

ISSN 2181-922X

LANGUAGE & CULTURE

O'ZBEKISTON

TIL VA MADANIYAT

**KOMPYUTER
LINGVISTIKASI**

2024 Vol. 2 (6)

www.compling.tsuull.uz

UZBEKISTAN

ISSN 2181-922X

O‘ZBEKISTON

TIL VA MADANIYAT

KOMPYUTER LINGVISTIKASI

2024 Vol. 2 (6)

compling.tsuull.uz

Alisher Navoiy nomidagi Toshkent davlat o'zbek tili va adabiyoti universiteti

Bosh muharrir:

Botir Elov

Bosh muharrir o'rinbosari:

Shahlo Hamroyeva

Mas'ul kotib:

Oqila Abdullayeva

Tahrir kengashi

Shuhrat Sirojiddinov (O'zbekiston), Eshref Adali (Turkiya), [Viktor Zaxarov] (Rossiya), Vladimir Benko (Slovakiya), Ayrat Gatiatullin (Tataristan), Rinat Gilmullin (Tataristan), Murat O'rxun (Turkiya), Suyun Karimov (O'zbekiston), Abduvali Qarshiyev (O'zbekiston), Muxammadjon Musayev (O'zbekiston), Kamoliddin Shukurov (O'zbekiston), O'tkir Hamdamov (O'zbekiston), Tal'at Zuparov (O'zbekiston), Bahodir Mo'minov (O'zbekiston), Faxriddin Nurullayev (O'zbekiston), Zulxumor Xolmanova (O'zbekiston), Muqaddas Abdurahmonova (O'zbekiston), Elova Dilrabo (O'zbekiston), Ruhillo Alayev (O'zbekiston), Rasuljon Atamuratov (O'zbekiston), Malika Abdullayeva (O'zbekiston), Mannon Ochilov (O'zbekiston), Xolisa Axmedova (O'zbekiston), Zilola Xusainova (O'zbekiston), Uldona Abdurahmonova (O'zbekiston).

Jurnal haqida ma'lumot

“O'zbekiston: til va madaniyat. Kompyuter lingvistikasi” seriyasi – Oliy attestatsiya komissiyasi ilmiy nashrlar ro'yxatidagi “O'zbekiston: til va madaniyat” akademik jurnalining ilovasi hisoblanib, unda professor-o'qituvchilar, doktorantlar, stajor-tadqiqotchilar, mustaqil izlanuvchilar, magistrantlarning kompyuter lingvistikasi, jumladan, tabiiy tilga ishlov berish (NLP), o'zbek tilining formal grammatikasi, korpus lingvistikasi, mashina tarjimasi, nutqni qayta ishlash tizimlari, intellektual tizimlar, kompyuter leksikografiyasi hamda lingvistik ontologiyalar kabi sohalarga oid tadqiqotlari nashr qilinadi.

Jurnal ilovasi bir yilda to'rt marta chop etiladi.

O'zbek, turk, rus va ingliz tillarida yozilgan maqolalar qabul qilinadi.

Jurnalda kitoblarga yozilgan taqrizlar, adabiyotlar sharhi, konferensiyalar hisobotlari va tadqiqot loyihalari natijalari ham e'lon qilinadi.

Mualliflar fikri tahririyat nuqtayi nazaridan farq qilishi mumkin.

“O'zbekiston: til va madaniyat. Kompyuter lingvistikasi” seriyasi 2023-yildan chiqa boshlagan.

Alisher Navoiy nomidagi Toshkent davlat o'zbek tili va adabiyoti universiteti. O'zbekiston, Toshkent, Yakkasaroy tumani, Yusuf Xos Hojib ko'chasi, 103-uy.

E-mail: kompling@navoiy-uni.uz

Website: kompling.tsuull.uz

Alisher Navoi'i Tashkent State University of the Uzbek Language and Literature

Chief editor:	Botir Elov
Deputy editor-in-chief:	Shahlo Hamroyeva
Responsible secretary:	Oqila Abdullayeva

Editorial board

Shukhrat Sirojiddinov (Uzbekiston), Eshref Adali (Turkiye), [Viktor Zakharov] (Russia), Vladimir Benko (Slovakia), Ayrat Gatiatullin (Tataristan), Rinat Gilmullin (Tataristan), Murat Orhun (Turkey), Suyun Karimov (Uzbekistan), Abduvali Karshiyev (Uzbekistan), Mukhammadjon Musayev (Uzbekistan), Kamoliddin Shukurov (Uzbekistan), O'tkir Hamdamov (Uzbekistan), Tal'at Zuparov (Uzbekistan), Bahadir Mo'minov (Uzbekistan), Fakhridin Nurullayev (Uzbekistan), Zulkhumor Kholmanova (Uzbekistan), Muqaddas Abdurakhmonova (Uzbekistan), Elova Dirlabo (Uzbekistan), Ruhillo Alayev (Uzbekistan), Rasuljon Atamuratov (Uzbekistan), Malika Abdullayeva (Uzbekistan), Mannon Ochilov (Uzbekistan), Kholisa Akhmedova (Uzbekistan), Zilola Khusainova (Uzbekistan), Uldona Abdurakhmonova (Uzbekistan).

