Semantic Space: between Continuity and Infinity

Valiaryan Ivashenka
Intelligent Information Department
Belarussian State University
of Informatics and Radioelectronics
Minsk, Belarus
ivashenko@bsuir.by

***Abstract.* The paper describes approaches to unified knowledge representation, providing the formation of semantic space models. On the basis of previously proposed models for the representation of non-numeric structures with denotational and operational semantics, quantitative attributes and approaches to their calculation are proposed. They allow to reveal additional topological and metric properties of semantic space structures. A model of multiversal (protomultiversal) numbers is proposed. This model is a formal basis for clarifying the meaning and revealing semantic properties of network structures and models such as artificial neural networks, Kolmogorov-Arnold networks and other models with operational semantics and using numerical features to solve problems within the semantic space.**

***Keywords:* semantic space, denotational semantics, operational semantics, topological space, metric space, pseudo metrics, canonical form, finite structure, countability, graph entropy, graph tensor multiplication, temporal models**

I. Introduction

One of the attributes of knowledge is the presence of a semantic metric [1]. Thus, if "metric" is understood in the mathematical sense then knowledge-based systems are closely related to the concepts of metric space [2, 3] and topological space [4]. The presence of an additional spatial signature on the knowledge set of an intelligent system is closely related to such another feature of knowledge as scaling [1].

Generalized formal languages [5] are used to represent knowledge in knowledge-based systems.

When representing knowledge, (formal) concepts [6] are formed, which form the basis of ontologies [5, 6]. Each concept can have a communicative designation (name) and a meaning (value) [5].

The approach to modeling the semantic space [7] can be referred to theoretical-synthetic (interior) approaches based on unification [5]. Unification assumes that for a set of designations $N = C \cup I$ (signs $S \subseteq C$, concepts C and names (terms) I), there is a set of meanings (senses) E (Fig. 1). Correspondence between designations and their meanings are defined:

$$v_N \subseteq N \times E . \quad (1)$$

Tolerance of meanings (values) is similarity relation t_N :

$$t_N \overset{def}{=} v_N \circ \left(v_N^{-1}\right). \quad (2)$$

Projection t_N onto the set of concepts C is a relation of equivalence of meanings (values):

$$\sim_C \overset{def}{=} t_N \cap (C \times C). \quad (3)$$

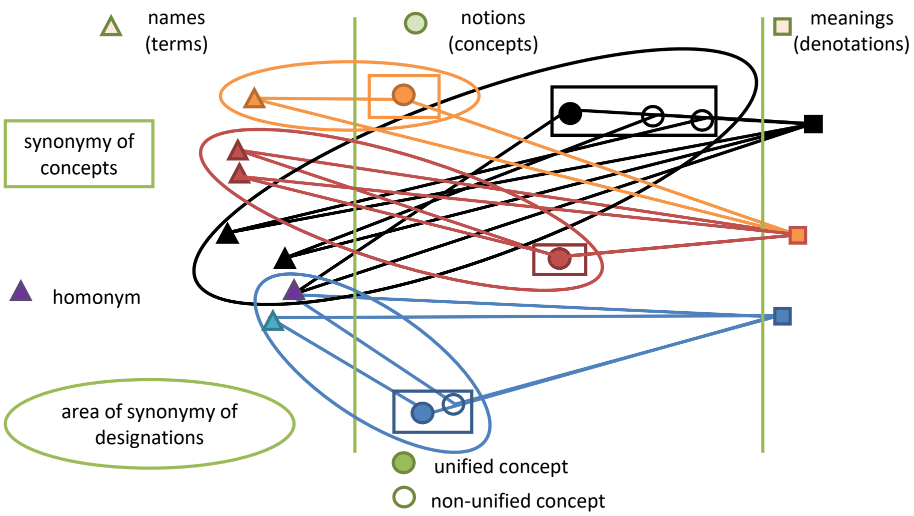

Fig. 1. Unification in the semantic space

The equivalence relation $\sim_C$ forms equivalence classes. There is a bijection between them and signs $S : C_{\sim_C} |\leftrightarrow| S$ with (denotation) $d_N \overset{def}{=} v_N \circ (C \times E)$ being an injection $S |\leftrightarrow E$.

Unification consists in the transition in the relation of similarity $t_N \circ (\sim_C \cap (C \times S))$ from designations N to signs *S*. Semantic normalization in the languages of the knowledge representation model used assumes that primary meanings $F \subseteq E$ are distinguished among the meanings $E \subseteq 2^{N \cup F}$.

