

Information granule system induced by a perceptual system

Anna Bryniarska Opole University of Technology Institute of Computer Science ul. Proszkowska 76, 45-758 Opole, Poland Email: a.bryniarska@po.opole.pl

Abstract-Knowledge represented in the semantic network, especially in the Semantic Web, can be expressed in attributive language AL. Expressions of this language are interpreted in different theories of information granules: set theory, probability theory, possible data sets in the evidence systems, shadowed sets, fuzzy sets or rough sets. In order to unify the interpretations of expressions for different theories, it is assumed that expressions of the AL language can be interpreted in a chosen relational system called a granule system. In this paper, it is proposed to use information granule database and it is also demonstrated that this database can be induced by the measurement system of the adequacy of information retrieval, called a perceptual system. It can simplify previous formal description of the information granule system significantly. This paper also shows some examples of inducing rough and fuzzy granule databases by some perceptual systems.

I. INTRODUCTION

T IS intuitively assumed that conceiving of information, represented by descriptions of something, is to discern, distinguish, and indentify this thing. Conceiving of information about an object is preceded by the perception of the description of this object. The perception consists of a degree of compliance between certain information resources about the object and precisely determined knowledge represented in the set of object descriptions called the thesaurus [9], [24]. Thus, object perception determines the weight, rank, and importance of object descriptions representing information about this object. This also applies to sources of information about objects, pointing to these objects, called in the computer science entities, i.e. such signs of these objects, which are different from their descriptions. Each such reference is called the information granule [24], [25] and its instance is called data about the object that this information is concerned with. The description of the information granule indicates what this information is about. In the Web, any description, and thus the description of the information granule, has the address of information in the memory of computers connected to this network. This address indicates the sign for human of what the information relates to, including a specific description of the object. These are, for example, natural language expressions describing these objects, data representative about these objects, their image or their sound characteristics. Granules are grouped into granule systems in which granular calculations are made, i.e. the information about objects is interpreted. For

well-established knowledge, granules are *data sets*. When, for knowledge representation, incorrect classification of objects is used, i.e. knowledge about them is probable, uncertain, unclear, or vogue, then information granules can not be described by abstract data sets. In such situation, to determine the information granules it is proposed to use the following nonstandard formalisms of the set theory [22]: interval analysis [14], fuzzy sets [29], [30], [31], rough sets [15], [16], [17], [18], [19] and shadowed sets [20], [21], [22]. In the papers mentioned above, as well as in other papers on granular calculations, there is a lack of uniform methods of using abstraction in order to interpret the expressions of the attributive language *AL* in the information granule theories.

It should be noted that in computer science, from the beginning of its existence, abstraction has been used to reduce the complexity of the problem and achieve greater transparency [27]. The elementary form of abstraction introduces a distinction between the level of a concrete (instance, instance of data) and its type. With abstraction, the class type of similar concretes can be specified. The lowest level of abstraction is one that does not require skipping (abstaining from) significant differences between objects. The abstraction model is an abstract set of data about objects, described in the terminology of the set theory in the Cantor sense. At present, the sets are understood as such abstractions in which the formal language of the Zermelo-Frenkel ZF set theory can be interpreted, e.g. sets of decidable data strings in the alternative Vopenka theory of sets [28], such as: extended sets [3] and multi sets [2]. For these sets, formal axiomatic set theories have been built, i.e. precise descriptions of these sets.

For any abstraction there is a relational structure called **a** granule system. In this system defines: the set of the universe elements of this structure, the atomic sets determined by the universe elements (singletons), relationships like membership of elements to sets, sets conclusion, sets equality. Furthermore, any set is a sum of atomic sets and analogously to standard one, all operations on sets are specified. In any granule system, the *ZF* set theory language is interpreted. However, not all axioms (except from the axioms of the granule system) and not all *ZF* theorems must be met in this system. The granule system, at a higher level of abstraction, specifies knowledge that is inaccurate, uncertain or unclear at a lower level of abstraction. For fifty years, for fuzzy sets, no such system has

been defined and no formal theory has been built. In 1981, Pawlak [15] described the approximated set at a higher level of abstraction as an abstract class of the relation of the equal approximation of sets and proposed to build an axiomatic rough set theory. A solution to this problem is proposed in the Bryniarski's papers [10], [11].

Recently, *the information retrieval IR* in the semantic network, especially in the Semantic Web, usually means looking for a reliable source of this information. So far, information retrieval systems and information interpretations have only indicated semantically the nearest, described in the thesaurus, sources of searched information. However, this is not always the case. Often, when searching for information about an object in a language deviating from the thesaurus, uncertain, unclear or inaccurate knowledge is obtained. Nevertheless, this uncertainty may lead to the unequivocal establishment of sources of knowledge about this object, i.e. precise knowledge. In this way, compliance with the description of the object model is obtained (compatibility with the thesaurus). The situation described above is called **the information disambiguation paradox** of information retrieval [5].

Searching information in the Semantic Web is to find data copies which are:

- one-argument values of attributes data representing knowledge about some features or types of objects,
- two-argument values of attributes data representing knowledge about some properties of objects or relations between two objects..

In first case, data are called *concepts*, and in the second one they are called *roles*. To describe concepts and roles, *the Description Logic (DL)* language [1], [4] is used. The *DL* language describing concepts and roles can be extended to some formulas of the first order logic. In the extended language, *a thesaurus* is created, which describes model concepts and roles, while *the ontology* is a language which describes searched concepts and roles. For the searched data described in the ontology and the recommendations (criteria and knowledge) of experts, there may be a certain degree of compliance of these data with the data described in the thesaurus. This is the assessment of the compatibility of data with the thesaurus accepted by experts. The conceiving rule determining the paradox of accuracy appearing here is called **the residuum rule** [5].

This paper presents a perception model of descriptions representing information in semantic networks. In this model, accepted methods to the description perceptions are used, in order to use the residuum rule. It is the perception of references information resources about the object to the degree of compliance of this information with the precisely determined knowledge represented in the set of object descriptions called the thesaurus. Such perceptual system for descriptions will be called **the residuum system**.

The model of information granule systems represented in semantic networks was formulated at the syntactic and semantic level in the papers [7], [9]. Continuing this research, only the method of inducing information granules by the residuum system will be presented in this paper.