Information about the magazine

"Uzbekistan: language and culture. "Computer Linguistics" series is an appendix of the academic journal "Uzbekistan: Language and Culture" in the list of scientific publications of the Higher Attestation Commission, in which computer linguistics, including natural language processing (NLP) of professors-teachers, doctoral students, intern-researchers, independent researchers, master's students, researches related to formal grammar of the Uzbek language, corpus linguistics, machine translation, speech processing systems, intelligent systems, computer lexicography and linguistic ontologies are published.

The magazine supplement is published four times a year.

Articles written in Uzbek, Turkish, Russian and English languages are accepted.

The journal also publishes book reviews, literature reviews, conference reports, and research project results.

The opinion of the authors may differ from the editorial point of view.

"Uzbekistan: language and culture. "Computer Linguistics" series has been published since 2023.

Tashkent State University of Uzbek Language and Literature named after Alisher Navoi. Yusuf Khos Hajib street, 103, Yakkasaray district, Tashkent, Uzbekistan.

E-mail: kompling@navoiy-uni.uz

Website: kompling.tsuull.uz

MUNDARIJA

Фархад Мирзаев Рамиз, Гюнель Новрузова Сиявуш

Компьютерное моделирование основные инструменты
исследования6

Xolisa Axmedova, Elbek Malikov

Bilinga asoslangan yondashuvlar asosida omonimiyani
farqlash.....15

Xolisa Axmedova, Shohnazar Sultonov

Statistik usullar yordamida polifunksional so'zlarni
semantik farqlash.....26

Oqila Abdullayeva, O'g'iloy Bozorqulova

Jahon tilshunosligida treebanklar tasnifi.....43

Zilola Xusainova, Surayyo Yangibayeva

Til korpusi turlari.....54

Shaxinabonu Mansurova

Son so'z turkumini grammatik pos teglashning
lingvistik modellari.....64

Botir Elov, Zilola Xusainova, Sarvinoz Qosimova

Katta til modellari.....78

Oqila Abdullayeva, Fotima O'tkirova

Jahon tilshunosligida Dependancy Parsingga oid
tadqiqotlar.....92

CONTENT

Farhad Mirzayev Ramiz, Gunel Novruzova Siyavush

Computer simulation basic research

tools.....13

Xolisa Axmedova, Elbek Malikov

Differentiating knowledge-based

Homonymy.....24

Xolisa Axmedova, Shohnazar Sultonov

Semantic differentiation of polyfunctional words using

statistikal methods.....40

Oqila Abdullayeva, O'g'iloy Bozorqulova

The classification of treebanks in world linguistics.....52

Zilola Xusainova, Surayyo Yangibayeva

Types of language corpus.....62

Shaxinabonu Mansurova

Linguistic methods of grammatical pos tagging

of the number word group.....76

Botir Elov, Zilola Xusainova, Sarvinoz Qosimova

Large language models.....90

Oqila Abdullayeva, Fotima O'tkirova

Dependancy Parsing in world

linguistics.....104

BILIMGA ASOSLANGAN YONDASHUVLAR ASOSIDA OMONIMIYANI FARQLASH

Xolisa Axmedova¹

Elbek Malikov²

Annotatsiya. Mazkur maqolada omonimiyani farqlash muammosining dolzarbligi va yechimlar yondashuvlari tahlil qilingan. Omonim soʻzlarning maʼnosini avtomatik aniqlash tabiiy tilni qayta ishlash sohasida muhim masala boʻlib, u qidiruv tizimlari, mashina tarjimasi va semantik tahlil kabi ilovalarda qoʻllaniladi. Bilimga asoslangan yondashuv va Lesk algoritmi orqali omonim soʻzlarning kontekst asosida toʻgʻri maʼnosini aniqlash usullari yoritilgan. Tadqiqot natijalariga koʻra, ushbu usullar koʻplab tillarda yuqori samaradorlik koʻrsatgan boʻlib, ular Oʻzbek tilida ham omonimiyani farqlash uchun mos kelishi mumkin.

