very abstract meaning and, hence, they can be treated in terms of alternative (i.e., not probabilistic) approaches to the description of uncertainty. In fact, the basic notions of probability theory and statistics, as well as the methodology and results, are easily extended to other theories dealing with uncertainty. In [13, 15], it is shown that a rather substantive decision theory may be constructed even on the very moderate basis of multivalued or fuzzy maps.

Approaches proposed in this work may provide a background for construction and study of new classes of ITs, in particular, dynamical nondeterministic ITs, which may provide an adequate description for information flows and information interactions evolving in time. Besides, a uniform approach to problems of information transformations may be useful for a better understanding of information processes that take place in complex artificial and natural systems.

1. Categories of Information Transformers

1.1. Common Structure of Classes of Information Transformers

It is natural to assume that, for any information transformer a, there are defined a couple of spaces: $\mathcal{A}$ and $\mathcal{B}$, the space of "inputs" (or input signals), and the space of "outputs" (results of measurement, transformation, processing, etc.). We say that a "acts" from $\mathcal{A}$ to $\mathcal{B}$ and denote this as $a: \mathcal{A} \to \mathcal{B}$. It is important to note that typically an information transformer not only transforms signals, but also introduces some "noise". In this case, it is *nondeterministic* and cannot be represented just by a mapping from $\mathcal{A}$ to $\mathcal{B}$.

It is natural to study information transformers of a similar type by aggregating them into families endowed by a fairly rich algebraic structure [5, 11]. Specifically, it is natural to assume that families of ITs have the following properties:

(a) If $a: \mathcal{A} \to \mathcal{B}$ and $b: \mathcal{B} \to \mathcal{C}$ are two ITs, then their *composition* $b \circ a: \mathcal{A} \to \mathcal{C}$ is defined.

(b) This operation of composition is *associative*.

(c) There are certain *neutral* elements in these families, i.e., ITs that do not introduce any alterations. Namely, for any space $\mathcal{B}$, there exists a corresponding IT $i_{\mathcal{B}}: \mathcal{B} \to \mathcal{B}$ such that $i_{\mathcal{B}} \circ a = a$ and $b \circ i_{\mathcal{B}} = b$.

Algebraic structures of this type are called *categories* [7, 9].

Furthermore, we assume that, to every pair of information transformers acting from the same space $\mathcal{D}$ to spaces $\mathcal{A}$ and $\mathcal{B}$, respectively, there corresponds a certain IT $a * b$ (called *product* of a and b) from $\mathcal{D}$ to $\mathcal{A} \times \mathcal{B}$. In a certain sense, this IT "represents" both ITs a and b simultaneously. Specifically, ITs a and b can be "extracted" from $a * b$ by means of projections $\pi_{\mathcal{A},\mathcal{B}}$ and $\nu_{\mathcal{A},\mathcal{B}}$ from $\mathcal{A} \times \mathcal{B}$ to $\mathcal{A}$ and $\mathcal{B}$, respectively, i.e., $\pi_{\mathcal{A},\mathcal{B}} \circ (a*b) = a$, $\nu_{\mathcal{A},\mathcal{B}} \circ (a*b) = b$. Note, that typically an IT c such that $\pi_{\mathcal{A},\mathcal{B}} \circ c = a$, $\nu_{\mathcal{A},\mathcal{B}} \circ c = b$ is not unique, i.e., a category of ITs does not have products (in category-theoretic sense [7–10]). Thus, the notion of a category of ITs demands for an accurate formalization.

Analysis of classes of information transformers studied in [5, 6, 11–18, 20, 21] gives grounds to consider these classes as categories that satisfy certain fairly general conditions.

1.2. Categories: Basic Concepts

Recall that a category (see, for example, [7–10]) $\mathbf{C}$ consists of a class of objects $\mathrm{Ob}(\mathbf{C})$, a class of morphisms (or arrows) $\mathrm{Ar}(\mathbf{C})$, and a composition operation $\circ$ for morphisms such that:

(a) To any morphism a, there corresponds a certain pair of objects $\mathcal{A}$ and $\mathcal{B}$ (the source and the target of a) which is denoted $a: \mathcal{A} \to \mathcal{B}$.

(b) To every pair of morphisms $a: \mathcal{A} \to \mathcal{B}$ and $b: \mathcal{B} \to \mathcal{C}$, their *composition* $b \circ a: \mathcal{A} \to \mathcal{C}$ is defined.

Moreover, the following axioms hold:

(c) The composition is *associative*:

$$c \circ (b \circ a) = (c \circ b) \circ a.$$

(d) To every object $\mathcal{R}$, there corresponds an (*identity*) morphism $i_{\mathcal{R}}: \mathcal{R} \to \mathcal{R}$ so that

$$\forall a: \mathcal{A} \to \mathcal{B} \qquad a \circ i_{\mathcal{A}} = a = i_{\mathcal{B}} \circ a.$$

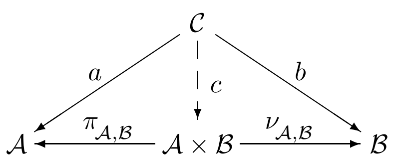
A morphism $a: \mathcal{A} \to \mathcal{B}$ is called *isomorphism* if there exists a morphism $b: \mathcal{B} \to \mathcal{A}$ such that $a \circ b = i_{\mathcal{B}}$ and $b \circ a = i_{\mathcal{A}}$. In this case, objects $\mathcal{A}$ and $\mathcal{B}$ are called *isomorphic*.

Morphisms $a: \mathcal{D} \to \mathcal{A}$ and $b: \mathcal{D} \to \mathcal{B}$ are called *isomorphic* if there exists an isomorphism $c: \mathcal{A} \to \mathcal{B}$ such that $c \circ a = b$.

An object $\mathcal{Z}$ is called *terminal* object if, for any object $\mathcal{A}$, there exists a unique morphism from $\mathcal{A}$ to $\mathcal{Z}$ which is denoted $z_{\mathcal{A}}: \mathcal{A} \to \mathcal{Z}$ in what follows.

A category $\mathbf{D}$ is called a *subcategory* of a category $\mathbf{C}$ if $\mathrm{Ob}(\mathbf{D}) \subseteq \mathrm{Ob}(\mathbf{C})$, $\mathrm{Ar}(\mathbf{D}) \subseteq \mathrm{Ar}(\mathbf{C})$, and a morphism composition in $\mathbf{D}$ coincides with the composition in $\mathbf{C}$.