In this paper firstly is presented the semantic network and the perception in the residuum systems in this network. Further is definition of the information granule database and its extension to the information granule system. The perceptual system defined by the information granule database is novel in this paper. At the end there are two examples of such system – rough and fuzzy one.

II. THE SEMANTIC NETWORK

The semantic network, or the Semantic Web, most commonly is considered to be a graph schema of knowledge representation. It can be identified with an ordered, indexed graph. In the semantic network the vertices and edges are described by some attributes: one or two-argument. In this paper, a more general graphical scheme of knowledge representation is given in which edges can have more than two vertices [7], [8], [9].

Definition 2.1: The semantic network is a system:

$$\mathbf{SN} = \langle U, AS, DS \rangle,\tag{1}$$

where:

- U is a finite set of individual names, object names of represented knowledge (in the Semantic Web it is a set of names which have the Web address). Elements of U are called vertices of the semantic network.
- DS is a family of nonempty sets of vertices descriptions, and also certain systems of these vertices, called **edges**. The number n is the largest number of vertices in edges.

Let card(U) = n and $U_{qen} = U \cup U^2 \cup ... \cup U^n$. Then:

$$AS \subseteq DS \times U_{gen},\tag{2}$$

Elements of the set AS are called **assertions**. AS includes all vertices U:

 $U = \{x : \text{ exists } (ds_k, (x_1, ..., x_i..., x_k)) \in AS, x = x_i\}$ (3)

When $ds \in DS$, then exists a set $X(ds) \subseteq U_{qen}$, such that:

$$\{ds\} \times X(ds) = (\{ds\} \times U_{gen}) \cap AS.$$
(4)

Such set and its any subset X is called a subject X with description ds (shortly: **subject**), and pair $\langle ds, X \rangle$ is called a conceiving subject of description ds. The subject X with description ds will be identified with the conceiving subject $\langle ds, X \rangle$ of description ds.

For example, let the conceiving subject of description ds be $\langle ds, X \rangle$, where R is k-th argument relation such that:

$$\{ds\} \times R = (\{ds\} \times U^k) \cap X(ds).$$
(5)

The set X(ds) is a relation or sum of relations R defined as above for any number of arguments. About subject X(ds), it is said that, it is *a maximum subject* of the description ds, and about the description ds, that it is *an instance of the subject* $X \subseteq X(ds)$.

SN is called *full*, if sum of all such sets is equal U_{qen} .

Any element of U_{gen} is called *an instance* of the **SN**, and when this element belongs to some subject X with some description, it is called *an instance occurrence in the* **SN** *network*.

Sets $\{ds\} \times X(ds) = (\{ds\} \times U_{gen}) \cap AS$ are called *attributes* of subjects $X \subseteq X(ds)$ with descriptions ds, and these descriptions will be identified with the instance of this attribute. The family of all such subjects X(ds) is denoted by C_0 . This set is sometimes a division of U_{gen} set. The elements of X(ds) are called *an instance occurrence of attribute ds*, and elements of AS are *assertions*. The following notary agreement is accepted:

For attribute ds the instance occurrence set about this attribute is denoted as:

- $||ds|| =_{df} X(ds).$
- the assertion occurrence (ds, u) ∈ AS, is denoted as u:
 ds, eg. instead of '6 < 9' it will be written '(6,9) :<'.

One-argument relations will be called *concepts*, and at least two-argument relations will be called *roles*, i.e. concepts and roles are subjects of conceiving certain descriptions.

Due to the fact that for any X subject, uniquely designated by the attribute $\{ds\} \times X$, the description ds corresponds only to one relationship $R \subseteq X$, $\{ds\} \times R = (\{ds\} \times U^k) \cap X$ with a given number k of arguments. Therefore, in the further part of this paper, with a fixed number of arguments of these relations, *concepts and roles will be identified with the corresponding descriptions*.

Occurrences of instances with some attribute, occurrence of attributes, subjects with this attribute, concepts, roles and assertions are described in attributive language AL. Basic syntax and semantic of AL language are formulated in the paper [1], and the generalized construction of this language is presented in papers [4], [5], [6], [7], [8], [9].

For example, a role *sonhood* connecting a person named *John* with a person named *Simon*, who is his father, leads to assertion: $\langle sonhood, John, Simon \rangle$, what can be denoted as: *sonhood(John, Simon)* or (*John,Simon)*: *sonhood*.

To join the concept *sonhood* with the time *current* year, we need two assertions (sonhood, John, Simon) and $(sonhood, current_year)$, what can be written as a set of descriptions $\{(John, Simon): sonhood, (current_year): sonhood\}$.

An assertion which is expressed in a sentence *Eva* sits between John and Simon can be denoted as: sit_between(Eva,John, Simon) or (Eva, John, Simon): sit_between. Roles which are functions, in terms of the last component, are called **operations**, for example in the assertion drive_to (John, New York).

It is significant to notice that in a triple (*Eva, John, Simon*) the cyclic inverse of names can be used, and then the following triple is created: (*John, Simon, Eva*), which is also an instance of some role. This new assertion can be expressed in a sentence *John and Simon sit next to Eva* and can be denoted as: (*John, Simon, Eva*): *sit_nextto*. The role *sit_nextto* is **cyclically inverse** to the role *sit_between*.

When a triple (Eva, John, Simon), which is an occurrence of an assertion $sit_between$, is reduced by the first component, then a pair (John, Simon) is also an instance of some assertion, for example expressed in a sentence: *someone sits between John and Simon* – (John, Simon): *someone_sits_between*. This role is called **a reduction** of a role *sit_between*.

Distinguished by experts subsystem of SN is denoted as SN^+ in which concepts and roles are considered to be accurate – experts have confidence in this knowledge. $SN^+ = \langle U^+, AS^+, DS^+ \rangle$ is called a **confidence range** for the SN. In the confidence range is $U_{gen}^+ = U^+ \cup (U^+)^2 \cup ... \cup (U^+)^n$. The set $SN_{tez} = DS^+ \cup SN_{inst}^+$ of all attribute descriptions and instances of these attributes in the SN^+ is called a **thesaurus** of the semantic network SN [8], [9].