Kalit soʻzlar: *Omonimiya, soʻz maʼnosi, bilimga asoslangan yondashuv, Lesk algoritmi, tabiiy tilni qayta ishlash.*

Kirish

Tabiiy tilda oʻqilishi bir xil, lekin turli kontesktlarda maʼnosi har xil boʻlgan soʻzlar uchraydi. Bu soʻzlar omonim soʻzlar hisoblanadi. Inson omonim soʻzni oson farqlab, maʼnosini tushunadi. Lekin bu mashina uchun qiyin vazifa, uzoq yillardan buyon odamlarga oʻxshab soʻzning maʼnosini avtomatik aniqlaydigan tizim ishlab chiqish masalasi oʻzining dolzarbligini saqlab kelmoqda.

Soʻz maʼnolarini aniqlash usullari omonim soʻzning toʻgʻri manisini topish va uni toʻgʻri oʻringa avtomatik joylashtirish vazifasidir. Soʻz maʼnolarini aniqlash tabiiy tilni qayta ishlash sohasining muhim, murakkab masalasidir. Bu mashina tarjimasi,

¹Axmedova Xolisxon Ilxomovna – texnika fanlari boʻyicha falsafa doktori (PhD), Alisher Navoiy nomidagi Toshkent davlat oʻzbek tili va adabiyoti universiteti.
E-pochta: xolisa9029@mail.ru

ORCID: 0000-0002-9828-1650

²Malikov Elbek Bahodir oʻgʻli – Alisher Navoiy nomidagi Toshkent davlat oʻzbek tili va adabiyoti universiteti magistranti.

E-pochta: malikovelbekk@gmail.com

ORCID: 0009-0000-0187-1965

semantik xaritalash, semantik izohlash va ontologiyani o'rganish kabi ko'plab ilovalari uchun zarur. Shuningdek, u ko'plab axborot qidirish, ma'lumot olish va nutqni aniqlash kabi ilovalarning ish faoliyatini yaxshilashda foydali bo'ladi. Ingliz [Eneko Agirre va boshqalar, 2007], hind [Pushpak Bhattacharyya, 2017], fransuz [Isabelle Tellier, 2019], ispan [Victoria Marrero Aguiar, 2008], xitoy [Huang Xiaorong, 2009] va boshqa ko'plab tabiiy tillar omonim so'zlarga ega tillar hisoblanadi.

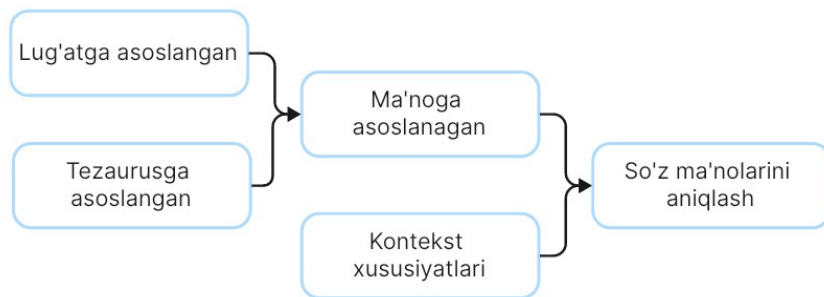
Asosiy qism

Omonimiyani aniqlash masalasi kompyuter lingvistikasining ko'plab dasturlarida, xususan, qidiruv tizimlarida ahamiyatli. Chunki u muayyan so'rov sinflarini qayta ishlash aniqligini oshirishi yoki saqlangan axborot miqdorini kamaytirishi mumkin. O'zbek tilida shunday omonim so'zlar borki, turli so'z turkumlari ichida ham bir so'z tukumiga oid bir nechta ma'nolarni anglatishi mumkin.

3-Jadval: Bir so'z turkumi doirasidagi omonim so'zlar

So'z	So'z turkumi	Ma'nosi
O't	Fe'l	O'tmoq fe'li
	O't	Maysa, o't-o'lan
	O't	Inson o'rgani
	O't	Olov
Oz	Ravish	Kam, miqdori nisbatan ko'p bo'lmagan.
	Fe'l	Oriqlamoq, etidan yo'qotmoq.
	Fe'l	Noxush bo'lmoq, kuchsizlanmoq, holsizlanmoq.
	Fe'l	Adashmoq, to'g'ri yo'ldan chetga chiqmoq
Bo'y	Ot	Uzunlik o'lchovi.
	Ot	Hid, is
...	...	...

3-jadvada keltirilgan o't omonim so'zi turli so'z turkumlari doirasida o'ziga xos ma'nolarni anglatadi. Shuni alohida ta'kidlash kerakki, shu so'z turkumlarining biri doirasida bir qancha ma'nolarni anglatuvchi so'zlar ham talaygina. Bunday omonim so'zlarni POS tegging masalasini yechish bilan aniqlab bo'lmaydi. Buning uchun alohida usullar, modellar zarur. Faraz qilaylik, T matn $W_1, W_2, W_3, \dots, W_n$ so'zlar ketma-ketligidan iborat. So'z ma'nolarini aniqlash usullari matndagi barcha omonim so'zlar uchun to'g'ri ma'noni belgilash vazifasidir. So'z ma'nolarini aniqlash usullarining nazariy modeli 2-rasmda ko'rsatilgan.



