ISNN 3030-3907

2026
JUNE
ISSUE - 6
VOLUME - 9

DEVELOPMENT OF SCIENCE

scientific journal



Development of science

Ilmiy jurnal

IYUN 2026
Issue - 6, VOLUME – 9

ISSN 3030 -3907

Ilmiy jurnal OAK rayosatining 2024 yil 27 sentabrdagi 361-son qaroriga asosan 02.00.00 kimyo fanlari hamda 2025 yil 12 fevral 367- son qaroriga asosan 05.00.00 texnika fanlari bo‘yicha falsafa doktori (PhD) va fan doktori (DSc) ilmiy darajasiga talabgorlarning dissertatsiya ishlari yuzasidan dissertatsiyalari asosiy ilmiy natijalarini chop etish tavsiya etilgan milliy ilmiy nashrlar ro‘yxatiga kiritilgan.

BOSH MUHARRIR:
Temirov Alisher Hoshim o‘g‘li – texnika fanlari falsafa doktori (PhD),Dotsent.
TAHRIRIYAT KENGASH RAISI:
Fozilov Sadriddin Fayzullayevich – texnika fanlari doktori, professor
TAHRIRIYAT KENGASH A’ZOLARI

TEXNIKA FANLARI/ TECHNICAL SCIENCES/ ТЕХНИЧЕСКИЕ НАУКИ

Bozorov G‘ayrat Rashidovich
Buxoro Davlat Texnika Universiteti, professor
Abdurahmonov Olim Rustamovich
Buxoro Davlat Texnika Universiteti, professor
Adizov Bobur Zamirovich
O‘zbekiston Respublikasi Fanlar akademiyasi Umumiy va noorganik kimyo instituti, professor
Hayitov Ruslan Rustamjonovich
Buxoro Davlat Texnika Universiteti, professor
Jumayev Qayum Karimovich
Buxoro Davlat Texnika Universiteti, professor.
Berdiyev Dorob Murodovich
Islom Karimov nomidagi Toshkent Davlat Texnika Universiteti, professor
Maxmudov Rafiq Amonovich
Buxoro Davlat Texnika Universiteti, professor.
Tilloyev Lochin Ismatilloyevich
Buxoro Davlat Texnika Universiteti, dotsent.
Raxmonov Qahramon Sanakulovich
Buxoro Davlat Texnika Universiteti, dotsent.
Barakayev Nusratullo Rajabovich
Islom Karimov nomidagi Toshkent Davlat Texnika Universiteti, professor.
Jo‘rayev Xayrullo Fayziyevich
Buxoro Davlat Texnika Universiteti, professor.
Panoyev Erali Rajabboyevich
Buxoro Davlat Texnika Universiteti, dotsent.
Ataullayev Sherzod Nabiullyevich
Buxoro Davlat Texnika Universiteti, dotsent.
Murodov Malik Negmurodovich
Buxoro Davlat Texnika Universiteti, dotsent.
Axmedova Ozoda Bahronovna
Islom Karimov nomidagi Toshkent Davlat Texnika Universiteti.
Safarov Bahri Jumayevich
Buxoro Davlat Texnika Universiteti, dotsent.
Xaydarov Axtam Amonovich
Buxoro Davlat Texnika Universiteti, dotsent.
Safarov Jasur Alijon o‘g‘li
Buxoro Davlat Texnika Universiteti, dotsent.

KIMYO FANLARI/ CHEMICAL SCIENCES/ ХИМИЧЕСКИЕ НАУКИ

Axmedov Voxid Nizomovich
Buxoro Davlat Texnika Universiteti, professor
Umarov Baqo Bafoyevich
Buxoro Davlat Universiteti, professor
Kodirov Abduaxad Abduraximovich
Qarshi davlat Universiteti, professor
Qurbonov Mingniqul Jumagulovich
Qarshi davlat Universiteti, dotsent
Turemuratov Sharibay Naurizbaevich
O‘zbekiston Respublikasi Fanlar akademiyasi Qoraqalpog‘iston bo‘limi Bosh ilmiy kotibi, professor
Tursunov Kaxor Shonazarovich
Qarshi davlat Universiteti, professor
Do‘stov Hamro
Buxoro Davlat Texnika Universiteti, professor
Panoyev Nodir Shavkatovich
Buxoro Davlat Texnika Universiteti, dotsent.
Olimov Bobur Bahodir o‘g‘li
Buxoro Davlat Texnika Universiteti, dotsent.
Kasimov Sherzod Abduazimovich
Termiz Davlat Universiteti, professor.
Aliqulov Rustam Valiyevich
Termiz Davlat Universiteti, professor
Tropov Xamza Tursunovich
Samarqand Davlat Universiteti, professor

