**O‘ZBEKCHA SO‘ZLARNING GRAMMATIK TAHLIL QILISH ALGORITMLARI**

**Muhammadsolih TURSUNOV**

Muhammad al-Xorazmiy nomidagi Toshkent axborot texnologiyalari universiteti Samarqand filiali katta o‘qituvchisi, PhD

E-mail: [muhammadsolih927@gmail.com](mailto:muhammadsolih927@gmail.com)

ORCID: 0000-0002-6485-3630

**Shirinboy MAXMIDOV**

Muhammad al-Xorazmiy nomidagi Toshkent axborot texnologiyalari universiteti Samarqand filiali assistenti

**Bekzod MUHAMMADIYEV**

Muhammad al-Xorazmiy nomidagi Toshkent axborot texnologiyalari universiteti Samarqand filiali magistri

**Annotatsiya**: Ushbu maqolada o‘zbekcha so‘zni grammatik tahlil qilish algoritmlari, usullarni tavsiflash va kodlashga umumiy yondashuv muhokama qilinadi. O‘zbek tilining boy morfologik va sintaktik tizimi, shu jumladan qo‘shimchalarning keng turlari va so‘z turkumlari grammatik tahlil qilish jarayonini murakkablashtiradi. Shuning uchun so‘zlarni tarkibiy qismlarga ajratish, ularning ma’nosini va grammatik vazifalarini aniqlash, so‘zning turkumi va grammatik tahlilini belgilash kabi jarayonlar uchun maxsus algoritmlar ishlab chiqilishi talab etiladi. Ishda so‘zni grammatik tahlil qilishda dastavval so‘z shaklidan so‘zning asosini va qo‘shimchalar qismini ajratib olish. Undan so‘ng qo‘shimchlarni bo‘laklarga ajratish hamda so‘z asosida mos turkumni aniqlash hamda qo‘shimchalar tahlilini bajarishdan iborat.

Tadqiqotda asosan so‘zlarning o‘zagini aniqlash, qo‘shimchalarning turli turlari va ularning so‘z ma'nosi hamda grammatik shakllariga ta'siri, shuningdek, so‘zlarning fe’l, ot, sifat, son kabi turkumlarga ajratilishi masalalari ko‘rilgan. Ushbu algoritmlar o‘zbek tilidagi matnlarni avtomatik qayta ishlash, tarjima tizimlari, matn tahrirlash va tahlil qilishda keng qo‘llanilishi mumkin.

**Kalit soʻzlar:** O‘zbek tili, so‘z turkumlari, qo‘shimchalar, so‘z shakli, so‘z asosi, model, algoritm, ma’lumotlar bazasi, lug‘at, grammatika, morfologik tahlil, avtomatik tahlil.

**ALGORITHMS FOR GRAMMATIC ANALYSIS OF UZBEK WORDS**

**Abstract**. This article discusses algorithms for grammatical analysis of Uzbek words, description of methods and a general approach to coding. The rich morphological and syntactic system of the Uzbek language, including a wide variety of affixes and word groups, complicates the process of grammatical analysis. Therefore, it is necessary to develop special algorithms for such processes as dividing words into components, determining their meaning and grammatical functions, determining the category and grammatical analysis of a word. At work, in the grammatical analysis of the word, first of all, it is necessary to separate the base of the word and the suffixes from the word form. After that, it consists of dividing the adverbs into parts and determining the appropriate category based on the word and performing the analysis of the adverbs.

The study focuses mainly on identifying the root of words, different types of suffixes and their influence on word meaning and grammatical forms, as well as the issues of dividing words into categories such as verbs, nouns, adjectives, and numbers. laughed These algorithms can be widely used in automatic processing of Uzbek texts, translation systems, text editing and analysis..

