ДИСКУРС. 2022. Т. 8, № 1. С. 133–141
DISCOURSE. 2022, vol. 8, no. 1, pp. 133–141

Языкознание
Linguistics

Original paper
УДК 81’37
http://doi.org/10.32603/2412-8562-2022-8-1-133-141

Linguistic Data Model for Natural Languages and Artificial Intelligence. Part 7. Internal Logic 1

Oleg M. Polyakov

Saint Petersburg State University of Aerospace Instrumentation, St Petersburg, Russia, road.dust.spb@gmail.com, https://orcid.org/0000-0002-8520-3936

Introduction. The article continues a series of publications on the linguistics of relations (hereinafter referred to as R-linguistics) and is devoted to the study of the functioning of logical connectives with verbs. The article is the first part of the discussion of internal logic, which examines the use of logical connectives within sentences. This research involves the formation of semantic logic, that is, logic that takes into account the semantics of sentences.

Methodology and sources. The results obtained in the previous parts of the series are used as research tools. To develop the necessary mathematical representations in the field of internal logic, the previously formulated semantic concepts and operations are used.

Results and discussion. The use of logical connectives with verbs is analyzed. It is shown that these connectives actually refer to external logic, although in some cases it is necessary to adjust part of sentence, taking into account the semantics of the linguistic model. The concept of semantic substitution is defined and the first rule of substitution for verbs is formulated and justified.

Conclusion. Abandoning the traditional view of natural language logic means abandoning logical operations and logical inference. This forces us to consider logical operations that now take into account semantics, since they are related to the structure of the linguistic model. Analysis of the functioning of logical connectives with verbs shows that they are related to the linguistic model, which leads to the need for various semantic transformations of the text when such connectives are used. In particular, the use of logical connectives can lead to the loss of the meaning of the text. The rejection of logical inference is compensated by the appearance of semantic substitution rules, one of which is considered in this paper.

Keywords: R-linguistics, ascription operation, interpretation operator, substitution rule, semantics

For citation: Polyakov, O.M. (2022), “Linguistic Data Model for Natural Languages and Artificial Intelligence. Part 7. Internal Logic 1”, *DISCOURSE*, vol. 8, no. 1, pp. 133–141. DOI: 10.32603/2412-8562-2022-8-1-133-141.

ДИСКУРС. 2022. Т. 8, № 1. С. 133–141
DISCOURSE. 2022, vol. 8, no. 1, pp. 133–141

Оригинальная статья

Лингвистическая модель данных для естественных языков и искусственного интеллекта. Часть 7. Внутренняя логика 1

Олег Маратович Поляков

Санкт-Петербургский государственный университет аэрокосмического приборостроения, Санкт-Петербург, Россия, road.dust.spb@gmail.com, https://orcid.org/0000-0002-8520-3936

Введение. Статья продолжает серию публикаций по лингвистике отношений (далее R-лингвистика) и посвящена изучению функционирования логических связок с глаголами. Статья представляет собой первую часть обсуждения внутренней логики, которая изучает использование логических связок внутри предложений. Это исследование предполагает формирование семантической логики, т. е. логики, учитывающей семантику предложений.

Методология и источники. В качестве инструментов исследования используются результаты, полученные в предыдущих частях серии. Для разработки необходимых математических представлений в области внутренней логики используются сформулированные ранее семантические понятия и операции.

Результаты и обсуждение. Проанализировано использование логических связок с глаголами. Показано, что эти связки фактически относятся к внешней логике, хотя в некоторых случаях требуется корректировка членов предложения, учитывающая семантику лингвистической модели. Определено понятие семантической подстановки и для глаголов сформулировано и обосновано первое правило подстановки.

Заключение. Отказ от традиционного взгляда на логику естественного языка означает отказ от логических операций и логического вывода. Это принуждает к рассмотрению логических операций, которые теперь учитывают семантику, поскольку связаны со структурой лингвистической модели. Анализ функционирования логических связок с глаголами показывает, что они связаны с лингвистической моделью, что приводит к необходимости различных семантических трансформаций текста, когда такие связки применяются. В частности, использование логических связок может приводить к потере смысла текста. Отказ от логического вывода компенсируется появлением правил семантической подстановки, одно из которых рассмотрено в данной работе.

Ключевые слова: R-лингвистика, операция приписывания, оператор интерпретации, правило подстановки, семантика

Для цитирования: Поляков О. М. Лингвистическая модель данных для естественных языков и искусственного интеллекта. Часть 7. Внутренняя логика 1 // ДИСКУРС. 2022. Т. 8, № 1. С. 133–141. DOI: 10.32603/2412-8562-2022-8-1-133-141.