IQTISODIYOT FANLARI/ ECONOMICS/ ЭКОНОМИЧЕСКИЕ НАУКИ

Nizamov Aslitdin Badreddinovich
Buxoro Davlat Texnika Universiteti, professor.
Boboyev Akmal Choriyevich
Buxoro Davlat Texnika Universiteti, professor.
Ergashxojayeva Shaxnoza Djasurovna
Toshkent davlat iqtisodiyot universiteti, professor
Saidaxmedova Nodira Ilxomovna
Toshkent davlat iqtisodiyot universiteti, professor
Muxsinov Bekzod Toxirovich
Buxoro Davlat Texnika Universiteti, dotsent.
Shakarov Qulmat Ashirovich
Renessans ta'lim universiteti,dotsent.
Jumaeva Zulfiya Qayumovna
Buxoro Davlat Texnika Universiteti, dotsent.
Norova Salomat Yusupovna
Buxoro Davlat Texnika Universiteti, dotsent.
Avezova Shaxnoza Maxmudjonovna
Buxoro Davlat Texnika Universiteti, dotsent.
Jalilov Jamshid G‘anijonovich
Toshkent davlat iqtisodiyot universiteti, dotsent
Xasanova Gulruh Djumanazarovna
Buxoro Davlat Texnika Universiteti, dotsent.

PEDAGOGIKA FANLARI/ PEDAGOGICAL SCIENCES / ПЕДАГОГИЧЕСКИЕ НАУКИ

Siddiqova Sadoqat G‘afforovna
Buxoro Davlat Texnika Universiteti, dotsent.
Zoyirov Erkin Xalilovich
Buxoro Davlat Texnika Universiteti, DSc, professor
Gaffarov Feruz Hasanovich
Puchon Universiteti, *professor*
Ruziyev Davron Yuldashevich
Buxoro Davlat Universiteti, professor
Muradova Firuza Rashidovna
Buxoro Davlat Texnika Universiteti, professor.
Yunusova Gulandon Samiyevna
Buxoro Davlat Texnika Universiteti, professor.
To‘xtayeva Zebo Sharifovna
Buxoro Davlat Texnika Universiteti, professor.
Navro‘zova Gulchixra Nigmatovna
Buxoro Davlat Texnika Universiteti, dotsent.
Uzoqova Laylo Polvonovna
Buxoro muhandislik – texnologiya instituti, dotsent.
To‘raqulova Marjona Qiyom qizi
Buxoro Davlat Texnika Universiteti, dotsent.
Muradova Zarina Rashidovna
Buxoro Davlat Texnika Universiteti, dotsent.
Murodov Sanjar Aslonovich
Buxoro Davlat Texnika Universiteti, dotsent.
Shodiyev Jahongir Jo‘raqulovich
Buxoro Davlat Texnika Universiteti, dotsent
Gaffarov Laziz Hasanovich
Buxoro muhandislik – texnologiya instituti, dotsent.

Tahririyat a’zolarning to‘liq ro‘yxati jurnalning www.devos.uz saytida keltirilgan.

Tahririyat hay’ati a’zolarining tarkibini o‘zgartirish hamda qo‘shimcha kiritish yoki tahririyat a’zoligidan chiqarish bosh muharrir tomonidan amalga oshiriladi.

ISSN 3030-3907

Development of science ilmiy jurnali 2024 yil 15 aprel kuni №256445 sonli guvohnoma bilan Davlat ro‘yxatidan o‘tgan.