II. Semantic space features and models

A. Static ontological structures

For ontological structures of knowledge bases with denotational semantics, the set-theoretic representation (using the membership relation as a basic one) and its generalization [5] are used to represent various (finite) classical and non-classical mathematical substructures [5] and to investigate their spatial properties including topology discrete structures (Fig. 2–4).

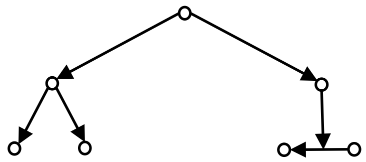

Fig. 2. Ontological structure

In this case, when modeling the semantic space, canonical representations of ontological structures are used. These allow us to identify their canonical spatial features and talk about canonical models of the meaning space. This works in cases when the structures are represented completely.

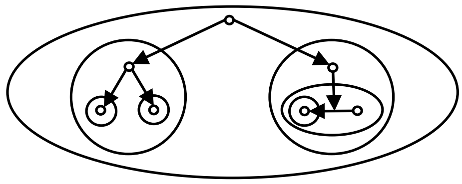

Fig. 3. Ontological structure and extensional closures of its elements

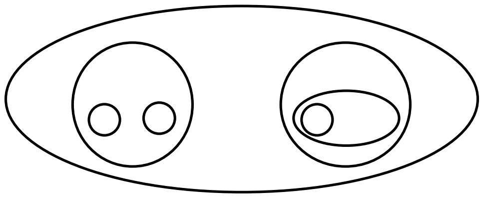

Fig. 4. Closed non-empty sets of the corresponding topological space

The algorithm consists of the following steps:

1. Form closed sets of ontology structure elements based on extensional closure.

2. Order the closures and their elements according to the order relation of the closed subsets, constructing a (topologically sorted) canonical form of the orgraph of this relation, obtaining a directed acyclic metagraph.

3. Starting from the lower vertices (from which there are no paths to other vertices) of the metagraph and up to the upper ones, in the order given by its canonical form, construct and complete the canonical forms of ontological structures included in the corresponding vertices of the metagraph.

In accordance with the canonical forms of ontological structures, canonical matrices and their embeddings can be constructed: adjacency, semantic distances, semantically bounded stable distances (due to the property of knowledge connectivity, the values cannot exceed 8), distance matrices of ontological structure elements in the minimum dimensionality basis.

The embedding of canonical adjacency matrices:

$$\left(\begin{array}{ccccc|ccccc|ccc}
(0) & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & (0) & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
1 & 0 & 0 & 0 & 3 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 1 & 0 & 0 & 3 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 3 & 3 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & (0) & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 1 & 0 & 3 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 3 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 3 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 3 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 3 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 3 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 3 & 3 & 0
\end{array}\right)$$

The embedding of canonical distance matrices:

$$\left(\begin{array}{ccccc|ccccc|ccc}
(0) & 8 & 1 & 7 & 4 & 17 & 16 & 19 & 15 & 12 & 5 & 11 & 8 \\
8 & (0) & 7 & 1 & 4 & 17 & 16 & 19 & 15 & 12 & 5 & 11 & 8 \\
1 & 7 & 0 & 6 & 3 & 16 & 15 & 18 & 14 & 11 & 4 & 10 & 7 \\
7 & 1 & 6 & 0 & 3 & 16 & 15 & 18 & 14 & 11 & 4 & 10 & 7 \\
4 & 4 & 3 & 3 & 0 & 13 & 12 & 15 & 11 & 8 & 1 & 7 & 4 \\
17 & 17 & 16 & 16 & 13 & (0) & 1 & 4 & 2 & 5 & 12 & 6 & 9 \\
16 & 16 & 15 & 15 & 12 & 1 & 0 & 3 & 1 & 4 & 11 & 5 & 8 \\
19 & 19 & 18 & 18 & 15 & 4 & 3 & 0 & 4 & 7 & 14 & 8 & 11 \\
15 & 15 & 14 & 14 & 11 & 2 & 1 & 4 & 0 & 3 & 10 & 4 & 7 \\
12 & 12 & 11 & 11 & 8 & 5 & 4 & 7 & 3 & 0 & 7 & 1 & 4 \\
5 & 5 & 4 & 4 & 1 & 12 & 11 & 14 & 10 & 7 & 0 & 6 & 3 \\
11 & 11 & 10 & 10 & 7 & 6 & 5 & 8 & 4 & 1 & 6 & 0 & 3 \\
8 & 8 & 7 & 7 & 4 & 9 & 8 & 11 & 7 & 4 & 3 & 3 & 0
\end{array}\right)$$