It is said that a category has (pairwise) products if, for every pair of objects $\mathcal{A}$ and $\mathcal{B}$, there exist their *product*, that is, an object $\mathcal{A} \times \mathcal{B}$ and a pair of morphisms $\pi_{\mathcal{A},\mathcal{B}}: \mathcal{A} \times \mathcal{B} \to \mathcal{A}$ and $\nu_{\mathcal{A},\mathcal{B}}: \mathcal{A} \times \mathcal{B} \to \mathcal{B}$ called projections such that for any object $\mathcal{D}$ and for any pair of morphisms $a: \mathcal{D} \to \mathcal{A}$ and $b: \mathcal{D} \to \mathcal{B}$, there exists a unique morphism $c: \mathcal{D} \to \mathcal{A} \times \mathcal{B}$ that yields a commutative diagram:



i.e., satisfies the following conditions:

$$\pi_{\mathcal{A},\mathcal{B}} \circ c = a, \qquad \nu_{\mathcal{A},\mathcal{B}} \circ c = b. \tag{1}$$

We call such a morphism c the *product of morphisms* a and b and denote it $a * b$.

It is easily seen that the existence of products in a category implies the following equality:

$$(a * b) \circ d = (a \circ d) * (b \circ d). \tag{2}$$

In a category with products, for two arbitrary morphisms $a: \mathcal{A} \to \mathcal{C}$ and $b: \mathcal{B} \to \mathcal{D}$, one can define the morphism $a \times b$:

$$a \times b: \mathcal{A} \times \mathcal{B} \to \mathcal{C} \times \mathcal{D}, \qquad a \times b \overset{\mathrm{def}}{=} (a \circ \pi_{\mathcal{A},\mathcal{B}}) * (b \circ \nu_{\mathcal{A},\mathcal{B}}). \tag{3}$$

This definition and (1) obviously imply that the morphism $c = a \times b$ satisfy the following conditions:

$$\pi_{\mathcal{C},\mathcal{D}} \circ c = a \circ \pi_{\mathcal{A},\mathcal{B}}, \qquad \nu_{\mathcal{C},\mathcal{D}} \circ c = b \circ \nu_{\mathcal{A},\mathcal{B}}. \tag{4}$$

Moreover, $c = a \times b$ is the only morphism satisfying conditions (4).

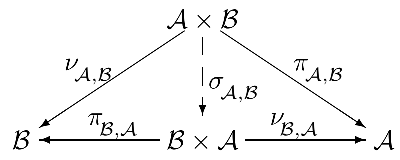
It is also easily seen that (2) and (3) imply the following equality:

$$(a \times b) \circ (c * d) = (a \circ c) * (b \circ d). \tag{5}$$

Suppose $\mathcal{A} \times \mathcal{B}$ and $\mathcal{B} \times \mathcal{A}$ are two products of objects $\mathcal{A}$ and $\mathcal{B}$ taken in different order. By the properties of products, the objects $\mathcal{A} \times \mathcal{B}$ and $\mathcal{B} \times \mathcal{A}$ are isomorphic and the natural isomorphism is

$$\sigma_{\mathcal{A},\mathcal{B}}: \mathcal{A} \times \mathcal{B} \to \mathcal{B} \times \mathcal{A}, \qquad \sigma_{\mathcal{A},\mathcal{B}} \overset{\mathrm{def}}{=} \nu_{\mathcal{A},\mathcal{B}} * \pi_{\mathcal{A},\mathcal{B}}, \tag{6}$$

i.e., a unique morphism that makes the following diagram commutative:



Moreover, for any object $\mathcal{D}$ and for any morphisms $a: \mathcal{D} \to \mathcal{A}$ and $b: \mathcal{D} \to \mathcal{B}$, the morphisms $a * b$ and $b * a$ are isomorphic, that is,

$$\sigma_{\mathcal{A},\mathcal{B}} \circ (a * b) = b * a. \tag{7}$$

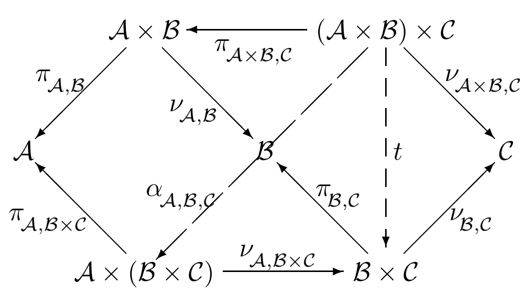
Similarly, by the properties of products, the objects $(\mathcal{A} \times \mathcal{B}) \times \mathcal{C}$ and $\mathcal{A} \times (\mathcal{B} \times \mathcal{C})$ are isomorphic. Let

$$\alpha_{\mathcal{A},\mathcal{B},\mathcal{C}}: (\mathcal{A} \times \mathcal{B}) \times \mathcal{C} \to \mathcal{A} \times (\mathcal{B} \times \mathcal{C})$$

be the corresponding natural isomorphism. Its "explicit" form is:

$$\alpha_{\mathcal{A},\mathcal{B},\mathcal{C}} \overset{\mathrm{def}}{=} (\pi_{\mathcal{A},\mathcal{B}} \circ \pi_{\mathcal{A}\times\mathcal{B},\mathcal{C}}) * \big((\nu_{\mathcal{A},\mathcal{B}} \circ \pi_{\mathcal{A}\times\mathcal{B},\mathcal{C}}) * \nu_{\mathcal{A}\times\mathcal{B},\mathcal{C}}\big). \tag{8}$$

It can be easily obtained with the following diagram:



Here $t = (\nu_{\mathcal{A},\mathcal{B}} \circ \pi_{\mathcal{A}\times\mathcal{B},\mathcal{C}}) * \nu_{\mathcal{A}\times\mathcal{B},\mathcal{C}}$

Then, for any object $\mathcal{D}$ and for any morphisms $a: \mathcal{D} \to \mathcal{A}$, $b: \mathcal{D} \to \mathcal{B}$, and $c: \mathcal{D} \to \mathcal{C}$, we have

$$\alpha_{\mathcal{A},\mathcal{B},\mathcal{C}} \circ \big((a * b) * c\big) = a * (b * c). \tag{9}$$

1.3. Elementary Axioms for Categories of Information Transformers

In this subsection, we set forward the main properties of categories of ITs. All the following study will rely exactly on these properties.