To‘plam elektron ko‘rinishda (pdf formatida) mualliflarga taqdim etiladi. To‘plamga kiritiladigan maqolalarning mazmuni, ulardagi statistik ma’lumotlar va me’yoriy hujjatlarning to‘g‘riligiga hamda har qanday fikr - mulohazalarga mualliflarning o‘zlari mas’ul hisoblanadilar. Talablarga javob bermaydigan maqolalar to‘plamga kiritilmaydi. Tashkiliy qo‘mita maqola matnini qisqartirish, qisman tuzatish kiritish, seksiyalarga joylashtirish huquqiga ega.

ISSN 3030-3907

DEVELOPMENT OF A NEURO-SYMBOLIC ALGORITHM FOR IDENTIFYING THE SUBJECT AND ATTRIBUTE IN UZBEK SENTENCES

Sharipov M.S.
Tashkent University of Information Technologies named after Muhammad al-Khwarizmi Urgench State University named after Abu Rayhan Beruni
E-mail: maqsbek72@gmail.com
Avezmatov I.D.
Tashkent University of Information Technologies named after Muhammad al-Khwarizmi Urgench State University named after Abu Rayhan Beruni
E-mail: ixtiyoravezmatov07@gmail.com
Bakhtiyorova M.N.
Tashkent University of Information Technologies named after Muhammad al-Khwarizmi Urgench State University named after Abu Rayhan Beruni
E-mail: madinabonubaxtiyorova448@gmail.com
Alimbaeva Y.A.
Tashkent University of Information Technologies named after Muhammad al-Khwarizmi Urgench State University named after Abu Rayhan Beruni
E-mail: alimbayevayulduzxon@gmail.com

Annotatsiya. Ushbu maqolada o‘zbek tilidagi matnlarda gapning asosiy sintaktik qismlari hisoblangan ega va aniqlovchini avtomatik aniqlash uchun samarali va ishonchli neyro-simvolik algoritm hamda dasturiy pipeline arxitekturasi ishlab chiqilishi batafsil bayon etiladi. Taklif etilayotgan hisoblash lingvistikasi yechimi uchta asosiy ketma-ket bosqichdan iborat: o‘zbek tilining agglyutinativ tabiatiga moslashtirilgan maxsus leksik tokenizatsiya tizimi; qo‘shimcha o‘qitilgan va nozik sozlangan BERT modeli asosidagi so‘z turkumlarini aniqlash (POS tagging) arxitekturasi; hamda aniq va murakkab lingvistik qoidalarga asoslangan sintaktik rollarni ajratish moduli. Matnda egani xatosiz aniqlash barqarorligini ta’minlash uchun tizim avvalgi bosqichda aniqlangan kesimga tayanadi. Aniqlovchini topishda esa qaratqich kelishigi va sifatlar kabi morfologik-mantiqiy zanjirlardan foydalaniladi. Morfologik tahlil asosida kesim aniqlangach, ega va aniqlovchini belgilovchi maxsus sintaktik qoidalar tizim tomonidan ishga tushiriladi. Baholash ishlari O‘zbekistondagi turli yangiliklar va ilmiy maqolalardan olingan, ekspertlar tomonidan qo‘lda annotatsiya qilingan 1000 ta gapdan iborat katta test to‘plamida amalga oshirildi. Natijada taklif etilgan tizim ega aniqlashda o‘rtacha 85.4% va aniqlovchini aniqlashda 87.2% F1-score natijani qayd etdi.

Kalit so‘zlar: tabiiy tilni qayta ishlash, o‘zbek tili, sintaktik tahlil, ega, aniqlovchi, neyro-simvolik yondashuv, BERT, agglyutinativ til morfologiyasi.

Аннотация. В данной статье подробно рассматривается разработка эффективного и надежного нейро-символического алгоритма и архитектуры программного конвейера для автоматического определения подлежащего и определения, которые считаются основными синтаксическими компонентами предложений в узбекских текстах. Предлагаемое решение в области компьютерной лингвистики состоит из трех основных последовательных этапов: специализированной системы лексической токенизации, адаптированной к сложной природе узбекского языка как агглютинативного; архитектуры разметки частей речи (POS tagging) с использованием дополнительно обученной и донастроенной модели BERT; а также модуля извлечения синтаксических ролей, основанного на точных и сложных лингвистических правилах. Для обеспечения стабильности безошибочного определения подлежащего система опирается на сказуемое, выявленное на предыдущем этапе. При определении определения используются морфолого-логические цепочки, такие как родительный падеж и прилагательные. После определения сказуемого на основе морфологического анализа система активирует специальные синтаксические правила для выявления подлежащего и определения. Оценка проводилась на большом тестовом наборе из 1000 предложений, собранных из различных новостных материалов и научных статей в Узбекистане и размеченных вручную экспертами. В результате предложенная система показала средний F1-score 85,4% для определения подлежащего и 87,2% для определения определения.

