

УДК 681.3:629.73

Г.Г. АБДУЛЛАЕВА, И.А. ИСМАИЛОВ

КОНСТРУКЦИЯ БАТАРЕИ ЭКСПЕРТНЫХ СИСТЕМ ДЛЯ УСТАНОВЛЕНИЯ ЭТИМОЛОГИЙ ЭТНОНИМОВ (НА ПРИМЕРЕ ОГУЗСКИХ ЭТНОНИМОВ)

Рассматривается идея разработки батареи экспертных систем (БЭС) для решения мультидисциплинарных задач – установления этимологий этнонимов, которая решается на примере 24 огузских этнонимов. Показана сложность таких задач и невозможность их решения с помощью специалистов узкого профиля. Предлагаемая батарея конструируется из элементов: тюркская ЭС, хотанская ЭС, китайская ЭС, блоки аргументов и анти-аргументов, блок анализа и принятия решений. Подробно описывается тюркская фонетическая ЭС и база знаний, с представлением некоторых экспертных правил. Представлены некоторые результаты тестирования прототипа батареи ЭС.

Ключевые слова: экспертная система, база знаний, правила, батарея экспертных систем

1. Введение. Как известно в конце 70-х годов двадцатого столетия в сфере Искусственного Интеллекта была принята новая концепция интеллектуализации компьютерных программ путём снабжения их множеством высококачественных специальных знаний в отдельных предметных областях. Такие компьютерные программы, ставшие экспертами в узких предметных областях, получили название экспертных систем (ЭС), которые консультируют и/или частично заменяют экспертов – специалистов в их трудовой деятельности. Основу ЭС составляет база знаний, которая представляет собою модель предметной области и состоит из комплекса предметных фактов – знаний в форме достоверно установленных утверждений и так называемых экспертных правил – эмпирических правил, упрощающих или ограничивающих поиск решений в сложных предметных областях.

В результате анализа современных экспертных систем нами установлено, что в настоящее время находят развитие батареи экспертных систем (БЭС), в качестве примеров можно упомянуть, например психодиагностическую БЭС «Ориентир» [1, с.253-270], батарею интегрированных экспертных систем (БИЭС) для задач определения коллекторских свойств и нефтегазоносности пластов [2, с.78-79]. Отличительной особенностью БЭС является то, что здесь в рамках единой интеллектуальной системы объединяются результаты, получаемые от отдельных ЭС различной модальности. Разработка БЭС предполагает, что здесь чётко различается несколько задач [1, с.254-255].

2. Постановка задачи. В настоящее время с помощью специалистов высокого уровня в одной определённой узкой предметной области невозможно решать сложные задачи, которые относятся ко многим смежным и не смежным областям науки и производства. Даже очень продвинутые экспертные системы, которые, как известно, отличаются своей конкретностью, узкой направленностью [3, с.14], в чём и состоит их сила (известно, что чем уже предметная область, тем эффективнее ЭС) не в состоянии решать мультидисциплинарные задачи, требующие привлечения знаний и навыков специалистов из разных отраслей науки и производства. Например, задача установления правильных этимологий этнонимов (этнонимы - наименования родов, племён, народов и наций). Этимологизировать – значит устанавливая первоначальное (истинное, основное) значение слова, т.е. отыскивать то исходное слово (этимон), от которого произошло рассматриваемое слово [4, с.10]. Не только зарождение, но и дальнейшее историческое изменение этнических имён подвержены ряду случайностей; это вызывает чрезвычайные трудности этимологизации в большинстве

случаев, часто усугубляемые недостаточностью источников [5, с.159]. Этимологическому исследованию свойственна множественность возможных решений, проблематичность, гипотетичность [4, с.13]. Несмотря на то, что такая сложная задача как установление этимологий этнонимов относится к задачам этнонимии, изучающей этнонимы, для решения такой задачи, кроме решения чисто этнонимических задач нужно решать исторические, лингвистические, географические и другие задачи, т.е. нужны знания историков, лингвистов по разным языкам, специалистов по фольклору, мифологии, топонимиков, литераторов, географов и т.д.