Introduction. This article continues a series of publications devoted to the introduction to the linguistics of relations (R-linguistics) – a formal direction in linguistics. Here we will continue the conversation about logic within the framework of R-linguistics representations.

In [1], the problems of the approach to the logic of natural language from the point of view of mathematical logic were considered. It was also shown that the traditional logical approach to language has many disadvantages, so a new approach is needed, which naturally follows from the nature of the linguistic model and language constructions. Abandoning the use of traditional logic means abandoning the concepts of truth and falsity, and with it the rejection of traditional logical operations and logical inference. Nevertheless, in [1] an attempt was made to formulate some

ДИСКУРС. 2022. Т. 8, № 1. С. 133–141
DISCOURSE. 2022, vol. 8, no. 1, pp. 133–141

logical representations based on the most general semantic view and natural properties of the text. There were also obtained some results about the conditions for the loss of text semantics. All those manifestations of logic that are associated with the operation of attributing sentences one after another, we call external logic, because it describes the interaction of the sentences themselves and leaves out of sight what happens inside the sentences.

In [2], the questions of external logic were considered by examples and on this basis some conclusions were made about the order of semantic interpretation of the text in the framework of R-linguistics. To understand the logic that operates within sentences (internal logic), it is necessary to consider the functioning of logical connectives (conjunctions) within sentences. To do this, we will consider the use of connectives for verbs and nouns with adjectives. Due to the volume, this review will have to be done in two stages. It should be noted at once that logical connectives perform other functions in sentences besides logical functions, for example, functions of communicative semantics [3]. This aspect of their functioning is well described and will not be considered here. In addition, we will define some semantic substitutions that are not related to truth, but rather to the semantic identity of phrases. These substitutions are in a sense a substitute for logical productions.

Methodology and sources. The results obtained in the previous parts of the series are used as research tools. To develop the necessary mathematical representations in the field of internal logic, the previously formulated semantic concepts and operations are used.

Results and discussion.

Logical operations with verbs.

In the previous article [2], we only touched on the question of compound sentences, when simple sentences are combined into one with the use of conjunctions And and OR. Do logical connectives (actions) in a language always mean only logical operations with language sentences? In other words, can logical operations in a language always be interpreted (translated) into operations between sentences? The use of these conjunctions concerns primarily verbs, since they are combined into some new compound verb, so in this article we will start with verbs, and in the next we will continue for other parts of speech. In cases where other members of the sentence are not affected by unions, such a union can simply be reduced to an attribution operation from [2] with or without branching. For example, the sentence “It rained and I stayed home” is semantically equivalent to two consecutive sentences (“It rained. I stayed at home”). But what happens when other members of the sentence are involved in the connection of two sentences (round and square brackets from article [2])?

As it was shown in [2], simple narrative affirmative sentences ultimately get the form of relations when interpreted in the model. If we look at relations as ordinary sets whose elements are pairs or triples of objects, it may seem that the logical connectives “AND”, “OR” correspond to the intersection or union of these relations, which at first glance should lead to classical logic with a classical set-theoretic model and operations of intersection, union and complement of relations. This, however, is not so, since verbs of different arity can be connected by logical connectives in a sentence.