III. DESCRIPTION PERCEPTION IN THE RESIDUUM SYSTEMS

In this paper, some aspects of the perceptual proximity theory are used in the context of the proximity of knowledge searched in the semantic Web to the adequate knowledge represented in the thesaurus. A certain view of nearness perception is accepted, combining the basic understanding of perception in psychophysics with the view of the perception described in the Merleau-Ponty paper [13]. This means that the perception of nearness of knowledge about reality to adequate knowledge – and as a result to human knowledge about objects – depends on the signals of sensors, i.e. signals of the senses or measuring systems [12].

But it is known that these signals from measuring systems are received by our senses, and then, as descriptions of objects, are analyzed in the mind. In this approach, our senses are compared to the sampling function. They mimic the impressions describing the features on the numerical values recognized by the mind. Human sensors (senses) collect data samples and measure the physical characteristics of objects in our environment. The physical properties of the object that are read are described and identified with the features of the object. It is our mind that identifies the relations between the values of the features of the object, creating the perception of the detected objects [13]. Object perception is a measure of the adequacy of the information resource that defines this object. As it was written earlier, such a measurement of the adequacy of the information resource will be hereinafter referred to as a granule of information. In this sense, by searching for information about certain objects in the Semantic Web, the perception of object descriptions and descriptions representing information about these objects are made. In this way a certain set of information granules is obtained [8], [9].

Thus, the definition of algorithms for granular calculations should begin with the definition that is a part of the definition of *the perceptual system*.

A. The residuum system

Definition 3.1: The system

$$\mathbf{S}_{\mathbf{P}} = \langle S_P, \bullet_P, \to_P, 0_P, 1_P \rangle \tag{6}$$

is a residuum system, where a set S_P is called the set of perception values, and it includes two different elements $0_P, 1_P$, called values of *truth* and *false*. An operation $\bullet_P: S_P \times S_P \to S_P$ is called the operation of perception *combination*, an operation $\rightarrow_P: S_P \times S_P \to S_P$ is called the operation of perception residuum.

Operations of the residuum system satisfy following conditions (for any $z \in S_P$):

if
$$z \neq 0_P$$
, then $(0_P \to_P z) = 1_P, (z \to_P 0_P) = 0_P$. (7)

In addition, there is such an operation $\sum_P : \wp(S_P) \rightarrow S_P, \wp(S_P) = \{X : X \subseteq S_P\}$, called *the generalized* combination of perception, such that for any $x \in S_P$ and for any nonempty disjoint sets $A, B \subseteq S_P$:

$$\sum_{P} \emptyset = 0_P, \tag{8}$$

$$\sum_{P} \{x\} = x, \tag{9}$$

$$\sum_{P} (A \cup B) = \sum_{P} (A) \bullet_{P} \sum_{P} (B), \tag{10}$$

Hence, for any $x, y \in S_P$:

$$\sum_{P} \{x, y\} = (x \bullet_P y). \tag{11}$$

A differentiated operation $(.^d)$: $S_P \rightarrow S_P$, such that $(0_P)^d = 1_P$ and $(1_P)^d = 0_P$, $(x^d)^d = x$, is called the dual value operation in the system $\mathbf{S}_{\mathbf{P}}$, such that if $x, y \in S_P$ and x < y, then $y^d < x^d$.

- If $(z \rightarrow_P z) = 1_P$, then the residuum operation is called the t-residuum.
- If $(z \rightarrow_P z) = 0_P$, then the residuum operation is called the s-residuum.

With the above definitions results:

Fact 3.1: Perception combination is a commutative and associative operation.

Let in the residuum system $\mathbf{S}_{\mathbf{P}} = \langle S_P, \bullet_P, \to_P, 0_P, 1_P \rangle$, for the operation $(.^d)$ of dual value in the system S_P , exist such operation $\sum_{P'} : \wp(S_P) \to S_P$, that (for any $x \in S_P$ and $A, B \subseteq S_P$):

$$\sum_{P'} \emptyset = 0_P, \tag{12}$$

$$\sum_{P'} \{x\} = x,$$
 (13)
($\sum_{P'} (A)^d = \sum_{P'} (D)^d A^d$ (14)

$$\sum_{P'} (A \cup B) = (\sum_{P'} (A)^d \bullet_P \sum_{P'} (B)^d)^d, \qquad (14)$$

Then, for any numbers $x, y \in S_P$ it is assumed that:

$$x \bullet_{P'} y =_{df} (x^d \bullet_P y^d)^d, \tag{15}$$

for any nonempty, disjoint sets $A, B \subseteq S_P$:

$$\sum_{P'} (A \cup B) = \sum_{P'} (A) \bullet_{P'} \sum_{P'} (B).$$
(16)

Let $x \to_{P'} y =_{df} (x^d \to_P y^d)^d$. If $z \neq 0_{P'}$, then $(0_P \to_{P'} z) = 1_P, (z \to_{P'} 0_P) = 0_P$. Therefore:

Fact 3.2: The algebra system $\mathbf{S}_{\mathbf{P}'} = \langle S_P, \bullet_{P'}, \rightarrow_{P'}, \rangle$ $0_P, 1_P$ is the residuum system.

Definition 3.2: The residuum system

 $\mathbf{S}_{\mathbf{P}'} = \langle S_P, \bullet_{P'}, \to_{P'}, 0_P, 1_P \rangle$ is called the dual system for the S_P system.

Fact 3.3: The operation $\rightarrow_{P'}$ in the residuum system $\mathbf{S}_{\mathbf{P}}$ is the s-residuum operation.

B. T-norm and s-norm systems in the partially ordered set

Theorem 3.1: The algebra system $\mathbf{S}_{\mathbf{t}} = \langle L, \bullet_t, \rightarrow_t, 0_L, 1_L \rangle$, is called **the t-norm system** in the set L partially ordered by the relation \leq , in which any subset has infimum and supremum, where $0_L = \inf L$, and $1_L = \sup L$. It is the residuum system, if the operation $\bullet_t : L \times L \to L$, called the t-norm in the L, satisfies following conditions (for any numbers $w, x, y, z \in L$):

boundary conditions

$$0_L \bullet_t y = 0_L, y \bullet_t 1_L = y \tag{17}$$

• uniform value increase, monotonicity

$$x \bullet_t y \le z \bullet_t y$$
, when $x \le z$ (18)

• uniform value limitation

$$w \le x \bullet_t y \le z$$
, when $w \le x \le z$ or $w \le y \le z$ (19)

commutativity

$$x \bullet_t y = y \bullet_t x \tag{20}$$

associativity

$$x \bullet_t (y \bullet_t z) = (x \bullet_t y) \bullet_t z \tag{21}$$

• and

exist
$$x \to_t y = \sup\{t \in L : x \bullet_t t \le y\}.$$
 (22)

Such described operation $\rightarrow_t : L \times L \rightarrow L$ is the t-residuum in the set L.