The embedding of distance matrices of ontological structure elements in the basis (bold on the diagonal) of minimal dimensionality:

$$\begin{pmatrix} (0) & 6 & 1 & 7 & 8 & 3 & 4 & 1 & 5 & 8 & 9 & 9 & 5 \\ 6 & (0) & 7 & 1 & 2 & 9 & 10 & 7 & 11 & 14 & 3 & 15 & 11 \\ 1 & 7 & 0 & 6 & 7 & 4 & 5 & 2 & 6 & 9 & 10 & 10 & 6 \\ 7 & 1 & 6 & \mathbf{0} & 4 & 10 & 11 & 8 & 12 & 15 & 4 & 16 & 12 \\ 8 & 2 & 7 & 4 & 0 & 6 & 7 & 6 & 8 & 11 & 6 & 12 & 8 \\ 3 & 9 & 4 & 10 & 6 & (0) & 1 & 2 & 2 & 5 & 4 & 6 & 4 \\ 4 & 10 & 5 & 11 & 7 & 1 & 0 & 3 & 1 & 4 & 5 & 5 & 3 \\ 1 & 7 & 2 & 8 & 6 & 2 & 3 & 0 & 4 & 7 & 6 & 8 & 6 \\ 5 & 11 & 6 & 12 & 8 & 2 & 1 & 4 & 0 & 3 & 6 & 4 & 2 \\ 8 & 14 & 9 & 15 & 11 & 5 & 4 & 7 & 3 & 0 & 9 & 1 & 3 \\ 9 & 3 & 10 & 4 & 6 & 4 & 5 & 6 & 6 & 9 & 0 & 10 & 6 \\ 9 & 15 & 10 & 16 & 12 & 6 & 5 & 8 & 4 & 1 & 10 & \mathbf{0} & 4 \\ 5 & 11 & 6 & 12 & 8 & 4 & 3 & 6 & 2 & 3 & 6 & 4 & 0 \end{pmatrix}$$

The previously formulated theorem [5] allows us to pass from an existing metric space to a new metric space by taking into account pseudometrics that can specify a dimension.

This is not enough to analyze topological properties of reflexive ontological structures. In order to highlight the topology of these structures, we propose to consider additional numerical features for a sign σ of the tuple μ:

$$incidensity(\sigma) \overset{def}{=} \frac{|\iota(\sigma)|+1}{|\upsilon(\sigma)|}\,;\; weight_{\lambda}(\mu) \overset{def}{=} \frac{\sum_{\varepsilon\in\lambda\mu}|C(\varepsilon)|}{\left|\bigcup_{\varepsilon\in\lambda\mu} C(\varepsilon)\right|}, \quad (4)$$

where $\upsilon(\sigma)$ is a set of signs in the proximity of σ, $\iota(\sigma)$ is a set of their incidences, $C(\varepsilon)$ is a proximity of λ-type ε.

On the basis of the introduced features and similar ones, it is possible to specify the weight of connectives which can be the basis of metric properties of the semantic space. These properties allow us to form a system of closed sets of the topological subspace which is formed by means of ascending hierarchical clustering.

B. *Dynamic ontological structures*

For ontological structures of knowledge bases with operational semantics, we use models of transition-state diagrams (graphs) (using the becoming relation as a basic one), as well as the proposed model of operational-information space [5] which allow us to represent different (finite) formal submodels (pseudo-orgraphs) of knowledge processing and investigate their spatial properties including topology of discrete structures.

Previously [5, 8], we considered the time-averaged entropy of wavefronts of the equilibrium state of a knowledge processing model (system) with a pseudo-orgraph structure of returnable operations. Returnability are one of the three key features for constructing a general classification of operations and problems solved with their help [5].

If a pseudo-orgraph is not strongly connected or does not correspond to non-redundant operations, i.e. it contains some features, vertices in which the total flow is not conserved (final or initial) then we can apply to it some meta-operation (surgery) which completes it to a connected one. The simplest of the operations augments such a pseudo-orgraph with a bipartite complete orgraph, the first partition of which is the set of all final vertices and the second partition is the set of all initial vertices. After that, one can compute the flows for such augmented pseudo-orgraph and the corresponding flows for the original pseudo-orgraph. As for the calculation of wave fronts, lengths and periods of such a pseudo-orgraph, it is more complicated. A theorem was formulated earlier that this entropy is a monotone measure [5] for the separability relation on equivalence classes of isomorphic strongly connected pseudoormultigraphs. In addition to this, the following theorem holds true.