Ключевые слова: обработка естественного языка, узбекский язык, синтаксический анализ, подлежащее, определение, нейро-символический подход, BERT, морфология агглютинативных языков.

Abstract. This article details the development of an efficient and reliable neuro-symbolic algorithm and software pipeline architecture designed for the automatic identification of the subject and attribute, which are considered the main syntactic parts of sentences in Uzbek texts. The proposed computational linguistics solution consists of three main sequential stages: a specialized lexical tokenization system directly adapted to the complex nature of the Uzbek language, characteristic of the agglutinative language family; a Part-of-Speech (POS) tagging architecture using a additionally trained and fine-tuned BERT model; and a syntactic role extraction engine based on precise and complex linguistic rules. To ensure the stability of error-free subject identification within the text, the system relies on the predicate identified in the previous stage. Furthermore, to identify the attribute, it refers to morphological-logical chains such as the genitive case and adjectives. Once the predicate is found based on morphological analysis, specific syntactic rules

ISSN 3030-3907

defining the subject and attribute are triggered by the system. Evaluation was carried out on a large test set consisting of 1000 sentences collected from various news reports and scientific articles in Uzbekistan, manually annotated by experts. As a result, the proposed system recorded an overall average F1-score of 85.4% for subject identification and 87.2% for attribute identification.

Keywords: natural language processing, Uzbek language, syntactic parsing, subject, attribute, neuro-symbolic approach, BERT, agglutinative language morphology.

1. INTRODUCTION.

The field of Natural Language Processing (NLP) has developed rapidly in recent years with the introduction of deep neural networks and transformer architectures (e.g., BERT, GPT) [1]. However, these successes have mostly been achieved for resource-rich languages where large amounts of data are continuously collected. Languages like Uzbek, which have complex agglutinative morphology and belong to the list of low-resource languages, remain a challenging problem for NLP experts [2].

Syntactic analysis is an initial stage that plays an important role in the practical pipeline of natural language processing, serving as a foundation for machine translation, AI-driven semantic analysis, and question-answering systems. In Uzbek syntax, the Predicate is considered the central element of the analyzed sentence and serves as the pivot. Discovering the Subject and the Attribute subordinate to nominals based on it is considered a crucial step [3]. The Subject is the principal part of the sentence that performs the main action or possesses a certain static state. The Attribute, on the other hand, is an important secondary part that always denotes a sign of a person, object, or event (color, degree, quantity, characteristic) as well as its ownership.

Although existing purely rule-based systems work quite accurately, they become rigid when sentence structures change unconventionally or when things are misspelled. Purely neural (Deep learning) approaches, however, cannot always perform correct extraction amidst the complex order of grammatical affixes in the Uzbek language [4]. Taking the above into consideration, this research work aimed to introduce a modern neuro-symbolic algorithm that equally relies on a neural network model and finely-crafted hand-written syntactic rules.

2. MORPHOLOGY AND SYNTAX OF THE SUBJECT AND ATTRIBUTE IN THE UZBEK LANGUAGE.

To find the subject and attribute in the Uzbek language, the program first makes rapid reference to morphological priorities hidden behind the word, and then to syntactic chain links with other words directly within the sentence context. In computational linguistics, the Subject generally refers to nominative case forms

ISSN 3030-3907

written without specific case affixes. It manifests with nouns that take plural (e.g., -lar) or possessive affixes, as well as personal and demonstrative pronouns.