**Keywords**: Uzbek language, word groups, suffixes, word form, word base, model, algorithm, database, dictionary, grammar, morphological analysis, automatic analysis*.*

**АЛГОРИТМЫ ГРАММАТИЧЕСКОГО АНАЛИЗА УЗБЕКСКИХ СЛОВ**

**Аннотация**. В данной статье рассматриваются алгоритмы грамматического анализа узбекских слов, описание методов и общий подход к кодированию. Богатая морфологическая и синтаксическая система узбекского языка, включающая большое разнообразие аффиксов и словосочетаний, усложняет процесс грамматического анализа. Поэтому необходимо разрабатывать специальные алгоритмы для таких процессов, как деление слов на компоненты, определение их значения и грамматических функций, определение категории и грамматический анализ слова. При работе, при грамматическом анализе слова, в первую очередь, необходимо отделить основу слова и суффиксы от словоформы. После этого он состоит из деления наречий на части и определения соответствующей категории на основе слова и проведения анализа наречий.

Основное внимание в исследовании уделяется выявлению корня слова, различных типов суффиксов и их влиянию на значение слова и грамматические формы, а также вопросам деления слов на такие категории, как глаголы, существительные, прилагательные и числа. laughed Данные алгоритмы могут быть широко использованы в автоматической обработке узбекских текстов, системах перевода, редактировании и анализе текстов..

**Ключевые слова:** Узбекский язык, словосочетания, суффиксы, словоформа, база слов, модель, алгоритм, база данных, словарь, грамматика, морфологический анализ, автоматический анализ.

O‘tgan asrning 90-yillarida morfologik tahlil *generativ model, paradigmatik model, ikki tarkibli morfologik model* [1] kabi metodlarga asoslanilgan. O‘ttiz yildan oshiqroq vaqt mobaynida kompyuter morfologiyasi uchun turli texnologiyalarni yaratishga doir tadqiqotlar olib borilmoqda. Tillar uchun morfologik tahlil qilishning o‘nlab algoritmlari bo‘lishi tabiiy hol. Algoritmning bir yo‘nalishi so‘z oxiridagi harfiy birikmalarni qamrab oluvchi ma’lumotlar bazasidan foydalanadi. Bunda ehtimollik darajasiga ko‘ra, lingvistik resursdan olingan natijalar statistik tahlil qilinadi. Avtomatik usulda morfologik tahlil qilish jarayoni quyidagi turlarga ko‘ra tasniflanadi [2]:

1. Mantiqiy ko‘paytirish metodi orqali morfologik analiz.
2. Lug‘atsiz jadval orqali morfologik analiz.
3. So‘zlar asosiga asoslangan morfologik tahlil (so‘zlar asosiga tayanilgan morfologik tahlil Yevropa tillari uchun moslashtirilgan. Unda qo‘shimcha yordamchi jadvallardan foydalaniladi).
4. So‘zshakllarga (morfotaktik qoidalarga ko‘ra) ko‘ra morfologik analiz.

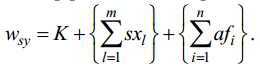
Yuqoridagilardan so‘zshakllariga ko‘ra morfologik tahlil qilish usuli keng qo‘llanadi. Ushbu algoritmga muvofiq so‘zlar morfemalarga ajratilganda qidirilayotgan so‘z lug‘atdan topilmaslik ehtimoli kuzatiladi. Bu holatda o‘ngdan chapga qarab tahlil qilish jarayonida ko‘makchi morfemalarning turi va modeliga qarab uning qaysi so‘z turkumiga oidligi aniqlanadi. U holda o‘sha so‘zning qaysi so‘z turkumiga tegishli ekanligi belgilanadi [3]. So‘z shakllarni tahlil qilish jarayonida quyidagi bosqichlar amalga oshiriladi:

1. So‘zshakl morfem segmentlarga ko‘ra zanjirsimon birikadi.
2. So‘zshakl morfem segmentlar zanjiri sifatida tartiblanmagan yangi elementlar to‘plami - morfemalarga almashtiriladi. So‘z shakllar morfemalar to‘plami sifatida ma’lumotlar yoziladi.
3. Ushbu ma’lumot so‘z shakllari haqidagi yakuniy natija bilan qayta to‘ldiriladi.
4. Mantiqiy ko‘paytirish metodi orqali morfologik tahlil qilish flektiv tillar uchun foydalaniladi

O‘zbek tilidagi so‘z shakllarini grammatik asosini aniqlash muhim ahamiyatga ega.