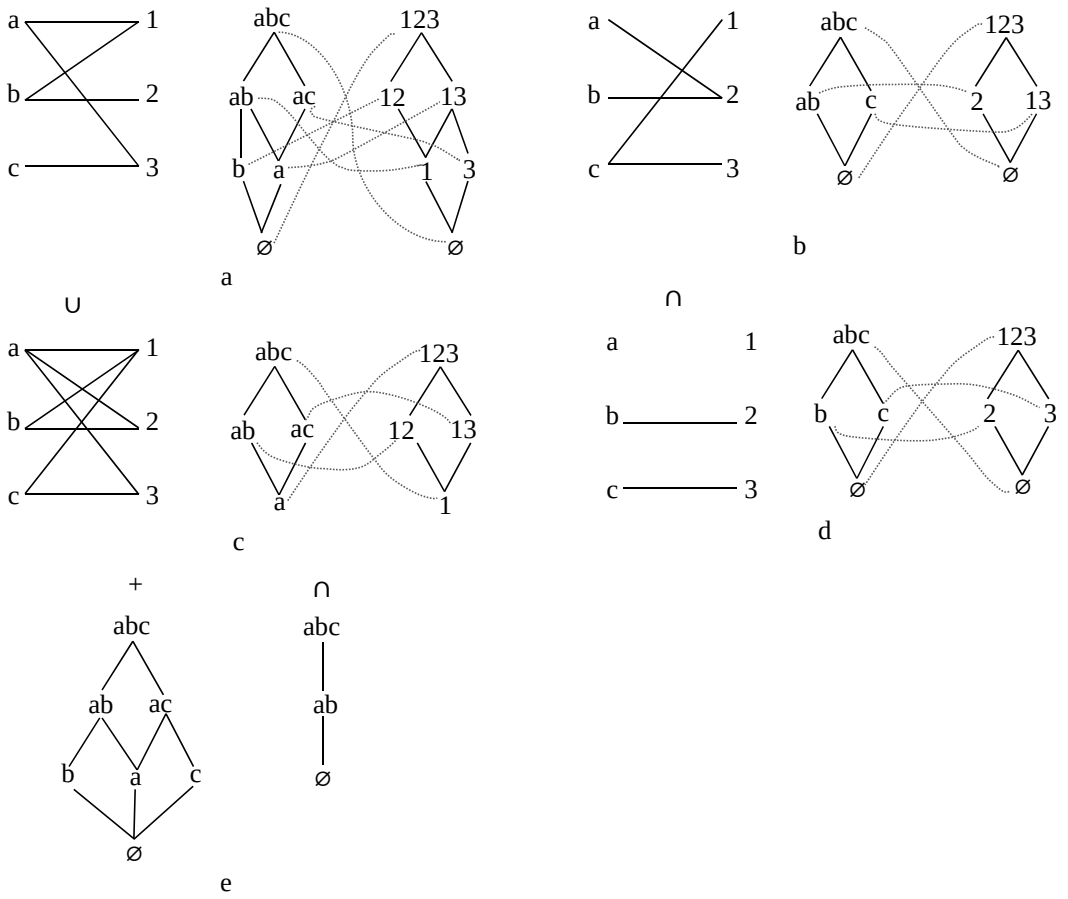
For example, in the sentence “the man was not young and in love with his neighbor on the porch”, one verb is unary (“the man was not young”), since it describes the man’s ability to be not young, and the other is binary (“the man was in love with his neighbor in entrance”). It is clear that classical actions cannot be performed with such relations.

ДИСКУРС. 2022. Т. 8, № 1. С. 133–141
DISCOURSE. 2022, vol. 8, no. 1, pp. 133–141

These two sentences are combined into one to show that we are talking about the same person in the MAN category. In these sentences, the common is the universe of subjects.

But even if these relationships were in the same universe, and then the classical approach has been called into question. In Figures a and b, respectively, show two relations of the same arity on the same universes, as well as their spaces and co-spaces. Figures c and d, respectively, show the union and intersection of these relations, as ordinary sets, and the corresponding spaces. Figure e shows the resulting spaces for mixing and intersecting spaces.

As can be seen from the figures, the space for relations union and the mixing space are significantly different. The same can be said about spaces of intersection of relations and intersection of spaces. This is due to the fact that during operations with relationships, new relationships are obtained, in which there is no memory of the previous relationship. On the contrary, when operating with spaces in the language, only temporary spaces are obtained, while the initial spaces are preserved, since their categories determine the model. Logical operations with verbs are defined by the systematization space [4] with the operations of mixing (which corresponds to the OR operation) and the intersection of spaces (which corresponds to the AND operation) specified on it. Moreover, due to Proposition 8 [4], division is defined in this space.



Actions with linguistic spaces in logical operations with verbs

The use of several verbs connected by logical connectives in a sentence means that verbs have one common universe, so actions in the systematization space are defined correctly regardless of the arity of relations corresponding to verbs.

ДИСКУРС. 2022. Т. 8, № 1. С. 133–141
DISCOURSE. 2022, vol. 8, no. 1, pp. 133–141

In the future, we will identify the open and closed forms of the language [5], bearing in mind that in the closed form, names are replaced by the values of signs and are usually not pronounced. Thus, there are no formal differences between these forms. After creating a recognition system, there is no need to literally store the linguistic space in the form of categories with an enumeration of the contents of each category. The structure of names (feature values) itself preserves the structure of the space, so the model actually stores structures with names (signs values) and verbal references to other names.

As was shown in [4], due to the independence of signs, their analysis has the form of a conjunction, so that the sequence of their analysis does not matter. They can be analyzed in parallel. We keep here the usual way of depicting linguistic space and recognition solely for the sake of convenience of presentation. Finally, we recall that from the definition of a sign and the way its values are assigned, it follows that a category can have only one non-zero value for each sign that characterizes it. Now consider the content of logical connectives with verbs.

