

ISSN 2181-922X

LANGUAGE & CULTURE

UZBEKISTON O'ZBEKISTON

TIL VA MADANIYAT

KOMPYUTER
LINGVISTIKASI

2024 Vol. 2 (6)

www.compling.tsuull.uz

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

JAHON TILSHUNOSLIGIDA DEPENDANCY PARSINGGA OID TADQIQOTLAR

Oqila Abdullayeva¹
Fotima O'tkirova²

Annotatsiya. Ushbu maqola Dependency parsing haqida umumiy ma'lumot, uning jahon tilshunosligida tutgan o'рни, ahamiyati, tamoyillari, amaliy qo'llanilishi hamda o'rganilishi haqida axborot beradi. Soha yuzasidan yangi yondashuvlarni va usullarni muhokama qiladi. Dependency parsingning tilshunoslikda qanday ishlatilishi, tabiiy tilni qayta ishlash (NLP) va boshqa ko'pgina sohalarda qanday yordam berishi haqida ma'lumot beradi. O'z navbatida maqola qaramlik tahlilining asosiy xususiyatlarini va uning lingvistik jihatlarini hamda bu sohada yaratilgan dasturiy ishlanmalarini taqdim etadi. Ularning afzalligi, istiqbollari va bu bo'yicha amalga oshirilgan ilg'or tadqiqotlar haqida ma'lumot beradi.

Kalit so'zlar: *Dependency parsing, sintaktik tahlil, parser, so'z birikmasi tobe bog'lanish, bosh so'z, tobe so'z, tamoyil, yondashuv.*

Kirish

Jahon tilshunosligida NLPning muhim va eng rivojlangan yo'nalishlaridan biri *Dependency Parsing* (yoki "qaramlikni tahlil qilish") hisoblanadi. Bu soha sintaktik tahlilning bir turi bo'lib, uning maqsadi gapdagi so'zlar o'rtasidagi grammatik bog'lanishlarni aniqlashdir. Tilshunoslikda qaramlik tahlili tilning strukturasini chuqur o'rganishga yordam beradi va mashina tarjimasini, tabiiy tilni qayta ishlash (NLP), nutqni tanlab olish va ko'plab boshqa sohalarda qo'llaniladi. Jahon tilshunosligida tobe bog'lanish borasida grammatika asosidagi yondashuvlar yillar davomida

¹Abdullayeva Oqila Xolmo'minovna - filologiya fanlari bo'yicha falsafa doktori. Alisher Navoiy nomidagi Toshkent davlat o'zbek tili va adabiyoti universiteti doktoranti (DSc).

E-pochta: abdullayeva.oqila@navoiy-uni.uz

ORCID: 0000-0002-2524-4832

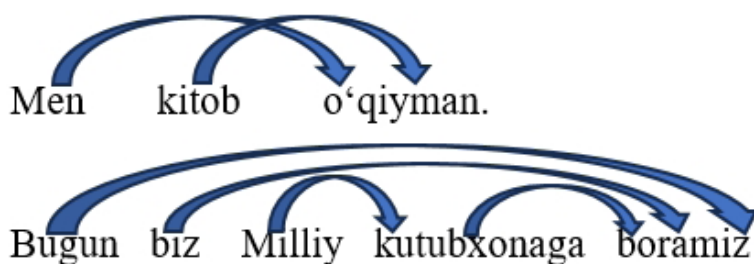
²O'tkirova Fotima Baxtiyorovna - Alisher Navoiy nomidagi Toshkent davlat o'zbek tili va adabiyoti universiteti Kompyuter lingvistikasi fakulteti 1-kurs magistranti.

E-pochta: fotimaotkirova@gmail.com

nazariy va amaliy tilshunoslikning asosiy e'tibor markazida bo'lgan. Tobelanishni tahlil qilish kompyuter lingvistikasining muhim jihati bo'lib, sintaktik tahlil bo'yicha chuqur tushunchalarni taqdim etadi. Bir necha o'n yillar mobaynida tobe bog'lanishni tahlil qilish nazariyalarining e'lon qilinishi kompyuter lingvistikasiga sezilarli ta'sir ko'rsatdi. Nazariy tilshunoslikda Link Grammar kabi dastlabki hisoblash modellari tabiiy tilni qayta ishlashda (NLP) tobelanishni tahlil qilishning amaliy tatbiqlarini ko'rsatdi.

