

УДК 519.7



И.В. Каменева

ХНУРЭ, г. Харьков, Украина, irina.kamenieva@gmail.com

АНАЛИЗ ФЛЕКСИИ ПИСЬМЕННОЙ И УСТНОЙ МОДЕЛЕЙ ПОЛНЫХ НЕПРИТЯЖАТЕЛЬНЫХ ИМЕН ПРИЛАГАТЕЛЬНЫХ РУССКОГО ЯЗЫКА

В данной статье рассматриваются разработки морфологических структур русского языка предшественников, проводится сравнительный анализ эффективности реляционной сети склонения полных непряжательных имен прилагательных русского языка.

МОЗГОПОДОБНЫЕ СТРУКТУРЫ, РЕЛЛЯЦИОННАЯ СЕТЬ, МАТЕМАТИЧЕСКАЯ МОДЕЛЬ,
ПОЛНЫЕ НЕПРИТЯЖАТЕЛЬНЫЕ ИМЕНА ПРИЛАГАТЕЛЬНЫЕ

Введение

Использование механизмов естественного языка, с которым непосредственно связан человек, обеспечивает создание усовершенствованных новейших средств и технологий обработки информации для интеллектуализации работы с текстовой информацией. На сегодняшний день проблема формализации естественного языка является актуальной, и интеллектуальные системы автоматизированной обработки естественно-языковой информации активно развиваются.

Механизм естественного языка невозможно охватить одним конкретным исследованием. В нем для исследования приходится выделять некоторую область, с которой предстоит упорная и без сомнения трудная работа. Существует огромное множество естественных языков, но исследования проводятся в недрах русского языка, в частности, раздела морфологии.

Как известно, одна из задач морфологии – изучение правил образования форм слова. «Слово – это «атом» языка, нечто первичное и простейшее в языке. И, тем не менее, даже отдельно взятое слово представляет собой чрезвычайно сложный объект для математического описания» [1]. Большое значение при исследовании взаимодействия термина и текста имеет морфологическая структура слова.

Обработка естественных языковых конструкций позволяет разрабатывать модели естественного языка и свидетельствует о том, что классическая грамматика может с успехом использоваться как основа для математического и технического моделирования механизма естественного языка.

Попытки формализации естественного языка привели к формальному описанию отношений, без которых в принципе работать со структурой языка невозможно. Поскольку в проделанной работе представить отношения формулами не удалось, была реализована алгебра предикатов, которая выражает предикаты заданными отношениям и наоборот. Создание алгебры предикатов помогло формально описать и автоматизировать мыслительную деятельность человека, но в процессе работы возникла преграда в записи и обработке

отношений, которые выходят за рамки классического определения понятия алгебры. В процессе долгой и трудоемкой работы была создана алгебраическая система предикатов. Понятие «алгебраическая система» впервые было представлено академиком А.И. Мальцевым [2]. В зависимости от условий поставленной задачи используется также алгебра предикатных операций [3].

Были и другие попытки формализации естественного языка. В частности Давид Гильберт, в своем курсе лекций по основам теоретической логики описал идею косвенного представления отношений, но в дальнейшем ни он сам, ни другие авторы к идее формализации естественного языка больше не обращались [4].

Цель данного исследования – приблизиться к усовершенствованию существующих письменной и звуковой математических моделей, разработанных для склонений полных непряжательных имен прилагательных русского языка.

1. Мозгоподобные структуры

Основным понятием для мозгоподобных структур является математическая структура. Математическая структура – родовое название, объединяющее понятия, общей чертой которых является то, что они применимы к множествам, природа элементов которых не определена. Чтобы определить структуру, задают отношения, в которых находятся элементы множеств (типовая характеристика структуры), а затем постулируют, что данные отношения удовлетворяют условиям – аксиомам структуры».

Характерной особенностью мозгоподобных структур является слияние памяти с обработкой данных: данные обрабатываются одновременно по всей памяти с максимально возможной степенью распараллеливания всех операций. Мозгоподобные структуры с параллельными процессами, управляемыми многими потоками данных и команд, несомненно, представляют собой высший уровень развития архитектур ЭВМ.

