International Science Journal of Education & Linguistics
2025; 4(6): 75-80
https://isg-journal.com/isjel/
doi: 10.46299/j.isjel.20250406.08
ISSN: 2720-684X

INTERNATIONAL SCIENCE GROUP

Vector modeling of atomic linguistic elements taking into account their dual algebraic essence

Nataliya Anatoliivna Yakimova
Odessa I.I. Mechnikov National University, Odessa, Ukraine
ORCID 0009-0005-8577-2775

Abstract: In modern linguistics and mathematics the task of mathematical formalization of natural language for the purpose of its further computer processing is becoming more and more relevant. This topic is in demand, primarily, in artificial intelligence systems, which are becoming increasingly widespread both in science and in everyday life. Therefore, the efforts of scientists are aimed at finding a mathematical apparatus that would meet the needs of an interdisciplinary research direction associated with the symbiosis of mathematical and humanitarian knowledge. Human thought processes, which are reflected in speech phenomena, are traditionally formalized by the apparatus of mathematical logic. But at the current level of development of linguistic research, the tools of classical logic are insufficient. The mathematical apparatus of vector logical algebra, which is a development of the algebra of finite predicates, makes it possible to develop computer linguistics. This article shows the fundamental possibility of mathematically formalizing the simplest speech constructions using the basic operations of vector logical algebra. Using the example of Ukrainian and English, which are representatives of different language groups, it is proven that for all simple word combinations of natural language the axioms of the logical field are fulfilled. It does not matter whether these word-combinations consist of main or secondary members of the sentence, nor what type of grammatical connection is used in these word combinations. Such formalization allows us to represent natural language in the form of algebraic formulas and consider its constructions as elements of a vector space.

Keywords: logical scalar, logical vector, sentential connections, disjunction, conjunction, inversion, simple word-combinations of natural language, grammatical connection, main and secondary members of a sentence, axioms of the logical field.

1. Introduction

The history of science spans more than one century. But since ancient times, it has historically been that the natural and humanities scientific directions have been clearly demarcated. Moreover, these directions did not even compete, because they did not have any common ground. However, over time, a fierce struggle between "physicists and lyricists" for leadership in the hierarchy of the scientific community gained momentum. Representatives of the exact sciences considered the humanities to be dreamers, disconnected from reality. And representatives of the humanities, and primarily philologists, considered mathematicians to be soulless rascals who were incapable of feeling beauty. But this feud between the scientific "Montagues and Capulets" was nevertheless destined to end. Today, the scientific community faces increasingly profound challenges that cannot be overcome by the means of any one scientific direction alone. This applies to all spheres of human life. Computerization and robotization are increasingly penetrating all layers of society. But for these systems to work adequately, they need an understanding of the laws of the functioning of the human psyche, human speech, and human thinking. And here they come in handy with the acquisition of humanitarian scientific thought: psychologists, philologists, historians, etc. But even those who consider themselves deep humanitarians can no longer do without the achievements of scientific and technological progress. Computers have long since become part of the everyday lives of even people far from mathematics. To facilitate the process of communication between humans and machines, "physicists and lyricists" must communicate and work together. And this interdisciplinary cooperation opens up new development prospects for science, which are interesting for representatives of the most diverse scientific fields, and which yields unexpected but very effective results.

2. Object and subject of research

A piece of text consists of sentences, and each sentence is composed of statements that may contain «sub-statements» and ultimately consist of words [1]. The grammar of each natural language defines the way in which words and expressions can be combined into sentences. In each simple word-combination, the grammatical presence of the main word and dependent word is provided. Each of them, in turn, can be represented by different parts of speech. The types of grammatical connection between these words can also be different. The grammars of different groups of natural languages provide for different types of such a connection. Thus, the atomic level of any natural language is a single word, and the simplest construction is a simple word-combination consisting of two words.

3. Target of research

The given arbitrary two words, which make up the simplest speech objects – simple word-combinations – can be arranged by attributing certain mathematical properties to them. If we draw an analogy with a vector space, the simplest algebraic objects in such a space are scalars and vectors. A scalar is an object of lower rank compared to an n-dimensional vector. A scalar object has no dimension or its dimension is smaller than that of the object that is taken as a vector. The elements of an algebraic vector space obey certain laws. In this article, the author sets himself the goal of proving that an analogy can be drawn between the elements of a vector space that has a logical structure and simple word-combinations of any natural language.