Let operation $d: L \to L$ be an operation of the dual values in the \mathbf{S}_t system, then the system $\mathbf{S}_s = \langle L, \bullet_s, \rightarrow_s, 0_L, 1_L \rangle$ is defined as follows (for any numbers $x, y \in L$):

$$x \bullet_s y = (x^d \bullet_t y^d)^d, \tag{23}$$

$$x \to_s y = (x^d \to_t y^d)^d, \tag{24}$$

Then:

Theorem 3.2: In the system $\mathbf{S}_{\mathbf{s}} = \langle L, \bullet_s, \rightarrow_s, 0_L, 1_L \rangle$, the following conditions are satisfied (for any numbers $w, x, y, z \in$ L):

• boundary conditions

$$0_L \bullet_s y = y, y \bullet_s 1_L = 1_L \tag{25}$$

• uniform value increase, monotonicity

$$x \bullet_s y \le z \bullet_s y$$
, when $x \le z$ (26)

• uniform value limitation

$$w \le x \bullet_s y \le z$$
, when $w \le x \le z$ or $w \le y \le z$ (27)

commutativity

$$x \bullet_s y = y \bullet_s x \tag{28}$$

associativity

$$x \bullet_s (y \bullet_s z) = (x \bullet_s y) \bullet_s z \tag{29}$$

$$x \to_s y = \inf\{t \in L : y \le x \bullet_s t\}.$$
 (30)

Definition 3.3: The system $\mathbf{S}_{\mathbf{s}} = \langle L, \bullet_s, \rightarrow_s, 0_L, 1_L \rangle$ is called **the s-norm system** in the set *L* partially ordered by the relation \leq , and the operation $\bullet_s : L \times L \to L$ is called *the s-norm* in the *L* for the operation $(.^d)$ of dual values in the system $\mathbf{S}_{\mathbf{t}}$.

Example 3.1: Let range of numbers L = [0, 1] be ordered by the relation \leq . The operation $\bullet_t : [0, 1] \times [0, 1] \rightarrow [0, 1]$, for any $x, y \in [0, 1]$, is the t-norm and is defined by formula:

$$x \bullet_t y = \inf\{x, y\} = \min\{x, y\}.$$
 (31)

Its generalized form determines the formula, for any set $X \subseteq [0, 1]$

$$\sum_{t} X = \inf X. \tag{32}$$

It can be determined that:

$$x \to_t y = \sup\{t \in [0,1] : \min\{x,t\} \le y\}.$$
 (33)

Example 3.2: Let range of numbers L = [0, 1] be ordered by the relation \leq . The operation $\bullet_s : [0, 1] \times [0, 1] \rightarrow [0, 1]$ for any $x, y \in [0, 1]$, is the s-norm and is defined by formula:

$$x \bullet_s y = \sup\{x, y\} = \max\{x, y\}$$
(34)

Its generalized form determines the formula, for any set $X \subseteq [0, 1]$:

$$\sum_{s} X = \sup X. \tag{35}$$

It can be determined that:

$$\begin{aligned} x \to_s y &= 1 - (1 - x) \to_t (1 - y) = \\ &= 1 - \sup\{t \in [0, 1] : \min\{1 - x, t\} \le 1 - y\} = \\ &= \inf\{t \in [0, 1] : y \le x \bullet_s t\}. \end{aligned}$$
(36)

Having t-norm and s-norm systems can be determined:

Definition 3.4: The system $\mathbf{S}_{\text{logic}} = \langle L, \bullet_s, \rightarrow_t, 0_L, 1_L \rangle$ is called **the logical residuum system**.

IV. THE INFORMATION GRANULE DATABASE

Firstly, the definition of the perceptual system is given in order to define the information granule database. Then, it is shown how this database is induced by this system.

A. The perceptual system

Definition 4.1: For the semantic network $SN = \langle U, AS, DS \rangle$, a perceptual object is an instance which was given a certain value in the residuum system $S_{logic} = \langle L, \bullet_s, \rightarrow_t, 0_L, 1_L \rangle$, i.e. perceptual objects are elements of some set $O = U_{gen}$, where the set $L_0 \subseteq L$ is a set of values of instances in the residuum system S_{logic} .

Giving certain values in the S_{logic} residuum system also has some interpretation in S_{logic} . This interpretation is defined by the definition: Definition 4.2: A probe function or a perception is a function that $\phi : O \to L_0$ represents a feature of a perceptual object [23], [26].

Further, for the **SN**, it is assumed that any description $\phi \in DS$ corresponds to a certain perception $\phi : O \to L_0$ determined by this description ϕ .

Extending the perceptual system definition [26], for the concept of the logical residuum system S_{logic} , it is assumed that:

Definition 4.3: A perceptual system $\mathbf{PS} = \langle O, F, \mathbf{S}_{logic} \rangle$ consists of a nonempty set O of sample perceptual objects and the set F of chosen perceptions $\phi : O \to L_0$, called *the perceptions of the* **PS** system. Elements of the L_0 are called then *the perception degrees*.

B. The granule information database induced by the perceptual system

Definition 4.4: The system:

If

y

$$\mathbf{G}_{\mathbf{base}} = \langle G, \{\}_G, \cup_G, \subseteq_G, =_G, \\ 0_G, 1_G, G_{inst}, G_{set}, G_0 \rangle, \quad (37)$$

is called **the granule information database**, where elements of the set G are called granules, G_{inst} is a set of *instance* granules, G_{set} is a set of set granules and G_0 is a set of singletons of granules, a granule 0_G is called an empty granule and a granule 1_G is called a full granule, the operations $\{\}_G, \cup_G$ and the relations $\subseteq_G, =_G$ are defined by following conditions:

$$G_0 \subseteq G_{set},\tag{38}$$

$$G = G_{inst} \cup G_{set},\tag{39}$$

$$G_{inst} \cap G_{set} = \emptyset. \tag{40}$$

There is some function $\{\}_G : G_{inst} \to G_0$, such that

for any
$$x \in G_{inst}, \{x\}_G \subseteq_G 1_G$$
, additionally, if
 $\{x_0\}_G \subseteq_G \{x\}_G$, then $\{x_0\}_G =_G \{x\}_G$, (41)

$$G_0 = \{ x \in G_{set} : \exists z \in G_{inst} (x = \{z\}_G) \}.$$
(42)

