

**ЎЗБЕКИСТОН РЕСПУБЛИКАСИ
РАҚАМЛИ ТЕХНОЛОГИЯЛАР ВАЗИРЛИГИ**

**МУҲАММАД АЛ-ХОРАЗМИЙ НОМИДАГИ ТОШКЕНТ АХБОРОТ
ТЕХНОЛОГИЯЛАРИ УНИВЕРСИТЕТИ**

**МАТЕМАТИК МОДЕЛЛАШТИРИШ,
АЛГОРИТМЛАШ ВА ДАСТУРЛАШНИНГ
ДОЛЗАРБ МУАММОЛАРИ**

**Республика илмий-техник анжумани материаллари тўплами
Тошкент, 2023-йил 14-15-сентябрь**

**Материалы республиканской научно-технической конференции
АКТУАЛЬНЫЕ ПРОБЛЕМЫ
МАТЕМАТИЧЕСКОГО МОДЕЛИРОВАНИЯ,
АЛГОРИТМИЗАЦИИ И
ПРОГРАММИРОВАНИЯ**

Ташкент, 14-15 сентября 2023 г.

**CURRENT PROBLEMS OF MATHEMATICAL
MODELING, ALGORITHMIZATION AND
PROGRAMMING**

**Materials of the republican scientific and technical conference
Tashkent, September 14-15, 2023**

Toshkent – 2023

UDK:56.162.4.49

КБК: 14.523.1

M19

“Математик моделлаштириш, алгоритмлаш ва дастурлашнинг долзарб муаммолари” мавзусидаги Республика илмий-техник конференцияси материаллари тўплами. 2023-йил 14-15 сентябрь. – Тошкент., 393 б.

Мазкур тўпланда математик моделлаштириш, алгоритмлаш ва дастурлашнинг долзарб муаммолари, мураккаб тизимларда моделлаштириш, ижтимоий-иқтисодий соҳаларда алгоритмлаштириш, маълумотларни интеллектуал таҳлил қилиш ва қарор қабул қилиш, дастурлаш технологиялари, соҳада замонавий инновацион ва педагогик технологияларни қўлланилиши, математик моделлаштириш, алгоритмлаш ва дастурлашни ривожлантиришга оид амалга оширилган илмий изланишлар ҳамда тадқиқотлар натижалари билан илмий жамоатчиликни таништиришга оид ташкилий, илмий-назарий ва методологик тадқиқот ишлари натижалари ёритилган.

Тўпламга киритилган мақолалар мазмуни, илмий салоҳияти ва далилларнинг ҳаққонийлиги учун муаллифлар мас’улдирлар.

Маъсул муҳаррирлар:

ф.-м.ф.д., профессор **М.Арипов**, т.ф.д., профессор **Ф.М.Нуралиев**

Тақризчилар:

т.ф.д., профессор **Н.Равшанов**, т.ф.д., профессор **Ҳ.Н.Зайниддинов**,
т.ф.д., профессор **А.В.Кабулов**

Муҳаррир *Ш.А.Анарова, Қ.С.Раҳманов*

Рассом *Э.Э.Сайфиев*

Рассом муҳаррир *Э.Э.Сайфиев*

Техник муҳаррирлар *Э.Ш.Назирова, Ш.Б.Абидова, О.Ш.Абдирозиқов*

©Муҳаммад Ал-Хоразмий номидаги Тошкент
Ахборот Технологиялари Университети
© “Zamon poligraf ” nashriyoti 2023

СТЕММАТИЗАЦИЯ И ГЕНЕРАЦИЯ СЛОВОФОРМ В УЗБЕКСКОМ ЯЗЫКЕ ДЛЯ СИСТЕМ АВТОМАТИЧЕСКОЙ ОБРАБОТКИ ТЕКСТОВ

Боймуродов Ф.Ф.

Докторант, Каршинский государственный университет

Аннотация. В качестве исследования намечена планируется разработать алгоритмы синтеза и анализа словоформ узбекского языка, основанные на принципах разбиения слов на флективные классы. Поскольку узбекский язык является агглютинативным, подключать словарь словоформ для автоматизации морфологического анализа нецелесообразно. Значительно эффективнее пользоваться словарями аффиксов и наборами правил. В процессе выполнения исследования была построена морфологическая модель узбекского языка, которая реализована алгоритмически. В частности, словари были структурированы по классам флексий для существительных, прилагательных и глаголов.

