

news transcription”, in *Proceedings of the V International scientific-practical conference information society*, pp. 48-50.

8. Kozhirbayev, Zh and Islam, Sh 2016, “A distributed platform for speech recognition research”, in *Proceedings of the V International scientific-practical conference information society*, pp. 38-40.

9. Makhambetov, O, Makazhanov, A, Yessenbayev, Zh, Matkarimov, B, Sabyrgaliyev, I and Sharafudinov, A 2013, “Assembling the Kazakh Language Corpus”, In *Proceedings of the 2013 Conference on Empirical Methods in Natural Language Processing*, pp. 1022–1031.

10. Can, D and Saraclar, M 2011, “Lattice indexing for spoken term detection”, *IEEE Trans. Audio Speech Lang. Process.*, pp. 2338–2347.

УДК 681.5

ТАБИГЫЙ ТИЛДЕГИ ТЕКСТТЕРДИ ОРУС ТИЛИНЕН КЫРГЫЗ ТИЛИНЕ МАШИНАЛЫК КОТОРУУДА СӨЗДӨРДҮ АНАЛИЗДӨӨНҮН АЛГОРИТМИН ТҮЗҮҮ

Кочконбаева Буажар Осмоналиевна, старший преподаватель, ОшТУ им. академика М.М. Адышева, 714018, г. Ош, ул. Исанова 81, e-mail: buajar@mail.ru

Морфологиялык талдоо жана сөз түзүү процесстери бири бирине жакын болуп эсептелет. Машиналык которуу процессинде бир эле учурда сөздөрдү морфологиялык жактан талдап аны жаны тилге которуу үчүн сөз жасоого туура келет. Бул процесстер бүгүнкү күндө актуалдуу проблемалардан болуп эсептелет. Анткени бардык эле улут өз тилинин унутулбай автоматташтырылышын жана интернет сайттарына кошулуусу үчүн аракет кылышууда. Ошондуктан мен бул макалада табигый тилдеги тексттерди орус тилинен кыргыз тилине которуучу программаны түзүүнүн алгоритмин карап чыктым.

Ачкыч сөздөр: машиналык которуу, морфологиялык анализатор, аффикстер, синтаксистик анализатор, семантикалык анализатор, программа, маалыматтар базасы

DEVELOPMENT OF ALGORITHM ANALYSIS OF WORDS IN NATURAL TEXT MACHINE TRANSLATION FROM RUSSIAN INTO KYRGYZ

Kochkonbaeva Buazhar Osmonalievna, senior lecturer, OshTU after named acad. M.M. Adishev, 714018, c. Osh, Isanov st. 81, e-mail: buajar@mail.ru

Morphological analysis of words and word-formation are very similar in how to conduct and are closely related to each other. Today these processes are very actual. Because all the nations seek to automate their language and add to the online translators. In this article I will consider the algorithm of a program algorithm machine translation of texts from Russian to Kyrgyz language.

Keywords: machine translation, Morphological analyzer, affixes, syntactical analyzer, semantic analyzer, program, database.

Киришүү

Жасалма интеллект багытындагы адам баласынын изилдоолору кун санап өсүп барат. Азыркы күндө компьютер бир гана эсептөө техникасы эмес, ар кандай багыттагы суроолорго жооп берүүчү эксперттик система, бир тилден экинчи тилге тексттерди которуучу каражат же болбосо интеллектуалдык оюндарды ойноодо экинчи тарап катары жасалма интеллекттин ролун жаратып келүүдө. Бул макалада жогоруда айтылган багыттардын ичинен табигый

тилдеги тексттерди орус тилинен кыргыз тилине машиналык которууда сөздөрдү анализдөөнүн компьютердик моделин (морфологиялык анализатор) түзүү маселеси каралды.

Тексттин үстүнөн иштөөнүн этаптары

Машиналык которуу – бул компьютердик программанын жардамында бир табигый тилдеги текстти экинчи табигый тилге которуу болуп эсептелет [1]. Түзүлгөн программанын сапаты которуунун тактыгынан көз каранды. Бүгүнкү күндө коммерциялык багыттагы көптөгөн программалар (PROMPT, Lingvo, Google сайтынын которуу кызматы ж.б.) атайын түзүлгөн сөз айкаштардын корпустук сөздүгүнүн жардамында тексттерди которуп келет. Бул программалардын ичинен Google сайтынын которуу сервистик программасына тексттерди кыргыз тилине которуу мүмкүнчүлүгү кошулду.

Машиналык которуу программасын түзүү өз ичине графометикалык, морфологиялык, синтаксистик жана семантикалык этаптарды камтыйт. Морфологиялык анализдөө модулуна кийирилген орус тилиндеги текст лексемаларга бөлүнөт жана алардын морфологиялык касиеттери (сөз түркүмү, жагы, чагы) аныкталат. Кийинки этап синтаксистик анализаторго берилет да, анда ар бир сөз сүйлөмдө өз ордуна жайгашат. Мисалы кыргыз тилинде сүйлөмдөр SOV (Subject – object - verb) тартибине баш ийет. Башкача айтканда ээлик состав сүйлөмдүн башында жайгашат жана баяндоочтук состав сүйлөмдүн аягынан орун алат. Мисалы: На полях работали новые машины. Бул сүйлөмдү кыргыз тилине которууда төмөнкү тартиптеги сүйлөм алынат: Жаңы машиналар талааларда иштешти. Кийинки этапта семантикалык анализатордун жардамында сүйлөмдөгү сөздөрдүн бири-бири менен байланышы каралат да сүйлөмдүн маанисине көңүл бурулат.