Asosiy qism

Dependency Parsing – bu sintaktik tahlilning bir turi bo'lib, unda har bir so'zning boshqa so'zga bo'lgan grammatik munosabati (qaramligi) tahlil qilinadi [Cakici, 2008. 116]. Bu tahlil usuli tilni gapda bog'lanishlar orqali modellashga qaratilgan. Misol uchun “Men kitob o'qiyman” jumlasida “o'qiyman” – markaziy fe'l bo'lib, “Men” subyekt sifatida, “kitob” esa obyekt sifatida qaraladi. Bu yerda “Men” so'zi “o'qiyman”ga bog'lanadi, “kitob” esa “o'qiyman” fe'lga bog'lanadi. Shu tarzda, bu tahlil orqali jumladagi so'zlarning o'zaro bog'lanishlari va ularning roli aniq belgilanadi. Buni quyidagi misollar orqali ham ko'rish mumkin:



Jahon tilshunosligida til texnologiyalarining dastlabki prototiplari XX asrning o'rtalarida paydo bo'lgan. Parsing deb nomlanuvchi yozma matnlarning avtomatik sintaktik tahlili ulardan biridir. Tabiiy tilni qayta ishlash (NLP) ga bo'lgan qiziqishlarning asosiy qismini ham parsing tashkil qiladi. Xususan, avtomatik sintaktik tahlil ma'lum turdagi semantik tasvir talab qiladigan dasturlar kabi chuqur lingvistik ma'lumotlar bilan ishlaydigan til ilovalariga dastlabki ishlov berish bosqichini ta'minlaydi. Natijada ko'plab til texnologiyalari sintaktik tahlil darajasidagi takomillashuvlardan bevosita foyda oladi [Buchholz, Marsi, 2006. 35]. Masalan, tahlil qilish NLPning boshqa sohalarida ham mavjud. Jumladan, semantik rolni belgilash, mashina tarjimai, ma'lumot ajratib olish, ma'lumotlarni qidirish hamda sentimental tahlil kabi yo'nalishlarda avtomatik

sintaktik tahlil muhim bosqichlarni egallaydi. Buning natijasida avtomatik sintaktik tahlil masalasi til taraqqiyoti va texnologiyalari bilan bog'lanib bormoqda.

Dependancy Parsing ildizlari XIX asr oxirlaridan boshlab tilshunoslikda faol ravishda izlanishlar olib borilgan davrga borib taqaladi. Birinchi marta u grammatikaning asosiy yondashuvi sifatida ishlatilgan. 1950-yillarda kompyuter texnologiyalarining rivojlanishi bilan birga, tilshunoslar so'zlar orasidagi bog'lanishlarni aniqlash uchun avtomatlashtirilgan tizimlar ishlab chiqa boshladilar. 1960-yillarda kompyuterlar yordamida sintaktik tahlil qilish bo'yicha dasturlarni ishlab chiqish xususida birinchi yondashuvlar ishlab chiqildi [Thesis, 2016. 50]. Ushbu dasturlar tillarning grammatik jihatlarini analitik tarzda o'rganishga imkon yaratdi. Bu davrda Dependancy Parsing sintaksisning asosiy tamoyillaridan biri sifatida shakllandi. Jahon tilshunosligida Dependancy parsing o'rganilishi nazariy va amaliy ahamiyatga ega. XX asr oxiri va XXI asr dastlabki yillarida mavzu yuzasidan, asosan, nazariy ishlar amalga oshirildi. Jumladan, tahlil qilish samaradorligini oshirish bo'yicha Marina Lloberes Salvatella, Ruket Cakici, Hindle va Rooth, Collins va Brooks, Stetina va Nagao, Merro va Ferrer, Agirre, Belinkov, Gala va Lafourcade, Hokkenmeyer, Nivre va boshqalar so'z birikmalari hamda tobe bog'lanishlarni aniqlashning muayyan yechim va algoritmlarini ishlab chiqqan hamda tadqiqotlarida qoidalarga asoslangan yondashuvdan borganlar.

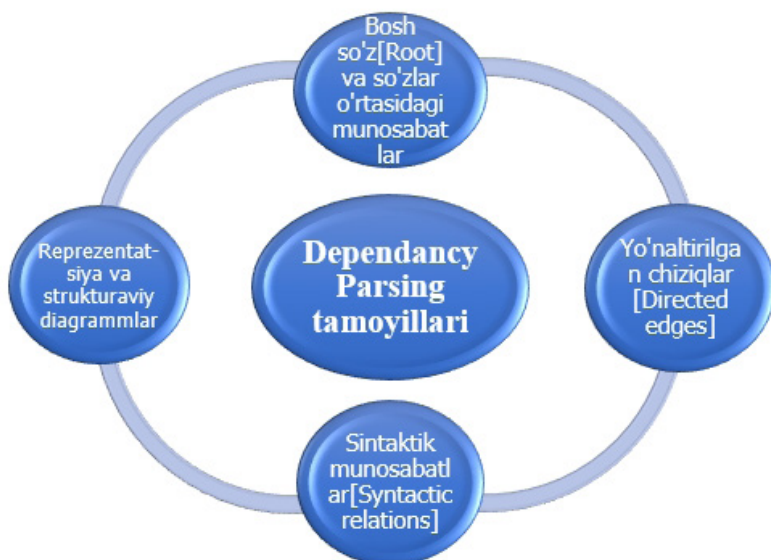
Buochholz va Marsi, Ballesteros va Nivre, Ruket Cakici, Foth, By, Menzel, Olteanu va Moldovan, Turney va Pantellar keyingi yillarda sohada ko'plab amaliy natijalarga erishdi. Xususan, Buochholz va Marsi 13 ta til treebanklarini ko'p tilli (multilingual) treebank sifatida shakllantirib, bir xil tobelik formatiga aylantirgan. Ballesteros va Nivre CoNLL-X ma'lumotlar bazasiga Penn Treebankni qo'shish algoritmini ishlab chiqib, Stanfordga aylantirdi. Joakim Nivre hozirgi kunda dependancy parsingning eng ilg'or yondashuvlari va modellarini ishlab chiqqan va o'rgangan olimlardan biridir. U MaltParserni yaratishda ishtirok etgan va dependancy parsingni statistik yondashuvlar bilan qo'llab-quvvatlashni rivojlantirgan. Ruket Cakici tomonidan mavjud treebank samaradorligini oshirishga qaratilgan tizimli tuzatishlar dasturdagi bir qator muammolarni hal etish imkonini yaratdi. Foth va Menzel tomonidan birlashtirilgan ma'lumotlarga asoslangan hamda qoidalarga asoslangan komponentlar kombinatsiyasi nemis tili sintaktik tahlilida tobe bog'lanish hajmini 92%ga oshirdi. Olteanu va Moldovanlar PP