Таким образом, для решения задачи установления этимологий этнонимов очень актуальная и для ее решения требуется объединение знаний, перечисленных выше наук. Учитывая вышеперечисленные, для установления этимологий этнонимов в работе предложена разработка батареи экспертных систем, которая объединяет знания, относящиеся к разным аспектам, в рамках единой интеллектуальной системы.

3. Решение. В статье предлагается конструкция батареи экспертных систем для решения задачи установления правильных этимологий этнонимов на примере этнонимов огузских (тюркских) племён общее число которых 24. Наименования этих 24 племён приводятся в нескольких источниках, самым ранним (70–е годы XI века нашей эры) и наиболее надёжным из которых ввиду личного посещения огузов, является список М. Кашгари [6, с.93-94], несмотря на то, что он приводит лишь 22 наименования из 24-х. Ввиду указанных преимуществ мы в данной работе будем опираться на его список, который в латинской и кириллической транскрипциях приведён в таблице 1.

Таблица 1
Огузские этнонимы из списка М. Кашгари

№	Латинская транскрипция	Кириллическая транскрипция	№	Латинская транскрипция	Кириллическая транскрипция
1	Qīñiq	Кынык	12	Qarabölük	Карабулук
2	Qayığ	Кайыг	13	Alqabölük	Алкабулук
3	Bayundur	Байандур	14	İgdir	Игдир
4	Yıva	Йива	15	Yüregir	Юрегир
5	Salğur	Салгур	16	Tutırqa	Тутирка
6	Afşar	Афшар	17	Ulayundluğ	Улаюндлуг
7	Begtili	Бектили	18	Tüger	Тюгер
8	Bügdüz	Бюгдюз	19	Beçenek	Беченег
9	Bayat	Баят	20	Çuvaldar	Чувалдар
10	Yazğır	Йазгир	21	Cebni	Джебни
11	Eymür	Эймур	22	Çaruqluğ	Чаруклуг

Для изучения этимологий огузских племён, многими учеными, такими как Баскаков Н.А., Кононов А.Н., Толстов С.П., Кумекон Б.Е., Плетнёва С.А., Махпиров В.У., Зуев Ю.А., были применены узконаправленные методы одной определенной науки (истории, лингвистики, этнонимии и т.д.). Поэтому, по мнению большинства других экспертов, эти исследования страдают однобокостью, не имеют достаточно веских аргументов в пользу предлагаемых этимологий.

Таким образом, современное состояние наук по этой проблеме, именно некоторого множества наук, а не одной отдельной науки (не исключительно этнонимии) обуславливает актуальность и необходимость разработки интеллектуальной системы как комплекса или батареи экспертных систем, структурная схема, которой представлена на рисунке 1.