Theorem 1: The entropy is an additive measure with respect to the tensor product operation of equivalence classes of isomorphic strongly connected pseudoormultigraphs.

The introduced entropy is an invariant of knowledge processing operations and it is also the basis for the formation of pseudometrics on elements of ontological structures and in accordance with the mentioned theorem is the basis for the transition from one metric space to another.

C. *Network problem solving models*

Due to the prevalence of neural network models that compute numerical feature values of input image data in linear vector space, the question arises about the relationship between input data, object images, and their secondary image features. To what extent the computed feature values correspond to some formal context and ontological structure of the corresponding formal concepts. Whether the corresponding images and concepts have an extensional, what are their denotates and designates. A separate difficulty is that the description of features uses models of continuous mathematics (including real numbers, the power of the set of which is uncountable), and the program implementation is discrete. However, only a countable subset of real numbers is computable. To overcome computational difficulties, such models as interval arithmetic [9], quasivector spaces [10] and others have been proposed, but they do not remove all the issues [9, 11].

III. PROTOMULTIVERSAL NUMBERS

A model of protomultiversal numbers is proposed to decide the above inconsistencies and provide answers to the questions posed.

The model of protomultiversal numbers is given by:

$$\langle N, L, R, F, M, A \rangle, \quad (5)$$

where N is the set of (canonical forms of) computable numbers, L is the set of places, R and F are the sets of measurement and re-generalization steps ($\grave{s}$ is the next step after s), M is the set of protomultiversal numbers, A is the set of conjugacy constraints that is satisfied:

$$A \subseteq \left(\left(\left(\left(\left(N^2 \right)^* \right)_+^{\left(2^M\right)} \right)^L \right)^R \right)^F . \quad (6)$$

Each constraint $\langle \underline{a}_t, \overline{a}_t \rangle \in N^2$ for x has the form:

$$\sum\nolimits_{i=0}^{n} \sum\nolimits_{j=0}^{q^i-1} \overline{a}_{\frac{q^i-1}{q-1}+j} * \prod\nolimits_{k=1}^{i} x_{1+\left(\left\lfloor \frac{j}{q^{k-1}} \right\rfloor \bmod q\right)} \le 0,$$

$$\sum\nolimits_{i=0}^{n} \sum\nolimits_{j=0}^{q^i-1} \underline{a}_{\frac{q^i-1}{q-1}+j} * \prod\nolimits_{k=1}^{i} x_{1+\left(\left\lfloor \frac{j}{q^{k-1}} \right\rfloor \bmod q\right)} \ge 0. \quad (7)$$

Each number m ($m.bottom \le \langle \top, x \rangle \le m.top$) of M:

$$m = \langle locations, low, hi, down, up, bottom, top \rangle, \quad (8)$$

where

$$\{low\} \cup \{hi\} \subseteq \left(\left(\left(\{\bot, \top\} \times N \right)^L \right)^R \right)^F$$

$$\{down\} \cup \{up\} \subseteq \left(\left(\{\bot, \top\} \times N \right)^R \right)^F \quad (9)$$

$$\{bottom\} \cup \{top\} \subseteq \left(\{\bot, \top\} \times N \right).$$

Other properties are also performed:

$$\forall p \exists q \left(lo(f)(r)(p) \ge lo(\grave{f})(r)(q) \right) \wedge$$
$$\wedge \left(hi(f)(r)(p) \le hi(\grave{f})(r)(q) \right);$$
$$\forall x \left(lo(f)(r)(p) \le \langle \top, x \rangle \le hi(f)(r)(p) \right) \to$$
$$\to \exists q \left(lo(\grave{f})(r)(q) \le \langle \top, x \rangle \le hi(\grave{f})(r)(q) \right);$$
$$\forall x \left(lo(\grave{f})(r)(p) \le \langle \top, x \rangle \le hi(\grave{f})(r)(p) \right) \to$$
$$\to \exists q \left(lo(f)(r)(q) \le \langle \top, x \rangle \le hi(f)(r)(q) \right);$$
$$\forall p \exists q \left(lo(f)(\grave{r})(p) \ge lo(f)(r)(q) \right) \wedge \quad (10)$$
$$\wedge \left(hi(f)(\grave{r})(p) \le hi(f)(r)(q) \right);$$
$$lo(h)(r)(p) = lo(f)(s)(p) = lo(f)(r)(p);$$
$$hi(h)(r)(p) = hi(f)(s)(p) = hi(f)(r)(p);$$
$$bottom \le down(f)(r) \le low(f)(r)(p) \le hi(f)(r)(p$$
$$\left(low(f)(r)(p) \ge hi(f)(r)(p) \right) \to \left(low(f)(r)(p) \right)_1$$
$$hi(f)(r)(p) \le up(f)(r) \le top,$$