birikmani to'g'ri aniqlash ustida izolyatsiyaga asoslangan tahlilga yondashib eng yaxshi natijaga erishdi (92,85%). M.L.Salvatella tadqiqotlarida sintaktik tahlil jarayonlarini takomillashtirishga qaratilgan ilmiy izlanishlar olib borgan. Bu esa parserlarning samaradorligini oshirishga xizmat qiladi. Turney va Pantel tabiiy tilning ba'zi jihatlarini ifodalashga qaratilgan distribyutiv gipoteza taqdim etdi. L.Tesniere fransuz tilshunosi va grammatikasi bo'yicha eng katta yutuqlarga erishgan olimlardan biridir. U 1950-yillarda "Elements desyntaxe structurale" ("Struktural sintaksis elementlari") nomli asarida so'zlar o'rtasidagi bog'lanishni strukturalar orqali ifodalab, tilni tahlil qilishda yangi metodlardan foydalanadi. Stanley Peters amerikalik tilshunos, sintaksis va grammatikalar bo'yicha ishlanmalar yaratgan olim bo'lib, ularda sintaktik tuzilmalar grammatik shakllar asosida tushuntirilgan va qaramlikning ba'zi metodlari tadqiq qilingan. Michael Collins mashinali o'rganish va tabiiy tilni qayta ishlash (NLP) sohalarida taniqli olim hisoblanib, u dependency parsingning statistik yondashuvlarini rivojlantirishda muhim rol o'ynagan. Ko'plab mashhur dependency parserlar va ularning modellari ustida ishlagan. Markus Dreyer J.Nivre bilan hamkorlikda MaltParser va Stanford Parser kabi yirik parserlar bilan ishlagan va ular orqali tilshunoslikdagi yirik masalalarni hal qilishda yordam bergan.

Dependency Parsing sohasida yuqorida keltirilgan olimlarning yutuqlari va ilmiy ishlari bu sohaning rivojlanishiga katta hissa qo'shgan. Ular grammatik tuzilmalar va so'zlar o'rtasidagi bog'lanishlarni tahlil qilishda yangi metodlarni ishlab chiqqan va shu orqali dependency parsingni nafaqat tilshunoslik, balki tabiiy tilni qayta ishlash (NLP) va sun'iy intellekt sohalarida qo'llash imkoniyatlarini kengaytirgan.

Dependency Parsingning asosiy tamoyillari

Bugungi kundagi tadqiqotlar qaramlikni tahlil qilishning 4 ta tamoyili mavjudligini ta'kidlaydi. Ular quyidagilar:



1-rasm. Dependency Parsing tamoyillari

Bosh soʻz (Root) va soʻzlar oʻrtasidagi munosabatlar

Dependency parsingda har bir soʻz boshqa biror soʻzga bogʻlanadi va bu bogʻlanish tizimi “root” soʻzi atrofida tashkil topadi. Root soʻzi – gapdagi asosiy soʻz boʻlib, barcha boshqa soʻzlar shu asosda bogʻlanadi. Masalan, “Men kitob oʻqidim” degan jumlada “oʻqidim” – root soʻz boʻlib, “men” va “kitob” soʻzlari unga bogʻlangan.

Yoʻnaltirilgan chiziqlar (Directed edges)

Dependency parsingda bogʻlanishlar yoʻnaltirilgan chiziqlar shaklida ifodalanadi. Har bir bogʻlanish bir soʻzni boshqasiga bogʻlaydi va bu bogʻlanish yoʻnaltirilgan boʻladi. Yaʼni, bir soʻz ikkinchisiga bogʻlanib, uni boshqaradi. Masalan, “Men oʻqidim” jumlasida “oʻqidim” soʻzi “men” soʻzini boshqaradi.

Sintaktik munosabatlar

Dependency Parsingda soʻzlar oʻrtasidagi bogʻlanishlar sintaktik munosabatlar orqali tasvirlanadi. Har bir bogʻlanish sintaktik munosabatni, masalan, subyekt, obyekt, predikat va boshqa sintaktik rollarni ifodalaydi. Bunday munosabatlar grammatik jihatdan gapning toʻgʻriligini tekshiradi.