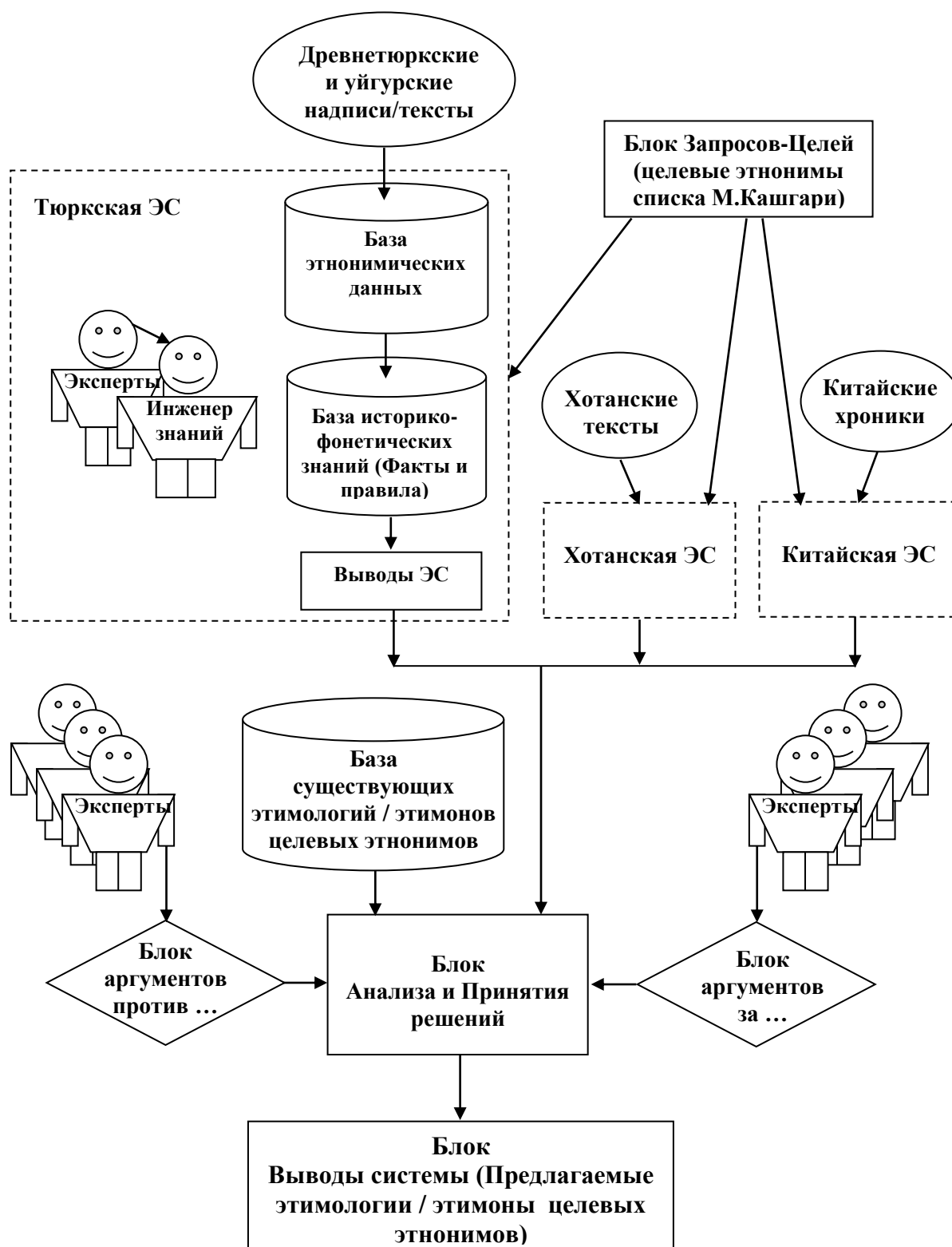


Рис. 1 Блок-схема системы

Опишем работу батареи в целом. В качестве исходных данных для батареи ЭС используются в основном древнетюркские и уйгурские тексты VIII-IX веков рунического и уйгурского письма, хотанские тексты, и китайские хроники. Извлечённые из этих текстов тюркские этнонимы являются предметными фактами для баз знаний батареи ЭС,

представленной в виде трёх непосредственно не связанных между собой локальных экспертных систем: Тюркской ЭС, Хотанской ЭС и Китайской ЭС.

Пользователи БЭС с помощью запросов-целей, в качестве которых используются 22 огузских этнонима списка М. Кашгари, на языке логического программирования Prolog, могут обращаться к локальным ЭС (тюркской, хотанской и китайской). Используя предметные факты – древнетюркские и уйгурские этнонимы и экспертные правила (некоторые из которых представлены далее), локальные ЭС делают логические заключения, которые поступают на вход Блока «Анализа и Принятия решений».

В блок «Анализа и Принятия решений» поступают также аргументы из блока «Аргументов» и анти-аргументы из блока «Анти-аргументов». Блок «Аргументов» представляет собой список исторических, лингвистических, логических, ассоциативных и иных аргументов за, а блок «Анти-аргументов» соответственно аргументы, вступающие против логических заключений – прототипов (этимонов) этнонимов предлагаемых локальными ЭС для каждого конкретного огузского этнонима из списка М. Кашгари.

Данные из базы существующих этимологий/этимонов огузских этнонимов списка М. Кашгари с целью сравнения с предлагаемыми локальными ЭС логическими заключениями также поступают на вход блока «Анализа и Принятия решений». Совпадение логического заключения локальной ЭС с существующей этимологией/этимонимом учитывается в системе как дополнительный аргумент за данное логическое заключение. Отсутствие совпадения - дополнительный аргумент против данного логического заключения.

После анализа заключений трёх ЭС, а также аргументов и анти-аргументов, в блоке «Анализа и Принятия решений» происходит принятие окончательного решения относительно этимологии или этимона каждого из 22 огузских этнонимов. Принятые окончательные решения отражаются в блоке «Выводы Системы».