In [5, 6, 11, 12, 14, 17, 18], it was shown that classes of information transformers can be considered as morphisms in certain categories. As a rule, such categories do not have products, which is a peculiar expression of the nondeterministic nature of ITs in these categories. However, it turns out that deterministic information transformers, which are usually determined in a natural way in any category of ITs, form a subcategory with products. This point makes it possible to define a "product" of objects in a category of ITs. Moreover, it provides an axiomatic way to describe an extension of the product operation from the subcategory of deterministic ITs to the whole category of ITs.

DEFINITION 1.1 We say that a category $\mathbf{C}$ is a *category of information transformers* if the following axioms hold:

1. There is a fixed subcategory of *deterministic* ITs $\mathbf{D}$ that contains all the objects of the category $\mathbf{C}$ $(\mathrm{Ob}(\mathbf{D}) = \mathrm{Ob}(\mathbf{C}))$.
2. The classes of *isomorphisms* in $\mathbf{D}$ and in $\mathbf{C}$ coincide, that is, all the isomorphisms in $\mathbf{C}$ are deterministic.
3. The categories $\mathbf{D}$ and $\mathbf{C}$ have a common *terminal object* $\mathcal{Z}$.
4. The category $\mathbf{D}$ has pairwise *products*.
5. There is a specified *extension of morphism product* from the subcategory $\mathbf{D}$ to the whole category $\mathbf{C}$, that is, for any object $\mathcal{D}$ and for any pair of morphisms $a: \mathcal{D} \to \mathcal{A}$ and $b: \mathcal{D} \to \mathcal{B}$ in $\mathbf{C}$, there is a certain information transformer $a * b: \mathcal{D} \to \mathcal{A} \times \mathcal{B}$ (which is also called a *product* of ITs a and b) such that
$$\pi_{\mathcal{A},\mathcal{B}} \circ (a * b) = a, \qquad \nu_{\mathcal{A},\mathcal{B}} \circ (a * b) = b.$$
6. Let $a: \mathcal{A} \to \mathcal{C}$ and $b: \mathcal{B} \to \mathcal{D}$ be arbitrary ITs in $\mathbf{C}$, then the IT $a \times b$ defined by Eq. (3) satisfies Eq. (5):
$$(a \times b) \circ (c * d) = (a \circ c) * (b \circ d).$$
7. Equality (7) holds not only in $\mathbf{D}$ but in $\mathbf{C}$ as well, that is, *product* of information transformers is "*commutative* up to isomorphism."
8. Equality (9) also holds in $\mathbf{C}$. In other words, *product* of information transformers is "*associative* up to isomorphism" too.

Now let us make several comments concerning the above definition.

We stress that, in the description of the extension of morphism product from the category $\mathbf{D}$ to $\mathbf{C}$ (Axiom 5), we *do not require the uniqueness* of an IT $c: \mathcal{D} \to \mathcal{A} \times \mathcal{B}$ that satisfies conditions (1).

Nevertheless, it is easily verified that Eqs. (4) are valid for $c = a \times b$ not only in the category $\mathbf{D}$, but in $\mathbf{C}$ as well, that is,
$$\pi_{\mathcal{C},\mathcal{D}} \circ (a \times b) = a \circ \pi_{\mathcal{A},\mathcal{B}}, \qquad \nu_{\mathcal{C},\mathcal{D}} \circ (a \times b) = b \circ \nu_{\mathcal{A},\mathcal{B}}.$$
However, the IT c that satisfy Eqs. (4) may be not unique. Note also that Eq. (2) does not hold in the category $\mathbf{C}$ in general.

Further, note that Axiom 6 immediately implies
$$(a \times b) \circ (c \times d) = (a \circ c) \times (b \circ d).$$

Finally, note that any category that has a terminal object and pairwise products can be considered as a category of ITs in which all information transformers are deterministic.

2. Category of Information Transformers as a Monoidal Category

As we have already mentioned above, in a category of ITs, there are certain "meaningful" operations of product for objects and for morphisms. However, these operations are not product operations in category-theoretic sense. Nevertheless, every category of ITs is a *monoidal category* (see, e.g., [7, 9]). Before introducing this notion, we need to recall the basic notions of a functor and a natural transformation.

2.1. Functors and Natural Transformations

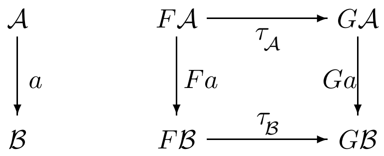
DEFINITION 2.1 A *functor* F from a category $\mathbf{C}$ to a category $\mathbf{D}$ is a function which maps all the objects and morphisms of $\mathbf{C}$ respectively to objects and morphisms of $\mathbf{D}$ that "preserves" a category structure in the following sense:

(a) If $a: \mathcal{A} \to \mathcal{B}$ in $\mathbf{C}$, then $Fa: F\mathcal{A} \to F\mathcal{B}$ in $\mathbf{D}$;

(b) $F(i_{\mathcal{A}}) = i_{F\mathcal{A}}$ for every $\mathcal{A} \in \mathrm{Ob}(\mathbf{C})$;

(c) $F(a \circ b) = Fa \circ Fb$ whenever $a \circ b$ is defined in $\mathbf{C}$.

Definition 2.2 Given two functors F and G from $\mathbf{C}$ to $\mathbf{D}$, a *natural transformation* (also called *functor morphism*) $\tau: F \to G$ is a function which assigns, to each object $\mathcal{A}$ of $\mathbf{C}$, a morphism $\tau\mathcal{A}$ (also denoted $\tau_{\mathcal{A}}$) in such a way that every $\mathbf{C}$-morphism $a: \mathcal{A} \to \mathcal{B}$ of $\mathbf{C}$ yields a commutative diagram:

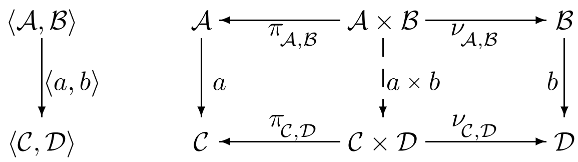


In addition, if $\tau_{\mathcal{A}}$ is an isomorphism for all $\mathcal{A} \in \mathrm{Ob}(\mathbf{C})$, then τ is called *natural equivalence* (or *natural isomorphism*).