Reprezentatsiya va strukturaviy diagrammlar

Dependency parsingda gapning strukturasi tasvirlash uchun turli diagrammlar (dependency trees) ishlatiladi. Har bir soʻz node (tugun) sifatida ifodalanadi va ular oʻrtasidagi bogʻlanishlar yoʻnaltirilgan chiziqlar bilan koʻrsatiladi. Diagrammalarni koʻrish orqali tilshunoslar va kompyuter tizimlari grammatik tuzilmalarni osongina tahlil qilish mumkin.

Dependancy Parsingning asosiy yondashuvlari

Yuqoridagi tamoyillardan tashqari, jahon tilshunosligida qaramlikni tahlil qilish uchun bir qancha yondashuvlarda tadqiqotlar olib boriladi. Dependency Parsingda ishlatiladigan asosiy yondashuvlar quyidagilardir:

- **Shift-Reduce Parsing:** Bu yondashuvda soʻzlar bir-biriga qoʻshilib, tahlil qilingan soʻzlar soni kamayadi. “Shift” – bu soʻzlarni yigʻish, “reduce” esa bu soʻzlarni bogʻlash jarayonini anglatadi. Shift-reduce parsing jarayonida soʻzlar bosqichma-bosqich “stack”ga qoʻshiladi (shift), soʻngra ular oʻrtasidagi sintaktik bogʻlanishlar aniqlanib, tugunlar birlashtiriladi (reduce). Bu yondashuv oddiy va samarali boʻlib, koʻplab sintaktik analizatorlar tomonidan ishlatiladi (1-jadval).

- **Graph-based Parsing:** Bu usulda grammatik bogʻlanishlar graf shaklida ifodalanadi. Soʻzlar oʻrtasidagi bogʻlanishlar uzluksiz tarzda bogʻlanadi va eng yaxshi bogʻlanish topiladi. Bu tahlil barcha bogʻlanishlarni ehtimoliy tahlil qiladi va eng toʻgʻri sintaktik tuzilma tanlanadi. Ushbu metod orqali, jumla tarkibidagi soʻzlar va ularning oʻrtasidagi munosabatlarni aniqroq va murakkabroq tahlil qilish mumkin boʻladi. Bu yondashuv sintaktik xatoliklarni kamaytirish va grammatik jihatdan aniqroq tuzilmalar olish imkonini yaratadi (1-jadval).

- **Transition-based Parsing:** Bu yondashuvda tizim soʻzlarni ketma-ket tarzda tahlil qiladi va har bir tahlil qilingan soʻzni avvalgi soʻzlar bilan bogʻlaydi(1-jadval). Bu tahlil turi tilni tahlil qilishning samarali usullaridan biri boʻlib, bosqichma-bosqich tahlil qilishi bilan shift-reduce parsingga oʻxshab ketadi. Chunki ushbu yondashuvda ham, asosan, “shift” hamda “reduce” amallaridan foydalaniladi.

1-jadval. Dependency Parsingdagi asosiy yondashuvlar

Shift-reduce Parsing	Graph-based Parsing	Transition-based Parsing
<pre> S /\ Men VP /\ kitob o'qiyman </pre>	<pre> o'qiyman /\ Men kitob </pre>	<pre> S /\ Men VP /\ kitob o'qiyman </pre>

Dependancy parsingning amaliy qoʻllanilishi

Avtomatik xulosa chiqarishda (Automatic summarization) matnlarini qisqartirish va umumlashtirish uchun Dependency

parsing tizimlari ishlatiladi. Bu tizimlar matnni tahlil qilib, uning asosiy fikrlarini aniqlashga yordam beradi. So'zlar o'rtasidagi bog'lanishlarni aniqlash, matnning mazmunini yaxshiroq tushunishga yordam beradi. Misol:

O'zbekistonning poytaxti Toshkent juda rivojlangan shahar. Toshkentda ko'plab zamonaviy infratuzilmalar va biznes markazlari mavjud. Shahar transporti ham yaxshi rivojlangan.

1-qadamda so'zlar o'rtasidagi grammatik munosabatlar aniqlanadi:

O'zbekiston → *poytaxti* (*O'zbekiston* – subyekt, *poytaxti* – obyekt)

poytaxti → *Toshkent* (*poytaxti* – ot, *Toshkent* – nom)

Toshkent → *rivojlangan* (*Toshkent* – nom, *rivojlangan* – sifat)

Toshkent → *mavjud* (*Toshkent* – nom, *mavjud* – predikat)

mavjud → *infratuzilmalar* (*mavjud* – predikat, *infratuzilmalar* – obyekt)

mavjud → *markazlari* (*mavjud* – predikat, *markazlari* – obyekt)

Shahar → *transporti* (*Shahar* – subyekt, *transporti* – obyekt)

transporti → *rivojlangan* (*transporti* – predikat, *rivojlangan* – sifat)

2-qadamda muqobil g'oyalar aniqlanadi. Dependency parsing natijasida tahlil qilingan so'zlar va ularning munosabatlari yordamida matndan eng muhim fikrlar ajratiladi:

1. *O'zbekistonning poytaxti Toshkent* – asosiy g'oya.