2-rasm. Soʻz maʼnolarini aniqlash usullarining nazariy modeli

Soʻz maʼnosini aniqlashda ishlatiladigan ikkita asosiy yondashuv - bu *Bilimga asoslangan yondashuv* va *Mashinali oʻqitishga asoslangan yondashuvlar*dir.

Ushbu yondashuvlarning har biri oʻz navbatida bir qator usullarni qamrab oladi.

Bilimga asoslangan yondashuvda omonim soʻzning barcha maʼnolari lugʻatdan olinadi. Keyin bu maʼnolar kontekstda qolgan barcha soʻzlarning lugʻat taʼriflari bilan taqqoslanadi. Biz soʻz maʼnosini aniqlash yondashuvlarini soʻz maʼnosini ajratish uchun chuqur va sayoz yondashuv sifatida tasniflashimiz mumkin. Bilimga asoslangan yondashuv kompyuter oʻqishi mumkin boʻlgan lugʻatlar, masalan, til korpusi, WordNet, tezaurus asosida ishlaydi. Soʻz maʼnosini aniqlash uchun grammatik qoidalardan ham foydalanishi mumkin. Bilimga asoslangan yondashuvning maqsadi (lugʻatga asoslangan yondashuv) kontekstdagi soʻzlarning maʼnosini aniqlash uchun bilim resurslaridan foydalanishdir. Bilim resurslari lugʻatlar, tezauruslar, ontologiyalar, birikmalar va boshqalardir. Yuqoridagi usullar boshqariladigan muqobil usullarga qaraganda unumdorligi pastroq, lekin ular kengroq diapazondagi afzalliklarga ega.

Bilimga asoslangan yondashuv

Bilimga asoslangan yondashuv mashinada oʻqiladigan lugʻat (MDR) qoidalarini talab qiladi. U kontekstdagi soʻzlarning xususiyatlari bilan bir qatorda noaniq boʻlgan soʻzlarning maʼnolarini, ularning turli xususiyatlarini aniqlashni oʻz ichiga oladi.

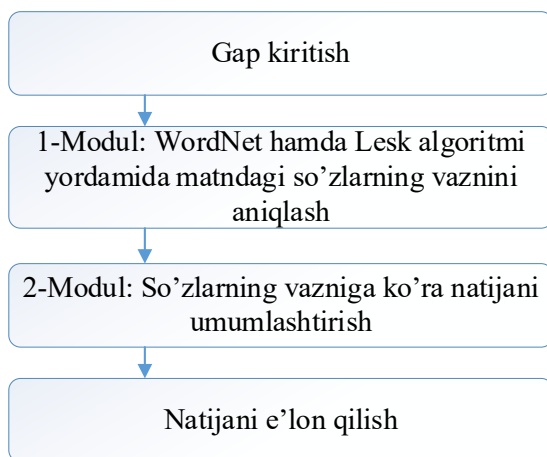
Lesk algoritmi: Korpus maʼlumotlari asosida soʻz manosi aniqlashda LESK algoritmidan foydalanish jarayonini koʻrib chiqamiz. **Lesk algoritmi** 1986-yilda **E.Maykl Lesk** tomonidan kiritilgan boʻlib, klassik WSD algoritmidir. Lesk algoritmi matnning maʼlum bir qismidagi soʻzlarga oʻxshash maʼnoga ega boʻladi, - degan fikrga asoslanadi. Soddalashtirilgan Lesk algoritmidan har

bir soʻz kontekstining toʻgʻri maʼnosi berilgan kontekst va uning lugʻat maʼnosi orasida **eng koʻp oʻxshash maʼnoni** olish orqali topiladi. Hind tilida soʻz maʼnosini aniqlashda ushbu algoritmidan foydalanilgan. Savol-javob tizimlari, sentiment tahlil qilish tizimlarini ishlab chiqishda foydalanilgan [Tripathi, 2021. 939-954]. Zouaghi, Merhbene va boshqalar tomonidan arab tilidagi soʻzlarning maʼnolarini aniqlashda ham Lesk algoritmi qoʻllanilgan [Zouaghi, Merhbene va boshqalar, 2012. 257–269]. Olib borilgan tadqiqodlari davomida 73% aniqlikka erishganlar. Basuki, Kholimi va boshqalar indonez tilidagi homograflarni semantik aniqlashda foydalanganlar. Natijada bir soʻz turkumidagi soʻzlarni aniqlashda 78.6%, ikki soʻz turkumi doirasidagi omonim soʻzlarni aniqlashda 62.5% aniqlikka erishilgan [Basuki, Kholimi va boshqalar, 2019. 8-15]. Lesk algoritmi qoidalarga asoslangan usul boʻlib, u tabiiy tilning WordNetiga asoslanadi. Lesk algoritmi yordamida omonimiyani aniqlash uchun quyidagi maʼlumotlar talab qilinadi:

Maʼnolar toʻplami: omonim soʻzning mavjud maʼnolaridan iborat toʻplam.