Ключевые слова. генерация, стемматизация, Портера, аффикс, суффикс.

Введение

Процесс информирования общественности становится требованием современности. Все сферы деятельности человека неразрывно связаны с процессами получения, обработки и усвоения информации. Сегодня компьютерные и информационные технологии становятся основой нашей повседневной жизни, поскольку они быстро обновляются, изменяются и развиваются. В то же время Узбекистан постепенно входит в систему оцифровки. В связи с этим развитие компьютерной лингвистики является актуальной задачей.

Целью статьи является введение узбекского языка в компьютерную лингвистику, разработка эффективного способа усвоения основных понятий, морфологического образования узбекского языка.

Морфологическая основа узбекского языка характеризуется большим разнообразием аффиксов, добавляющихся к основе слова в определенной последовательности и придающих ему самые разные характеристики, а также огромным количеством глагольных форм, позволяющих создавать до 150-200 вариантов, отличающихся по смыслу, порой, кардинально, но при этом образованных от одного глагола. Программные средства, которые автоматически находят нужные словоформы в изучаемых текстах, являются важным помощником в проведении лингвистических исследований.[5]

Диалект - помогает людям, живущим в определенном районе, взаимодействовать друг с другом. В узбекском языке диалект отличается от литературного языка более или менее фонетически, лексически и грамматически.

Есть только устная форма народных диалектов.

Набор тесно связанных диалектов называется диалектом (диалект - это сочетание диалектов и диалектов).

В узбекском языке три диалекта:

1. Карлукский диалект (юго-восточная группа).
2. Кипчакский диалект (юго-западная группа).
3. Огузский диалект (северо-западная группа).

Узбекский - тюркский язык карлукской группы, который относится к категории синтетических агглютинативных языков и имеет богатую и сложную морфологию. Как и в других тюркских языках, слово состоит из основы, к которой прикреплены суффиксы, обозначающие различные грамматические особенности.[4]

Карлукский диалект в основном включает городские диалекты (Ташкентский, Андижанский, Ферганский, Бухарский). Важными фонетическими и морфологическими особенностями этих диалектов являются:

- 1) Буква “к” в конце слова произносится как “у” : elak - elay, terak - teray.
- 2) есть деление: aka - oka, Akram - Akrom.
- 3) в форме винительного падежей, “ –ni ”: ukamni daftari - тетрадь моего брата.

Узбекский язык включает ряд диалектов, такие диалекты с этнолингвистической точки зрения он объединен в три основных диалекта - карлукский, кипчакский и огузский. Естественно, что эти три диалекта приняли участие в формировании современного узбекского литературного языка.

При изучении языка большое значение имеет способность правильно понимать структуру слова и определять его компоненты. Слово отражает структурные особенности языка, его лексико-семантические и функционально-грамматические законы. По своей типологии и морфологической структуре узбекский язык значительно шире обычных выражений и отличается относительной регулярностью, позиционной и грамматической устойчивостью морфологической структуры различных словоформ. Слова образуются путем добавления грамматических частиц к основанию слов – аффиксов. [Барахнин, 2017, с. 435].

Результатом этой статьи является электронная система для изучения и обучения узбекскому языку, которая может работать в локальных и глобальных сетях и должна выполнять следующие задачи:

- Знакомство с курсами узбекской грамматики;
- Разработать системную базу данных на основе собранных данных;
- Разработка интерфейса пользователя и менеджера;
- Разработка пользовательских и управляющих модулей;
- Организация системного управления;
- Разработка системного дизайна;

Когда мы говорим, что программа для создания слов на узбекском языке является научной новинкой, мы имеем в виду, что она была создана с использованием новых технологий, а также имеет следующие преимущества при использовании морфологических систем. Система обладает следующими удобствами и преимуществами:

- Любой пользователь системы может пользоваться системой через Интернет;
- Не требует программного обеспечения от пользователя, использующего систему;
- удобный интерфейс, позволяющий пользователям без специальной подготовки;
- Предоставить любому пользователю хотя бы некоторые навыки узбекской грамматики;

Вид приложения показан на рисунке ниже. На разделе нашего меню посвящен созданию словоформ имена существительные (Генератор существ-ных). Пользователи должны ввести слово при использовании приложения. Наша программа предлагает все формы слова, в зависимости от значения слова. В основе нашей программы - суффиксы, составляющие целую словосочетание узбекского языка. Это наиболее распространенные случаи на узбекском языке. На узбекском языке большинство имен существительных создаются в основном с использованием определенных суффиксов.