There is some function $\cup_G : \wp(G_{set}) \to G_{set}$, such that

$$\cup_G \varnothing = 0_G, \tag{43}$$

$$\cup_G G_{set} = 1_G,\tag{44}$$

$$\cup_G \{z\} = z, \text{ for } z \in G_{set}, \tag{45}$$

$$x \subseteq_G x, \tag{46}$$

$$0_G \subseteq_G x, \tag{47}$$

$$x \neq 0_G$$
, then, it is not true that $x \subseteq_G 0_G$, (48)

$$= \bigcup_G \{ x \in G_0 : x \subseteq_G y \}, \tag{49}$$

for any
$$x, y \in G_{set}, (x =_G y) \Leftrightarrow_{df}$$

 $\cup_G \{z \in G_0 : (z \subseteq_G x) \Leftrightarrow (z \subseteq_G y)\} = 1_G.$ (50)

It is assumed, in the sense of the set theory, that the $Inst : U_{qen} \rightarrow G_{inst}$ function assigns a certain instance

granule to each instance. Any such function allows to enter a notation agreement for the granule $Inst(\langle v_1, v_2, ..., v_k \rangle)$, when $\langle v_1, v_2, ..., v_k \rangle \in U^k$:

$$(Inst)\langle x_1, x_2, ..., x_k \rangle = Inst(\langle v_1, v_2, ..., v_k \rangle)$$

iff $\langle v_1, v_2, ..., v_k \rangle \in U^k$ and
 $x_i = Inst(v_i), \text{ for } i = 2, ..., k.$ (51)

Theorem 4.1: Let, for any perceptual system $\mathbf{PS} = \langle O, F, \mathbf{S}_{logic} \rangle$, $\mathbf{S}_{logic} = \langle L, \bullet_s, \rightarrow_t, 0_L, 1_L \rangle$, symbols $G, \{\}_G, \cup_G, \subseteq_G, =_G, 0_G, 1_G, G_{inst}, G_{set}, G_0$ be interpreted as follows (for any $\phi_1, \phi_2 \in G_{set}, \phi \in G_0, C_0$ is a family of maximum subjects in the semantic network \mathbf{SN}):

$$G_{inst} = U_{gen} \times L_0 \cup C_0, \tag{52}$$

where $C_0 \subseteq \wp(U_{gen})$ and $L_0 = \{r : \text{exists } a \in U_{gen} \text{ and } exists \phi \in F \text{ such that } r = \phi(a)\},\$

$$G_{set} = F, (53)$$

$$G = G_{inst} \cup G_{set},\tag{54}$$

$$\phi_1 \subseteq_G \phi_2$$
 iff for any $x \in L, \phi_1(x) \to_t \phi_2(x) = 1_L$, (55)

$$\{(a,r)\}_G = \mu \in F,\tag{56}$$

where μ is such perception that for $(a, r) \in G_{inst}, \mu(a) = r \neq 0_L$, and $\mu \subseteq_G 1_G$, additionally, if $\mu_0 \in F$ and $\mu_0 \subseteq_G \mu$, then $\mu_0 = \mu$,

$$\{K\}_G = \mu \in F,\tag{57}$$

where μ is such perception that for $K \in C_0, \mu(a) = 1_L$ for any $a \in K, \mu(a) \neq 0_L$ for any $a \notin K$ and $\mu \subseteq_G 1_G$, additionally, if $\mu_0 \in F$ and $\mu_0 \subseteq_G \mu$, then $\mu_0 = \mu$,

$$G_0 = \{ \phi \in F : \text{ exists } u \in G_{inst},$$

such that $\phi = \{u\}_G \}.$ (58)

Additionally, if $A \subseteq F$, then:

$$(\cup_G A)(x) = \sum_P \{ y \in L : y = \phi(x) \land \phi \in A \},$$
 (59)

$$(\phi_1 =_G \phi_2) \Leftrightarrow_{df} \cup_G \{ \phi \in G_0 : (\phi \subseteq_G \phi_1) \Leftrightarrow \\ \Leftrightarrow (\phi \subseteq_G \phi_2) \} = 1_G.$$
 (60)

Then, the system $\langle G, \{\}_G, \cup_G, \subseteq_G, =_G, 0_G, 1_G, G_{inst}, G_{set}, G_0 \rangle$ is the granule information database.

Definition 4.5: The granule database described in above theorem is called *the granule information database induced* by the perceptual system **PS**.

V. EXTENDING THE INFORMATION GRANULE DATABASE TO THE INFORMATION GRANULE SYSTEM

Definition 5.1: For any $x \in G_{inst}$ and $y \in G_{set}$,

$$x \in_G y \text{ iff } \{x\}_G \subseteq_G y. \tag{61}$$

The relation \in_G is called the relation of belonging instance granule to the granule set.

Hence, and from the conditions describing the information granule database:

Theorem 5.1:

ł

$$\{x \in G_{inst} : x \in_G 0_G\} = \emptyset, \tag{62}$$

$$\{x \in G_{inst} : \exists z \in G_0(x \in_G z)\} = G_{inst}, \tag{63}$$

$$\{z \in G_{set} : \exists x \in G_{inst}(x \in_G z)\} = G_{set}, \tag{64}$$

$$y = \bigcup_G \{ \{x\}_G : x \in_G y \}, \tag{65}$$

Theorem 5.2: Granules y'^G , $y \cap_G z$, $y \cup_G z$, and $y \setminus_G z$ exist in the granule database, and are defined by formulas:

$$y'^G = \bigcup_G \{ z \in G_0 : \exists x \in G_{inst} (x \in_G z \land \neg x \in_G y) \},$$
(66)

$$y \cap_G z = \bigcap_G \{y, z\} = \bigcup_G \{t \in G_0 :$$

$$\exists x \in G_{inst} (x \in_G t \land x \in_G y \land x \in_G z)\}, \quad (67)$$

$$y \cup_G z = \bigcup_G \{y, z\} = \bigcup_G \{t \in G_0 :$$

$$\exists x \in G_{inst} (x \in_G t \land (x \in_G y \lor x \in_G z))\}, \quad (68)$$

$$y \setminus_G z = \bigcup_G \{t \in G_0 : \\ \exists x \in G_{inst} (x \in_G t \land x \in_G y \land \neg x \in_G z) \}.$$
(69)

In order to unify the expressions of the attributive language AL in various theories of information granules, formulated in theories: set theory, probability theory, possible data sets in the evidence systems, fuzzy set theory, rough sets theory and shadowed sets theory, the expressions of the attributive language are assumed to be interpreted in a chosen relational system **G** [8], [9], [11] given below. A distinction is made between the set of *granule instances* and the set of *granule set instances*. In the first set, attribute instances are interpreted and in the second - set of instances (concepts and roles). The important thing is that *the granule set instances determine sets of instances*. Granule instances are interpreted as elements of granule set instances, analogically to some classical **G**⁺ algebra.