Теперь рассмотрим, как функционирует одна из основных компонентов БЭС – «Тюркская ЭС» более подробно. Извлечённые из древнетюркских и древнеуйгурских текстов этнонимы аккумулируются в базе этнонимических данных, которые затем после обработки (разделения на собственно этнонимы и аффиксы или этнонимобразующие форманты, а также если этнонимы дву или более составные, то разделение их на составляющие) поступают в виде предметных фактов-этнонимов в базу знаний тюркской ЭС.

База знаний снабжена тюркскими историко-фонетическими правилами, которые были извлечены из экспертов, в роли которых выступают известные тюркологи, начиная от М. Кашгари (XI век) до современных тюркологов. После извлечения экспертные знания были формализованы инженерами знаний, в качестве которых выступили авторы данной статьи, в виде тюркских историко-фонетических правил, некоторые из которых представлены далее. На основе предметных фактов и экспертных правил, механизм логического вывода ЭС формирует логические заключения – предполагаемые древнетюркские/древнеуйгурские прототипы для 22 – х огузских этнонимов списка М. Кашгари.

Этноним – это слово, и как любое слово, подчиняется законам языка [7, с.25], поэтому разработка тюркской лингвистической (с разделами фонетика, грамматика, словообразование) базы знаний имеет для нашей системы громадное значение. Как известно из науки языкознания, язык никогда не стоит на месте, а изменяется согласно определённым языковым закономерностям. Особенно сильны и быстры эти изменения, когда этносы – носители языка вступают в длительное взаимодействие с другими этносами (носителями совершенно иных языков и культур), а именно в такой ситуации находились в частности огузские племена в IX – XI веках.

Как известно языковыми изменениями занимается специальная наука - историческая лингвистика или историческая фонетика. Согласно этой науке существуют следующие

видоизменения звуков в потоке речи:

I. Комбинаторные изменения. В зависимости от соседства других звуков;

II. Позиционные изменения. Связанные с положением в неударном слоге, в конце слова и т.д.

К комбинаторным изменениям относятся, например: приспособление артикуляции (произношения) согласных под влиянием гласных и гласных под влиянием согласных; ассимиляция с её видами – уподобление согласного согласному или гласного гласному; диссимиляция – обратное ассимиляции – расподобление артикуляции двух одинаковых или подобных звуков; метатеза (греч. перестановка) – взаимная перестановка звуков или слогов в пределах слова и т.д. К позиционным изменениям относятся, например: редукция – изменение (ослабление) звуков по качеству и количеству; отпадение звуков; оглушение – потеря звонкости звуков, паразитические звуки и т.д. [8, с.114-118].

Историко-фонетические явления с их последствиями в тюркских языках в общем и в огузском в частности с особым вниманием учтены при разработке тюркской фонетической базы знаний системы.

Как известно, знанием о некоторой предметной области является аналитическая обработка информации об этой области, которая отвечает критериям структурированности, лаконичности, непротиворечивости, полноты и наличия процедур обработки. Экспертные знания в базе знаний должны быть каким-либо образом представлены. В настоящее время применяются в основном следующие способы (модели) представления экспертных знаний: логические методы (язык предикатов); семантические сети; фреймы; продукции; нечёткая логика и другие [9, с.147-156]. В работе для представления знаний нами выбрана логическая модель. В ЭС, основанных на логических моделях представления знаний (также известных под названием модели на основе «правил»), базы знаний представляются набором экспертных правил, которые проверяются на группе фактов. Когда часть правила «ЕСЛИ» удовлетворяет фактам, то действие, указанное в части «ТО», выполняется. Процесс сопоставления с фактами частей «ЕСЛИ» правил может породить цепочку выводов, по которой можно проследить то, как система, используя правила, выводит заключение. Существует два способа использования правил:

- прямая цепочка рассуждений <ЕСЛИ «посылки» → ТО «следствие»>;
- обратная цепочка <ТО «следствие» → ЕСЛИ «посылки»> [3, с.71-77].

Прямой порядок означает, что рассуждения отталкиваются от данных (условий) к гипотезам. Обратный порядок означает, что рассуждения отталкиваются от цели (гипотезы) к условиям, при которых возможно достижение цели [10, с.118-120]. Обратная цепочка рассуждений (обратный логический вывод), как более рентабельная [3, с.75], используется нами при разработке базы знаний ЭС.