Definition 2.3 Given two categories $\mathbf{C}$ and $\mathbf{D}$, their *product* is defined as a category whose objects are the ordered pairs $\langle \mathcal{C}, \mathcal{D} \rangle$ of objects $\mathcal{C} \in \mathrm{Ob}(\mathbf{C})$ and $\mathcal{D} \in \mathrm{Ob}(\mathbf{D})$ and for which morphisms $\langle \mathcal{C}, \mathcal{D} \rangle \to \langle \mathcal{C}', \mathcal{D}' \rangle$ are just pairs $\langle a, b \rangle$ with $a: \mathcal{A} \to \mathcal{A}'$ and $b: \mathcal{B} \to \mathcal{B}'$ with the identities and composition induced by $\mathbf{C}$ and $\mathbf{D}$:

$$i_{\langle \mathcal{C}, \mathcal{D} \rangle} = \langle i_{\mathcal{C}}, i_{\mathcal{D}} \rangle, \qquad \langle a, b \rangle \circ \langle a', b' \rangle = \langle a \circ a', b \circ b' \rangle .$$

It is easily seen that a binary product in a category $\mathbf{C}$ may be considered as a functor (called *product functor*) $\times : \mathbf{C} \times \mathbf{C} \to \mathbf{C}$ if its action on a morphism $\langle a, b \rangle$ is defined by Eq. (3), i.e., as a unique morphism $a \times b$ that yields a commutative diagram:



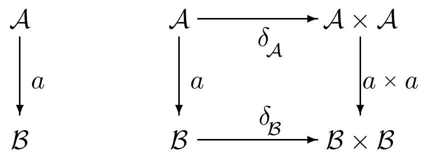
It is also seen from this diagram that π and ν are natural transformations from the product functor to the obvious *projection* functors $\mathbf{C} \times \mathbf{C} \to \mathbf{C}$.

In addition, one can easily see that σ and α defined by Eqs. (6) and (8) are natural equivalences, respectively, of the obvious functors $\mathbf{C} \times \mathbf{C} \to \mathbf{C}$ and $\mathbf{C} \times \mathbf{C} \times \mathbf{C} \to \mathbf{C}$.

Let us also note here that there is a "diagonal" natural transformation δ defined on $\mathbf{D}$:

$$\delta_{\mathcal{C}} \stackrel{\mathrm{def}}{=} i_{\mathcal{C}} * i_{\mathcal{C}}: \mathcal{C} \to \mathcal{C} \times \mathcal{C}.$$

Note that "naturality" of δ is expressed by the following diagram



meaning that

$$(a \times a) \circ \delta_{\mathcal{A}} = \delta_{\mathcal{B}} \circ a \tag{10}$$

and immediately follows from (2) and (5).

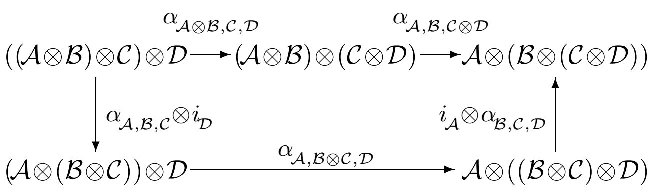
It is easily seen that, with the help of a "diagonal" natural transformation, the product of morphisms $a * b$ may be expressed through their "functorial product" $a \times b$, i.e.,

$$a * b = (a \times b) \circ \delta_{\mathcal{C}}.$$

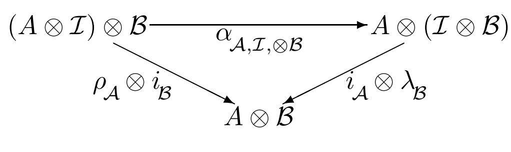
2.2. Monoidal Category

Let us recall briefly the definition and basic properties of monoidal category. Further details can be found in [6, 7, 9].

Definition 2.4 A *monoidal category* is a 6-tuple $\langle \mathbf{C}, \otimes, \mathcal{I}, \alpha, \lambda, \rho \rangle$ such that $\mathbf{C}$ is a category, $\otimes: \mathbf{C} \times \mathbf{C} \to \mathbf{C}$ is a functor, $\mathcal{I}$ is an object of $\mathbf{C}$ and $\alpha_{\mathcal{A},\mathcal{B},\mathcal{C}}: (\mathcal{A} \otimes \mathcal{B}) \otimes \mathcal{C} \to \mathcal{A} \otimes (\mathcal{B} \otimes \mathcal{C})$, $\lambda_{\mathcal{A}}: \mathcal{I} \otimes \mathcal{A} \to \mathcal{A}$, $\rho_{\mathcal{A}}: \mathcal{A} \otimes \mathcal{I} \to \mathcal{A}$ are natural equivalences of functors $\mathbf{C} \times \mathbf{C} \times \mathbf{C} \to \mathbf{C}$, $\mathbf{C} \to \mathbf{C}$, $\mathbf{C} \to \mathbf{C}$ such that the following *coherence axioms* hold:



$$\mathcal{I} \otimes \mathcal{I} \xrightarrow{\lambda_{\mathcal{I}}} \mathcal{I} \;=\; \mathcal{I} \otimes \mathcal{I} \xrightarrow{\rho_{\mathcal{I}}} \mathcal{I}$$



It can be easily seen that if a category $\mathbf{D}$ has pairwise products $\times$ and a terminal object $\mathcal{Z}$, then $\langle \mathbf{D}, \times, \mathcal{Z}, \alpha, \lambda, \rho \rangle$ is a monoidal category, where $\times: \mathbf{D} \times \mathbf{D} \to \mathbf{D}$ is the product functor and $\alpha_{\mathcal{A},\mathcal{B},\mathcal{C}}: (\mathcal{A} \times \mathcal{B}) \times \mathcal{C} \to \mathcal{A} \times (\mathcal{B} \times \mathcal{C})$, $\lambda_{\mathcal{A}}: \mathcal{Z} \times \mathcal{A} \to \mathcal{A}$, and $\rho_{\mathcal{A}}: \mathcal{A} \times \mathcal{Z} \to \mathcal{A}$

are the "obvious" natural equivalences that are explicitly expressed through π and ν, i.e.,

$$\lambda_{\mathcal{A}} \stackrel{\text{def}}{=} \pi_{\mathcal{Z},\mathcal{A}}, \qquad \rho_{\mathcal{A}} \stackrel{\text{def}}{=} \nu_{\mathcal{A},\mathcal{Z}},$$

$$\alpha_{\mathcal{A},\mathcal{B},\mathcal{C}} \stackrel{\text{def}}{=} (\pi_{\mathcal{A},\mathcal{B}} \circ \pi_{\mathcal{A}\times\mathcal{B},\mathcal{C}}) * \Big((\nu_{\mathcal{A},\mathcal{B}} \circ \pi_{\mathcal{A}\times\mathcal{B},\mathcal{C}}) * \nu_{\mathcal{A}\times\mathcal{B},\mathcal{C}}\Big).$$