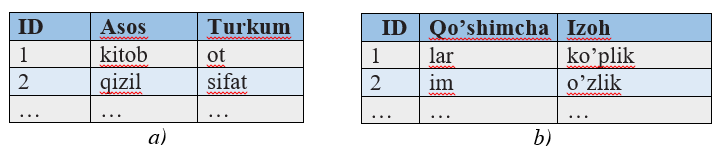
O‘zbek tili so‘z shakllarini grammatik tahlil qilish morfotaktik qoidalarga asoslangan holda amalga oshiriladi.

O‘zbek tilidagi so‘z shakllarini morfologik asosini aniqlashda so‘zning strukturasi muhim ahamiyat kasb etadi. Shu munosabat bilan quyidagi matematik ifodalar orqali tasvirlangan lingvistik qoidalar asosidagi affiksasiya usuli bilan yasalgan so‘z shakllarining morfemik modellari taklif etildi:



Bu yerda K- sodda tub so‘z (o‘zak), sx - so‘z yasovchi affiks (suffiks), af - affikslar, - ketma-ket konkatenasiya operasiyasi, l – suffikslar miqdori indeksi, i – affikslar miqdori indeksi - sodda yasama so‘z. Fikrimizni dalillash uchun ot so‘z turkumidagi «gul-don-lar-imiz-ni» so‘z shaklini tahlil qilamiz. Bu misolda K- «gul», sx - «-don», af - «-lar «, «-imiz»,»-ni «, «gul-don-lar-imiz-ni» - wsy so‘z shakli [4].

Bunda so‘z shaklidan so‘zning ajratib olinadi, hamda keyingi qadamda qo‘shimchalardan tashkil topga bo‘lagi tahlil qilinadi. So‘z shaklini grammatik tahlil qilish jarayonida albatta ma’lumotlar bazasiga murojaat qilinadi. Biz ilmiy ishimiz uchun ma’lumotlar bazasida ikkita jadvaldan foydalanganmiz: 1) so‘zlar asosi lug‘ati, 2) qo‘shimchalar lug‘ati (1-rasm). So‘zlar asosi lug‘ati 3ta maydondan iborat: 1) tartib raqam (So‘z asosining IDsi), 2) So‘z asosi, 3) So‘z asosining turkumi. Qo‘shimchalar lug‘ati ham 3ta maydondan iborat: 1) tartib raqam (qo‘shimchaning IDsi), 2) Qo‘shimchaning nomi, 3) Qo‘shimchaning grammatik ma’noga ega izohi.



1-rasm. a) So‘zlar asosi lug‘ati, b) Qo‘shimchalar lug‘ati

So‘zlarni ma’lumotlar bazasidan o‘qib olish uchun ularning simvollar ketma- ketligiga amal qilinadi. Belgilash o‘sha so‘z yaxlit birlik sifatida aniqlanmagunicha bu holat takrorlanaveradi. Morfologik tahlil uchun komponent tahlil metodidan ham foydalanish mumkin. Unga ko‘ra so‘zning uzunligi aniqlanadi. Shu tarkibdagi simvollar soniga qarab ma’lumotlar bazasidan qidiriladi. So‘zning asosini topish va uni grammatik tahlilini qililshda ishlab chiqilgan algoritm quyidagi ketma ketlikdan iborat:

1-qadam. Dastur ishini boshlash.

2-qadam. *<input>* tegiga so‘z shakli kiritiladi.

3-qadam. Input tegiga kiritilgan so‘z shakli *“suz”* nomli o‘zgaruvchiga ta’minlanadi.

4-qadam. *“suz”* o‘zgaruvchining qiymati, *“suz1”* nomli o‘zgaruvchiga ta’minlanadi, ya’ni so‘zning nusxasi yaratiladi.