where

$$\left(\langle \alpha, \chi \rangle \ge \langle \beta, \gamma \rangle \right) \Leftrightarrow \left(\langle \beta, \gamma \rangle \le \langle \alpha, \chi \rangle \right)$$
$$\left(\langle \alpha, \chi \rangle > \langle \beta, \gamma \rangle \right) \Leftrightarrow \left(\langle \beta, \gamma \rangle < \langle \alpha, \chi \rangle \right) \quad (11)$$
$$\left(\langle \alpha, \chi \rangle < \langle \beta, \gamma \rangle \right) \Leftrightarrow \left((\chi < \gamma) \vee \left((\chi = \gamma) \wedge \left((\neg \alpha) \wedge \beta \right) \right) \right)$$
$$\left(\langle \alpha, \chi \rangle \le \langle \beta, \gamma \rangle \right) \Leftrightarrow \left((\chi < \gamma) \vee \left((\chi = \gamma) \wedge (\alpha \to \beta) \right) \right)$$
$$\left(\langle \alpha, \chi \rangle = \langle \beta, \gamma \rangle \right) \Leftrightarrow \left(\left(\langle \alpha, \chi \rangle \le \langle \beta, \gamma \rangle \right) \wedge \left(\langle \alpha, \chi \rangle \ge \langle \beta, \gamma \rangle \right) \right).$$

The operation max (and similarly min) is defined:

$$\max\left(\{\langle \alpha, \chi \rangle\} \cup \{\langle \beta, \gamma \rangle\} \right) \overset{def}{=}$$
$$= \left\langle \left(\left(\alpha \wedge (\chi \ge \gamma) \right) \vee \left(\beta \wedge (\chi \le \gamma) \right) \right), \max\left(\{\chi\} \cup \{\gamma\} \right) \right\rangle. \quad (12)$$

Also:

$$expanse(m) \overset{def}{=} m.top - m.bottom; \quad (13)$$

$$latitude(m)(f)(r)(p) \overset{def}{=} m.hi(f)(r)(p) - m.low(f)(r)(p).$$

The regular protomultiversal numbers are satisfied:

$$m.bottom \overset{def}{=} \min\left(\left\{ m.down(f)(r) \middle| \langle f, r \rangle \in F \times R \right\} \right)$$
$$m.top \overset{def}{=} \max\left(\left\{ m.up(f)(r) \middle| \langle f, r \rangle \in F \times R \right\} \right)$$
$$m.down \overset{def}{=} \min\left(\left\{ m.low(f)(r)(p) \middle| p \in m.locations(f)(r) \right\} \right)$$
$$m.up \overset{def}{=} \max\left(\left\{ m.hi(f)(r)(p) \middle| p \in m.locations(f)(r) \right\} \right)$$
$$m.locations(f)(\grave{r}) \subseteq m.locations(f)(r). \quad (14)$$

There exist three cases to define relations *expand*, *extend*, *lapse* and operations on the numbers from *M*: partially defined ($Q = O = \exists$), an unary completely defined ($\langle Q, O \rangle = \langle \forall, \exists \rangle$) and a completely defined ($Q = O = \forall$) operations. For example, for the sum operations:

$$\forall g \forall t \forall o Q h Q s Q q O f O r O p \langle l.low(g)(t)(o), l.hi(g)(t)(o) \rangle \overset{def}{=}$$
$$\langle m.low(h)(s)(q), m.hi(h)(s)(p) \rangle + \quad (15)$$
$$+ \langle n.low(f)(r)(p), n.hi(f)(r)(p) \rangle$$

where $\langle \alpha_{low}, \alpha_{hi} \rangle + \langle \beta_{low}, \beta_{hi} \rangle = \langle \lfloor] \alpha_{low} + \beta_{low} [\rfloor, \lceil] \alpha_{hi} + \beta_{hi} [\rceil \rangle$;

