A formal framework for Data Fusion

Ildar Baymuratov ITMO University
Saint-Petersburg, Russia
Email: baimuratov.i@gmail.com
Nataly Zhukova ITMO University
Saint-Petersburg, Russia
Email: nazhukova@mail.ru

***Abstract*—In order to develop a formal framework for data fusion field main data fusion models are observed and a logical model of data fusion is suggested. It is shown that considered data fusion models are representable in the logical model. The logical model being insufficiently general to compare different methods of data fusion is reformulated on category theory language. After that a number of information theoretic measures defined on morphisms are suggested as universal criteria for evaluation of data fusion methods.**

***Index Terms*—data fusion, predicte logic, category theory, information theory.**

I. Introduction

There is a research area named 'Data fusion'. Data fusion is integration of different types of information from different sources, directed to decision-making. Aims of a typical data fusion system are processing of continuously incoming raw data, evaluation of current situation, prediction of forthcoming events and recommendation considering corresponding actions.

Consider biometric data as an example. A system that uses the only biometric factor has high error probability. For instance, if that factor is fingerprints, then for people with extremely fat or dry or injured fingers it is problematic to get fingerprints of acceptable quality. But if we use iris recognition as well, error probability will significantly decrease. Another example of data fusion is a moving object, such as an aircraft, observed by both a pulsed radar and an infrared sensor. The radar accurately determines the aircraft's range, but it hardly determines the angular direction of the aircraft. By contrast, the infrared sensor accurately determines the aircraft's angular direction, but is unable to measure range. The combination of these two observations provides an improved determination of the aircraft's location.

Data fusion process is analogous to the process during which humans and animals combine data from different senses, experience and reasoning ability to improve their survive chances. For example, when vision is limited, the sense of hearing can provide advanced warning of impending danger. Though the concept of data fusion is not new, the emergence of new sensors and processing techniques makes data fusion increasingly viable. Data fusion has advanced from collection of related techniques to a comprehensive engineering discipline with established terminology, mathematical techniques and design principles. Data fusion techniques are drawn from such disciplines as digital signal processing, statistical estimation, control theory and artificial intelligence. In our time, data fusion systems are used for target tracking, automated identification, reasoning applications and so on. The extensive research of data fusion field can be found in, for example, Mitchell's 'Data Fusion: Concepts and Ideas' [14] or Hall's and McMullen's 'Mathematical Techniques in Multisensor Data Fusion' [9].

By now numerous methods of data integration were developed. The data fusion processes include different estimation, feature extraction, classification and inference techniques. Some of them are more elaborated, such as image recognition or classification, and some are hardly articulated. Data fusion methods were developed apart from each other and it is difficult to put them in any common context. As result, data fusion field dissolves to set of poorly correlated subproblems. There is a lack of theoretical framework for designing data fusion systems. Existing classifications of data fusion processes, such as JDL-model, are expressed mainly in natural language rather than in formal and it is problematic to formally prove that one fusion system is more preferable than another.

The purpose of this work is to develop a formal framework of data fusion. That allows:

- represent the process of data fusion in more systematic way;
- develop universal criteria for estimating and comparing data fusion methods.
- automate designing of data fusion systems.

In other words, the purpose is to develop a formal theory of data fusion. Examples of such formalization already exist. In 'Mathematics of Data Fusion' by I. R. Goodman, Ronald P. S. Mahler and Hung T. Nguen [7] fuzzy logic, random set theory and conditional event algebra are used as such mathematical foundation. In this work data fusion process is reduced, first, to logic then to category theory. Finally, a set of criteria for estimating and comparing data fusion methods is suggested and verified.

II. Review of data fusion models

1) JDL model: In 1986 the Joint Directos of Laboratories (JDL) Data Fusion Working Group began codifying data fusion terminology. As result, data fusion process model, named JDL-model, was created. The JDL-model with subsequent revisions is the most widespread system for categorizing data fusion methods. The JDL-model was designed as a functional model, i.e. a set of definitions that comprises any data fusion system, therefore, the model was intended to be very general and useful across multiple application areas. It includes:

1) Object refinement: data integration for getting improved representation of individual objects.
2) Situation refinement: description of relations between objects and environment.
3) Threat refinement: projecting of current situation in future and predicting of consequences.
4) Process refinement: meta-process that monitors data-fusion processes to maintain real-time performance.

For each of these processes JDL-model associates specific functions and techniques. Besides of that processes JDL-model also considers sensor inputs, human-computer interaction, database management and source preprocessing. A corresponding terminology was developed to provide consistent definitions.

In 1998 the JDL-model was revisited by [18] to provide a categorization representing logically different types of problems, which are solved by different techniques, and to update terminology. The updated model is following:

1) Sub-object data assessment: signal level data association and characterization.
2) Object assessment: estimation and prediction of object states.
3) Situation assessment: estimation and prediction of relations among objects.
4) Impact assessment: estimation and prediction of participants actions.
5) Process Refinement: different processes supporting mission objectives.