5-qadam. *“l”* o‘zgaruvchisiga *“suz1”* o‘zgaruvchisi qiymatining uzunligi, ya’ni so‘z shaklining belgilar soni ta’minlanadi.

6-qadam. *“suz1”* o‘zgaruvchisiga ta’minlangan so‘z shaklidan so‘z asosini aniqlash algoritmi boshlanadi. Bunda berilgan so‘z shakli ma’lumotlar bazasidan qidiriladi. Agar mavjud bo‘lsa 11-qadamga o‘tiladi, aks holda so‘zning oxiridan boshlab bittadan belgi kesib tashlash tartibida lug‘atdan qayta qidirish shaklida siklimiz davom etadi. Bunda sikl aylanishi toki so‘z asosi topilgunga qadar yoki *“suz1”* o‘zgaruvchining qiymati bo‘sh bo‘lib qolgunga qadar davom etadi. *“suz1”* o‘zgaruvchisining qiymati bo‘sh bo‘lib qolsa, u holda so‘z shaklidan uning asosini aniqlashni imkoni bo‘lmagan, va “so‘z asosini aniqlab bo‘lmadi” degan xabarni berib, 12-qadamga o‘tiladi. *“suz1”* o‘zgaruvchisidagi so‘z shaklidan uning asosi topilsa siklimiz rost qiymatni olib, davom etish to‘xtatiladi va 7-qadamga o‘tiladi.

7-qadam. *“suz1”* o‘zgaruvchisidagi so‘z shaklidan uning asosi o‘zgaruvchining o‘ziga ta’minlangan bo‘ladi. “suz” o‘zgaruvchisi qiymatidan *“suz1”* o‘zgaruvchisi qiymati kesib olinib, *“qush1”* o‘zgaruvshisiga ta’minlanadi, ya’ni *“suz1”* o‘zgaruvchisida so‘z shaklining asosi, *“qush1”* o‘zgaruvchisida so‘z shaklining qo‘shimchalar qismi yaratiladi.

8-qadam. *“qush1”* o‘zgaruvchisi qiymati sifati so‘z shaklini qo‘shimchalar qismini olgan va undan qo‘shimchalarni aniqlash algoritmi boshlanadi. Bunda qo‘shimchalar lug‘atidan qo‘shimchani izlaydi. Agar topilsa 11-qadamga o‘tiladi, aks holda qo‘shimchalarni Grammatik tahlil qilish uchun bo‘laklash boshlanadi