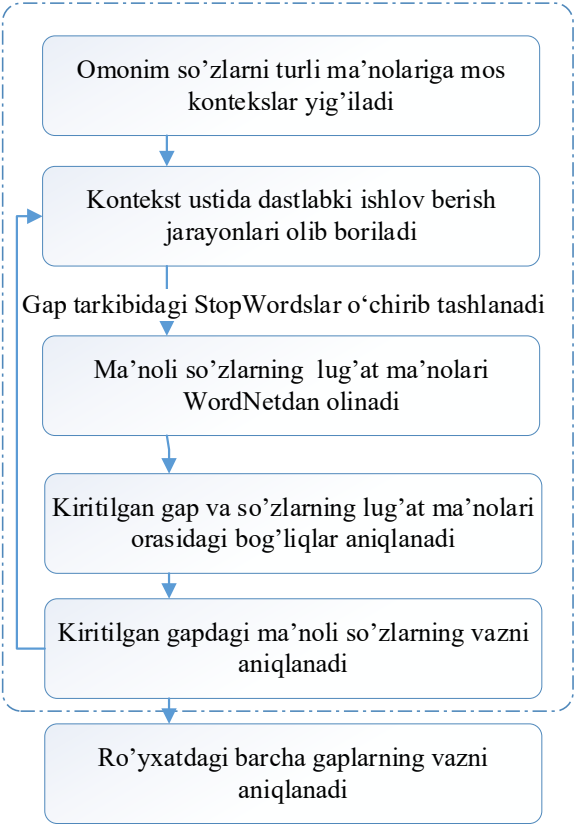
Kontekstlar toʻplami: soʻzning har bir maʼnosini anglatuvchi kontekstlar toʻplami.

Omonimiyani Lesk algoritmi yordamida aniqlash jarayoni ikkita moduldan iborat.



3-rasm: Lesk algoritmidagi modullar

1-modul: Keyingi qadamda ushbu soʻzlarning har bir maʼnosiga mos kontekst hosil qilinadi. Buning uchun korpus maʼlumotlaridan foydalaniladi. Dastlab, tarkibida omonim soʻz mavjud boʻlgan kontekstlar ajratiladi hamda ular omonim soʻzning maʼnolariga qarab ajratib chiqiladi. Ajratilgan kontekstlar ustida dastlabki ishlov berish jarayonlari bajariladi. (4-rasm).



4-rasm: 1-Modulning vazifalari

Lesk algoritmi yordamida omonim so'zlarni semantik farqlash ketma- ketligini quyidagi gap yordamida ko'rib chiqsak.

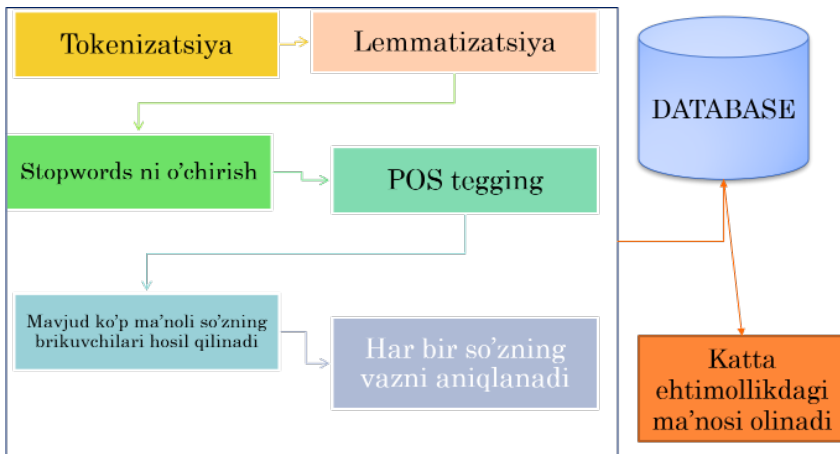
*"Mahallamiz boshidagi hashamatli uy **yoqib** yuborildi."*

Bu so'zlarning ma'nolaridan iborat bo'lgan to'plam ma'lumotlar bazasida kiritilgan, ya'ni tayyor.