Сегодня уже появилась возможность создания вычислительных структур с производитель-

ностью, сравнимой с производительностью мозга человека [5].

Для того чтобы не возникало сложностей с некоторыми выражениями в следующем разделе, определим, что собой представляет математическая модель. Математическая модель – это формализованное описание каких-либо отношений во внешнем мире, выраженных математическими символами.

На сегодняшний день научной школой Юрия Петровича Шабанова-Кушнаренко построены модели спряжения глаголов [6], склонения имен существительных [7], склонения полных непряжательных имен прилагательных [8], которые реализованы в виде реляционных сетей для создания мозгоподобных ЭВМ.

Благодаря работе с непряжательными именами прилагательными, появилась идея сравнить письменную модель полных непряжательных имен прилагательных и звуковую модель склонения потенциальных слов и их потенциальных форм полных непряжательных имен прилагательных для морфологических структур.

2. Анализ процесса склонения полных непряжательных имен прилагательных

Остановимся более подробно на сравнении двух математических моделей (письменной и звуковой), разработанных для полных непряжательных имен прилагательных [9, 10]. Для этих моделей ниже формально представлены предметные переменные: x_1 – род формы слова со значениями М – мужской, Ж – женский, С – средний; x_2 – число формы слова со значениями Е – единственное, М – множественное; x_3 – падеж формы слова со значениями И – именительный, Р – родительный, Д – дательный, В – винительный, Т – творительный, П – предложный; x_4 – признак одушевленности формы слова со значениями О – одушевленный, Н – неодушевленный; x_5 – признак современности формы слова со значениями С – современный, А – архаичный; r – номер влияния контекста, s – тип склонения слова; y_1 – последняя буква основы слова, y_2 – признак ударности слова, y_3 – признак смягчения слова, z – окончание формы слова, где z_l – левая часть окончания, $z_{п}$ – правая часть окончания. Для письменной математической модели полных непряжательных имен прилагательных при развертке окончания z по буквам используется знак * при переменной z_0 , что означает отсутствие знака ударения над первой буквой окончания, а при развертке окончания z по буквам знак * при переменной z_3 означает отсутствие третьей буквы в окончании. Итак, математическую модель склонения полных письменных непряжательных имен прилагательных представляет собой предикат модели, который состоит из 17 конечных бинарных отношений:

$$P(x_1, x_2, x_3, x_4, x_5, r, s, y_1, y_2, y_3, z_l, z_{п}, z, z_0, z_1, z_2, z_3) = \\ = P_1(x_1, r) \wedge P_2(x_2, r) \wedge P_3(x_3, r) \wedge P_4(x_4, r) \wedge P_5(x_5, r) \wedge$$

$$\wedge P_6(z_l, r) \wedge P_7(z_{п}, r) \wedge P_8(z_l, z) \wedge P_9(z_{п}, z) \wedge P_{10}(z_l, z_0) \wedge \\ \wedge P_{11}(z_l, z_1) \wedge P_{12}(z_{п}, z_2) \wedge P_{13}(z_{п}, z_3) \wedge P_{14}(s, z_l) \wedge \\ \wedge P_{15}(y_1, s) \wedge P_{16}(y_2, s) \wedge P_{17}(y_3, s).$$

Предикату данной модели соответствует отношение модели P , связывающее между собой предметные переменные

$$x_1, x_2, x_3, x_4, x_5, r, s, y_1, y_2, y_3, z_l, z_{п}, z, z_0, z_1, z_2, z_3.$$

Отношение модели P можно наглядно изобразить в виде сети, называемой логической. Логическую сеть можно построить для любой модели. В конкретном случае логическая сеть склонения полных непряжательных имен прилагательных имеет вид, представленный на рис. 1.