According to [18], it is possible to generalize JDL-model levels. Sub-object data assessment is a special case of object assessment, where objects are signals, and impact assessment is a special case of situation assessment, where relations are relations to agents. Also the authors mention that process refinement is not data fusion, but a kind of resource management, therefore, only two levels of fusion remain: object assessment and situation assessment.

2) Dasarathy Functional Model: Dasarathy [5] categorize data fusion functions according to the types of input and output data. The types are:

- Data.
- Features.
- Objects.

And the functions:

- DAI-DAO: Signal detection.
- DAI-FEO: Feature extraction.
- FEI-FEO: Feature refinement.
- FEI-DEO: Object characterization.
- DEI-DEO: Object refinement.

In [17] Darathy's model was augmented by other possible combinations. The resulting model is represented in the Table I.

3) Omnibus model: The Omnibus model was proposed by [1]. It is posed to be a generalization of the Boyd Control Loop [2], also known as the Observe, Orient, Decide, Act (OODA), of the U.S. intelligent services' intelligence cycle and the commonly used waterfall model [8]. The model is a four stage cycle describing the main activities in a fusion system. The stages are following:

- Observe: sensing and signal processing.
- Orient: feature extraction and pattern processing.
- Decide: decision making and context processing.
- Act: control and resource tasking.

The cycle is represented by the following diagram:

$$\begin{tikzcd}
\text{Observe} \arrow[r] & \text{Orient} \arrow[d] \\
\text{Act} \arrow[u] & \text{Decide} \arrow[l]
\end{tikzcd}$$

Elements and relations of these models are quite similar and overlapping. However, due to informal definitions and different structure, JDL-model is hierarchical, augmented Dasarathy's model is a matrix and Omnibus model is cyclical, it is not possible to compare and generalize them, which is why the following formal model is suggested. Nevertheless, overview of these models presents a list of main categories and relations that have to be expressible in the suggested model, if all of them are representable, it allows to consider the suggested model as valid.

III. LOGICAL DATA FUSION MODEL

The first assumption is that the problem of data fusion model development is identical to the old issue of defining cognitive processes and its hierarchy from psychology and cognitive science. If we agree that concept of data fusion is applicable to human cognitive processes, and human cognitive processes are, in turn, defined in different psychological and epistemological models, that theories must be at least included in the data fusion model.

One of the first examples of such epistemological model is Kant's faculties of cognition [10]:

- Sensibility: the faculty of intuitions and sense perception.
- Understanding: the faculty of concepts and thoughts.
- Imagination: mediates between understanding and sensibility, the source of all sorts of synthesis.
- Reason: produces logical inferences and decisions, imposes coherence and consistency and implements modal concepts such as necessity or obligation.
- Self-consciousness: imposes a higher-order unity into all lower-order faculties.

Kant holds that human mind has two basic cognitive faculties: understanding and sensibility. The essential difference between concepts and intuitions is that concepts are general representations, the logical form of objects, while intuitions

TABLE I

		Output		
		Data	Features	Objects
Input	Data	Signal detection	Feature extraction	Gestalt-based object characterization
	Features	Model-based feature extraction	Feature refinement	Feature-based object characterization
	Objects	Model-based estimation	Model-based feature extraction	Object refinement

are singular, sense-related, immediate object representations. Understanding and sensibility both serve to the faculty of imagination and the faculty of reason. Finally, the faculty of apperception or rational self-consciousness plays executive role in the organisation of the mind.

As we see, Kant's faculties of cognition look quite similar to data fusion concepts. Sensibility, being object representation, can be associated with object assessment, understanding, as extracting logical form of objects, can be associated with situation assessment, reason — with impact assessment and, finally, self-consciousness — with process refinement.

Numerous other categorizations of cognition process have been made after Kant, but still it does not solve the problem, as these categorizations are not formal. The second assumption is that the formal representation of cognition and thinking in general is the subject matter of logic, which is why the first suggested model is logical.

In the following, a logical model of data fusion process is presented. First of all, define what logic is meant. This logical model is based on first order predicate logic, extensive elaboration of which and its semantics can be found, for example, in Chang's and Keisler's 'Model Theory' [3], but has the following differences:

- for shortness, individual constants are considered as reflexive individual functions and propositions are considered as reflexive propositional operator.
- due to discreteness of the data the universe for this logic is finite.

The language of suggested logical model consists of:

- individual variables: $x_1, x_2, ...$;
- individual functions: $f_1, f_2, ...$;
- predicate functions: $P_1, P_2, ...$;
- quantifiers: $\forall, \exists$;
- propositional operators: $\neg, \wedge, \vee, \rightarrow$.

Each predicate P is assumed to be an n-placed function, if $n = 0$, predicate P is a proposition. Similarly, each function f is an m-placed function, if $m = 0$, f is a constant.

Certain strings of symbols from the above list are called terms. Definition of a term:

- a variable x is a term,
- if f is an m-placed function and $t_1, ..., t_m$ are terms, then $f(t_1, ..., t_m)$ is a term,
- a string of symbols is a term only if it is a term due to a finite number of applications of previous rules.