The Attribute, on the other hand, occupies a broader morphological spectrum. In the Uzbek language, there are three main forms of expressing the attribute: adjective (e.g., "red apple"), numeral ("five people"), and genitive attributes using the genitive case ("my book"). The confusing part for computational analysis is that sometimes adjectives or numerals can turn completely into nouns (through the process of substantivization) and take on the role of the subject ("Good people are honored"). Alternatively, hidden attributes that come without the genitive marker require complex logic from the computer ("school yard"). Therefore, combining the power of contextual models like BERT with grammatical rules can provide a sufficient solution to this problem.

3. RESEARCH METHODOLOGY AND NEURO-SYMBOLIC ALGORITHM.

The architecture we propose in practice is a three-part software processing pipeline. In the first preparatory stage, a Custom Tokenizer designed for the Uzbek language is applied, which can process complex and compound words, unconventional punctuation marks, and particularly the apostrophe modifier sign (tutuq belgisi). The second stage consists of determining which word belongs to which part-of-speech contextually (POS-tagging) based on the BERT (transformers) neural network model. The third operational stage is a symbolic rule engine that refers to explicit logical rules. This is why this approach was named Neuro-symbolic (Neural Network + Symbolic Rules).

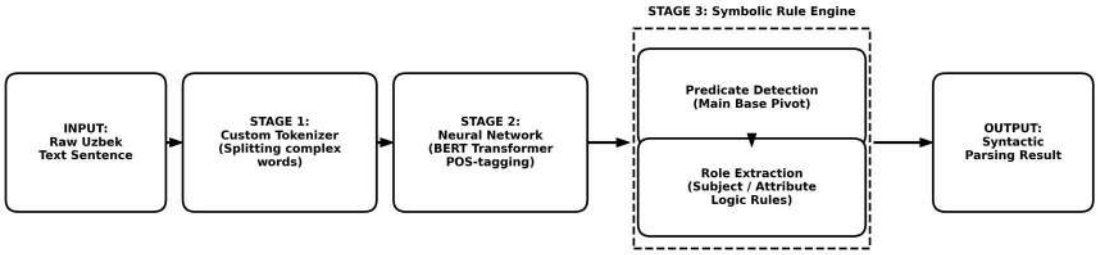


Figure 1. Neuro-symbolic 3-stage pipeline architecture of the system.

We robustly integrated the fast predicate detection module presented in our previous studies into the system models [5]. Once the predicate is determined, other arguments are identified through backtracking. Below are some symbolic rules that differentiate the subject and attribute, presented as examples:

Primary Rule 1 (Noun-based main pivot - Subject): If a probable word in the text is identified via the neural network as NOUN or PRON (pronoun), and if it does not appear in accusative, dative, or locative-temporal cases, and if the person

ISSN 3030-3907

and number of that word perfectly correspond with the identified predicate variable with absolute confirmation, the word in this region is directly tagged as the Subject.

Production Rules (Formal Logic representation):

Rule 1 (Subject Identification):
IF Token.POS ∈ {NOUN, PRON}
AND Token.Case ∉ {ACC, DAT, LOC}
AND Match(Token.PersonNumber, Predicate.PersonNumber) == True
THEN SyntacticRole(Token) = SUBJECT

Rule 2 (Adjectival Attribute):
IF Token.POS ∈ {ADJ, NUM}
AND Next(Token).POS == NOUN
THEN SyntacticRole(Token) = ATTRIBUTE

Rule 3 (Genitive Attribute):
IF Token.Case == GENITIVE
AND CheckRelation(Token, FollowingNominals) == True
THEN SyntacticRole(Token) = ATTRIBUTE

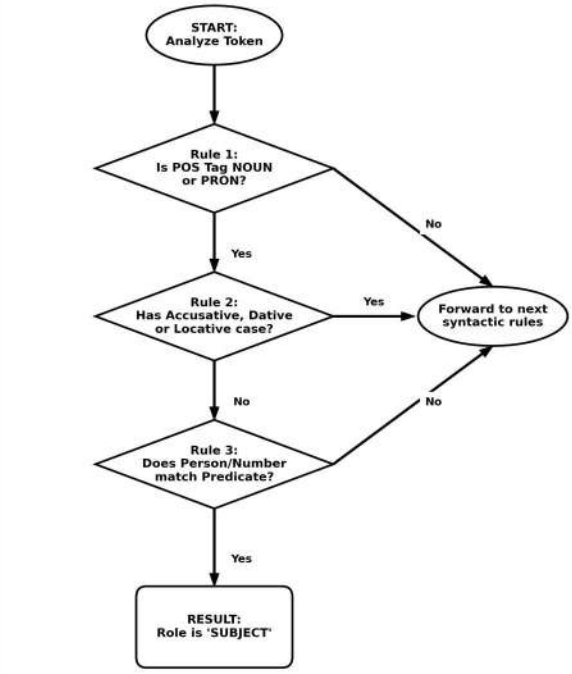


Figure 2. Flowchart illustrating the logical rules for Subject identification.