Система проверяет в базе основ корень слова, полученного из предыдущего шага, если результат успешный, то алгоритм заканчивает свою работу. А если нет, то он сохраняет эту комбинацию аффиксов и возвращается к шагу №2. В шаге №2 он уже не будет использовать те аффиксы, то были сохранены в шаге №3.

После нормализации для каждого слова осуществляется вычисление его морфологических характеристик и определяется морфологический класс основы и представляет пользователю.

Основные этапы применения метода.

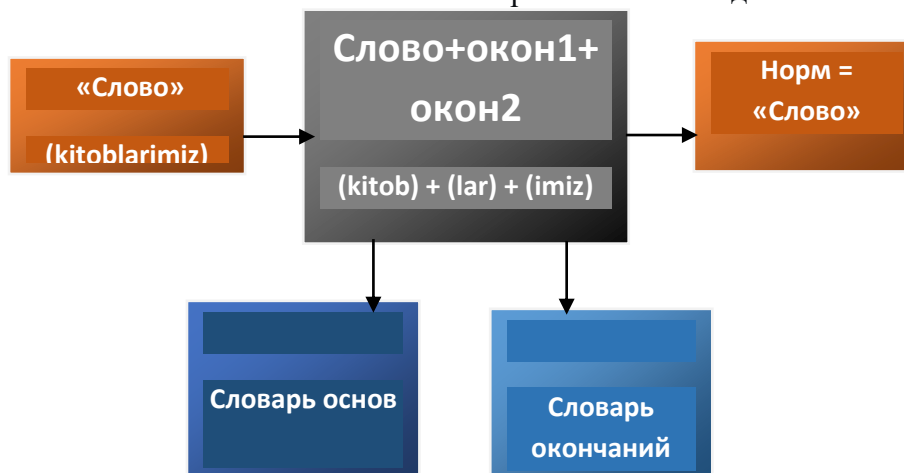


Рис. 1. Процесс определения нормальной формы слова и его морфологических параметров.

На вход подается словоформа, происходит поиск в словарях аффиксов *-lar*, *-imiz* и основы слова - *kitob*. На основе морфологического класса основы (существительное) и аффиксов вычисляем морфологическую информацию: *-lar* «множественное число», *-imiz* «притяжательное окончание».

Пример работы приложения:

1. Генератор существительных: В этом разделе мы можем создавать словоформы существительных в узбекском языке. От пользователя требуется ввести корневую форму существительных в соответствующем месте.

давайте попробуем ввести существительное, “*kitob*” – “книга”:

Генератор глаголов Стемматизатор глаголов **Генератор существ-ных** Стемматизатор существ-ных Прилагательное

 Анализация

Конфигурация слово образования	Форма слова
Начальный падеж	kitob
Притяжательный падеж	kitobning
Винительный падеж неод	kitobni
Направительный падеж	kitobga
Местно-временной падеж	kitobda
Исходный падеж	kitobdan
Множественное число	kitoblar
1 лицо. ед. число	kitobim
1 лицо. мн. число	kitobimiz
2 лицо. ед. число	kitobing
2 лицо. мн. число	kitoblaring
2 лицо. (ув) ед. число	kitobingiz
2 лицо. (ув) мн. число	kitoblaringiz

2. Стемматизатор существительных: в этом разделе можно определить основную часть существительных, то есть корневую форму существительных.