So'zlar	Turkum	Ma'nosi
yoqmoq	Fe'l	1. Yondirib, o't, alanga hosil qilmoq, yondirmoq; kuydirmoq. 2. Nur, yorqin hosil qilmoq, yoritmoq. 3. O't, olov yordamida qizdirmoq, isitmoq. 4. Qizdirib harakatga keltirmoq; o't oldirmoq.
		1. Ma'qul bo'lmoq; did, xohish, ko'ngilga to'g'ri kelmoq.

kechmoq	Fe'l	1. Biror narsa ustidan bosib (botib) o'tmoq, yurmoq. 2. Yuz bermoq; sodir bo'lmoq. 3. O'tmoq. 4. Afv etmoq, kechirmoq.
		5. Bahridan o'tmoq. 6. Yuz o'girmoq, aloqa, munosabatni uzmoq.

So'zlarning vaznini aniqlashda har bir ma'nosi bo'yicha so'zlarning sonini sanash orqali amalga oshiriladi. Kontekstlar bo'yicha olingan ma'lumotlar ma'lumotlar bazasiga saqlanadi.



5-rasm: Leks algoritmidagi jarayonlar

Keyingi qadamda yangi kiritilgan gapni tahlil qilish jarayonini ko'rib chiqamiz.

"Korxonaning asosiy sektoridagi barcha buyumlar yoqib yuborildi."

Gapidagi omonim so'zlarning ma'nosini Lesk algoritmi yordamida aniqlaymiz. Kiritilgan gap ustida quyidagi amallar bajariladi.

Tokenizatsiya;

Lemmatizatsiya;

Nomuhim so'zlarni (StopWords) olib tashlash;

POST tagging;

Korxona, asosiy, sektor, barcha, buyum, yoq, yubor

Keyingi qadamda kiritilgan gap tarkibidagi omonim so'z aniqlanadi hamda uning birikuvchilari ajratib olinadi. Ajratilgan birikuvchilarning vazni ma'lumotlar bazasidan aniqlab olinadi. Masalan, *yoq* so'zining birikuvchilari va ularning vaznidan iborat dataset 2-jadvalda keltirilgan.

Birikuvchilar	1-ma'nosi	2-ma'nosi
Yubormoq	10	2
Buyum	15	1
Barcha	0	25
Sector	35	0
Korxona	18	8
Asosiy	10	12
...	...	...

Keltirilgan ma'lumotlar orasidan omonim so'zning gapdagi birikuvchilari va ularning vazni aniqlanadi. Aniqlangan ma'lumotlardan eng yuqori vaznga ega bo'lgan birikuvchining ma'nosi tanlab olinadi.

Olingan vektorlar odatda 200-2000 o'lchamga ega va semantik jihatdan o'xshash so'zlarni o'xshash yo'nalishli vektorlar sifatida ifodalaydi. Bunday vektorlardan kirish funksiyalari sifatida foydalanish bizga katta teglanmagan korpuslardan ma'lumot olish va bitta tegdan kontekstdan semantik jihatdan o'xshash so'zlarni o'z ichiga olgan barcha kontekstlarga umumlashtirish imkonini beradi. Biz individual so'z vektorlarining o'rtacha vaznini olamiz: vaznlar har bir so'z kontekstning ma'nosiga qay darajada ta'sir qilishini ko'rsatadi.

Xulosa

Omonimiyani farqlash tabiiy tilni qayta ishlash sohasida muhim va dolzarb masalalardan biridir. Ushbu masala qidiruv tizimlari, mashina tarjimasi, semantik tahlil va sentiment aniqlash kabi ko'plab amaliy sohalarda qo'llaniladi. Bu jarayon so'zlarning ma'nosini kontekstual tahlil qilishni talab qiladi, chunki bir so'z turli kontekstlarda turli ma'nolarni anglatishi mumkin.

Bilimga asoslangan yondashuv va Lesk algoritmi kabi usullar omonimiyani farqlashning asosiy vositalaridan hisoblanadi. Bilimga asoslangan yondashuv lug'at va semantik resurslardan foydalanib, so'zning ma'nosini aniqlashda samarali natijalar beradi. Ayniqsa, mavjud resurslardan foydalanishda bu yondashuv nisbatan kam xarajatli va aniqdir. Lesk algoritmi esa kontekst va lug'at ta'riflari o'rtasidagi o'xshashlikni o'lchash orqali ma'no tanlash imkonini beradi. Ushbu usullar boshqa tillarda muvaffaqiyatli qo'llanilgan va o'zbek tiliga ham moslashtirish imkoniyatlari mavjud.

O'zbek tilida omonimiyani farqlashni avtomatlashtirish uchun maxsus lug'atlar, tezauruslar va katta hajmdagi kontekstual ma'lumotlar jamlanmalarini yaratish zarur. Shuningdek, mashina o'qitishi va chuqur o'rganish kabi zamonaviy yondashuvlarni qo'llash orqali algoritmlar samaradorligini oshirish mumkin.