The logical connective OR. So, mixing spaces corresponds to the operation OR with verbs of the language. By the definition of mixing spaces, the missing intersections of the categories of the two spaces must be added to the resulting space in addition to the categories of the original spaces. This means that for some categories X and Y from the original spaces, when connecting verbs with a connective OR, it is necessary to make a refinement (take the intersection) of the original categories. The intersection of categories corresponds to the addition of signs (category names). Thus, the newly arising category cannot be Σ-generator. In addition, by theorem 5 [5] in the language space of Σ-generators and ∪-forming the same, so the characteristics for a new category is obtained simply by union the meanings of categories signs that contains this intersection. To recognize the new category, you can take the name of the X category and add the missing signs values to it. In other words, for example, to the name (set of values of signs) of category X, it is necessary to add several adjectives characterizing the missing characteristics from category Y, or vice versa. So, during operations with verbs, new spaces are not specially constructed in the language. The categories are simply clarified online: by adding the necessary adjectives, the matter is limited. Before proceeding to the example, let us draw the reader's attention: without the mentioned property of Theorem 5 [5], this would be impossible! It is language, by equalizing the generators, that opens the way to these logical representations.

As an example, consider the phrase: “foreign workers must find work or be deported from the country”. This sentence is based on the action of two verbs in the phrases “workers must find work” and “immigrants will be deported out of the country”. Here there is a mixing of spaces with the categories of WORKERS and IMMIGRANTS in the universe of PEOPLE, so the category in the general phrase requires clarification. We can to clarify the category of WORKERS or the category of IMMIGRANTS. For the category of WORKERS, such a clarification is the adjective FOREIGN. After clarification of the category of WORKERS narrowed, and the phrase acquires a finished look. Of course, it is not always necessary to use a apply connective OR refinements. This follows from the definition of mixed spaces. Since spaces are mixed only on the universe of PEOPLE, we consider action OR only on this universe. The co-spaces remain unchanged, keeping the complements in their original meaning.

ДИСКУРС. 2022. Т. 8, № 1. С. 133–141
DISCOURSE. 2022, vol. 8, no. 1, pp. 133–141

The logical connective AND corresponds to the intersection of spaces representing verbs connected by the connective. As follows from the definition of this operation, only the categories that are present simultaneously in the intersected spaces remain in the resulting space. As a result, some of the categories of source spaces drop out of the list of acceptable for use, since the sentence losing sense. This is something like the phrase: “students died and went to work out on simulators”.

When using the connective AND, no qualifying signs are added to the sentence. As for OR, the speaker also does not do any special actions to form a new space. He does not even need to memorize the categories that are “out of the game” in the resulting space. This problem is again solved by the recognition system. The values of the parameters and signs of the recognition system in dead students and cheerful jocks contradict each other, which is a signal that the category is unsuitable. For example, these two categories may have different nonzero values of the same sign (and this is impossible by the definition of a sign), or they may not have common signs at all.

The fact that the nature of the appearance of logical connectives in the language is the operation of attribution [1] is manifested in the fact that these connectives, in contrast to the logical interpretation, are non-commutative. Our comic example is meaningless AND makes sense if the verbs are rearranged, since this actually changes the order of writing simple sentences in the text, and therefore the semantics of the text. Also, note that the sentence “students went to exercise on simulators” in the semantic sense is a negation of the first sentence, since Ψ (⌀, “Students died” * “Students went to exercise on simulators”) = ⌀. In other words, in order to be a negation, it is not necessary to contain the verbal “not” in its composition. In addition, it follows that if $s*s^{-1} = e$, then it is not necessarily $s^{-1}*s = e$.

In [6], Noam Chomsky laid the foundations of the theory of generative grammars. But already in the fifth section, he formulates problems due to which even the power of the grammar of immediate components may not be enough to describe the syntax of a natural language. In particular, he shows this on the example of the operation of the verbal AND. Combining sentences with the help of the verbal AND, he discovers that sometimes the result turns out to be meaningful (the students drank the juice and went to exercise on the simulators) and, therefore, grammatically correct, and sometimes it turns into some kind of nonsense (the students died and went to exercise on simulators).