2. *Toshkentda ko'plab zamonaviy infratuzilmalar va biznes markazlari mavjud* – Toshkentning rivojlantirish jihatlari.

3. *Shahar transporti ham yaxshi rivojlangan* – Toshkentning transport tizimi.

3-qadamda asosiy fikrlarni tahlil qilib, matnni qisqartirish amalga oshiriladi: "*Toshkent – O'zbekistonning poytaxti. Shaharda zamonaviy infratuzilmalar, biznes markazlari va yaxshi rivojlangan transport tizimi mavjud*".

Mashina tarjimasida ham dependancy parsingdan keng foydalaniladi. Til o'rtasidagi grammatik va sintaktik farqlarni hisobga olib, tarjimon tizimining to'g'ri va aniq tarjima qilishiga yordam beradi. So'zlar o'rtasidagi bog'lanishlar aniqlangach, tarjima jarayonini yana-da samarali qilish mumkin. Bunda 1-qadam subyekt va predikatni ajratish bo'ladi. *Men* so'zi subyekt sifatida, *o'qiyman* so'zi predikat sifatida aniqlanadi. *Kitob* obyekt sifatida predikat bilan bog'lanadi. 2-qadamda bog'lanishlar tarjima qilinadi.

Men (subyekt) "I" ga tarjima qilinadi.

o'qiyman (fe'l) "read" ga tarjima qilinadi.

kitob (obyekt) "book" ga tarjima qilinadi.

Tarjimada grammatik munosabatlar (masalan, subyekt-predikat-obyekt) saqlanadi. Tarjimada so'zlar o'rtasidagi bog'lanishlar aniqlanadi va tarjima jarayonida to'g'ri tartib va ma'no saqlanadi.

Semantik tahlilning maqsadi matnning ma'nosini tushunishdir. Dependency Parsing semantik tahlilda ham ishlatiladi, chunki so'zlar o'rtasidagi grammatik bog'lanishlar, jumlaning ma'nosini anglash uchun muhimdir. Buni quyidagi gap orqali izohlaymiz: *O'zbekiston poytaxti Toshkentda ko'plab yirik infratuzilmalar mavjud*. Ushbu gapda 1-qadam jumlani qaramlik aniqlash asosida tahlil qilish hisoblanadi (2-rasm).



2-rasm. Jumlaning sintaktik tahlili

2-qadamda semantik tahlilni aniqlash uchun dependency parsingda aniqlangan bog'lanishlardan foydalangan holda, so'zlar o'rtasidagi ma'no aloqalari tushunilishi kerak:

O'zbekiston poytaxti – O'zbekistonning poytaxti haqida gapirilmoqda.

Poytaxti → *Toshkent* – Toshkent poytaxt sifatida aniqlanadi.

Mavjud → *infratuzilmalar* – Toshkentda mavjud infratuzilmalar haqida gapirilmoqda.

Infratuzilmalar → *ko'plab* – Toshkentda ko'plab yirik infratuzilmalar mavjud.

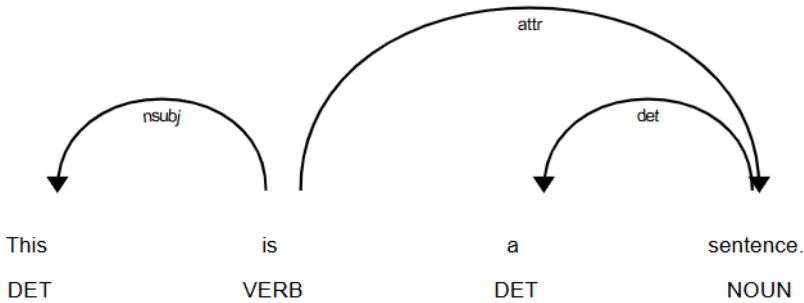
Tahlillardan shu ma'lum bo'ladiki, semantik tahlilda dependency parsingdan foydalanish grammatik bog'lanishlarni aniqlash, noaniqliklarni bartaraf etish hamda jumlaning ma'nosini to'liq tushunishga xizmat qiladi. Chunki so'zlar o'rtasidagi bog'lanishlarni tahlil qilish semantik tahlilni to'g'ri va aniq amalga oshirish uchun zarurdir.

Chatbotlar va dialog tizimlari foydalanuvchidan tushunarli javoblar olish, uning niyatini anglash va to'g'ri javoblarni taqdim etish uchun dependency parsing usulidan keng foydalanadi. Dependency parsing yordamida chatbotlar foydalanuvchining so'zlarini sintaktik jihatdan tahlil qiladi, so'zlar o'rtasidagi grammatik munosabatlarni aniqlaydi va so'zlar ma'nosini aniq tushunishga yordam beradi.