Definition 5.2: Let **the granule system** for the *attributive language* be:

$$\mathbf{G} = \langle G, M_G, \cup_G, \cap_G, \setminus_G, \prime_G, \in_G, \subseteq_G, \\ =_G, 0_G, 1_G, G_{inst}, G_{set}, G_0 \rangle.$$
(70)

where:

• $G = G_{inst} \cup G_{set}$ is a sum of sets: G_{inst} – is a set of granule instances and G_{set} is a set of granule set instances,

• M_G is a set of functions $m_G: G \to G$,

- operations ∪_G, ∩_G are generalized operations of sum and product described on the subsets of the granules family G,
- \backslash_G is an operation of granules difference,
- $'_G$ is an operation of granules closure,
- for an empty set, a value of these generalized operations is an empty granule 0_G and for the G set it is a full granule 1_G ,

- \in_G is a relation of being a granule element,
- \subseteq_G is a relation of granules inclusion for instance set granules,
- $=_G$ is a relation of granules closeness for instance set granules,
- G₀ is a set of chosen granules.

 G_0 is a set of granules called granules of instances singletons such that there is some function $\{\}_G : G_{inst} \to G_0$ and $G_0 = \{x \in G : \exists z \in G_{inst} (x = \{z\}_G)\}.$

VI. EXAMPLES OF THE GRANULE SYSTEM

Let consider two examples of the granule system the rough and the fuzzy granule databases.

A. The rough granule database

Let in the system $\langle U_{gen}, C \rangle$, where C is a partition of U_{gen} , operations be defined (for any $X \subseteq U_{gen}$):

$$C^{-}(X) = \bigcup \{ K \in C : K \subseteq X \}, \tag{71}$$

$$C^+(X) = \bigcup \{ K \in C : K \cap X \neq \emptyset \}, \tag{72}$$

Any sets $X, Y \subseteq U_{gen}$ are indiscernibility, what is written: $X \sim Y$ iff

$$C^{-}(X) = C^{-}(Y),$$
 (73)

$$C^+(X) = C^+(Y).$$
 (74)

The relation \sim is a relation of equivalence. The abstraction classes $[X]_{\sim}$ of this relation for the representative X is denoted as X_C . \varnothing_C is denoted by 0_C and $(U_{gen})_C$ is denoted by 1_C .

The abstract classes of the relations \sim are called **rough sets** in the system $\langle U_{gen}, C \rangle$.

The set inclusion relations and the rough membership relation [10] is defined as follows (for any $X, Y \subseteq U_{gen}$):

$$X \subseteq_C Y_C \text{ iff } C^-(X) \subseteq C^-(Y)$$

and $C^+(X) \subseteq C^+(Y)$, (75)

$$X \in_C Y_C$$
 iff $X \neq \emptyset$ and exists such $K \in C$,
that $X \subseteq K, C^-(X) \subseteq C^-(Y)$ and $K \subseteq C^+(Y)$. (76)

The expression $X \in_C Y_C$ is read: X is an element of the rough set Y_C . Hence:

$$X \in_C Y_C \text{ iff } X \neq \emptyset \text{ and exists such } K \in C,$$

that $X \subseteq K$ and $X \subseteq_C Y_C.$ (77)

Intuitively, due to the fact that the description of $x \in Y$ cannot be precisely determined, this description is interpreted as follows: indistinguishable from x elements of the indistinguishable elements from the sets Y. With the relation \in_C , the conclusion of rough sets can also be defined. For any $Y \subseteq U_{gen}$,

$$X_C \subseteq_C Y_C$$
 iff for any $Z \subseteq U_{gen}$,
if $Z \in_C X_C$, then $Z \in_C Y_C$. (78)

Using the theorems given by Bryniarski [10], [11], in the family approximate sets, analogically to the classical set theory, the following operations can be defined: the addition \cup_C , the multiplication \cap_C , the difference \setminus_C and the complement ${}^{\prime C}$ of the rough sets.

For any rough sets X_C, Y_C , and any $Z \subseteq U_{gen}$,

$$Z \in_C X_C \cup_C Y_C \text{ iff } Z \in_C X_C \text{ or } Z \in_C Y,$$
(79)

$$Z \in_C X_C \cap_C Y_C \text{ iff } Z \in_C X_C \text{ and } Z \in_C Y,$$
(80)

$$Z \in_C X_C \setminus_C Y_C \text{ iff } Z \in_C X_C \text{ and not } Z \in_C Y, \quad (81)$$

$$Z \in_C (X_C)^{\prime C} \text{ iff } Z \in_C (U_{gen})_C \setminus_C X_C, \qquad (82)$$

Operations \cup_C , \cap_C can be generalized and used in the same way as in the set theory.

If the system:

$$\mathbf{G_{rough}} = \langle G, M_G, \cup_G, \cap_G, \backslash_G, {}^{\prime G}, \{\}_G, \in_G, \subseteq_G, \\ =_G, 0_G, 1_G, G_{inst}, G_{set}, G_0 \rangle, \quad (83)$$

is interpreted as follows:

$$G_{set} =_{df} \{ X_C : X \subseteq U_{gen} \}, \tag{84}$$

$$G_{inst} =_{df} \{ X \subseteq U_{gen} : X = \{ x \} \text{ or } X \in C \},$$
 (85)

$$G =_{df} G_{inst} \cup G_{set}, \tag{86}$$

$$\cup_G =_{df} \cup_C, \tag{87}$$

$$\cap_G =_{df} \cap_C, \tag{88}$$

$$\setminus_G =_{df} \setminus_C, \tag{89}$$

$$C = df \quad (90)$$

$$\in_G = {}_{df} \in_C, \tag{91}$$

$$=_{G} =_{df} =, \tag{93}$$

$$0_G =_{df} \varnothing_C = \{\varnothing\}, \tag{94}$$

$$1_G =_{df} (U_{gen})_C, \tag{95}$$

An operation $\{\}_G : G_{inst} \to G_0$ such that for any $X \in G_{inst}, \{X\}_G = X_C, G_0 = \{X_C : X \in G_{inst}\}$, and the set of operations M_G is empty.