Kelajakdagi izlanishlar o'zbek tilining o'ziga xos xususiyatlarini hisobga olgan holda yanada chuqur tahlillarni talab qiladi. Bunday tadqiqotlar nafaqat tilshunoslikni rivojlantiradi, balki qidiruv tizimlari va boshqa texnologiyalar uchun yangi imkoniyatlar eshigini ochadi. Shu sababli, omonimiyani farqlash sohasidagi izlanishlar til texnologiyalarini rivojlantirishda muhim ahamiyatga ega bo'lib, O'zbek tili uchun innovatsion yechimlar yaratishda muhim poydevor bo'lib xizmat qiladi.

Foydalanilgan adabiyotlar

- Блумфилд Л. Язык. – М.: Прогресс, 1968. – 608 с.
- Fries Ch.C. The structure of English. An introduction to the construction of English sentences. – L., 1969.
- Bongers H. The history and principles of Vocabulary control. – Woerden: WOCOPI, 1947.
- Po'latov A. Kompyuter lingvistikasi. – Toshkent: Akademnashr, 2011.
- Muhammedova S. Harakat fe'llari asosida kompyuter dasturlari uchun lingvistik ta'min yaratish. Metodik qo'llanma. – Toshkent, 2006.
- Jumanazarova G.U. Fozil Yo'ldosh o'g'li dostonlari tilining lingvopoetikasi: Fil. fan. dok. dis...avtoref. – Toshkent, 2017.
- Abduraxmonova N.Z. Inglizcha matnlarni o'zbek tiliga tarjima qilish dasturining lingvistik ta'minoti (Sodda gaplar misolida). Filol. fan. bo'yicha falsafa doktori (PhD)...dis. avtoref. – Toshkent, 2018. – 52 b.
- Abdurahmonova M., Raxmanova A. Korpus lingvistikasida polisemiya // "Kompyuter lingvistikasi: muammolar, yechim, istiqbollar" Respublika I ilmiy-texnikaviy konferensiya. – Vol. 1 №. 01 (2021).
- Abdurahmonova M., Raxmanova A. Ўзбек тили омонимлари учун миллий теглар тўплами // Ўқув услубий ишлар тўплами, 2021. – 159-б.
- Abdurahmonova M. O'zbek va turk tilidagi leksik omonimlarning lingvokognitiv xususiyatlari: Filol.fan.doktori (DcS)...diss. -Toshkent, 2024.-245-b.

- Elov B. Axmedova X. Business Process Modeling That Distinguishes Homonymy Within Three Parts of Speeches in Uzbek Language// 2022 7th International Conference on Computer Science and Engineering (UBMK) Turkiya. Sep.14-16, 2022. 289-294-p.
- Gulyamova Sh. O'zbek tili semantik analizatorining lingvistik asoslari: filol.fan.doktori (DcS)... dis. -Farg'ona, 2021. -278 b.
- Хамроева Ш. Ўзбек тили муаллифлик корпусини тузишнинг лингвистик асослари: Филол.фан.бўйича фалсафа доктори (PhD)...дис. – Бухоро, 2018. – 250 б.
- Zouaghi A., Merhbene L., Zrigui M. Combination of information retrieval methods with LESK algorithm for Arabic word sense disambiguation. Artificial Intelligence Review, 38(4), 257-269., 2012 <https://doi.org/10.1007/s10462-011-9249-3>
- Basuki S., Kholimi A.S., Minarno A.E., Sumadi F.D.S., Effendy M.R.A. Word Sense Disambiguation (WSD) for Indonesian homograph word meaning determination by LESK Algorithm Application. In Proceedings of 2019 International Conference on Information and Communication Technology and Systems, ICTS 2019 (pp. 8-15). Institute of Electrical and Electronics Engineers Inc., 2019. <https://doi.org/10.1109/ICTS.2019.8850957>

DIFFERENTIATING KNOWLEDGE-BASED HOMONYMY

Xolisa Axmedova¹
Elbek Malikov²

Abstract. This article analyzes the significance and solutions for word sense disambiguation (WSD). Automatically determining the meanings of homonyms in context is a critical task in natural language processing, applicable in search engines, machine translation, and semantic analysis. The study highlights knowledge-based approaches and the Lesk algorithm for accurate contextual sense disambiguation. The results demonstrate high efficiency in various languages, suggesting that these methods could also effectively address homonymy in the Uzbek language.