$$\alpha + \beta = \langle \alpha_1 + \beta_1, \alpha_2 + \beta_2 \rangle; \quad (16)$$

$$\alpha_1 + \beta_1 = (\alpha_1 \wedge \beta_1) \vee \left((|\alpha_2| = +\infty) \wedge \alpha_1 \right) \vee \left((|\beta_2| = +\infty) \wedge \beta_1 \right),$$

and $\lfloor] x [\rfloor$ with $\lceil] x [\rceil$ are rounding down and up on N ($\{-\infty\} \cup \{+\infty\} \subseteq N$). Other operations such as multiplication, subtraction and division can be specified similarly and according to the operations of interval arithmetic [10].

Depending on *N*, the algebraic systems of protomultiversal numbers can extend fields or linear vector spaces, exhibiting properties of rings or modules.

For each operation or function expressed by their superposition, as well as for the sum operation, we can define at least three variants: partially, unary completely and completely defined functions.

Then we can treat a concept as an image (feature value) of a name (sign), i.e. an image of the initial pattern (a set of protomultiversal numbers). Denotation (referent) is the image (value of the sign) of the concept and the image of the name (set of protomultiversal numbers). Designate is the set of all protomultiversal numbers partially or completely defined by operations (functions) between them.

IV. Conclusion

On the basis of previously proposed models for representing non-numeric structures possessing denotational and operational semantics, the steps of the algorithm and examples of identifying topological and metric features of static structures of the semantic space are considered. Semantically stable features are considered. These features allow analyzing topological and metric properties of strongly connected ontological structures. The measure of information entropy additive on the set of strongly connected pseudo-orgraphs of operations is considered for dynamic structures of the semantic space. The theorem on its additivity is formulated. On the basis of this measure as an invariant, an approach to the construction of pseudometrics specifying the metric properties of the semantic space is proposed. The model of protomultiversal numbers, which is a formal basis for the specification of meaning and revealing semantic properties of network structures and models, such as artificial neural networks, etc., is proposed. This model suggests additional studies of its practical capabilities.

References

[1] D. A. Pospelov, “Situational management: theory and practice”, Moscow, Nauka, 1986, p. 288.

[2] L. Collatz, “Functional analysis and numerical mathematics”, New York, San Francisco, London, Academic Press, p. xx + 1966, 473 p.

[3] J. Heinonen, “Lectures on analysis on metric spaces”, New York, Springer, p. x+141, 2001.

[4] M. A. Armstrong, “Basic topology [1979]”, Undergraduate Texts in Mathematics, Springer, 1983, p. 263.

[5] V. Ivashenko, “Approaches to the study of semantic space and integrated logical inference models using similarity, difference and other measures,” Open Semantic Technologies for Intelligent Systems (OSTIS), Minsk, vol. 9, 2025, pp. 97–112.

[6] J. Poelmans, S. Kuznetsov, D. Ignatov and G. Dedene, “Review: Formal concept analysis in knowledge processing: a survey on models and techniques,” Expert Systems with Applications: An International Journal, 2013, vol. 40, no. 16, pp. 6601–6623.

[7] Yu. Manin, M. Marcolli, “Semantic spaces,” Mathematics in Computer Science, vol. 10, Dec 2016. DOI:10.1007/s11786-016-0278-9.

[8] A. Mowshowitz, M. Dehmer, “Entropy and the complexity of graphs revisited”, Entropy, vol. 14, no. 3, 2012, pp. 559–570.

[9] M. Lechner, D. Žikelić, K. Chatterjee, T. A. Henzinger and D. Rus, “Quantization-aware interval bound propagation for training certifiably robust quantized neural networks,” vol. 37. AAAI Press, 2023, Article 1678, pp. 14964–14973. [Proceedings of the Thirty-Seventh AAAI Conference on Artificial Intelligence and Thirty-Fifth Conf. on Innovative Applications of Artificial Intelligence and Thirteenth Symp. on Educational Advances in Artificial Intelligence (AAAI'23/IAAI'23/EAAI'23)]. DOI:10.1609/aaai.v37i12.26747.

[10] S. Markov, “On the Algebra of intervals,” Reliable Computing 21, 2016, pp. 80–108.

[11] L. Darlow, C. Regan, S. Risi, J. Seely and L. Jones. “Continuous thought machines,” ArXiv, 2025. ArXiv, abs/2505.05522.