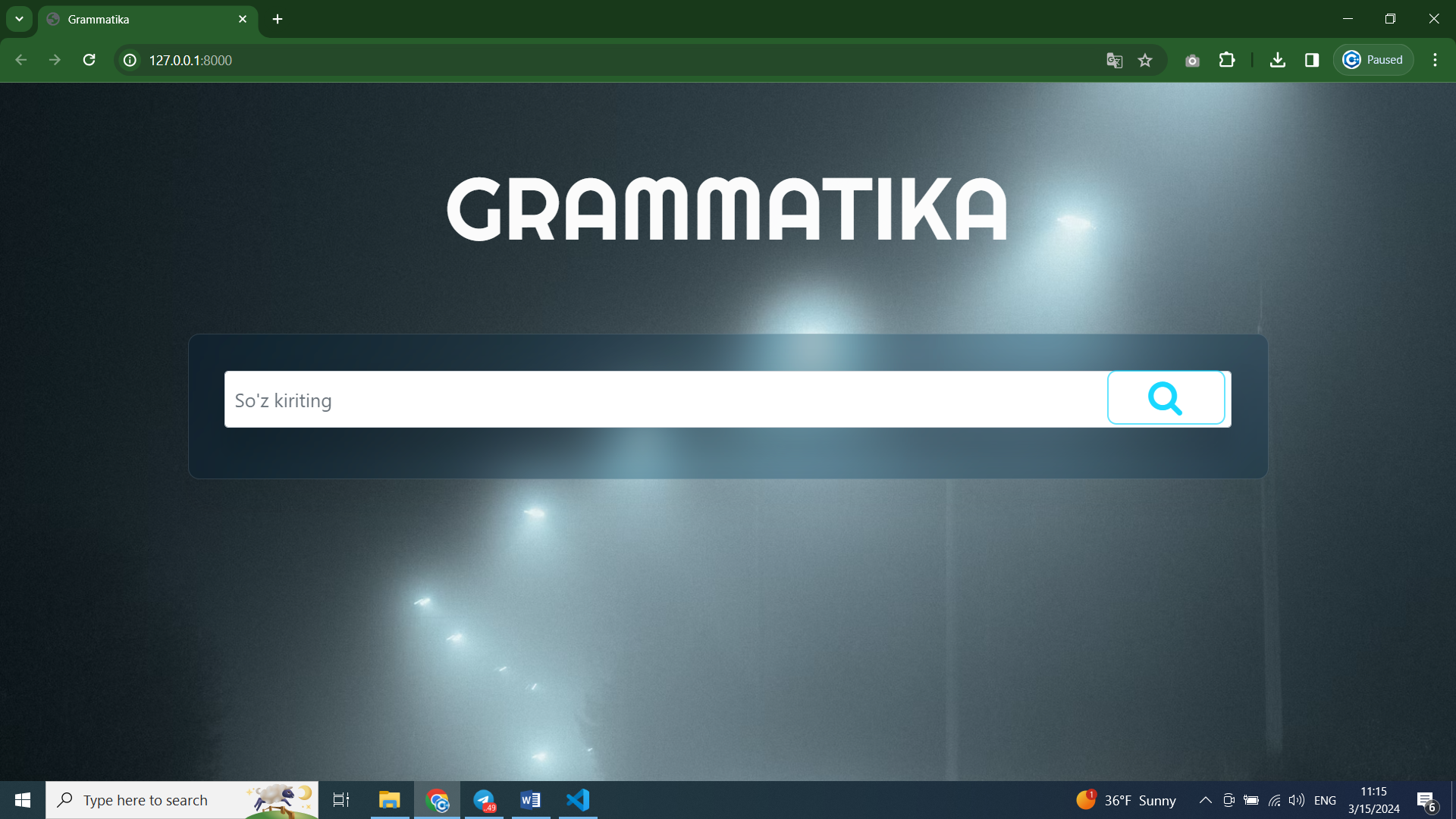
9-qadam. *“qush1”* o‘zgaruvshisi qiymatining birinchi belgisidan boshlab “qo‘shimchalar lug‘atidan qidiriladi. Agar lug‘atdan topilmasa keyingi belgini avvalgi belgiga qo‘shgan holatda lug‘atdan qidirish davom etadi. Bu amal toki *“qush1”* ning qiymatidagi belgilarning oxirigacha davom etadi. Agar topilmasa 12-qadamga o‘tiladi, aks holda dastur ishini davom ettiradi.

10-qadam. *“qush1”* qo‘shimchalar ketma-ketligidan iborat satrda birinchi belgidan boshlab tekshirish (9-qadamda) davomida birinchi qo‘shimcha aniqlanganda ushbu qo‘shimcha va uning izohini massivga yozadi va *“qush1”* o‘zgaruvchisining qiymatidan aniqlanga qo‘shimcha kesib tashlanadi. Masalan: -larining, -lar+-i-ning. Bunda –lar qo‘shimchasi aniqlanganda “qush1”ning qiymati –ining bo‘ladi va 9-qadamdagi algoritmga qaytadi.

11-qadam. Bunda morfolig tahlil qilingan ma’lumotlar so‘z sahklining asosi, turkumi, qo‘shimchalar ro‘yxati va ularning izohi foydalanuvchiga taqdim etiladi.