Dependency Parserlar

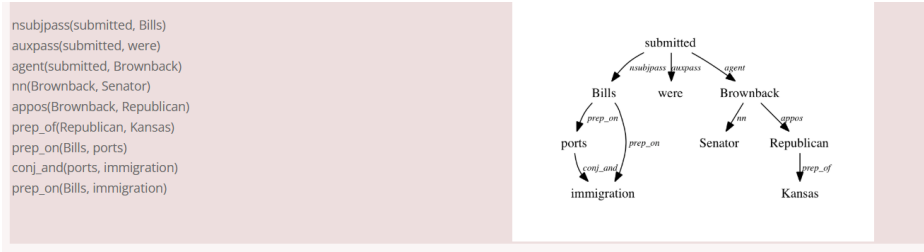
Dependency Parsingni amalga oshirishda ilg'or yondashuvlar va algoritmlar mavjud. Sun'iy intellekt va chuqur o'rganish (deep learning) texnologiyalarining rivojlanishi bilan dependency parsing yana-da aniqroq va samaraliroq bo'ldi. Ayniqsa, transformer arxitekturalari (masalan, BERT va GPT) tilning sintaktik va semantik strukturalarini tushunishda katta yordam beradi. Jahon tilshunosligida Dependency parsing uchun turli xil vositalar va parserlar ishlab chiqilgan. Bu parserlar turli tillarda so'zlar o'rtasidagi grammatik bog'lanishlarni aniqlashga yordam beradi va turli maqsadlarda ishlatiladi. Quyida ba'zi Dependency Parserlar haqida ma'lumot beramiz.

- *SpaCy* – bu Python dasturlash tilida yozilgan, yuqori samarali va tezkor NLP kutubxonasi bo'lib, u ko'plab tabiiy tilni qayta ishlash vazifalarini bajarishi mumkin. SpaCy bir nechta tilni qo'llab-quvvatlaydi va uning dependency parseri juda aniq ishlaydi. Tez ishlash va katta hajmdagi ma'lumotlar bilan ishlash imkoniyati, ma'lum bir tilga moslashtirilgan modellar, kengaytirilgan API va interfeyslar mavjudligi parserning afzalliklari hisoblanadi.



3-rasm. SpaCy parserining ishlatilgan holati

- *Stanford dependency parser* – Bu Stanford universiteti tomonidan ishlab chiqilgan dependency parsing vositasi. U mashhur va ilmiy jihatdan eng ishonchli parserlardan biridir. Stanford parseri eng ilg'or statistik modellarni va ilmiy yondashuvlarni qo'llaydi, shuning uchun u ko'plab tilshunoslik tadqiqotlarida foydalaniladi. Yuqori aniqlik va keng qamrovli tillar qo'llab-quvvatlanishi, zich va aniq sintaktik bog'lanishlarni aniqlash hamda bir necha tilni qo'llab-quvvatlash (masalan, ispan, ingliz, fransuz va boshqalar) parserning afzalliklarini tashkil qiladi.



4–rasm. Stanford Dependency Parserining ishlatilgan holati

• *Maltparser* – bu dastlabki, asosan, statistik yondashuvga asoslangan dependency parsing vositasidir. U yuqori aniqlikda ishlaydi va sintaktik tahlil uchun bir nechta statistik modellardan foydalanadi. Maltparserni ishlatish oson va u ishlab chiqish uchun juda mos keladi. Intuitivligi, qayta oʻrganishga va moshlashtirishga imkoniyatning mavjudligi hamda turli tillar uchun maxsus modellarning mavjudligi parserning asosiy afzalligidir.

Resources

Pre-trained models
Publications
CoNLL Shared Task
MaltEval
MaltOptimizer
Contact

Data Consortium to find out which conditions apply.

Running engmalt.poly and engmalt.linear

Download [engmalt.poly-1.7.mco](#) or [engmalt.linear-1.7.mco](#) into your working directory and execute the following command:

```
prompt> $ java -Xmx1024m -jar maltparser-1.7.jar -c engmalt.poly-1.7 -i infile.conll -o outfile.conll -m parse
```

or

```
prompt> $ java -Xmx1024m -jar maltparser-1.7.jar -c engmalt.linear-1.7 -i infile.conll -o outfile.conll -m parse
```

where `infile.conll` and `outfile.conll` should be replaced by the names of your input and output files.