We will not discuss here his attempts to explain this situation. From the point of view of R-linguistics, the essence of the matter is that the semantic interpretation of a language in a linguistic model obeys the principles of building this model, so that sometimes there is a meaning, and sometimes it is not. As a result, it turns out that for the same grammatical rule outside the grammar there is some information that determines the applicability of this rule. This fact in relation to natural languages destroys the very basis of generative grammar, since, according to this theory, all grammatical effects must be explained (generated) through the grammar itself.

We are faced here with the fundamental impossibility of applying generative grammars to natural languages without taking into account the peculiarities of the formation of meanings. This does not depend in any way on our resourcefulness or lack of sophistication in such matters. And the reason for this lies in the fact that language does not exist by itself. It exists only insofar as it can be interpreted in a model that simply does not exist in formal grammars.

ДИСКУРС. 2022. Т. 8, № 1. С. 133–141
DISCOURSE. 2022, vol. 8, no. 1, pp. 133–141

The logical connective NOT does deny the very relationship between the categories, that is, all pairs (triples) included in the Cartesian product of the subject and complement. It was defined in the previous part of this series [1, 2] when discussing the semantics of the text.

Since, by virtue of Proposition 8 [4], there is a division operation for verbs in the systematization space, then, at first glance, if the largest space (Boolean of the universe) is chosen as the dividend, then the quotient, in a sense, could be considered as a candidate for negating the divisor. However, such a definition of negation would require the use of the entire systematization space of verbs. At the same time, the considered connective have a local character in the sense that their result is determined without referring to the entire systematization space. That is why they turn out to be applicable in the language. The verbal NOT (or NOT external logic, which is the same), in contrast to classical logic, does not mean that within the framework of the systematization space, instead of one verb, some other verb should be considered – its negation: it is simply the absence of action between two categories or objects.

Verbal substitution. Let us remind that linguistic (semantic) substitution ("if…, then") means the following statement: "if we can say so…, then we can say so…". We will represent it with the symbol ⇒. It has nothing to do with the truth of judgments, and only outwardly resembles a logical syllogism. In particular, for example, the substitution A ⇒ B is not equivalent to B ⇒ A, where A and B are some sentences. This happens because linguistic verbal negation means the absence of a connection of a certain kind, which destroys the equivalence with the substitution "by contradiction". Substitutions are related to the correctness of language transformations in terms of the meaning they generate. Likewise, linguistic identity (⇔) means: "say so ... also what to say so ...". Identical expressions can be substituted one for the other, since they have an unconditional identical meaning, and not truth. In addition, linguistic identity is expressed through two linguistic substitutions in one direction and the other (by definition of linguistic identity). From the point of view of [2] for the semantics of single sentences, linguistic syllogism means that the set of pairs ("triplets" for ternary verbs) described by one phrase belongs to the set of pairs (triplets) described by another phrase.

Verbal Substitution Rule 1 looks like this:

If the space and co-space of verb 1 is less than the space and co-space of verb 2, then any meaningful sentence with verb 1 has the same meaning as a sentence with verb 2.

In other words, if $\mathbb{P}_1$ is the space of the verb S_1, and $\mathbb{P}_2$ is the space of the verb S_2, and $\mathbb{P}_1 \leq \mathbb{P}_2$, then for any categories X, Y from U and V, respectively, we have $XS_1Y \Rightarrow XS_2Y$. Indeed, by the definition of rule 1, if the category X belongs to the space $\mathbb{P}_1$, then $X \in \mathbb{P}_2$. The same for Y. Thus, if $X \times Y \subseteq S_1$, then $X \times Y \subseteq S_2$.

For example, the verb ПРИСТРАИВАТЬ is less than the verb СТРОИТЬ, since any category in the space of the verb ПРИСТРАИВАТЬ is included in the space of the verb СТРОИТЬ. Thus, the phrases "рабочие пристраивают веранду" and "рабочие строят веранду" have the same meaning. In other words, if we can say "рабочие пристраивают веранду", then we can say "рабочие строят веранду". Conversely, if we can say "рабочие строят веранду", then we can say "рабочие пристраивают веранду" provided that the categories "рабочие" and "веранда" make sense in the space and co-space of the verb ПРИСТРАИВАТЬ.

ДИСКУРС. 2022. Т. 8, № 1. С. 133–141
DISCOURSE. 2022, vol. 8, no. 1, pp. 133–141

Conclusion. So, we looked at various options for using logical connectives for verbs and the first verbal semantic substitution. Logical connectives for verbs refer to external logic, which means that these connectives describe the order of attribution and its interpretation. At the same time, as we have seen, the appearance of these bundles within sentences may require some modifications of the sentence members following from the essence of the linguistic model, and even lead to the loss of meaning. We will look at the reasons why logical connectives are carried inside sentences in the next article. In addition, we introduced the first rule of semantic substitution, which allows you to transform text without breaking semantics.