2.3. Second Definition for a Category of Information Transformers

First note that every category $\mathbf{D}$ with pairwise products and with terminal object $\mathcal{Z}$ constitutes a monoidal category $\langle \mathbf{D}, \times, \mathcal{Z}, \alpha, \lambda, \rho\rangle$, where $\times: \mathbf{D} \times \mathbf{D} \to \mathbf{D}$ is the product functor and $\alpha_{\mathcal{A},\mathcal{B},\mathcal{C}}: (\mathcal{A} \times \mathcal{B}) \times \mathcal{C} \to \mathcal{A} \times (\mathcal{B} \times \mathcal{C})$, $\lambda_{\mathcal{A}}: \mathcal{Z} \times \mathcal{A} \to \mathcal{A}$, and $\rho_{\mathcal{A}}: \mathcal{A} \times \mathcal{Z} \to \mathcal{A}$ are the obvious natural equivalences. Besides, as a category with products, the category $\mathbf{D}$ has a natural equivalence $\sigma_{\mathcal{A},\mathcal{B}}: \mathcal{A} \times \mathcal{B} \to \mathcal{B} \times \mathcal{A}$, which interchanges components in a product, and a "diagonal" natural transformation $\delta_{\mathcal{C}}: \mathcal{C} \to \mathcal{C} \times \mathcal{C}$.

DEFINITION 2.5 We say that a category $\mathbf{C}$ is a category of information transformers over a subcategory $\mathbf{D}$ if all the objects of $\mathbf{C}$ are contained in $\mathbf{D}$ and the following three axioms hold.

Axiom 1. Category $\mathbf{D}$ has binary products and a terminal object $\mathcal{Z}$.

Thus, $\mathbf{D}$ is equipped with a product functor $\times$ and natural transformations π, ν, α, λ, ρ, σ, and δ. Besides, $\langle \mathbf{D}, \times, \mathcal{Z}, \alpha, \lambda, \rho\rangle$ is a monoidal category.

Axiom 2. Object $\mathcal{Z}$ is a terminal object in $\mathbf{C}$, and there exist an extension $\times: \mathbf{C} \times \mathbf{C} \to \mathbf{C}$ of the functor $\times$ from $\mathbf{D}$ to the whole category $\mathbf{C}$ such that $\langle \mathbf{C}, \times, \mathcal{Z}, \alpha, \lambda, \rho\rangle$ is also a monoidal category with the functor $\times: \mathbf{C} \times \mathbf{C} \to \mathbf{C}$, object $\mathcal{Z}$, and natural equivalences α, λ and ρ defined on a level of the subcategory $\mathbf{D}$.

Axiom 3. Natural transformations π, ν and σ (in the category $\mathbf{D}$) are natural transformations in the whole category $\mathbf{C}$ as well.

We will refer to morphisms of the category $\mathbf{C}$ as *information transformers*. Morphisms of the subcategory $\mathbf{D}$ will be called *deterministic* information transformers.

Let us note that since all the mentioned natural transformations π, ν, α, λ, ρ, σ, and δ are defined on the level of the subcategory $\mathbf{D}$, all their components are deterministic ITs. Besides, they satisfy all the mutual commutativity relations that they satisfied in $\mathbf{D}$. Thus, the nontrivial part, say, of Axiom 2 is that α, λ, and ρ are natural transformations.

More precisely, we can substitute Axioms 2 and 3 by the following ones:

Axiom 2'. Object $\mathcal{Z}$ is a terminal object in $\mathbf{C}$, and there exists an extension $\times: \mathbf{C} \times \mathbf{C} \to \mathbf{C}$ of the functor $\times$ from $\mathbf{D}$ to the whole category $\mathbf{C}$ (in fact, to all morphisms of $\mathbf{C}$).

Axiom 3'. Natural transformations π, ν, α, λ, ρ, and σ (defined within $\mathbf{D}$) are natural transformations in the whole category $\mathbf{C}$ as well.

Let us also note that, strictly speaking, we did not define yet a product operation $*$ for all compatible information transformers. Indeed, it is well defined only for deterministic ITs (morphisms in $\mathbf{D}$).

However, we can extend the product operation for morphisms from the subcategory $\mathbf{D}$ to $\mathbf{C}$. Specifically, we define a product $*$ in $\mathbf{C}$ by the following relation which is true in $\mathbf{D}$:

$$a * b \stackrel{\text{def}}{=} (a \times b) \circ \delta_{\mathcal{C}}. \tag{11}$$

Since Eq. (11) is always true in the category with products $\mathbf{D}$, it provides a consistent extension of the product operation $*$ on morphisms from the subcategory $\mathbf{D}$ to $\mathbf{C}$.

Let us stress here that we do not require that δ be a natural transformation on the whole category $\mathbf{C}$. Furthermore, typically, in many important examples of categories of ITs, δ is not a natural transformation. Such categories do not have products in category-theoretic sense. We will present a simple example of non-naturality of δ in the category of multivalued ITs.

THEOREM 2.1 *Definitions 1.1 and 2.5 for a category of information transformers are equivalent.*

3. IT-Category as a Kleisli Category

3.1. Concept of Distribution. Kleisli Category

Two equivalent definitions presented above provide the minimal conceptual background for studying the categories of ITs, e.g., for definition and analysis of informativeness, semantic informativeness, decision problems, etc. [1–5,11,14,19]. However, these definitions do not provide any tools for constructing the categories of ITs on the basis of more elementary concepts. The concept of *distribution* is one of the most important and it plays a critical role in the uniform construction of a wide spectrum of IT-categories. Its importance is connected to the observation that, in many important IT-categories, an information transformer $a: \mathcal{A} \to \mathcal{B}$ may be represented by a morphism from $\mathcal{A}$ to the "object of distributions" over $\mathcal{B}$. For example, a probabilistic transition distribution (an IT in the category of stochastic ITs) may be represented by a certain measurable mapping $\mathcal{A}$ to the space of distributions on $\mathcal{B}$.