Definition of a sentence:

- if ϕ is an n-placed predicate P and $t_1, ..., t_n$ are terms then $P(t_1, ..., t_n)$ is a sentence,
- if x is a variable and ϕ is a formula containing x then $\forall x\phi$ and $\exists x\phi$ are sentences,
- if ϕ is a sentence then $\neg\phi$ is a sentence,
- if ϕ and ψ are sentences, then $\phi \wedge \psi, \phi \vee \psi, \phi \rightarrow \psi$ are sentences.
- a finite sequence of symbols is a sentence only if it is a sentence due to a finite number of applications of previous rules.

The third assumption is that the process of interpretation of a sentence ϕ corresponds to the process of data fusion, therefore, the formal definition of interpretation of a sentence ϕ corresponds to a formal model of data fusion. A model $\mathcal{M}$, or a 'possible world', for the suggested language consists of a universe A, a nonempty finite set of objects. The correspondence between the language and the universe is given by an interpretation function $\mathcal{I}$ defined by the following inductive procedure:

1) interpretation of a term:
   - $\mathcal{I}(f) = id : A \rightarrow A$,
   - $\mathcal{I}(f(t_1, ..., t_m)) = g : A^m \rightarrow A$,
2) $\mathcal{I}(\phi) = 1 \Leftrightarrow \mathcal{M} \models \phi$. If ϕ is an atomic sentence:
   - $\mathcal{M} \models P(t_1, ..., t_n) \Leftrightarrow R(a_1, ..., a_n) \in A$,
3) if ϕ is a complex sentence:
   - $\mathcal{M} \models \forall x_i\phi \Leftrightarrow$ for every $a \in A$ it is true that $R(a) \in A$,
   - $\mathcal{M} \models \exists x_i\phi \Leftrightarrow$ exists $a \in A$, such that $R(a) \in A$,
   - $\mathcal{M} \models P \Leftrightarrow \mathcal{M} \models P(t_1, ..., t_n)$,
   - $\mathcal{M} \models \neg\phi \Leftrightarrow \mathcal{M} \not\models \phi$
   - $\mathcal{M} \models \phi \wedge \psi \Leftrightarrow \mathcal{M} \models \phi$ and $\mathcal{M} \models \psi$,
   - $\mathcal{M} \models \phi \vee \psi \Leftrightarrow \mathcal{M} \models \phi$ or $\mathcal{M} \models \psi$,
   - $\mathcal{M} \models \phi \rightarrow \psi \Leftrightarrow \mathcal{M} \not\models \phi$ or $\mathcal{M} \models \psi$,

otherwise $\mathcal{I}(\phi) = 0$. A model $\mathcal{M}$ is a pair $< A, \mathcal{I} >$. Besides the interpretation function there is a valuation function V that corresponds each variable x to an object $a \in A$.

The principal difference between these three types of interpretation is the types of domain and codomain. Analyzing the definition of interpretation, two types are extracted:

- Objects (O), the elements of the universe A.
- Truth-values (TV): $\{0, 1\}$.

Therefore, three types of interpretation can be represented as follows:

- Terms (T): $A^m \rightarrow A$.
- Atomic sentences (AS): $A^n \rightarrow \{0, 1\}$.
- Complex sentences (CS): $\{0, 1\} \rightarrow \{0, 1\}$.

and the logical model of data fusion can be represented by the following diagram:

$$\begin{tikzcd} O \arrow[loop left, "T"] \arrow[r, "AS"] & TV \arrow[loop right, "CS"] \end{tikzcd}$$

Let us compare the suggested logical models with considered existing data fusion models. To do this, we are going to translate terms of the existing data fusion models to the logical language.

JDL-model considers signals, objects, features, situations, decisions and control. From the logical point of view, signals, objects and features are objects of the universe, situations are relations between objects and decisions are relations between truth-values, therefore, sub-object data assessment and object assessment correspond to terms, situation assessment corresponds to atomic sentences, impact assessment corresponds to complex sentencesn. Process refinement has no correspondence in the logical model, but as authors of [18] mentioned, process refinement may not be considered as part of the data fusion process.

In the augmented Dasarathy's Functional model three categories are considered: data, features and objects. As was previously said, these three categories are objects of the universe, therefore, different relations described in the augmented Dasarathy's model are different kinds of terms.

As JDL-model, Omnibus model considers signals, features, decisions and control as well, therefore, 'Observe' and 'Orient' correspond to terms and 'Decide' corresponds to complex sentences. 'Act' has no correspondence.

The result of comparing is represented in the table II. The augmented Dasarathy's Model is not included in the resulting table due to its matrix structure.

Summing up, the existing data fusion models are mostly representable in the suggested logical model. The only issue is control. But according to some authors, control is not a part of data fusion, it may be considered as a part of executing expert system instead.

However, at current stage it is still problematic to define universal criteria applicable to terms, atomic and complex sentences, thus, in the next section the presented logical model is reformulated on the language of more abstract mathematical theory.

IV. Category theoretic data fusion model

As such more abstract mathematical theory category theory is used. There are examples of data fusion representation in category theory terms. Kokar, Tomasik and Weyman in 'Formalizing Classes of Information Fusion Systems' [12] consider two classes of data fusion systems: data fusion and decision fusion, formalize them in terms of category theory and show that decision fusion is a subclass of data fusion.