Direct Rule 2 (Adjectival Attribute chain): If there is a high probability that the analyzed token belongs to the ADJ (adjective) or NUM (numeral) part of speech according to the BERT model, and if the absolute nearest first word lying to its right is a noun (even if it comes with some case or possessive affix), the preceding token is conditionally transferred to the "Attribute" role.

ISSN 3030-3907

Morphological Rule 3 (Genitive Attribute): If a symbolic rule witnesses that the last affix of a token takes the genitive case morpheme (-ning), and it appears to be connected to subsequent nominals in the text, such as the subject, object, or adverbial modifier of the sentence, it is strictly registered as an Attribute.

4. Established experimental conditions and their scientific results

In order to conclude and fully evaluate the stable functioning capability of the intelligent syntactic analyzer system for the Uzbek language under real conditions, we conducted separate trials. In doing so, we created a high-quality test corpus consisting of a total of 1000 complex sentences collected from popular social networks, open research lectures, blogs, and official news sites. Within each sentence, the locations of the subject and attribute were manually annotated by an expert group, transforming it into a "Gold Standard" test set. The research evaluation results were consistently measured using globally recognized classic NLP metrics - Precision, Recall, and the generalized F1-score.

Table 1.

Experimental results of the subject and attribute identification model based on the neuro-symbolic algorithm in the Uzbek language.

Metric	Subject	Attribute
Precision (%)	86.4	88.1
Recall (%)	84.4	86.3
F1-Score (%)	85.4	87.2

It is evident from the collected experimental results that the developed algorithm was able to record good accuracy in finding both syntactic members - the subject and the attribute. In particular, due to the presence of specific morphemes in adjectives and genitive cases, the Attribute indices came out slightly more accurate than the Subject results.

5. CONCLUSION.

This article presented a modern neuro-symbolic system developed for the automatic in-memory identification of the Subject with a central property and the secondary Attribute attaching to the base of nominals in any context within ambiguous grammatical texts of the complex Uzbek language. The reliable system directly combining transformer models with specific linguistic logic rules is a practical expression of hybrid technology. The created analysis model was able to record high results under conditions utilizing genuine real data, flawlessly achieving a generalized average F1-score of 85.4% for the Subject and 87.2% for the Attribute. These ongoing investigations signify an essential step towards transforming our thorough algorithm into a highly intellectual, comprehensive text parsing pipeline being developed for the Uzbek language.

ISSN 3030-3907

REFERENCES

1. Devlin, J., Chang, M. W., Lee, K., & Toutanova, K. (2018). BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding. In Proceedings of NAACL-HLT 2019 (pp. 4171-4186).
2. Kuriyozov, E., Matlatipov, S., Alonso, H. M., & Gómez-Rodríguez, C. (2020). Construction and evaluation of BPE embeddings for the Uzbek language. In Proceedings of the 7th Workshop on NLP for Similar Languages, Varieties and Dialects.
3. Sharipov M., Avezmatov I., Adinaev H. (2025). Development of a rule-based model and algorithm for predicate identification in Uzbek language texts. 10th International Conference on Computer Science and Engineering (UBMK). IEEE. pp. 594-598.
4. D'Avila Garcez, A., & Lamb, L. C. (2020). Neurosymbolic AI: The 3rd wave. Artificial Intelligence Review, 56(11), 12387-12406.
5. Sharipov M., Kuriyozov E., Yuldashev O., Sobirov O. (2024). UzbekVerbDetection: Rule-based detection of verbs in Uzbek texts. Proceedings of LREC-COLING 2024. pp. 17343-17347.

ISSN 3030-3907