В результате процесса инженерии тюркских историко-фонетических знаний [11, с.128-191; 12, с.184-373] нами были сформированы фонетические экспертные правила для базы знаний, некоторые из которых представлены ниже:

Правило 1:

THEN $b \sim m$

IF (“ b ” OR “ m ”) AND “середина слова”.

где “ \sim ” соответствие или чередование звуков.

Правило 2:

THEN “ \check{g} исчезает” OR “ g исчезает”

IF (“ \check{g} ” OR “ g ”) AND “конец слова”.

Правило 3:

THEN $d > y$

IF “d” AND “середина слова” AND “d после гласного”.

где “>” переход звука.

Правило 4 (Характерно для туркменских диалектов):

THEN $l > y$

IF “l” AND “середина слова”.

Правило 5 (Во многих современных и древних тюркских языках):

THEN $\ddot{i} \sim a$

IF (“a” OR “ \ddot{i} ”) AND “первый слог”.

где “ \ddot{i} ” представляет кириллическую “ы”.

Правило 6:

THEN $u \sim \ddot{i}$

IF (“u” OR “ \ddot{i} ”) AND “первый слог”.

Правило 7:

THEN $\ddot{a} \sim i$

IF (“ \ddot{a} ” OR “i”).

Правило 8:

THEN $r > l$

IF “r” AND “конец слова”.

Правило 9 (У Махмуда Кашгари):

THEN $t > d$

IF “t”.

Правило 10:

THEN $\ddot{a} \sim a$

IF (“ \ddot{a} ” OR “a”).

Для моделирования явления метатезы в исторической фонетике база знаний снабжена специальной процедурой, которая получает в качестве аргумента этноним и возвращает в качестве результата массив всех возможных, с точки зрения исторической фонетики, предполагаемых слов-этнонимов с перестановками звуков и слогов, имитируя, как звуковую, так и слоговую метатезу.

В процессе этнонимообразования активно принимают участие и так называемые этнонимообразующие аффиксы или суффиксы (форманты) [7, с.26-28].

Системный анализ рассматриваемых этнонимов также обнаружил, что большая часть рассматриваемых этнонимов образована с помощью специальных формантов. Так 9 из 24-ёх этнонимов, т.е. 37.5% образованы с помощью формантов –ар/-ур/-ир/-ыр/-ер, а 8 этнонимов, т.е. 33.3% образованы с помощью формантов –ык/-ыг/-юк/-уг/-ег. Таким образом, 70.8% рассматриваемых огузских этнонимов образованы с помощью всего двух формантов с модификациями (гласный звук + согласный звук «р») и (гласный звук + согласные звуки «к/г»).

Этнониμοобразующие форманты огузских этнонимов учитываются в базе знаний с помощью специальных процедур обработки этнонимов.

В процессе тестовых испытаний прототипа предлагаемой БЭС были получены представленные в таблице 2, следующие предварительные результаты для трёх этнонимов из списка М. Кашгари.

Таблица 2
Предварительные результаты тестирования прототипа БЭС

Порядковый номер этнонима в списке М. Кашгари	Целевые этнонимы из списка М. Кашгари	Предлагаемый БЭС этимон (прототип - этноним из древнетюркских или уйгурских текстов)	Используемые БЭС для принятия решения правила Базы знаний	Участие этнониμοобразующего форманта
1	Qiniq / Кынык	Qunı / Куны	Правило 6	-q
5	Salğur / Салгур	Sir / Сир	Правило 7 Правило 10 Правило 8	-ğur
14	İgdir / Игдир	İgdär / Игдяр	Правило 7	-

Предлагаемая батарея экспертных систем может быть использована для установления правильных этимологий/этимонов этнонимов не только тюркских, но и любых других народов, племён или родов. Достаточно всего лишь заменить предметные факты и экспертные правила базы знаний. Более того, предлагаемая БЭС при соответствующей адаптации может быть применена для задачи установления правильных этимологий не только этнонимов, но и слов какого-либо языка вообще

4. Выводы.