The presented category theoretic model is derived from logical one. It is shown, for instance, in R. Goldblatt's 'Topoi, the Categorial Analysis of Logic' [6] that logic can be described in terms of category theory.

Let us specify some basic elements of category theory. A category is a mathematical structure that consists of objects and morphisms and the following expressions are satisfied:

- each morphism f is associated with objects $dom(f)$ and $cod(f)$. If $dom(f) = a$ and $cod(f) = b$, it is denoted as
$$f : a \to b;$$
- each pair of morphisms $< f, g >$ is associated with a composition $g \circ f$, such that
$$g \circ f : dom(f) \to cod(g);$$
- associativity: if there are morphisms $f : a \to b, g : b \to c$ and $h : c \to d$, then
$$(f \circ g) \circ h = f \circ (g \circ h);$$
- identity: for every object x, there is an identity morphism $1 : x \to x$.

For instance, in the category Set the objects are sets and morphisms are functions between sets.

Analyzing the structure of the logical model in category theoretic terms, there are two categories:

- Objects: a universe $A = \{X_1, ..., X_n\}$.
- Truth-values: $TV = \{0, 1\}$.

And three morphisms:

- Terms: $T : A^n \to A$.
- Atomic statements: $AS : A^n \to TV$.
- Complex statements: $CS : TV^m \to TV$.

According to category theory, an identity morphism exists for each category, thus, two identity morphisms are defined:

- for category A identity morphism $id : A \to A$ forms a set of constants;
- for category TV identity morphism $id : TV \to TV$ forms a set of propositions.

Therefore, each logical function and, in turn, data fusion process can be represented as a morphism. From this point, it becomes possible to formulate universal criteria for data fusion algorithms. But first we are going to show using the developed framework that there is a gain of knowledge during the data fusion process.

A. Morphisms quantitative characteristics

Given a morphism $f : X \to Y$, there is a set $\mathcal{F} : X \to Y$ of alternative morphisms for the domain X and codomain Y. The cardinality $|\mathcal{F}|$ of the set $\mathcal{F}$ is defined as quantitative characteristic of the morphism f. In this section characteristics for each morphism of the present model will be given.

1) Terms: Given a morphism $f : X \to Y$, how many other possible morphisms from X to Y there are? First, there are $|Y|$ different valuations of $f(x)$ for some x. Then, there are $|Y|$ different values for each $x \in X$, thus, for a set of morphisms $\mathcal{F} : X \to Y$:
$$|\mathcal{F} : X \to Y| = |Y|^{|X|}.$$

This result can be generalized in two ways: first, morphism f can have n arguments, i.e. f is $f : X^n \to Y$. Then
$$|f : X^n \to Y| = |Y|^{|X_1|...|X_n|}.$$
Second, some $f'(X)$ can have a different codomain, thus, $\mathcal{F}(X) = \{\mathcal{F}(X) = Y_1\} \cup ... \cup \{\mathcal{F}(X) = Y_n\}$ and
$$|\mathcal{F}(X)| = |Y_1|^{|X|} + ...|Y_n|^{|X|}.$$
Therefore,
$$|T : X^n \to Y| = \sum_i |Y_i|^{\prod_j |X_j|},$$
for each $Y_i \subseteq Y$, such that $T_i : X_1 \times ... \times X_j \to Y_i$.

TABLE II

Input type	Output type	JDL	Omnibus Model	Logical Model
Signals	Objects	Sub-object data assessment	Observe	Terms
Objects	Features	Object assessment	Orient	
Features	Situations	Situation assessment		Atomic sentences
Situations	Decisions	Impact assessment	Decide	Complex sentences
Control		Process refinement	Act	

2) Atomic sentences: According to the definitions, an atomic sentence is a morphism $AS : A^n \to TV$ and $TV = \{0, 1\}$, therefore,

$$AS : A^n \to \{0, 1\}.$$

and

$$|AS : A \to TV| = 2^{\prod_i |A_i|}.$$

3) Complex sentences: As complex sentence is a morphism $CS : TV^m \to TV$ and $TV = \{0, 1\}$,

$$CS : \{0, 1\} \times ... \times \{0, 1\} \to \{0, 1\},$$

therefore,

$$|CS : TV^m \to TV| = 2.$$

Summing up, with each level of data fusion degrees of freedom decrease, that correlates with data integrity being increased during the data fusion process.

B. Morphism criteria

In this section the set of universal morphism criteria considering its effectiveness as data fusion algorithms is defined.

For each morphism $f : X \to Y$ it is possible to characterize:

- some value $y \in Y$ of morphism f;
- morphism f itself;
- morphism f in relation with the whole set of morphisms $\mathcal{F} : X \to Y$.

Therefore, a measure will be defined for each of this items. Considered measures origins from information theoretical ones. An extensive elaboration of information theoretic measures can be found, for example, in 'Elements of Information Theory' [4] by Cover and Thomas.

1) Information: Given a morphism $f : X \to Y$, a measure for some value $y \in Y$ of that morphism is defined.