Keywords: *Homonymy, word sense, knowledge-based approach, Lesk algorithm, natural language processing.*

References

- Blumfield L. Yazik. – M.: Progress, 1968. – 608 s.
- Fries Ch.C. The structure of English. An introduction to the construction of English sentences. – L., 1969.
- Bongers H. The history and principles of Vocabulary control. – Woerden: WOCOPI, 1947.
- Po'latov A. Kompyuter lingvistikasi. – Toshkent: Akademnashr, 2011.
- Muhammedova S. Harakat fe'llari asosida kompyuter dasturlari uchun lingvistik ta'min yaratish. Metodik qo'llanma. – Toshkent, 2006.
- Jumanazarova G.U. Fozil Yo'ldosh o'g'li dostonlari tilining lingvopoetikasi: Fil. fan. dok. dis...avtoref. – Toshkent, 2017.
- Abduraxmonova N.Z. Inglizcha matnlarni o'zbek tiliga tarjima qilish dasturining lingvistik taminoti (Sodda gaplar misolida).

¹*Axmedova Xolisxon Ilxomovna* – doctor of philosophy of technical sciences (PhD). Tashkent State University of Uzbek Language and Literature named after Alisher Navo'i.

E-pochta: xolisa9029@mail.ru

ORCID: 0000-0002-9828-1650

²*Malikov Elbek Bahodir o'g'li* – Graduate student of Computer Linguistics. Tashkent State University of Uzbek Language and Literature named after Alisher Navo'i.

E-pochta: malikovelbekk@gmail.com

ORCID: 0009-0000-0187-1965

- Filol. fan. bo'yicha falsafa doktori (PhD)...dis. avtoref. – Toshkent, 2018. – 52 b.
- Abdurahmonova M., Raxmanova A. Korpus lingvistikasida polisemiya // “Kompyuter lingvistikasi: muammolar, yechim, istiqbollar” Respublika I ilmiy-texnikaviy konferensiya. – Vol. 1 №. 01. 2021.
- Abdurahmonova M., Raxmanova A. O'zbek tili omonimlari uchun milliy teglar to'plami // O'quv uslubiy ishlar to'plami, 2021. – 159-b.
- Abdurahmonova M. O'zbek va turk tilidagi leksik omonimlarning lingvokognitiv xususiyatlari: Filol.fan.doktori (DcS)... disser.-Toshkent, 2024.-245-b.
- Elov B., Axmedova X. Business Process Modeling That Distinguishes Homonymy Within Three Parts of Speeches in Uzbek Language// 2022 7th International Conference on Computer Science and Engineering (UBMK) Turkiya. Sep. 14-16, 2022. 289-294-p.
- Gulyamova Sh. O'zbek tili semantik analizatorining lingvistik asoslari: filol.fan.doktori (DcS)... dis.-Farg'ona, 2021. –278 b.
- Xamroyeva Sh. O'zbek tili mualliflik korpusini tuzishning lingvistik asoslari: Filol.fan.bo'yicha falsafa doktori (PhD)...dis. – Buxoro, 2018. – 250 b.
- Zouaghi A., Merhbene L., Zrigui M. Combination of information retrieval methods with LESK algorithm for Arabic word sense disambiguation. Artificial Intelligence Review, 38(4), 257–269. 2012. <https://doi.org/10.1007/s10462-011-9249-3>
- Basuki S., Kholimi A. S., Minarno A. E., Sumadi F. D. S., Effendy M. R. A. Word Sense Disambiguation (WSD) for Indonesian homograph word meaning determination by LESK Algorithm Application. In Proceedings of 2019 International Conference on Information and Communication Technology and Systems, ICTS 2019 (pp. 8–15). Institute of Electrical and Electronics Engineers Inc. 2019. <https://doi.org/10.1109/ICTS.2019.8850957>

Jurnal 2017-yil 26-oktyabrda O'zbekiston Respublikasi Matbuot
va axborot agentligi tomonidan 0936-raqam bilan ro'yxatdan o'tgan.

Jurnal O'zbekiston Respublikasi Oliy Attestatsiya Komissiyasi tomonidan
filologiya fanlari bo'yicha falsafa doktori (PhD) va fan doktori (DSc)
dissertatsiyalari asosiy ilmiy natijalari chop etilishi lozim bo'lgan ro'yxatga
kiritilgan (30.10.2021. № 308/6).

Tahririyatga kelgan maqolalar mualliflarga qaytarilmaydi.

Manzil: Toshkent shahri, Yakkasaroy tumani, Yusuf Xos
Hojib ko'chasi 103-uy.
Telefonlar: +99871 281-45-11, +99871 281-41-93.
Website: compling.tsuull.uz
E-mail: kompling@navoiy-uni.uz

Bosishga **.**.****-yilda ruxsat etildi.
Bichimi 70x100 1/16, Ofset bosma. "Cambria" garniturasida.
Shartli b.t. 7,51. Nashr b.t. 7,62.

"O'zbekiston: til va madaniyat" jurnali tahririyatida
tayyorlandi va sahifalandi.
"YASHNOBOD NASHR" bosmaxonasida chop etildi.
Adadi 300 nusxa. Buyurtma №2.
Bosmaxona manzili: Toshkent shahar Yashnobod tumani,
58-a harbiy shaharcha.