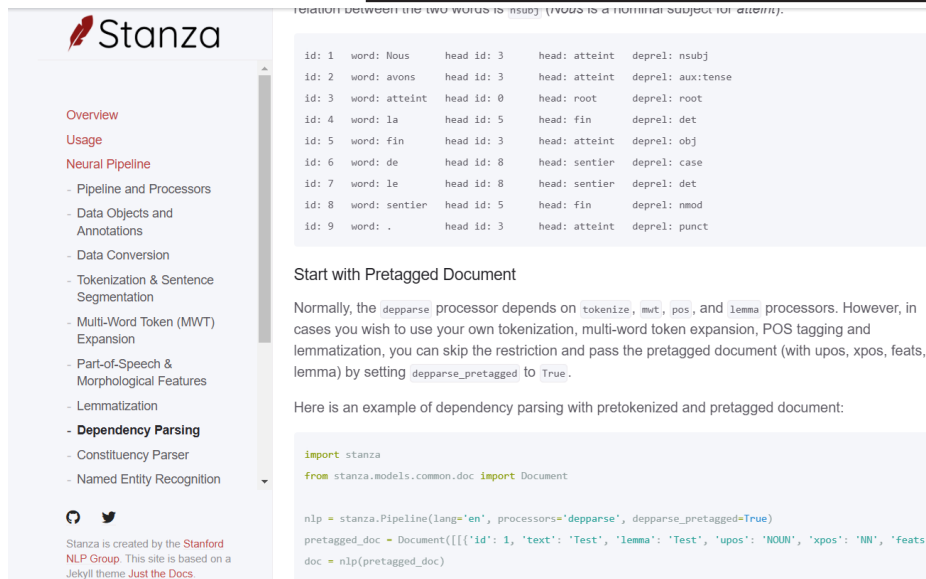
For more information, see the [MaltParser user guide](#).

Here is an example of an input sentence:

```
1  Pierre      -   NNP   NNP   -
2  Vinken     -   NNP   NNP   -
3  ' , - ' , -
4  61         -   CD    CD    -
5  years      -   NNS   NNS   -
6  old        -   JJ    JJ    -
7  ' , - ' , -
8  will       -   MD    MD    -
9  join       -   VB    VB    -
10 the        -   DT    DT    -
11 board      -   NN    NN    -
12 as         -   IN    IN    -
13 a          -   DT    DT    -
14 nonexecutive - JJ    JJ    -
15 director   -   NN    NN    -
16 Nov.       -   NNP   NNP   -
17 29         -   CD    CD    -
18 ' , - ' , -
```

5–rasm. MaltParserning dependency parsing qismi

• *Stanza* – bu Stanford NLP guruhi tomonidan ishlab chiqilgan dependency parsing dasturi. Bu dastur tilni qayta ishlashda qoʻllaniladigan juda yaxshi moslashuvchan va ochiq manba dasturdir. U koʻplab tillarni qoʻllab-quvvatlaydi va tilning sintaktik-semantik strukturasi haqida batafsil maʼlumot beradi. U BERT va boshqa modellardan foydalanib dependency parsingni amalga oshiradi. Stanza yuqori aniqlikda ishlaydi va juda intuitiv interfeysga ega.



The screenshot shows the Stanza NLP library website. On the left is a navigation menu with links: Overview, Usage, Neural Pipeline, Pipeline and Processors, Data Objects and Annotations, Data Conversion, Tokenization & Sentence Segmentation, Multi-Word Token (MWT) Expansion, Part-of-Speech & Morphological Features, Lemmatization, Dependency Parsing, Constituency Parser, and Named Entity Recognition. The main content area displays a dependency parse tree for the sentence "The relation between the two words is nsujg (verbs is a nominal subject for attaint)". The tree is shown as a table with columns: id, word, head id, head, and deprel. Below the table, there is a section titled "Start with Pretagged Document" which explains that the `depparse` processor depends on `tokenize`, `mwt`, `pos`, and `lemma` processors. It also provides an example of dependency parsing with pretokenized and pretagged document using Python code.

id	word	head id	head	deprel
1	Nous	3	attaint	nsubj
2	avons	3	attaint	aux:tense
3	attaint	0	root	root
4	la	5	fin	det
5	fin	3	attaint	obj
6	de	8	sentier	case
7	le	8	sentier	det
8	sentier	5	fin	nmod
9	.	3	attaint	punct

```

import stanza
from stanza.models.common.doc import Document

nlp = stanza.Pipeline(lang='en', processors='depparse', depparse_pretagged=True)
pretagged_doc = Document([['id': 1, 'text': 'Test', 'lemma': 'Test', 'upos': 'NOUN', 'xpos': 'NN', 'feats':
doc = nlp(pretagged_doc)

```

6-rasm. Stanza parserining kodli tahlili

Bundan tashqari Dependency Parsing sohasida ko'plab yirik va samarali parserlar mavjud: ClearNLP, AllenNLP, Neural Network-based parsers va h.k. Ko'plab zamonaviy **deep learning** va **transformer** asosidagi metodlar dependency parsingni yana-da samarali va tezroq qilishga yordam berdi. Bu parserlar sintaktik tahlilni takomillashtirish va tabiiy tilni qayta ishlashning boshqa vazifalarini amalga oshirishda juda muhim vositalar bo'lib xizmat qiladi.

Xulosa

Dependency parsing tilshunoslikda muhim o'rin tutadi, chunki u so'zlar orasidagi grammatik bog'lanishlarni aniqlashga yordam beradi. Bu yondashuvning rivoji, tilshunoslikni chuqurroq tushunishga imkon beradi va tabiiy tilni qayta ishlashda (NLP) samarali foydalaniladi. Bugungi kunda texnologiyalarning rivojlanishi bilan dependency parsing yana-da kuchayib, tilshunoslik va sun'iy intellektni birlashtirgan yangi imkoniyatlar yaratmoqda. Sohada yuqorida keltirilgan olimlarning yutuqlari va ilmiy ishlari bu tarmoqning rivojlanishiga katta hissa qo'shgan. Ular grammatik tuzilmalar va so'zlar o'rtasidagi bog'lanishlarni tahlil qilishda yangi metodlarni ishlab chiqqan va shu orqali dependency parsingni nafaqat tilshunoslik, balki tabiiy tilni qayta ishlash (NLP) va sun'iy intellekt sohalarida qo'llash imkoniyatlarini kengaytirgan.