As such measure Shannon's [16] information measure is used. For each $x_j \in X$ and $y_i \in Y$ it is possible to count frequency m_i of $f(x_j) = y_i$ being satisfied. Then for each $y_i \in Y$ it is possible to count information $I(y_i)$:

$$I(y_i) = -\log_{|X|} \frac{m_i}{|X|},$$

in other words,

$$I(y_i) = 1 - \log_{|X|} m_i.$$

This formula differs from Shannon's one in that cardinality $|X|$ of the set X is used as logarithm base. Due to this fact,

$$0 \leq I(y_i) \leq 1,$$

and $I(y_i) = 1$ means that after the transformation from $f^{-1}(y_i)$ to y_i all information is reserved. Vice versa, if $I(y_i) = 0$, it means that after the transformation $f : X \to Y$, all information presented in X is lost.

2) Knowledge: Another assumption is that considering the data fusion process, information loss has positive aspects. Relieving from extra details reveals common tendencies. In respect to human thinking this process is named abstraction. Abstraction results in knowledge, thus, the measure $K(y_i)$ of knowledge amount contained in some value $y_i \in Y$ of morphism $f : X \to Y$, as opposite to information measure, is introduced:

$$K(y_i) = \log_{|X|} m_i.$$

The measure $K(y_i)$ also satisfies

$$0 \leq K(y_i) \leq 1,$$

but if $K(y_i) = 1$, it means that after the transformation $f : X \to Y$ all information, contained in X, is generalized. Vice versa, if $K(y_i) = 0$, it means that transformation $f : X \to Y$ yields no knowledge.

Given some value y, it can be shown that

$$K(y) = 1 - I(y).$$

As $I(y) = 1 - \log_{|X|} m$,

$$1 - I(y) = 1 - (1 - \log_{|X|} m),$$

that is

$$1 - I(y) = \log_{|X|} m,$$

but $\log_{|X|} m = K(y)$.

Therefore, each value $y \in Y$ of some morphism $f : X \to Y$ can be characterized considering data fusion process in two ways. If we are interested in details and exceptional cases, they could be found by calculating information amount $I(y)$, or if we are more interested in common tendencies and typical cases, they could be found by calculating knowledge amount $K(y)$ contained in value y.

3) Entropy and negentropy: Values of $I(y_i)$ and $K(y_i)$ can be considered as random variables, thus, statistical methods are applicable to them. In this work only expected value is considered.

For some morphism $f : X \to Y$ expected value of information is entropy $H(f)$. According to present modifications, entropy $H(f)$ of morphism $f : X \to Y$ can be calculated as follows:

$$H(f) = -\sum_i \frac{m_i}{|X|} \log_{|X|} \frac{m_i}{|X|},$$

If some y_i satisfies $I(y_i) \geq H(f)$, y_i can be considered as informative and if $I(y_i) < H(f)$, y_i is uninformative.

If $H(f) = 0$, it means that in every case the morphism $f : X \to Y$ has one and the same value y, and, vice versa, if $H(f) = 1$, for each x_j morphism $f(x_j)$ has some different value y_i, thus, entropy $H(f)$ can be considered as uncertainty measure for morphism $f : X \to Y$.

For expected value of knowledge, denoted as $N(f)$, the term 'negentropy' is used. It was introduced by E. Schrdinger [15] as opposite to entropy. The measure called 'negentropy' mathematically coincides with relative entropy, or Kullback-Leibler divergence [13], however, in the present model this measure has its own interpretation.

Negentropy is calculated as follows:

$$N(f) = \sum_i \frac{m_i}{|X|} \log_{|X|} m_i.$$

If some y_i satisfies $K(y_i) \geq N(f)$, y_i can be considered as meaningful, and, vice versa, if $K(y_i) < N(f)$, y_i is meaningless. If $N(f) = 0$, it means that for each x_j the morphism $f(x_j)$ has different value, and, vice versa, if $N(f) = 1$, the morphism $f : X \to Y$ in every case has the only value y, therefore, negentropy, as opposition to entropy, can be considered as certainty measure for morphism $f : X \to Y$.

Since $I(y)$ for some value y is calculated as $1 - \log_{|X|} m$, entropy $H(f)$ can also be calculated as

$$H(f) = \sum_i \frac{m_i}{|X|}(1 - \log_{|X|} m_i).$$

Since $\sum_i \frac{m_i}{|X|} 1 = 1$,

$$H(f) = 1 - \sum_i \frac{m_i}{|X|} \log_{|X|} m_i,$$

and since $N(f) = \sum_i \frac{m_i}{|X|} \log_{|X|} m_i$, entropy can be calculated as

$$H(f) = 1 - N(f).$$

Therefore, entropy $H(f)$ and negentropy $N(f)$ characterize some morphism $f : X \to Y$ as data fusion algorithm. The more negentropy $N(f)$ is, the more data morphism f integrates.