4. Literature analysis

In [2, 3, 4], the mathematical apparatus of finite predicate algebra was developed. It combines the algebraic means of classical mathematical logic and the constructive possibilities of Boolean functions [5]. These works laid the foundation for creating the mathematical foundation of artificial intelligence systems. Such systems use predicate logic constructs to model not only thought processes, but also to adequately understand human behavior in general [6, 7]. For such understanding, all functions of the human personality, both biological and cognitive, are subject to modeling. But special attention when creating artificial intelligence systems is still paid to the cognitive characteristics of the human personality, because any biological aspects of life are still subject to understanding and are reflected in one or another model of behavior of each specific individual. A generalization of the developed methods was carried out in [8]. In this work, based on the mathematical apparatus of finite predicates, the functioning of the fundamental and applied algebra of finite predicates is demonstrated, and their application to the algebra of ideas is also extended. These mathematical algorithms were developed in [9]. In this work, a mathematical apparatus of logical spaces is developed, created by analogy with linear algebra [10]. But unlike the classical vector space, it is proposed to create a hierarchical structure of vector spaces, each of which can simultaneously be both a scalar field for predicates of higher arity and a system of vectors for predicates of lower arity. The linguistic principles of the conducted research are described in detail in [11, 12]. Based on the combination of these linguistic and algorithmic techniques, semantic networks are created, which is also the methodological basis for the application of mathematics in modeling the cognitive component of artificial intelligence [13].

5. Research methods

The following methods were used in conducting the study, the results of which are presented in this article:

- method of generalizing scientific literature. Both scientific developments in mathematical disciplines (mathematical logic, linear algebra) and linguistic literature were investigated;
- a method of scientific observation, thanks to which mathematical patterns in the functioning of linguistic objects were discovered;
- the method of systems analysis, which helped to make interdisciplinary generalizations of the obtained results and extend the algorithm of their functioning to the studied groups of natural languages as a whole.

6. Research results

6.1. Execution of the axioms of the logical field for secondary members of the sentence

Let's consider sentences from a mathematical point of view, that is, as some constructions of vector logical algebra [9]. We will conduct research on the examples of the Ukrainian and English languages, which are representatives of different groups of natural languages, namely Slavic and Romance-Germanic groups. As an example, let's take the object «***flower***» («***квітка***»), which has the quality «***blue***» («***синя***»). Let the object «***flower***» correspond to the vector $\boldsymbol{l}$ of some logical space $\boldsymbol{L}$, and the quality «***blue***» to the scalar α [14, 15]. Then the expression «***blue flower***» («***синя квітка***») corresponds to the product of the vector $\boldsymbol{l}$ by the scalar α. That is, this statement is formalized by the expression $\alpha\, \boldsymbol{l}$. Let's check the fulfillment of the axioms of logical algebra for the selected vector space and the selected scalar field. In this example, the set of all objects is taken as a vector space, and the set of all qualities of objects is taken as a scalar field. Boolean operations of inversion, disjunction, and conjunction are specified on a set of vectors. Then for word-combinations of the Ukrainian language:

$$\text{«не квітка»} = \overline{\text{«квітка»}} = \bar{\boldsymbol{l}},$$

and for English word-combinations:

$$\text{«not a flower»} = \overline{\text{«flower»}} = \bar{\boldsymbol{l}}.$$

Let the object «***handkerchief***» («***хустка***») correspond to some vector $\boldsymbol{g} \in \boldsymbol{L}$. Then in the Ukrainian language we have:

$$\text{«квітка або хустка»} = \text{«квітка»} \vee \text{«хустка»} = \boldsymbol{l} \vee \boldsymbol{g},$$

$$\text{«квітка та хустка»} = \text{«квітка»} \wedge \text{«хустка»} = \boldsymbol{l} \wedge \boldsymbol{g}.$$

For the English language, these formulas look like this:

$$\text{«a flower or a handkerchief»} = \text{«flower»} \vee \text{«handkerchief»} = \boldsymbol{l} \vee \boldsymbol{g},$$

$$\text{«a flower and a handkerchief»} = \text{«flower»} \wedge \text{«handkerchief»} = \boldsymbol{l} \wedge \boldsymbol{g}.$$

The same Boolean operations are specified on the set of elements of the scalar field. Then for the Ukrainian language

$$\text{«не синій»} = \overline{\text{«синій»}} = \bar{\alpha},$$

and for the English language

$$\text{«not blue»} = \overline{\text{«blue»}} = \bar{\alpha}.$$

Let the quality «*large*» («*великий*») correspond to the logical scalar β. Then for the Ukrainian language

$$«синій\ або\ великий» = «синій» \vee «великий» = \alpha \vee \beta,$$

$$«синій\ та\ великий» = «синій» \wedge «великий» = \alpha \wedge \beta.$$

For the English language, we have the following results:

$$«blue\ or\ large» = «blue» \vee «large» = \alpha \vee \beta,$$

$$«blue\ and\ large» = «blue» \wedge «large» = \alpha \wedge \beta.$$

Therefore, the operations of disjunction, conjunction and inversion of objects and their qualities given in this way satisfy all axioms of the logical field [1]. Sentential relations are the names of the corresponding operations. The auxiliary parts of speech «**or**» («**або**»), «**and**» («**і** (**й**, **та**)») and «**not** (**no**)» («**не** (**ні**)») are considered respectively as the names of the operations of disjunction, conjunction and inversion [15].