1. Предлагается разработка интеллектуальной системы, которая конструируется как комплекс или батарея экспертных систем, состоящую из отдельных лингвистических ЭС: (тюркской, хотанской, китайской), для задачи установления правильных этимологий огузских этнонимов, которая не может быть решена без использования преимуществ технологии экспертных систем;
2. Сформированы экспертные правила и специальные процедуры базы знаний для предлагаемой системы, моделирующие фонетические явления изменения тюркских звуков в потоке речи, а также учитывающие наличие этнониμοобразующих формантов.
3. Представлены предварительные результаты тестирования прототипа БЭС для трёх целевых этнонимов из списка М. Кашгари.
4. Предлагаемая БЭС может быть использована для установления правильных этимологий этнонимов не только тюркских, но и любых других народов, племён или родов, а также слов языка вообще.

Литература

1. Гаврилова Т.А., Хорошевский В.Ф. Базы знаний интеллектуальных систем. СПб.: Питер, 2001, 384 с.
2. Исмаилов И.А. Интегрированная экспертная система оценки свойств пластов // Западный Университет. Научный вестник. Серия социальных и технических наук, 2014, № 1, с. 75 – 85.
3. Уотермен Д. Руководство по экспертным системам. М.: Мир, 1989, 388 с.
4. Введенская Л.А., Колесников Н.П. Этимология. - СПб.: Питер, 2004, 221 с.
5. А.И. Попов, Названия народов СССР, Ленинград, «Наука», 1973, 173 с.
6. Махмуд ал-Кашгари. Диван Лугат ат-Турк / Перевод, предисловие и комментарии З.-А. М. Ауэзовой. – Алматы: Дайк-Пресс, 2005. – 1288 с.
7. Никонов В.А. Этнонимия // Этнонимы. М., 1970. с. 5-32.
8. Бондаренко М.А. Курс лекций «Введение в языкознание», Тула, 2007, 391 с.
9. Башмаков А.И., Башмаков И.А. Интеллектуальные информационные технологии. М.: МГТУ им. Н.Э. Баумана, 2005, 304 с.
10. Джексон П. Введение в экспертные системы. М.: Вильямс, 2001, 623 с.
11. Ряснен Н. Материалы по исторической фонетике тюркских языков. Москва, из-во иностранной литературы, 1955, 222 с.
12. Теницев Э.Р. Сравнительно-историческая грамматика тюркских языков. Фонетика. Москва: «Наука», 1984, 485 с.

UOT 681.3:629.73

G.G. Abdullayeva, İ.A. İsmaylov

Etnonimlərin etimologiyasını təyin edən ekspert sistemlər batareyasının işlənməsi

(oğuz etnonimləri təmsalında)

Multidissiplinar məsələlərin həlli üçün ekspert sistemlər (ES) batareyasının yaradılması ideyasına həsr olunub. Məsəl üçün 24 oğuz etnonimləri araşdırılıb. Bu sinif məsələlərin dar çərçivədə və bir sahənin mütəxəssisi ilə həllinin çətinliyi göstərilib. Təklif edilən ES batareyası aşağıdakı elementlərdən layihələndirilir: türk ES, hotan ES, çin ES, arqumentlər və anti-arqumentlər blokları, analiz və qərar qəbuletmə blokları. Türk fonetika ES və onun biliklər bazası dolğun şəkildə təsvir olunub. ES batareyasının prototipi işlənilib və məqalədə testləşdirmənin bəzi nəticələri nümayiş etdirilir.

Açar sözlər: ekspert sistemi, biliklər bazası, qaydalar, ekspert sistemləri batareyası

G.G. Abdullayeva, I.A. Ismayilov

Construction of the battery of expert systems for determining the etymology of ethnonyms

(on the example of Oghuz ethnonyms)

In the paper, the idea of developing the battery of the expert systems (BES) for solving multidisciplinary problems of ethnonymics on the example of Oghuz ethnonyms is considered. The complexity of ethnonymics problems and the impossibility of solution by subject matter experts are shown. The deep survey of researches concerning ethnonyms of 24 Oghuz tribes is carried out. The proposed battery is designed from different elements. The principles of constructing Turkic phonetics ES and knowledge bases, with some heuristic rules are described. Some results of the testing of the prototype ES battery are presented.

Keywords: expert system, knowledge base, rules, battery of expert systems