4) Partition information/knowledge amount: Given morphisms set $\mathcal{F} : X \to Y$, each morphism $f_i \in \mathcal{F}$ yields some partition $Part_j(Y)$ of Y. If each morphism f in some subset $\mathcal{F}_i \subseteq \mathcal{F}$ yields one and the same partition $Part_i(Y)$, then for each $\mathcal{F}_i$ it is possible to count frequency m_j, thus, it is possible to calculate information $I(Part_i(Y))$ and knowledge $K(Part_i(Y))$ for each partition $Part_i(Y)$ of the set Y:

$$I(Part_i(Y)) = 1 - \log_{|\mathcal{F}|} m_j$$

and

$$K(Part_i(Y)) = \log_{|\mathcal{F}|} m_j.$$

As result, there are two new random variables and it is possible to calculate entropy and negentropy for them. Entropy $H(\mathcal{F})$ of the set of morphisms $\mathcal{F}$:

$$H(\mathcal{F}) = 1 - \sum_j \frac{m_j}{|\mathcal{F}|} \log_{|\mathcal{F}|} m_j$$

and negentropy $N(\mathcal{F})$:

$$N(\mathcal{F}) = \sum_j \frac{m_j}{|\mathcal{F}|} \log_{|\mathcal{F}|} m_j.$$

If some partition $Part_i(Y)$ satisfies $I(Part_i(Y)) \geq H(\mathcal{F})$, it means that $Part_i(Y)$ is informative, analogously, if $K(Part_i(Y)) \geq N(\mathcal{F})$, partition $Part_i(Y)$ is meaningful.

5) Expected entropy/negentropy: Given some set of morphisms $\mathcal{F} : X \to Y$, it is possible to characterize some morphism $f \in \mathcal{F}$ in respect to the whole set $\mathcal{F}$. Thereby, expected entropy and negentropy of the set $\mathcal{F}$ are defined. Let some subset $\mathcal{F}_i \subseteq \mathcal{F}$ yields partition $Part_i(Y)$ that has frequency m_i, then expected entropy $\bar{H}(\mathcal{F})$ of the set $\mathcal{F}$ is calculated as follows:

$$\bar{H}(\mathcal{F}) = \sum_{ij} \frac{m_i}{|\mathcal{F}|} H(f_j)$$

for each $f_j \in \mathcal{F}_i$. Expected negentropy $\bar{N}(\mathcal{F})$ is calculated as follows:

$$\bar{N}(\mathcal{F}) = \sum_{ij} \frac{m_i}{|\mathcal{F}|} N(f_j).$$

If some $f_i : X \to Y$ satisfies $H(f_i) \geq \bar{H}(\mathcal{F})$, the morphism f_i is considered as informative with respect to the set of morphisms $\mathcal{F}$, analogously, if $N(f_i) \geq \bar{N}(\mathcal{F})$ is satisfied, the morphism f_i is considered as meaningful with respect to the set $\mathcal{F}$.

6) Internal and external morphism criteria: Given some set of morphisms $\mathcal{F} : X \to Y$ and some subset $\mathcal{F}_i \subseteq \mathcal{F}$ that yields partition $Part_i(Y)$, for some morphism $f_j \in \mathcal{F}_i$ it is possible to calculate, on the one hand, information $I(f_j)$ or knowledge $K(f_j)$ with respect to the whole set $\mathcal{F}$, on the other hand, it is possible to calculate entropy $H(f_j)$ and negentropy $N(f_j)$ of the morphism f_j itself. As result, one pair of measures characterize morphism f_j with respect to other morphisms, thus, these measures are considered as external criteria, while other pair of measures characterize the morphism f_j with respect to the way it divides the set Y, thus, these measures are considered as internal criteria. Therefore, products

$$I(f_j)H(f_j)$$

and

$$K(f_j)N(f_j)$$

characterize morphism f_j both from inner and outer sides. If we are interested in comprehensive estimation of a morphism, it is products $I(f_j)H(f_j)$ and $K(f_j)N(f_j)$ to be considered.

C. Application of the criteria to the model

In this section the set of developed criteria is applied to morphism classes of the presented model. Given some identity morphism $id : X \to X$, as any identity morphism is bijection, it satisfies

$$H(id) = 1$$

and

$$N(id) = 0.$$

TABLE III

ID	$x_1, ..., x_n$	M_1	M_2
1	...	A	C
2	...	A	C
3	...	B	D
4	...	B	E

TABLE IV

	A	B	C	D	E
P	$0,5$	$0,5$	$0,5$	$0,25$	$0,25$
I	$0,5$	$0,5$	$0,5$	1	1
K	$0,5$	$0,5$	$0,5$	0	0

Since there are identity morphisms for objects and truth-values categories that constitute constants and propositions, constants and propositions are the most informative and the least meaningful.

As constants are presented in every data set, the morphism T, terms, has lower bound:

$$min(N(T)) = 0.$$

Given some atomic statement $as : A \to TV$, on the one hand, it is possible that $|A| = |TV|$, on the other hand, it is possible that $|TV| = 1$. Therefore, the morphism AS satisfies

$$0 \leq N(AS) \leq 1.$$

As every $cs \in CS$ has the only value $tv \in TV$, each cs satisfies

$$H(cs) = 0$$

and

$$N(cs) = 1.$$

In other words, complex statements are the least informative and the most meaningful.