давайте попробуем ввести существительное, “*bolalarimizning*” – “наших детей”. коренное слово, “*bola*” – “ребёнок”:

[Генератор глаголов](#)
[Стемматизатор глаголов](#)
[Генератор существ-ных](#)
[Стемматизатор существ-ных](#)
[Прилагательное](#)

Конфигурация слово образования	Форма слова
stem	bola

3. Генератор глаголов: В этом разделе мы можем создавать словоформ глагола в узбекском языке. От пользователя требуется ввести инфинитивную форму глагола в соответствующем месте: например, мы можем увидеть результаты, введя следующий пример:

[Генератор глаголов](#)
[Стемматизатор глаголов](#)
[Генератор существ-ных](#)
[Стемматизатор существ-ных](#)
[Прилагательное](#)

Конфигурация слово образования	Форма слова
Отрицание	ketma
Вопрос	ketdimi?
Результативно прошедшее время	ketgan
Конкретное настоящее время	ketyapti
Переходное время	ketadi
Прошедшее время	ketdi
Будущее продолжительное время	ketadi
Будущее время намерения	ketmoqchi
Переходное прошедшее время	ketgan edi
Давнопрошедшее время	ketibdi
Результативно прошедшее время. + Личное окон. 1л ед.ч.	ketganman
Результативно прошедшее время. + Личное окон. 2л ед.ч.	ketgansan
Результативно прошедшее время. + Личное окон. 2ул.л ед.ч.	ketgansiz

4. Стемматизатор глаголов: в этом разделе можно определить основную часть глагольных, то есть инфинитивную форму глагола. давайте попробуем ввести глагол, “ketmoqchisiz” – “Вы хотите уйти”:

[Генератор глаголов](#)
[Стемматизатор глаголов](#)
[Генератор существ-ных](#)
[Стемматизатор существ-ных](#)
[Прилагательное](#)

Конфигурация слово образования	Форма слова
stem	ketmoq

ЛИТЕРАТУРА

1. Барахнин В.Б., Федотов А.М., Бакиева А.М., Бакиев М.Н. Тажибаева С.Ж., Батура Т.В., Кожемякина О.Ю., Тусупов Д.А., Самбетбаева М.А., Лукпанова Л.Х. Алгоритмы генерации и стемматизации словоформ казахского языка // Cloud of Science. – 2017. – Т. 4. – № 3. – С. 434-449. <https://cyberleninka.ru/article/n/algorithmy-generatsii-i-stemmatizatsii-slovoform-kazahskogo-yazyka/viewer>

2. Барахнин В.Б., Федотов А.М., Бакиева А.М., Бакиев М.Н., Тажибаева С.Ж., Батура Т.В., Кожемякина О.Ю., Тусупов Д.А., Самбетбаева М.А., Лукпанова Л.Х. Алгоритмы генерации и стемматизации словоформ казахского языка // Cloud of Science. - 2017. - Т.4. - № 3. - С.434-449. – EISSN 2409-031X. https://cloudofscience.ru/sites/default/files/pdf/CoS_15_434.pdf

3. Эшкараева Н.Г, Боймуродов Ф.Ф Разработка и реализация алгоритма

стемматизации и генерации словоформ для глаголов узбекского языка
http://qarshidu.uz/source/JURNAL/Aniq_fanlar_5-son_2022.pdf

4. Боймуродов.Ф., Algorithm development and implementation stemmatization and generation of word forms of uzbek language for automatic processing systems texts, https://www.researchgate.net/publication/347995316_algorithm_development_and_implementation_stemmatization_and_generation_of_word_forms_of_uzbek_language_for_automatic_processing_systems_texts

5. Эшкараева Н.Г, Боймуродов Ф.Ф Создание словоформ существительных в узбекском языке с использованием алгоритма портера. <http://qarshidu.uz/uz/page/ilmijurnal>

6. Эшкараева Н.Г, Боймуродов Ф.Ф. Разработка программу словоформ узбекского языка с использованием алгоритма портера.