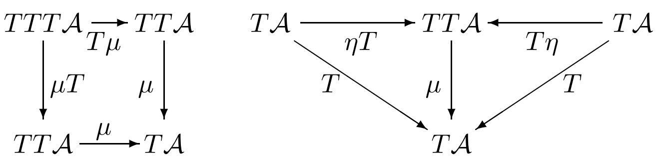
Thus, we suppose that, on some fixed "base" category **D** (category of deterministic ITs), there is defined a functor T, which takes an object $\mathcal{A}$ to the object $T\mathcal{A}$ of "distributions" on $\mathcal{A}$. Besides, we assume that there are two natural transformations connected to this functor: $\eta: I \to T$ and $\mu: TT \to T$. Informally, $\eta_{\mathcal{A}}: \mathcal{A} \to T\mathcal{A}$ takes an element of $\mathcal{A}$ to a "discrete distribution concentrated on this element", and $\mu_{\mathcal{A}}: TT\mathcal{A} \to T\mathcal{A}$ "mixes" (averages) a distribution of distributions on $\mathcal{A}$, by transforming it to a certain distribution on $\mathcal{A}$. Besides, there are natural "coherence" conditions for η and μ:

$$\mu_{\mathcal{A}} \circ T\mu_{\mathcal{A}} = \mu_{\mathcal{A}} \circ \mu_{T\mathcal{A}}$$

and

$$\mu_{\mathcal{A}} \circ T\eta_{\mathcal{A}} = i_{T\mathcal{A}} \qquad \mu_{\mathcal{A}} \circ \eta_{T\mathcal{A}} = i_{T\mathcal{A}}$$

that may be presented by the following commutative diagrams:



Commutativity of the square means that, for any "third-order distribution" on $\mathcal{A}$ (i.e. distribution on a collection of distributions on a family of distributions on $\mathcal{A}$), the result of "mixing" of distributions does not depend on the order of "mixing". More precisely, the result of mixing over the "top" (third order, element of $TTT\mathcal{A}$) distribution first and mixing the resulting second-order distribution next should give the same result as for mixing over "intermediate" (second-order, elements of $TT\mathcal{A}$) distributions first and then mixing the resulting second-order distribution. Commutativity of the left triangle means that mixing of a second order distribution, concentrated in one element (which is itself a distribution on $\mathcal{A}$) gives this distribution. Finally, commutativity of the right triangle means if we take some distribution on $\mathcal{A}$, transform it to "the same" distribution of singletons and then mix the resulting second-order distribution, we will obtain the original distribution.

It is well known that a collection $\langle T, \eta, \mu \rangle$, satisfying the two commutative diagrams above, is called a *triple* (monad) [7, 9, 23] on the category **D**.

The concept of triple provides an elegant technique of constructing a category of ITs **C** on the basis of the category of deterministic ITs, as a *Kleisli* category [6, 7, 22, 23]. In this construction, each morphism $a: \mathcal{A} \to \mathcal{B}$ in the category **C** is determined by a morphism $a': \mathcal{A} \to T\mathcal{B}$ of the category **D**. The composition $a \circ b$ of ITs $a: \mathcal{A} \to \mathcal{B}$ and $b: \mathcal{B} \to \mathcal{C}$ in **C** is represented by the morphism

$$(b \circ a)' \stackrel{\text{def}}{=} \mu_{\mathcal{C}} \circ Tb' \circ a'$$

$$(b \circ a)' = \mathcal{A} \xrightarrow{a'} T\mathcal{B} \xrightarrow{Tb'} TT\mathcal{C} \xrightarrow{\mu} T\mathcal{C}$$

in **D**, and any deterministic IT $c: \mathcal{C} \to \mathcal{D}$ (in **C**) is determined by the morphism

$$c' \stackrel{\text{def}}{=} \eta_{\mathcal{D}} \circ c$$

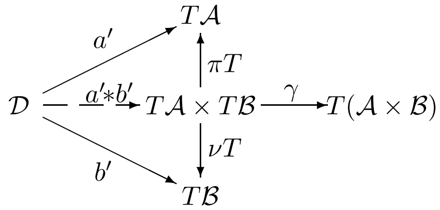
$$c' = \mathcal{C} \xrightarrow{c} \mathcal{D} \xrightarrow{\eta} T\mathcal{D}$$

in **D**.

3.2 Independent Distribution. Monoidal Kleisli Category

The main factor in the construction of the category of ITs as a Kleisli category is the equipping of it with a structure of monoidal category. For this purpose, we introduce a natural transformation $\gamma: \times T \to T\times$, $\gamma_{\mathcal{A},\mathcal{B}}: T\mathcal{A} \times T\mathcal{B} \to T(\mathcal{A} \times \mathcal{B})$, which "takes" a pair of distributions to their "independent joint distribution" (see also [24]). Then the product $c = a*b$ of ITs $a: \mathcal{D} \to \mathcal{A}$ and $b: \mathcal{D} \to \mathcal{B}$ (in **C**) is determined by the morphism

$$c' \stackrel{\text{def}}{=} \gamma_{\mathcal{A},\mathcal{B}} \circ (a' * b')$$



in **D**. Note that $a' * b'$ here exists and is uniquely defined since **D** is a category with products.