Машиналык которууда ар түрдүү маани берүүчү сөздөрдү анализдөө

Бардык түрк тилдери сыяктуу эле кыргыз тили да агглютинативдик тилдерге кирет. Тексттин толуктугу жана сөздөрдүн мааниси жактан байланышы унуга аффиксттердин уланышы менен ишке ашат. Колдонмо лингвистикалык процесстерде сөз түшүнүгүнүн жалпы моделин түзүү өтө татаал, анткени ар бир тилде өзүнө тиешелүү касиеттер жана чектөөлөр кабыл алынган. Сөз жасоонун негизги эрежелерине таянып төмөнкү туюнтманы жазууга болот:

$$\begin{aligned} <\text{сөздүн грамматикалык формасы}> ::= <\text{сөз}> <\text{грамматикалык көрсөткүчтөрү}> \\ <\text{грамматикалык көрсөткүчтөр}> ::= <\text{аффикс}> | <\text{предлог}> | <\dots> \\ <\text{сөз}> ::= <\text{зат атооч}> | <\text{сын атооч}> | <\dots> \end{aligned}$$

Изилдөөлөрдүн негизинде Delphi чөйрөсүндө морфологиялык анализатордун программасы түзүлдү. Сөздөрдүн морфологиялык касиеттерин камтыган таблица Access чөйрөсүндө түзүлүп, төмөнкүдөй мамычалардан турат (1-сүрөт).

ID	word_r	w_id	ch_r	id_chislo	id_padej	id_rod	koren	word_k
315	будут	332	2	2	0	0	буд	
316	голос	333	1	1	1	2	голос	үн
317	лица	334	1	2	1	3	лиц	жүздөр
318	солнце	335	1	1	1	3	солн	күн

1-сүрөт. Морфологиялык таблица

Мында сөздөр морфологиялык касиеттерине жараша бир канча мамычаларга бөлүнүп көрсөтүлгөн.

Морфологиялык анализатордун программасын түзүүдө орус тилиндеги көп колдонулган сөздөрдүн ар кандай морфологиялык абалынан турган 10000 ге жакын сөздөр маалыматтар базасына кийирилип, Delphi программалоо чөйрөсүнүн жардамында колдонуучу үчүн интерфейс түзүлдү.



2-сүрөт. Программанын интерфейси

Сөздөрдү жаратуунун ар кандай жолдорун техникалык жактан мүнөздөө өтө оор процесс, анткени ал көптөгөн эрежелерден жана чектөөлөрдөн турат. Бир канча унгудан турган сөздөрдү бир тилден экинчи тилге которуу кийирилген сөзгө бир дагы окшобогон башка морфологиялык сөздүн жаралуусуна алып келет. Мисалы: ВУЗ- ЖОЖ, малоисследованный- аз изилденген ж.б.

Төмөндөгүдөй белгилөөлөрдү кийирип алабыз: w-сөз, p-префикстер, r-сөздүн унгусу, a-аффикстердин көптүгү.

$$\langle w \rangle ::= \langle r \rangle | \langle p \rangle \langle r \rangle | \langle r \rangle \langle a \rangle | \langle p \rangle \langle r \rangle \langle a \rangle$$

Мисалы: игра-оюн, игрушка – оюнчук, играть-ойноо, выиграть – утуу, проиграть- утулуу.

Бул мисалдарда сөздөрдүн бир эле унгуга морфологиялык бирдиктердин уланышынын негизинде маанилик жактан өзгөрүшү көрсөтүлдү.

Бир табигый тилден экинчи тилге машиналык которууда көп маанилүүлүктү чагылдыруу методун пайдалануу которуунун эффективдүүлүгүн жогорулатып, тексттин маанисине жараша көп маанилүүлүктөн бир мааниге өтүүгө жардам берет.

Корутунду

Негизинен машиналык которуу багытында иштөөдө төмөнкүдөй амалдар аткарылды:

- көп колдонулуучу сөздөрдүн маалыматтар базасы түзүлдү;
- сөздөрдүн морфологиялык өзгөчөлүктөрү базага кийирилди;
- морфологиялык анализатордун программасы иштелип чыкты;
- жөнөкөй тексттерди сүйлөмдөргө ажыратуу жана которуу амалдары аткарылды.

Мындан сырткары көп маанилүүлүктү чагылдыруу методун программада пайдалануу маселеси каралды.

Колдонулган адабияттар

1. Ж.Ч. Джанузакова «Лингвокультурологические аспекты машинного перевода» [электрондук булак].
2. К.Х. Ким, А.П. Савинов «Синтаксический анализатор для вопросно-ответной системы»
3. А. Леонтьева, И. Кагиров «Автоматический синтаксический анализ русских текстов». Труды 10-й Всероссийской библиотеки: перспективные методы и технологии, электронные коллекции»-RCDL'2008, Дубна, Россия, 2008.

УДК 004.415.2

ПРОЕКТИРОВАНИЕ АВТОМАТИЗИРОВАННОЙ СИСТЕМЫ ПРОВЕРКИ ОЛИМПИАДНЫХ ЗАДАНИЙ ПО ПРОГРАММИРОВАНИЮ

Макиева Замира Джумакматовна, доцент КГТУ им. И.Раззакова, Кыргызстан, 720044, г.Бишкек, пр. Ч.Айтматова 66, e-mail: z.makieva@gmail.com