It is obvious that the axioms connecting the product of logical scalars and logical vectors also hold. In particular, on the example of the Ukrainian and English languages:

associativity

$$(\alpha\ \beta)\ l = «(синя\ та\ велика)\ квітка» = «синя»\ та\ («велика\ квітка») = \alpha\ (\beta\ l);$$

$$(\alpha\ \beta)\ l = «(blue\ and\ large)\ flower» = «blue»\ and\ («large\ flower») = \alpha\ (\beta\ l);$$

distributivity about the disjunction of scalars

$$(\alpha \vee \beta)\ l = «(синя\ або\ велика)\ квітка» = «синя\ квітка»\ або\ «велика\ квітка» = \alpha\ l \vee \beta\ l;$$

$$(\alpha \vee \beta)\ l = «(blue\ or\ large)\ flower» = «blue\ flower»\ or\ «large\ flower» = \alpha\ l \vee \beta\ l;$$

distributivity about the disjunction of vectors

$$\alpha\ (l \vee g) = «синя\ (квітка\ або\ хустка)» = «синя\ квітка»\ або\ «синя\ хустка» = \alpha\ l \vee \alpha\ g,$$

$$\alpha\ (l \vee g) = «a\ blue\ (flower\ or\ handkerchief)» =$$
$$= «a\ blue\ flower»\ or\ «a\ blue\ handkerchief» = \alpha\ l \vee \alpha\ g.$$

Thus, the set of objects can be considered as a vector logical space defined over a scalar field, the elements of which are the qualities of objects.

If we take a set of objects as a scalar field, and a set of qualities of objects as a space of vectors given above it, then, obviously, all axioms of the logical space will also be fulfilled. That is, the set of qualities of objects can also be considered as a vector logical space over a scalar field, the elements of which are objects. Therefore, the mathematical formalization of these word-combinations can occur in any direction: both from the main word to the dependent one, and vice versa, from the dependent word to the main one. The considered examples illustrate the case when the dependent word in the phrase is an adjective, and the main one is a noun. However, regardless of which part of the language the main and dependent words in a simple word-combination are given, according to the specified algorithm, this word-combination can be represented by a mathematical formula, namely a formula of the vector logical algebra [9], regardless of the direction of formalization. Also, there are different types of grammatical connection between the main and dependent words in word-combinations. According to the considered scheme of formalization, any grammatical relationship between the main and dependent words can be presented.

6.2. Execution of the axioms of the logical field for the main members of the sentence.

Let us now consider the word-combination as a part of the sentence, bearing in mind that the word-combination is the result of dismembering the sentence into units that have some meaningful integrity. Recently, this direction has gained popularity in linguistics. In every language there are algorithms for selecting the main members of a sentence [14]. Each sentence describes some relation expressed by a predicate. At the same time, the subject defines the subject, that is, some object of the real world. The predicate, in turn, expresses some feature (action, state, property, quality) of the object described by the subject. But in any case, the relationship between the subject and the predicate, both of which are the main members of the sentence, can be formalized similarly to the example discussed above, regardless of which of these members of the sentence is to be taken as the main word in the studied phrase, and which one is the dependent word.

As an example, consider the connection between the subject and a simple verbal predicate. The set of words that can be a subject is the set of «objects». It is a logical field, as well as a set of words that can be a simple verb predicate. Let the vectors ***h*** and ***d*** correspond to the objects «***boy***» («***хлопчик***») and «***girl***» («***дівчинка***»), and the scalars **μ** and **τ** to the verbs «***play***» («***грати***») and «***draw***» («***малювати***»). Then, obviously, for the Ukrainian language, which assumes a free order of words in a sentence:

(μ τ) *l* = «(*грає* та *малює*) *хлопчик*» = «*грає*» та («*малює хлопчик*») = μ (τ *l*);

(μ ∨ τ)*l* = «(*грає* або *малює*) *хлопчик*» = «*грає хлопчик*» або «*малює хлопчик*» = μ *l* ∨ τ *l*;

μ (*l* ∨ *g*) = «*грає* (*хлопчик* або *дівчинка*)» = «*грає хлопчик*» або «*грає дівчинка*» = μ *l* ∨ μ *g*,

At the same time, for the English language, taking into account the strict order of words in the sentence and the commutative properties of disjunction and conjunction operations [8]:

(μ τ) *l* = «*a boy is* (*playing* and *drawing*) » = («*a boy is drawing*») and «*playing*» = μ (τ *l*);

(μ ∨ τ)*l* = «*a boy* (*playing* or *drawing*)» = «*boy is playing* » or «*boy is drawing*» = μ *l* ∨ τ *l*;

μ (*l* ∨ *g*) = «(*a boy* or *a girl*) *is playing*» = «*a boy is playing*» or «*a girl is playing*» = μ *l* ∨ μ *g*,

Therefore, all axioms of vector logical algebra are also fulfilled in this case. Similarly, if predicates are taken as vectors, and subjects are taken as elements of the scalar field, then the axioms of the logical space will also be fulfilled in this case. Therefore, the connection between the subject and the simple verb predicate can also be formalized in any direction. Similarly, it can be shown that if the subject and predicate in a sentence are represented by some other acceptable grammar of the studied language by parts of speech, then their relationship can also be formalized in any direction by means of vector logical algebra.

7. Prospects for further research development

This article is devoted exclusively to the analytical study of the fundamental possibility of representing a simple phrase of natural language using vector algebra. In the future, we should consider the graphical representation of speech structures, taking into account their hierarchy, as a specific property of predicate vector spaces.

8. Conclusions

From all that has been said, it follows that any word-combination of natural language can be formalized in any direction by means of vector logical algebra. This indicates the dual nature of the simplest speech element – a word, which can simultaneously act as both a measured and a dimensionless object of mathematical modeling. The results obtained on the example of the Ukrainian language can undoubtedly be extended to the group of Slavic languages with a grammar close to Ukrainian. But, as it was shown, similar research methods can be implemented for any other languages. These studies show that by means of vector logical algebra, which is based on the apparatus of algebra of finite predicates, it is possible to formalize an arbitrary syntagm of any natural language.

References:

1) Yakimova N. A. (2025). Mathematical formalization of simple word-combinations using the algebra of finite predicates (on the example of the Ukrainian language). *Modern engineering and innovative technologies*. Karlsruhe, Germany. Issue № 3, part 3. Pp. 138 – 148. doi: 10.30890/2567-5273.2025-37-03-017. https://www.moderntechno.de/index.php/meit/issue/view/meit37-03/meit37-03

2) Шабанов-Кушнаренко Ю.П. (1984). Теорія інтелекту. Математичні засоби. Харків: Вища школа, 144с.

3) Шабанов-Кушнаренко Ю.П. (1986). Теорія інтелекту. Технічні засоби. Харків: Вища школа, 136с.

4) Шабанов-Кушнаренко Ю.П. (1987). Теорія інтелекту. Проблеми та перспективи. Харків: Вища школа, 160с.

5) Якімова Н.А. (2023). Дискретна математика. Частина 2. Булеві функції. Одеса: ОНУ ім. І.І. Мечникова, 126с.

6) Russel Stuart J., Norvig P. (2003). Artificial Intelligence: a modern approach. Ney Jersey: Prentice Hill Upper Saddle River, 1406p.

7) Luger George F. (2002). Artificial Intelligence: structures and strategies for complex problem solving. Boston: Addison Wesley, 864p.

8) Якімова Н.А. (2025) Алгебра скінченних предикатів. Одеса: ОНУ ім. І.І. Мечникова, 133с.

9) Якімова Н.А. (2025) Векторна логічна алгебра. Одеса: ОНУ ім. І.І. Мечникова, 126с.

10) F.R. Gantmacher. (1960). The Theory of matrices. AMS Chelsea Publishing, New York.

11) Левицький В.В. (2004). Квантитативні методи в лінгвістиці. Чернівці: «Рута», 190с.

12) Варбанець П. Д., Якімова Н. А. (2021). Лінгвостатистика. Одеса: ОНУ ім. І.І. Мечникова, 182 с.

13) Haykin S. (2006). Neural networks. Ney Jersey: Prentice Hill Upper Saddle River, 1104p.

14) Yakimova N.A. (2024). Operation of reversing of logical matrices. *SWorldJournal*. Svishtov, Bulgaria. Issue № 27. Part 1. Pp. 165 – 172. doi: 10.30888/2663-5712.2024-27-00-017. URL: https://www.sworldjournal.com/index.php/swj/issue/view/swj27-01/swj27-01

15) Якімова Н.А. (2023). Операції над блочними предикатними матрицями. *Вісник Одеського національного університету ім. І.І. Мечникова. Дослідження в математиці і механіці*. Том 28. Випуск 1 – 2 (41 – 42). С.185-199. doi: 10.18524/2519-206X.2023.1-2(41-42).305269. URL: http://liber.onu.edu.ua/pdf/RMM_2023_28_1-2.pdf