THEOREM 3.1 *Suppose that* $\mathbf{D}$ *is a category with pairwise products and with terminal object* $\mathcal{Z}$*;* π*,* ν*,* α*,* σ *are the corresponding natural transformations, and* $\langle T, \eta, \mu \rangle$ *is a triple on* $\mathbf{D}$ *with* $\eta_{\mathcal{B}}$ *monomorphic for every* $\mathcal{B}$*. Then the generated Kleisli category* $\mathbf{C}$*, equipped with a natural transformation* γ*, is a category of information transformers if and only if the following compatibility conditions of* γ *with the natural transformations* π*,* ν*,* α*,* σ*,* η*, and* μ *hold:*

$\boldsymbol{\pi}$-$\boldsymbol{\gamma}$ *and* $\boldsymbol{\nu}$-$\boldsymbol{\gamma}$ *conditions:*

$$T\pi_{\mathcal{A},\mathcal{B}} \circ \gamma_{\mathcal{A},\mathcal{B}} = \pi_{T\mathcal{A},T\mathcal{B}}, \qquad T\nu_{\mathcal{A},\mathcal{B}} \circ \gamma_{\mathcal{A},\mathcal{B}} = \nu_{T\mathcal{A},T\mathcal{B}}$$

$\boldsymbol{\sigma}$-$\boldsymbol{\gamma}$ *condition:*

$$T\sigma_{\mathcal{A},\mathcal{B}} \circ \gamma_{\mathcal{A},\mathcal{B}} = \gamma_{\mathcal{B},\mathcal{A}} \circ \sigma_{T\mathcal{A},T\mathcal{B}};$$

$\boldsymbol{\alpha}$-$\boldsymbol{\gamma}$ *condition:*

$$T\alpha_{\mathcal{A},\mathcal{B},\mathcal{C}} \circ \gamma_{\mathcal{A}\times\mathcal{B},\mathcal{C}} \circ (\gamma_{\mathcal{A},\mathcal{B}} \times i_{T\mathcal{C}}) =$$
$$= \gamma_{\mathcal{A},\mathcal{B}\times\mathcal{C}} \circ (i_{T\mathcal{A}} \times \gamma_{\mathcal{B},\mathcal{C}}) \circ \alpha_{T\mathcal{A},T\mathcal{B},T\mathcal{C}};$$

$\boldsymbol{\mu}$-$\boldsymbol{\gamma}$ *condition:*

$$\mu_{\mathcal{A}\times\mathcal{B}} \circ T\gamma_{\mathcal{A},\mathcal{B}} \circ \gamma_{T\mathcal{A},T\mathcal{B}} = \gamma_{\mathcal{A},\mathcal{B}} \circ (\mu_{\mathcal{A}} \times \mu_{\mathcal{B}});$$

$\boldsymbol{\eta}$-$\boldsymbol{\gamma}$ *condition:*

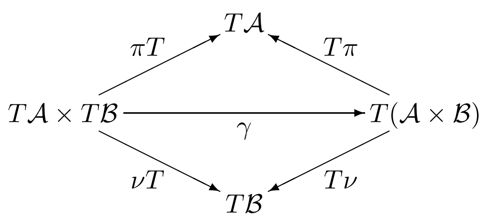
$$\gamma_{\mathcal{A},\mathcal{B}} \circ (\eta_{\mathcal{A}} \times \eta_{\mathcal{B}}) = \eta_{\mathcal{A}\times\mathcal{B}}.$$

Thus, the construction of the categories of ITs is, in effect, reduced to the selection of a base category $\mathbf{D}$, a functor $T: \mathbf{D} \to \mathbf{D}$, and a natural transformation $\gamma: \times T \to T\times$.

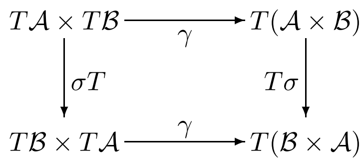
All these conditions have rather transparent meaning that we will try to comment below.

For better understanding, we also provide the corresponding commutative diagrams in which we omit the obvious indices for the sake of readability:

$\boldsymbol{\pi}$-$\boldsymbol{\gamma}$ a n d $\boldsymbol{\nu}$-$\boldsymbol{\gamma}$ c o n d i t i o n s. Marginal distributions extracted from an independent joint distribution coincide with the original distributions:



$\boldsymbol{\sigma}$-$\boldsymbol{\gamma}$ c o n d i t i o n. Transposition of components of an independent joint distribution leads to the corresponding transformation of the joint distribution, i.e., an independent joint distribution is "invariant" with respect to transposition of its components. More precisely, we can say that the independent distribution morphism for transposed components $\gamma_{\mathcal{B},\mathcal{A}}: T\mathcal{B} \times T\mathcal{A} \to T(\mathcal{B} \times \mathcal{A})$ is naturally isomorphic to the original morphism $\gamma_{\mathcal{A},\mathcal{B}}: T\mathcal{A}\times T\mathcal{B} \to T(\mathcal{A}\times\mathcal{B})$. The corresponding isomorphism (of morphisms) is provided by the pair $\left\langle \sigma_{T\mathcal{A},T\mathcal{B}}, T\sigma_{\mathcal{A},\mathcal{B}} \right\rangle$:



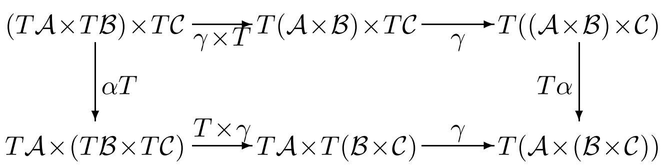
$\boldsymbol{\alpha}$-$\boldsymbol{\gamma}$ c o n d i t i o n. Independent joint distribution for three components is "naturally invariant" with respect to the order of parentheses. More precisely, the morphisms

$$\gamma_{\mathcal{A},\mathcal{B}\times\mathcal{C}} \circ (i_{T\mathcal{A}} \times \gamma_{\mathcal{B},\mathcal{C}}): T\mathcal{A}\times(T\mathcal{B}\times T\mathcal{C}) \to T(\mathcal{A}\times(\mathcal{B}\times\mathcal{C}))$$

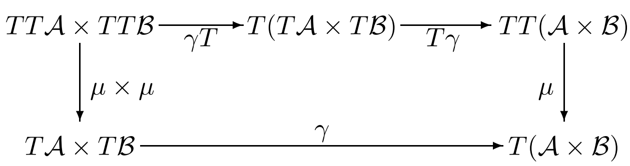
and

$$\gamma_{\mathcal{A}\times\mathcal{B},\mathcal{C}} \circ (\gamma_{\mathcal{A},\mathcal{B}} \times i_{T\mathcal{C}}): (T\mathcal{A}\times T\mathcal{B})\times T\mathcal{C} \to T((\mathcal{A}\times\mathcal{B})\times\mathcal{C})$$