To sum up, it is seen that data fusion processes could be ordered as follows:

$$N(T) \leq N(AS) \leq N(CS)$$

and data fusion process presents conversion of information to knowledge.

1) Application of the criteria to terms: According to the suggested logical model, different methods of signal, object refinement, feature extraction and pattern recognition can be represented as individual functions, i.e. terms. Let us consider an application of the suggested criteria to an abstract instance of such methods.

Suppose that we have a data set consisting of parameters $x_1, ..., x_n$ and four records and two different methods, M_1 and M_2, that associate each record with some class. The partitions are presented in the Table III.

Due to the developed criteria it is possible to compare these two methods. First, calculate frequency P, information I and knowledge K for each of the classes, see Table IV. Then calculate for each method entropy H and negentropy N, see Table V.

TABLE V

	M_1	M_2
H	$0,5$	$0,75$
N	$0,5$	$0,25$

Let us see if they produce enough knowledge or information in general. To do this, we shall consider every possible partitions of the four elements set and calculate frequency, information, knowledge, entropy and negentropy for each partition, see Table VI.

The expected entropy and negentropy are:

$$\bar{H} \approx 0,66, \bar{N} \approx 0,33.$$

It means that M_2 is informative and M_1 produce enough knowledge. Finally, calculate products $I(M_i)H(M_i)$ and $K(M_i)N(M_i)$ for M_1 and M_2, see Table VII. According to this results, method M_1 produces more information and more knowledge than M_2.

2) Application of the criteria to atomic sentences: The suggested criteria allow us to compare different relations between objects. Suppose we are questioning which of two relations, $R_1(x, y)$ or $R_2(x, z)$, to consider in the situation, described on the Table VIII, where 1 means that a relation holds and 0 means that it does not.

Calculating criteria for each of the cases see on the Table IX. Calculating criteria for partitions see on the Table X. The expected entropy and negentropy:

$$\bar{H} \approx 0,27, \bar{N} \approx 0,72.$$

It means that both of the relations are informative and none of them is meaningful. Calculating products $I(R_i)H(R_i)$ and $K(R_i)N(R_i)$ for R_1 and R_2 on the Table XI. It means that relation R_2 produces more information and relation R_1 produces more knowledge.

3) Application of the criteria to complex sentences: Suppose we have two facts, A_1 and A_2, which conclusion will be more informative or meaningful: $A_1 \vee A_2$ or $A_1 \not\equiv A_2$?

Calculation of the criteria for each case on the Table XII. Calculation of the criteria for partitions on the Table XIII. Expected entropy $\bar{H}$ and expected negentropy $\bar{N}$ are:

- $\bar{H} = 0,3875$,
- $\bar{N} = 0,6125$.

It means that both conclusions are informative but none of them is meaningful.

Calculation products $I * H$ and $K * N$ for each operator is presented on the Table XIV. According to the results, conclusion $A_1 \not\equiv A_2$ provides more information but the conclusion $A_1 \vee A_2$ provides more knowledge.

V. CONCLUSION

Summing up, two general models of data fusion are presented. The first model is logical. It provides the structure of data fusion process and the list of categories and morphisms

TABLE VI

	$Part_1$	$Part_2$	$Part_3$	$Part_4$	$Part_5$
Partition	$\{1\},\{1\},\{1\},\{1\}$	$\{2\},\{1\},\{1\}$	$\{2\},\{2\}$	$\{3\},\{1\},$	$\{4\},$
P	$0,09375$	$0,5625$	$0,140625$	$0,1875$	$0,015625$
I	$\approx 0,43$	$\approx 0,1$	$\approx 0,35$	$\approx 0,3$	$0,75$
K	$\approx 0,57$	$\approx 0,9$	$\approx 0,65$	$\approx 0,7$	$0,25$
H	1	$0,75$	$0,5$	$\approx 0,4$	0
N	0	$0,25$	$0,5$	$\approx 0,6$	1

TABLE VII

	M_1	M_2
$I(M_i)H(M_i)$	$0,175$	$0,075$
$K(M_i)N(M_i)$	$0,325$	$0,225$

TABLE VIII

X	Y	$R_1(x,y)$	X	Z	$R_2(x,z)$
x_1	y_1	1	x_1	z_1	1
x_1	y_2	0	x_1	z_2	1
x_1	y_3	0	x_1	z_3	1
x_2	y_1	1	x_2	z_1	0
x_2	y_2	0	x_2	z_2	0
x_2	y_3	0	x_2	z_3	0

TABLE IX

	$R_1(x,y)$	$R_2(x,z)$
P(0)	$\approx 0,67$	$0,5$
P(1)	$\approx 0,33$	$0,5$
I(0)	$\approx 0,23$	$\approx 0,39$
I(1)	$\approx 0,61$	$\approx 0,39$
K(0)	$\approx 0,77$	$\approx 0,61$
K(1)	$\approx 0,39$	$\approx 0,61$
H	$\approx 0,35$	$\approx 0,39$
N	$\approx 0,64$	$\approx 0,61$