MUNDARIJA

<i>IV SHO'BA. DASTURLASH TEXNOLOGIYALARI</i>	
СТРУКТУРА ПРОГРАММНОГО КОМПЛЕКСА ГЕОМЕТРИЧЕСКИ НЕЛИНЕЙНЫХ ЗАДАЧ ИЗГИБА ТЕРМОУПРУГИХ ПЛАСТИН СО СЛОЖНОЙ КОНФИГУРАЦИЕЙ Анарова Ш.А., Абдирозиқов О.Ш.	93-98
MILLIY H-INDEKSINI ANIQLOVCHI PLATFORMANING IDEF MODELLARI Nazirova E.Sh., Xojiyev S. A., Erkinov Sh., Sh.	99-103
YANGI AVLOD O'QUV LUG'ATLARI VA ULARNING MOBIL ILOVALARINI YARATISH Abidova Sh.B., Uzakova M.A.	103-107
TITLE: TRANSFORMING UZBEKISTAN'S EDUCATION LANDSCAPE: THE FUSION OF IT AND TESTING SYSTEMS Sadullayev S.A.	108-110
DFIT TIZIMINI ISHLAB CHIQISH JARAYONINI MODELLASHTIRISH Nishanov A.X., Akbaraliyev B.B., Xoliqnazarov R.H.	110-114
MATNLARNI TO'LDIRISH ALGORITMLARIDA QO'LLANILADIGAN METODLARI Bakayev I.I.	114-118
KOMPYUTER SAVODXONLIGI MAKTAB O'QUVCHILARINING FUNKSIONAL SAVODXONLIGINING TARKIBIY QISMLARIDAN BIRI Bakayeva R.I.	118-123
BIOTIBBIY SIGNALARNI VIZUALLASHTIRISH Zayniddinov H.N., Aliqulov A.X.	124-129
AQLLI UY XAVFSIZLIK TIZIMIDA IoT SENSORLARINI O'RNI Seungjik Lee Professor, Maxmudjanov S.U., Mamirov X.X.	129-133
TABIIY TILLI MATNLARNI AVTOMATIK QAYTA ISHLASH TIZIMLARI TAHLILI Nazirova E.SH, Boymurodov F.F	133-137
LINEAR PREDICTORS FOR NONLINEAR FED-BATCH SYSTEMS USING KOOPMAN OPERATORS Shuang Zhao, Jinlong Yuan	137-143
НЕСТАЦИОНАРНОЕ ВРАЩЕНИЕ ЖЕСТКОГО ШАРА ВБЛИЗИ ДРУГОГО ЖЕСТКОГО ШАРА В ПОРИСТО-УПРУГОМ ПРОСТРАНСТВЕ, НАСЫЩЕННОМ ЖИДКОСТЬЮ Мусурмонова М.О.	143-147
ATMOSFERADA ZARARLI MODDALARNING KO'CHISH VA TARQALISH JARAYONINI IFODALOVCHI MATEMATIK MODELNI ISSIQLIK ENERGIYASINI HISOBGA OLGAN HOLDA TAKOMILLASHTIRISH Muradov F.A., Narzullayeva N.U., Tashtemirova N.N.	147-153
TASVIR SEGMENTATSIYASI VA TASVIRNI QAYTA ISHLASH Muxamadiyev A.Sh., Abdullayev Y.Y.	153-157
НАВИГАЦИЯ В ДОПОЛНЕННОЙ РЕАЛЬНОСТИ МЕТОДОМ SLAM Артикова М.А., Сайфиев Э.Э., Талипова О.Х.	157-161
ЎҚУВ-ЖАНГОВАР Л-39 САМОЛЁТИНИНГ СИМУЛЯТОР-ТРЕНАЖЕРИНИНГ ДАСТУРИЙ ТАЪМИНОТ ВА АППАРАТ ВОСИТАЛАРИНИ МОДЕЛЛАШТИРИШ ПЛАТФОРМАСИ Дусматов Х.А.	161-164
СТЕМАТИЗАЦИЯ И ГЕНЕРАЦИЯ СЛОВОФОРМ В УЗБЕКСКОМ ЯЗЫКЕ ДЛЯ СИСТЕМ АВТОМАТИЧЕСКОЙ ОБРАБОТКИ ТЕКСТОВ Боймуродов Ф.Ф.	164-168
ЎЗБЕКЧА МАТНЛАРНИ АВТОМАТИК ҚАЙТА ИШЛАШДА ГЕНЕРАТИВ ГРАММАТИКАНИНГ АҲАМИЯТИ Норов А.М., Тоғаев И.Б., Сафаров Л.С., Жорабеков Т.К.	168-173
НЕЧЕТКО-СИТУАЦИОННЫЙ АНАЛИЗ И УПРАВЛЕНИЕ ТЕХНОЛОГИЧЕСКОЙ БЕЗОПАСНОСТЬЮ НЕФТЕХИМИЧЕСКИХ УСТАНОВОК И КОМПЛЕКСОВ Ганиев А.А.	173-177
LANDSHAFT LOYIHALARINI GEOMETRIK MODELLASHTIRISHDA DASTURIY PAKETLARDAN FOYDALANISH Muxamadiyev A.Sh., Bahromov A.A.	178-181