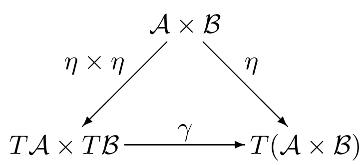
(that take independent joint distributions for three components with different order of parentheses) are naturally isomorphic via $\left\langle \alpha_{T\mathcal{A},T\mathcal{B},T\mathcal{C}}, T\alpha_{\mathcal{A},\mathcal{B},\mathcal{C}} \right\rangle$:



$\boldsymbol{\mu}$-$\boldsymbol{\gamma}$ c o n d i t i o n. Independent joint distribution for results of the mixing of two second-order distributions may also be obtained by mixing the corresponding second-order independent distributions:



$\boldsymbol{\eta}$-$\boldsymbol{\gamma}$ c o n d i t i o n. Independent joint distribution for two "singleton" distributions is just the corresponding "singleton" distribution on a product space:



Conclusions

The most developed theory of uncertainty in statistical physics is the probability theory. Certainly, mathematical methods of statistical physics accumulated a rich conceptual experience. At the same time, it appears that all these concepts have very abstract meaning and, hence, they can be treated in terms of category theory. We have shown in this article that the basic notions of probability theory and statistical physics are easily extended to the concept of information transformers.

Results of this article may provide a background for construction and study of new classes of ITs in dynamical nondeterministic systems.

For this purpose we have formulated a set of "elementary" axioms for a category of ITs, which would be sufficient for an abstract expression of the basic concepts of statistical physics.

The author wishes to express his deepest gratitude to Yu. P. Pyt'ev, who initiated and inspired the author's studies of informativeness, and S. S. Moskaliuk for fruitful discussions and encouraging the author in writing this paper.

1. *Sacksteder R.*// Ann. Math. Stat. – 1967. – **38**, N3. – P. 787–795.
2. *Morse N., Sacksteder R.* //Ibid. – 1966. – **37**, N2. – P. 203–214.
3. *Chentsov N.N.* Statistical Decision Rules and Optimal Inference. – Moscow: Nauka, 1972 (in Russian).
4. *Chentsov N.N.*// Dokl. Akad. Nauk SSSR. –1965. – **164**, N3. – P. 511–514.
5. *Golubtsov P.V.*// Pattern Recognition and Image Analysis. – 1991. – **1**, N1. – P. 77–86.
6. *Golubtsov P.V., Moskaliuk S.S.*// Hadronic J. – 2002. – **25**, N2. – P. 179–238 (http://arXiv.org/abs/math-ph/0211067).
7. *MacLane S.* Categories for the Working Mathematician. – New York: Springer, 1971.
8. *Herrlich H., Strecker G.E.* Category Theory. – Boston: Allyn and Bacon, 1973.
9. *Arbib M.A., Manes E.G.* Arrows, Structures and Functors. – New York: Academic Press, 1975.
10. *Goldblatt R.* Topoi. The Categorial Analysis of Logic. – Amsterdam: North-Holland, 1979.
11. *Golubtsov P.V.*// Probl. Inf. Transm. – 1992. – **28**, N2. – P. 125–140.
12. *Golubtsov P.V.*// Ibid. – 1995. – **31**, N3. – P. 195–215.
13. *Golubtsov P.V., Filatova S.A.*// Mat. Model. – 1992. – **4**, N7. – P. 79–94.
14. *Golubtsov P.V.*// Probl. Inf. Transm. – 1998. – **34**, N3. – P. 259–276.
15. *Golubtsov P.V.*// Ibid. – 1994. – **30**, N3. – P. 232–250.
16. *Golubtsov P.V.*// Applications of Fuzzy Logic Technology II. SPIE Proc. – **2493**. – Orlando, Florida, 1995. – P. 228–239.
17. *Golubtsov P.V.*//Proc. Intern. Conf. on Informatics and Control (ICI&C'97). – **2**. – St. Petersburg, Russia, 1997. – P. 512–517.
18. *Golubtsov P.V., Moskaliuk S.S.* Bayesian Decisions and Fuzzy Logic. – (Prepr./ESI Num. 626 Vienna, Austria, 1998). – P. 1–8.
19. *Golubtsov P.V.*// Probl. Inf. Transm. – 1999. – **35**, N3. – P. 80–99.
20. *Golubtsov P.V., Moskaliuk S.S.* // Hadronic J. – 2003. – **26** (in print).
21. *Moskaliuk S.S.*// Ibid. – 2003. – **26** (in print).
22. *Kleisli H.*// Proc. Amer. Math. Soc. – 1965. – **16**. – P. 544–546.
23. *Barr M., Wells C.* Toposes, Triples and Theories. – New York: Springer, 1984.
24. *Manes E.G.*// J. Math. Anal. and Appl. – 1982. – **85**. – P. 409–451.

СТРУКТУРА КЛАСІВ ПЕРЕТВОРЮВАЧІВ ІНФОРМАЦІЇ ЯК МОНОЇДАЛЬНА КАТЕГОРІЯ КЛЕЙСЛІ

П. В. Голубцов

Р е з ю м е

Пропонується розглядати будь-який однорідний клас систем перетворення інформації як сукупність морфізмів деякої категорії — категорії перетворювачів інформації (ПІ). Композиція ПІ відповідає їх "послідовному застосуванню". Пропонується аксіоматика для категорії ПІ як моноїдальної категорії, що містить підкатегорію (детермінованих ПІ) зі скінченними добутками і задовольняє певним аксіомам. Крім того, показано, що багато категорій ПІ можуть бути побудовані як категорії Клейслі.

СТРУКТУРА КЛАССОВ ПРЕОБРАЗОВАТЕЛЕЙ ИНФОРМАЦИИ КАК МОНОИДАЛЬНАЯ КАТЕГОРИЯ КЛЕЙСЛИ

П. В. Голубцов

Р е з ю м е

Предлагается рассматривать любой однородный класс систем преобразования информации как совокупность морфизмов некоторой категории — категории преобразователей информации (ПИ). Композиция ПИ отвечает их "последовательному применению". Предлагается аксиоматика для категории ПИ как моноидальной категории, содержащей подкатегорию (детерминированных ПИ) с конечными произведениями и удовлетворяющей определенным аксиомам. Кроме того, показано, что многие категории ПИ могут быть построены как категории Клейсли.