Then the system G_{base} (equation 37 from definition 4.4) is a granule information database, and the elements of G are called **the approximate granules**. Moreover, the system G_{rough} is **the rough granule system**.

B. The fuzzy granule database

Let $\mathbf{PS} = \langle O, F, \mathbf{S}_{logic} \rangle$ be a perceptual system, in which $O = U_{gen}$, and the instance values $L_0 \subseteq [0, 1]$ are in the residuum system $\mathbf{S}_{logic} = \langle [0, 1], \bullet_s, \to_t, \{0, 1\} \rangle$. In the range [0, 1], the s-norm $\bullet_s : [0, 1] \times [0, 1] \to [0, 1]$ is defined by formula (for any $x, y \in [0, 1]$):

$$x \bullet_s y = \sup\{x, y\} = \max\{x, y\}.$$
 (96)

Its generalized form determines the formula (for any set $X \subseteq [0, 1]$):

$$\sum_{s} X = \sup X. \tag{97}$$

It can be determined that:

$$x \to_t y = \sup\{t \in [0,1] : \min\{x,t\} \le y\}.$$
 (98)

F is a set of symbols μ_A of a function $\mu_A : U_{gen} \rightarrow [0, 1]$, and also a symbol of some fuzzy sets [29], [30], [31], described for any $A \subseteq U_{gen}$ as follows (for any $x \in K$) [17]:

$$[x] = K \Leftrightarrow K \in C \text{ and } x \in K, \tag{99}$$

$$\mu_A(x) = |A \cap [x]| / |[x]|, \tag{100}$$

i.e. $F = \{\mu_Y : Y \subseteq U_{gen}\}$, where |X| denotes the cardinality of the X.

Hence the result:

Theorem 6.1: There are (for $A, B \subseteq U_{qen}$):

$$C^{-}(A) = \{ x \in U_{gen} : \mu_A(x) = 1 \},$$
(101)

$$C^{+}(A) = \{ x \in U_{gen} : \mu_A(x) > 0 \},$$
(102)

 $A_C = B_C \Leftrightarrow \text{ for any } x \in U_{gen}, \mu_A(x) = \mu_B(x) \Leftrightarrow \\ \Leftrightarrow \mu_A = \mu_B, \quad (103)$

$$A_C \subseteq_C B_C \Leftrightarrow \text{ for any } x \in U_{gen}, \mu_A(x) \to_t \\ \to_t \mu_B(x) = 1, \quad (104)$$

$$\mu_A(x) \to_t \mu_B(x) = 1 \Leftrightarrow \mu_A(x) \le \mu_B(x), \tag{105}$$

$$A_C \subseteq_C B_C \Leftrightarrow$$
 for any $x \in U_{gen}, \mu_A(x) \leq \mu_B(x)$. (106)

The perceptual system $\mathbf{PS} = \langle O, F, \mathbf{S}_{logic} \rangle$, induces the **fuzzy granule database** \mathbf{G}_{base} (equation 37 from definition 4.4), where symbols $G, \{\}_G, \cup_G, \subseteq_G, =_G, 0_G, 1_G, G_{inst}, G_{set}, G_0$ are interpreted as follows:

$$G_{inst} = U_{gen} \times L_0 \cup C, \tag{107}$$

where C is a partition of U_{gen} and $L_0 = \{r : \text{exisits } u \in U_{gen} \text{ and exists } \mu \in F \text{ such that } r = \mu(u)\},\$

$$G_{set} = F, (108)$$

$$G = G_{inst} \cup G_{set},\tag{109}$$

 $\mu_1 \subseteq_G \mu_2 \text{ iff for any } x \in U_{gen}, \mu_1(x) \to_t \\ \to_t \mu_2(x) = 1, \quad (110)$

$$\{(a,r)\}_G = \mu \in F,$$
 (111)

where μ is such perception that $\mu(a) = r \neq 0$ and $\mu \subseteq_G 1_G$, additionally, if $\mu_0 \in F$ and $\mu_0 \subseteq_G \mu$, then $\mu_0 = \mu$,

$$\{K\}_G = \mu \in F,\tag{112}$$

where μ is such perception that for $K \in C$, $\mu(a) = 1$ for any $a \in K$, $\mu(a) \neq 0$ for any $a \notin K$ and $\mu \subseteq_G 1_G$, additionally, if $\mu_0 \in F$ and $\mu_0 \subseteq_G \mu$, then $\mu_0 = \mu$,

$$G_0 = \{ \phi \in F : \text{ exists } u \in G_{inst},$$

such that $\phi = \{u\}_G \},$ (113)

Additionally, if $A \subseteq F$, then

$$(\cup_G A)(x) = \sup\{y \in [0,1] : y = \mu(x) \land \mu \in A\}, \quad (114)$$

$$(\mu_1 =_G \mu_2) \Leftrightarrow_{df} \cup_G \{ \mu \in G_0 : (\mu \subseteq_G \mu_1) \Leftrightarrow \\ \Leftrightarrow (\mu \subseteq_G \mu_2) \} = 1_G.$$
(115)

VII. CONCLUSION

Presented application of abstraction methods in creating concepts allows to describe and solve more complex problems of knowledge representation in the semantic network, especially in the Semantic Web. Following issues are presented in this paper:

- A semantic network having a more general graph representation of the knowledge representation has been specified, i.e. one in which the edges of the network can have more than two vertices [5], [7], [8]. Roles in attribute language *AL* can join more than two vertices of the network. In addition, all currently used methods of knowledge representation can be implemented in a certain semantic network understood as in this paper.
- The theory of information granule databases in the semantic network has been formulated, in which axioms meet the standard theorems of the set theory defining the concept of a set. The model of this theory is the granule system.
- It has been shown that very complex constructs of the interpretation of the *AL* language expressions in the granule systems [9], at a higher level of abstraction, can be simplified by reducing them to interpretation in the granule databases induced by the perceptual system.
- Due to the fact that any granules are sums of singletons, calculations in granule systems can be simplified by performing them only on certain selected representatives of the elements of these granules. That allows to implement such computational procedures for the most frequently occurring in the processing of knowledge large data sets represented in the Semantic Web.