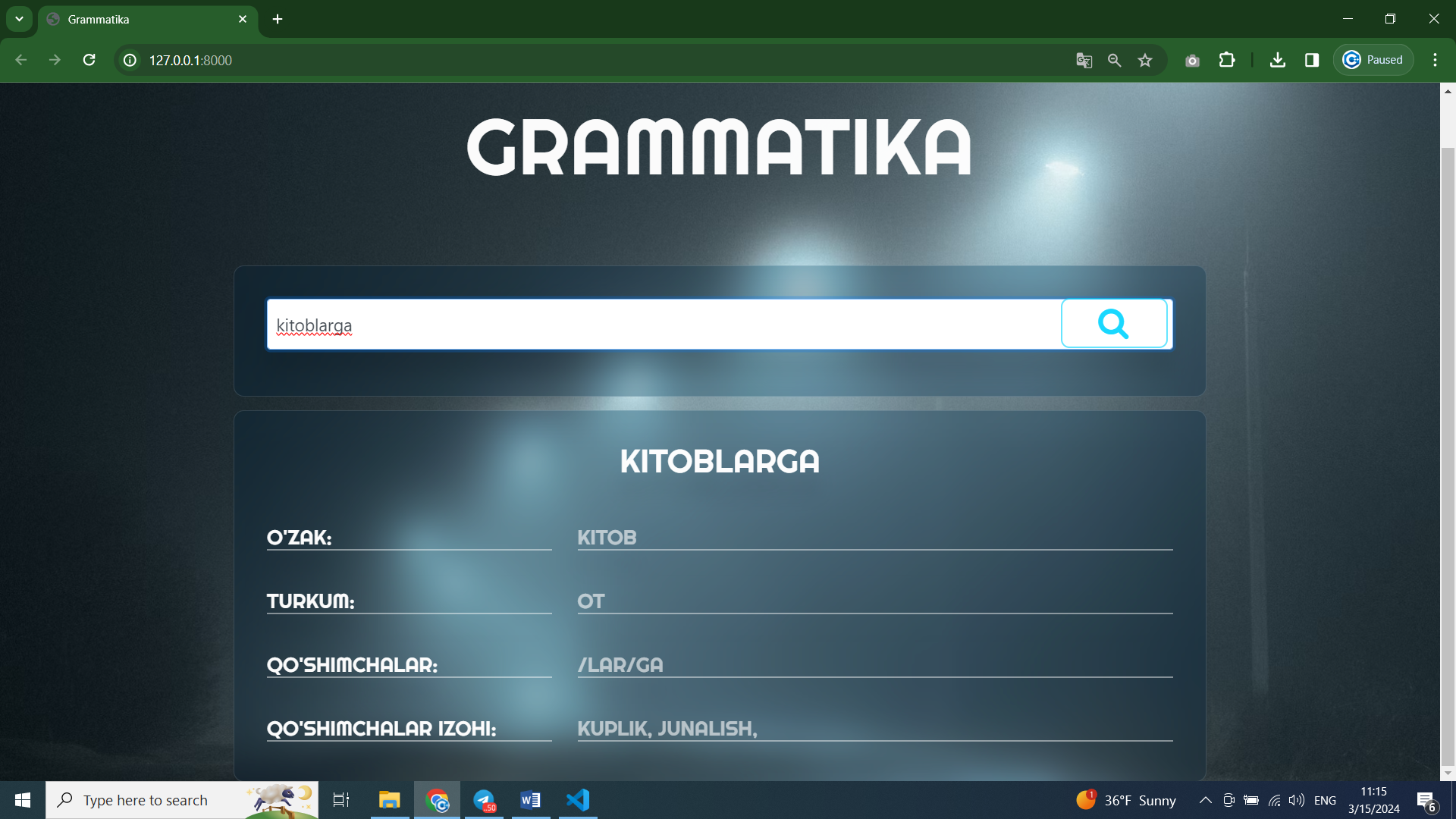
12-qadam. Dastur ishini tugatadi.

Ushbu algoritm asosida morfologik razmetkalash dasturi ishlab chiqilgan.