Foydalanilgan adabiyotlar

- Salvatella M.L. Parsing ad Evaluation. Improving dependency grammars accuracy. PhD. diss – Universitat de Barcelona, 2016.
- Abdullayeva O., Xudayarova S. O'zbek tilshunosligida so'z birikmasiga ta'rif, tavsif va tasnif masalasi. // International scientifictheoretical conference on the topic: "Problems of research and education of the Uzbek language".
- Hindle D., Rooth M. Structural ambiguity and lexical relations // Computational Linguistics. –AT&T Bell Labratories., 1993. –P. 229-236.
- Collins Michael, James Brooks. Prepositional attachment through a backed-off model. –Cambridge., 1995.
- Hockemeyer. A relational adaptive tutoring hypertext WWW-environment based on knowledge space theory. 4th International conference on Computer Aided Learning and Instruction in Science and Engineering, Göteborg, Sweden., 1998. –P 417-423.
- Nivre et al. Universal dependancy annotation for multilingual parsing. –Sofia, Bulgaria., 2013. –P 92-97.
- Buchholz S., Erwin Marsi. CoNLL-X Shared task on Multilingual Dependency Parsing. –NY., 2006. –P. 149-164.
- Ballesteros M., Nivre J. MaltOptimizer: A system for MaltParser Optimization // Istanbul, Turkey.: European Language Recources Association (ELRA), 2012. –P. 2757-2763.
- Foth K., Menzel W., By. Learning the constraints weights of a dependency grammar using genetic algorithms // 3th International conference on Domain Decomposition Methods, 2001. –P. 243-247.

<https://spacy.io/api/dependencyparser>

<https://nlp.stanford.edu/software/lex-parser.shtml>

<https://www.maltparser.org/>

DEPENDANCY PARSING IN WORLD LINGUISTICS

Oqila Abdullayeva¹

Fotima O'tkirova²

Abstract. This article provides an overview of dependence parsing, its role in world linguistics, its significance, principles, practical application, and study. Discusses new approaches and methods in the field. It explains how dependency parsing is used in linguistics and how it helps in natural language processing (NLP) and many other areas. In turn, the article presents the main features of dependence analysis and its linguistic aspects, as well as the software developments created in this area. Provides information on their advantages, prospects, and advanced research.

Keywords: *Dependence parsing, syntactic analysis, parser, word combination dependent connection, main word, dependent word, principle, approach.*

References

- Salvatella M.L. Parsing and Evaluation. Improving dependency grammars accuracy. PhD. diss – Universitat de Barcelona, 2016.
- Abdullayeva O., Xudayarova S. O'zbek tilshunosligida so'z birikmasiga ta'rif, tavsif va tasnif masalasi. // International scientific-theoretical conference on the topic: "Problems of research and education of the Uzbek language".
- Hindle D., Rooth M. Structural ambiguity and lexical relations // Computational Linguistics. – AT&T Bell Laboratories., 1993. – P. 229-236.
- Collins Michael, James Brooks. Prepositional attachment through a backed-off model. – Cambridge., 1995.
- Hockemeyer. A relational adaptive tutoring hypertext WWW-environment based on knowledge space theory. 4th International conference on Computer Aided Learning and

¹Abdullayeva Oqila Xolmo'minovna - PhD, post-doctorate student at Tashkent State University of Uzbek Language and Literature named after Alisher Navoi.

E-mail: abdullayeva.oqila@navoiy-uni.uz

ORCID: 0000-0002-2524-4832

²O'tkirova Fotima Baxtiyorovna – Master of degree. Alisher Navoi Tashkent State University of Uzbek Language and Literature.

E-mail: fotimaotkirova@gmail.com

- Instruction in Science and Engineering, Göteborg, Sweden, 1998. –P 417-423.
- Nivre et al. Universal dependancy annotation for multilingual parsing. –Sofia, Bulgaria., 2013. –P 92-97.
- Buchholz S., Erwin Marsi. CoNLL-X Shared task on Multilingual Dependency Parsing. –NY., 2006. –P. 149-164.
- Ballesteros M., Nivre J. MaltOptimizer: A system for MaltParser Optimization // Istanbul, Turkey.: European Language Recources Association (ELRA), 2012. –P. 2757-2763.
- Foth K., Menzel W., By. Learning the constraints weights of a dependency grammar using genetic algorithms // 3th International conference on Domain Decomposition Methods, 2001. –P 243-247.
- <https://spacy.io/api/dependencyparser>
- <https://nlp.stanford.edu/software/lex-parser.shtml>
- <https://www.maltparser.org/>

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.