In further work presented information granule system will be designed also for extended sets, multiset, Borel field of sets and the system of conceiving will be defined.

REFERENCES

- Baader F., Calvanese D., McGuinness D., Nardi D., Patel-Schneider P. (eds.): The Description Logic. Handbook Theory, Implementation and Application. Cambridge University Press, Cambridge, 2003.
- [2] Blizard W, D.: Multiset Theory. Notre Dame Journal of Formal Logic, vol. 30, Number 1, pp. 36-66, 1989.
- [3] Blass, A. C., Childs D. L.: Axioms and Models for an Extended Set Theory. University of Michigan, Mathematics Dept: 2011.
- [4] Bobillo F., Straccia U.: Fuzzy Description Logics with general t-norms and datatypes, Fuzzy Sets Systems, vol. 160(23), pp. 3382-3402, 2009.
- [5] Bryniarska, A.: The Paradox of the Fuzzy Disambiguation in the Information Retrieval. (IJARAI) International Journal of Advanced Research in Artificial Intelligence, pp. 55-58, Volume 2 No 9, 2013.
- [6] Bryniarska A.: The Model of Possible Web Data Retrieval. Proceedings of 2nd IEEE International Conference on Cybernetics CYBCONF 2015, pp. 348-353, 2015.

- [7] Bryniarska A., Bryniarski E.: Rough search of vague knowledge. In: G. Wang, A. Skowron, Y. Yao, D. Slezak, L. Polkowski (eds.), Thriving Rough Sets-10th Anniversary Honoring Professor Zdzisław Pawlak's Life and Legacy & 35 years of Rough Sets, Studies in Computational Intelligence, Springer, Berlin Heidelberg New York, pp.283-310, 2017.
- [8] Bryniarska A.: Autodiagnosis of Information Retrieval on the Web as a Simulation of Selected Processes of Consciousness in the Human Brain. In: Biomedical Engineering and Neuroscience, W. P. Hunek, S. Paszkiel eds., Advances in Intelligent Systems and Computing 720, pp. 111-120, Springer, 2018.
- [9] Bryniarska A.: Certain information granule system as a result of sets approximation by fuzzy context, International Journal of Approximate Reasoning, Volume 111, pp. 1-20, August 2019, in press.
- [10] Bryniarski E.: A calculus of rough sets of the first order. Bull. Pol. Ac.: Math. 37, pp. 109-136, 1989.
- Bryniarski E.: Formal conception of rough sets. Fund. Infor. 27(2-3), pp.103-108, 1996.
- [12] Fahle M., Poggio T.: Perceptual Learning. Cambridge, MA: The MIT Press, 2002.
- [13] Merleau-Ponty M.: Phenomenology of Perception. Paris and New York: Smith, Callimard, Paris and Routledge & Kegan Paul, 1945.
- [14] Moore R.: Interval Analysis, Prentice-Hall, Englewood Clifis, NJ, 1966.[15] Pawlak Z.: Rough sets. Intern. J. Comp. Inform. Sci. 11, pp. 341-356,
- 1982.[16] Pawlak Z.: Rough Sets. Theoretical Aspects of Reasoning about. Data.
- Kluwer Academic Publishers, Dordrecht, 1991.
 [17] Pawlak Z, Skowron A.: Rough membership function, in: R. E Yeager, M. Fedrizzi and J. Kacprzyk (eds.), Advaces in the Dempster-Schafer of Evidence, Wiley, New York, 251-271, 1994.
- [18] Pawlak Z., Skowron A.: Rudiments of rough sets. Information Sciences, 177,1, 1, pp. 3-27, 2007.
- [19] Pawlak Z., Skowron. A.: Rough sets and Boolean reasoning. Information Sciences, 177, 1, pp. 41-73, 2007.

- [20] Pedrycz W.: Shadowed sets: representing and processing fuzzy sets, IEEE Transactions on Systems, Man, and Cybernetics - Part B 28 pp. 103-109, 1998.
- [21] Pedrycz W.: Knowledge-Based Clustering: From Data to Information Granules, J. Wiley, Hoboken, NJ, 2005.
- [22] Pedrycz W.: Allocation of information granularity in optimization and decision-making models: towards building the foundations of Granular Computing. Eur J Oper Res 232(1),pp. 137-145, 2014. doi:10.1016/j.ejor.2012.03.038
- [23] Peters J. F., Skowron A., Stepaniuk J.: Nearness of objects: Extension of approximation space model. Fundamenta Informaticae, vol. 79, no. 3-4, pp. 497–512, 2007.
- [24] Peters J. F.: Discovery of perceptually near information granules. In: Novel Developments in Granular Computing: Applications of Advanced Human Reasoning and Soft Computation, J. T. Yao, Ed. Hersey, N.Y., USA: Information Science Reference, 2009.
- [25] Peters J. F., Ramanna S.: Affinities between perceptual granules: Foundations and Perspectives. In: Human-Centric Information Processing Through Granular Modelling, A. Bargiela and W. Pedrycz, Eds. Berlin: Springer-Verlag, pp. 49–66, 2009.
- [26] Peters J. F., Wasilewski P.: Foundations of near sets. Elsevier Science, vol. 179, no. 1, pp. 3091–3109, 2009.
- [27] Tsichritzis D. C., Lochovsky F.: Data models, Published by Prentice Hall, Inc. Englewood Clis, New Jersey, USA, 1982.
- [28] Vopenka P.: Mathematics in the Alternative Set Theory. Leipzig: Teubner, 1979.
 - [29] Zadeh L.A.: Fuzzy sets. Inf Control 8(3): pp. 338-353, 1965.
 - [30] Zadeh L.A.: Towards a theory of fuzzy information granulation and its centrality in human reasoning and fuzzy logic, Fuzzy Sets and Systems 90, pp. 111-117, 1997.
 - [31] Zadeh L.A.: Toward a generalized theory of uncertainty (GTU) an outline, Information Sciences 172, pp. 1-40, 2005.