REFERENCES

1. Polyakov, O.M. (2020), "Linguistic Data Model for Natural Languages and Artificial Intelligence. Part 5. Introduction to Logic", *DISCOURSE*, vol. 6, no. 3, pp. 107–115. DOI: 10.32603/2412-8562-2020-6-3-109-117.

2. Polyakov, O.M. (2021), "Linguistic Data Model for Natural Languages and Artificial Intelligence. Part 6. The External Logic", *DISCOURSE*, vol. 7, no. 2, pp. 127–134. DOI: 10.32603/2412-8562- 2021-7-2-127-134.

3. Bezyaeva, M.G. (2002), *Semantika kommunikativnogo urovnya zvuchashchego yazyka* [Semantics of the communicative level of the sounding language], Izd-vo MGU, Moscow, RUS.

4. Polyakov, O.M. (2019), "Linguistic Data Model for Natural Languages and Artificial Intelligence. Part 2. Identification", *DISCOURSE*, vol. 5, no. 5, pp. 99–113. DOI: 10.32603/2412-8562-2019-5-5-99-113.

5. Polyakov, O.M. (2020), "Linguistic Data Model for Natural Languages and Artificial Intelligence. Part 4. Language", *DISCOURSE*, vol. 6, no. 2, pp. 107–114. DOI: 10.32603/2412-8562-2020-6-2-107-114.

6. Chomsky, N. (1962), "Syntactic Structures", *Novoe v lingvistike* [New in linguistics], vol. II, Izd-vo inostr. literatury, Moscow, USSR, pp. 412–527.

Information about the author.

Oleg M. Polyakov – Can. Sci. (Engineering) (1982), Associate Professor at the Department of Information Technology of Entrepreneurship, Saint Petersburg State University of Aerospace Instrumentation, 67 Bol'shaya Morskaya str., St Petersburg 190000, Russia. The author of over 35 scientific publications. Areas of expertise: linguistics, artificial intelligence, mathematics, database design theory, philosophy.

No conflicts of interest related to this publication were reported.
Received 07.01.2021; adopted after review 19.02.2021; published online 24.02.2022.

СПИСОК ЛИТЕРАТУРЫ

1. Polyakov O. M. Linguistic Data Model for Natural Languages and Artificial Intelligence. Part 5. Introduction to Logic // DISCOURSE. 2020. Vol. 6, № 3. P. 107–115. DOI: 10.32603/2412-8562-2020-6-3-109-117.

2. Polyakov O. M. Linguistic Data Model for Natural Languages and Artificial Intelligence. Part 6. The External Logic // DISCOURSE. 2021. Vol. 7, № 2. P. 127–134. DOI: 10.32603/2412-8562- 2021-7-2-127-134.

3. Безяева М. Г. Семантика коммуникативного уровня звучащего языка. М.: Изд-во МГУ, 2002.

4. Polyakov O. M. Linguistic Data Model for Natural Languages and Artificial Intelligence. Part 2. Identification // DISCOURSE. 2019. Vol. 5, № 5. P. 99–113. DOI: 10.32603/2412-8562-2019-5-5-99-113.

5. Polyakov O. M. Linguistic Data Model for Natural Languages and Artificial Intelligence. Part 4. Language // DISCOURSE. 2020. Vol. 6, № 2. P. 107–114. DOI: 10.32603/2412-8562-2020-6-2-107-114.

ДИСКУРС. 2022. Т. 8, № 1. С. 133–141
DISCOURSE. 2022, vol. 8, no. 1, pp. 133–141

6. Хомский Н. Синтаксические структуры // Новое в лингвистике. Вып. II. М.: Изд-во иностр. литературы, 1962. С. 412–527.

Информация об авторе.

Поляков Олег Маратович – кандидат технических наук (1982), доцент кафедры информационных технологий предпринимательства Санкт-Петербургского государственного университета аэрокосмического приборостроения, ул. Большая Морская, д. 67, лит. А, Санкт-Петербург, 190000, Россия. Автор более 35 научных публикаций. Сфера научных интересов: лингвистика, искусственный интеллект, математика, теория проектирования баз данных, философия.

О конфликте интересов, связанном с данной публикацией, не сообщалось.
Поступила 07.01.2021; принята после рецензирования 19.02.2021; опубликована онлайн 24.02.2022.