2-rasm. Dasturning asosiy interfeysi

So‘z shaklini Grammatik razmetkalash uchun so‘z shakli kiritiladi. Misol sifatida “kitoblarga” so‘zni tahlil qilaylik. Bunda 3-rasmda ko‘rsatilgandek so‘z shaklini grammatik tahlil qilib beradi.



3-rasm. Grammatik ma’lumotlar taqdimoti

Grammatik tahlil xulosalarini quyidagicha ko‘rish mumkin:

So‘z shakli – “*kitoblarga”;*

So‘z asosi – *“kitob”;*

So‘z turkumi – *“ot so‘z turkumi”;*

Qo‘shimchalar ro‘yxati – *“ -lar, -ga”;*

Qo‘shimchalarga berilgan izoh – *“ko‘plik, jo‘nalish kelishigi”;*

Bundan o‘rganayotgan kalit so‘zimiz ot so‘z turkumiga taalluqligi, unda ko‘plik va jo‘nalish kelishiklaridagi qo‘shimchalar to‘g‘risidagi lingvistik ma’lumotlar olingan.

Xulosa qilib aytganda, o‘zbekcha so‘zlarning grammatik tahlil qilish algoritmlarini yaratish zamonaviy lingvistika va kompyuter texnologiyalari kesishmasida muhim ahamiyat kasb etadi. Ushbu maqolada o‘zbek tilidagi so‘zlarning morfologik va sintaktik xususiyatlarini tahlil qilish orqali ularni avtomatik ravishda qayta ishlash algoritmlarini ishlab chiqishga qaratilgan tadqiqot natijalari ko‘rib chiqildi.

Tadqiqot jarayonida o‘zbek tilining morfologik tizimi, so‘z turkumlari va qo‘shimchalarning grammatik tahlil qilishda qanchalik murakkabligini aniqlashga e’tibor qaratildi. Buning natijasida so‘zlarni o‘zakka ajratish, qo‘shimchalarni tahlil qilish, so‘z turkumlarini aniqlash uchun samarali algoritmik yondashuvlar ishlab chiqildi. Natijalar shuni ko‘rsatdiki, ishlab chiqilgan algoritmlar o‘zbek tilidagi so‘zlarning grammatik tahlil qilish jarayonini sezilarli darajada yengillashtiradi.

Ushbu algoritmlar nafaqat ilmiy tadqiqotlarda, balki amaliy loyihalarda ham keng qo‘llanishi mumkin, masalan, matnlarni avtomatik tahrirlash, tarjima qilish tizimlari va til o‘rgatuvchi ilovalar yaratishda. Shuningdek, bu algoritmlar o‘zbek tili grammatikasini o‘rganishga va uni kompyuter texnologiyalariga samarali integratsiya qilishga hissa qo‘shadi.

Kelgusi ilmiy tadqiqoylarda ushbu ishlar davom ettirilib, algoritmlar yanada takomillashtirilib, yanada murakkab lingvistik strukturalarni tahlil qilishga moslashtirilishi mumkin. Shu bilan birga, boshqa turkiy tillar uchun ham shunga o‘xshash algoritmlar yaratish imkoniyatlari mavjud.

Foydalanilgan adabiyotlar roʻyxati

1. *Mahlov C., Piotrowski M.,* JSLIM – Computational Morphology in the Framework of the SLIM, Theory Language/ State of Art in Computational Morphology, Zurich, 2009, 15-p.
2. *A.Po‘latov, S.Muhamedova.* Kompyuter lingvistikasi.-Т.,2008
3. *Tursunov M.S., Qarshiyev A.B.,* “O‘zbekcha matnlarni morfologik va leksik tahlil qilish algoritmlari va dasturlari”, International scientific conference “Actual problems of applied mathematics and information technologies”. –Toshkent, 14-15.11.2019, 290-291 betlar.
4. *Baqayev I.I.,* “O‘zbek tili so‘z shakllarini morfologik tahlil qilish modellari va algoritmlari”, avtoreferat, Toshkent: TATU, 2021, 11-bet.