TABLE X

	$Part_1$	$Part_2$	$Part_3$	$Part_4$
Partition	$\{3\},\{3\}$	$\{4\},\{2\}$	$\{5\},\{1\}$	$\{6\}$
P	$0,3125$	$0,46875$	$0,1875$	$0,03125$
I	$\approx 0,28$	$\approx 0,18$	$\approx 0,4$	$\approx 0,83$
K	$\approx 0,72$	$\approx 0,82$	$\approx 0,6$	$\approx 0,17$
H	$\approx 0,39$	$\approx 0,35$	$\approx 0,25$	0
N	$\approx 0,61$	$\approx 0,64$	$\approx 0,75$	1

TABLE XI

	R_1	R_2
$I(R_i)H(R_i)$	$\approx 0,065$	$\approx 0,1$
$K(R_i)N(R_i)$	$\approx 0,53$	$\approx 0,44$

TABLE XII

	$A_1 \vee A_2$	$A_1 \not\equiv A_2$
$P(0)$	0,25	0,5
$P(1)$	0,75	0,5
$I(0)$	1	0,5
$I(1)$	0,2	0,5
$K(0)$	0	0,5
$K(1)$	0,8	0,5
H	0,4	0,5
N	0,6	0,5

TABLE XIII

	$Part_1$	$Part_2$	$Part_3$
Partition	$\{4\}$	$\{3\},\{1\}$	$\{2\},\{2\}$
P	$0,125$	$0,5$	$0,375$
I	$0,75$	$0,25$	$0,35$
K	$0,25$	$0,75$	$0,65$
H	0	$0,4$	$0,5$
N	1	$0,6$	$0,5$

TABLE XIV

	$A_1 \vee A_2$	$A_1 \not\equiv A_2$
$I * H$	$0,1$	$0,175$
$K * N$	$0,45$	$0,325$

that constitute the process. There are two categories: objects and truth-values, and three morphisms: terms, atomic statements and complex statements.

The second model is category theoretic and it provides the set of universal criteria for each possible morphism. According to this criteria it is possible for each morphism to define, first, is some value of the morphism meaningful or informative, second, is the morphism meaningful or informative by itself, third, is the morphism meaningful or informative in respect to alternative morphisms.

REFERENCES

[1] M. Bedworth and J. OBrien. *The Omnibus model: a new model of data fusion?* IEEE Aerospace and Electronic Systems Magazine, April 2000, pp. 3036.

[2] J. R. Boyd. *The essence of winning and losing.* Unpublished lecture notes, 1996.

[3] C.C. Chang, H.J. Keisler *Model Theory.* Elsevier, 1990. pp. 649.
[4] Thomas M. Cover, Joy A. Thomas. *Elements of Information Theory, 2nd Edition.* Wiley-Interscience, 2006. pp. 792.
[5] Dasarathy, B. *Decision Fusion.* IEEE Computer Society Press, 1994.
[6] R. Goldblatt. *Topoi: The Categorial Analysis of Logic.* Elsevier Science Publishers, 1984. pp. 551.
[7] I. R. Goodman, Ronald P. S. Mahler, Hung T. Nguyen. *Mathematics of Data Fusion.* Springer Netherlands, 1997. pp. 346.
[8] Harris C. J., Bailey A., Dodd, T. J. *Multi-sensor data fusion in defence and aerospace* // Aeronautical Journal. 102 (1015): 229-244.
[9] David L. Hall, Sonya A. H. McMullen. *Mathematical Techniques in Multisensor Data Fusion, Second Edition.* Norwood, MA: Artech House, 2004. pp. 466.
[10] Kant I. *Critique of Pure Reason*. Translated/edited by P. Guyer and A. Wood, Cambridge: Cambridge University Press, 1997.
[11] M. M. Kokar, M. D. Bedworth, and C. B. Frankel. *Reference Model for data fusion Systems* // AeroSense 2000, pp. 191202, 2000.
[12] M. Kokar, J. Tomasik, and J. Weyman. *Formalizing classes of information fusion systems* // Information Fusion. 5(3), 2004. pp. 189-202.
[13] S. Kullback, R.A. Leibler. *On information and sufficiency* // The Annals of Mathematical Statistics. 22(1), 1951. pp. 79-86.
[14] H.B. Mitchell. *Data Fusion: Concepts and Ideas.* Springer-Verlag Berlin Heidelberg, 2012. pp. 346.
[15] E. Schrödinger. *What is life? The Physical Aspect of the Living Cell.* Cambridge University Press, 1944. pp. 91.
[16] C. Shannon. *A Mathematical Theory of Communication* // Bell System Technical Journal. 27(3). pp. 379-423.
[17] Steinberg, A., Bowman, C. *Revisions to the JDL Data Fusion Model* in *Handbook of Multisensor Data Fusion*, ed. by J. Llinas, Boca Raton, FL: CRC press, 2001.
[18] Steinberg, A.N., Bowman, C.L., and White, Jr., F.E. *Revisions to the JDL Data Fusion Model* // Proc. 3rd NATO/IRIS Conf., Quebec City, Canada, 1998.