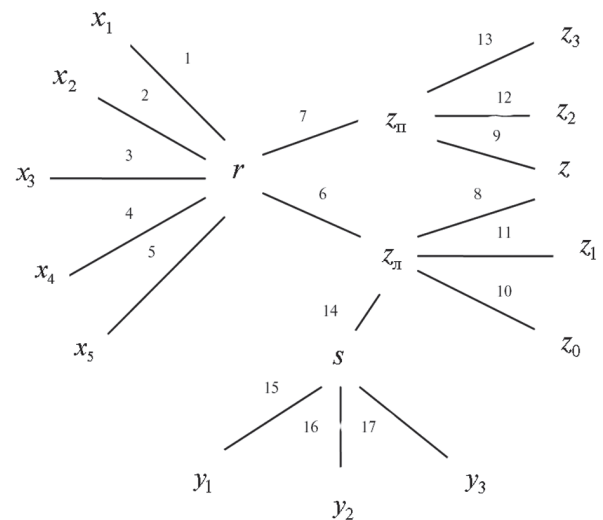


Рис. 1

Математическую модель склонения полных звуковых (устных) непряжательных имен прилагательных представляет собой предикат модели, который состоит из 11 конечных бинарных отношений:

$$P(x_1, x_2, x_3, x_4, x_5, r, s, y_1, y_2, y_3, z_l, z_{п}) = \\ = P_1(x_1, r) \wedge P_2(x_2, r) \wedge P_3(x_3, r) \wedge P_4(x_4, r) \wedge P_5(x_5, r) \wedge \\ \wedge P_6(y_1, s) \wedge P_7(y_2, s) \wedge P_8(y_3, s) \wedge P_9(s, z_{п}) \wedge \\ \wedge P_{10}(r, z_l) \wedge P_{11}(s, z_l),$$

где аргументы типов склонения: y_1 – последний звук основы слова, y_2 – признак ударности со значениями: У – ударный, Б – безударный, y_3 – признак смягчения со значениями: Т – твердый, М – мягкий. Модель характеризуется системой бинарных отношений $P_1 - P_{11}$, задаваемых формулами соответствующих предикатов. Образую конъюнкцию всех этих предикатов, получаем предикат модели. А также данная модель представляется в виде логической сети (рис. 2).

Каждая из представленных моделей имеет свои достоинства и недостатки. Речь идет о том, что в письменной модели не была рассмотрена потенциальная форма слова, а именно – способность человека склонять не только употребляемые в языке слова, но также и многие псевдослова т.е. словообразные

буквосочетания. Соответственно возникла сложность при склонении несуществующих в грамматике слов. А в звуковой модели такая проблема была решена, и возможность склонения потенциальных слов осуществлена в полном объеме. Недостаток звуковой модели в том, что в ней не рассмотрено разбиение правой $z_{\text{п}}$ и левой $z_{\text{л}}$ частей окончания на буквы окончания и признак ударности, тогда как в письменной модели окончания были успешно рассмотрены и внедрены в модель.

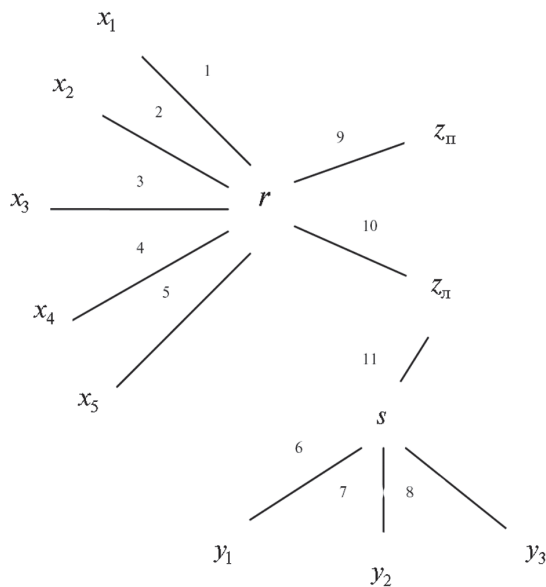


Рис. 2

Результаты моделирования устного словоизменения заставляют внести коррективы в известную модель письменного словоизменения. Например, в словах «длинношей», «-шуй», в качестве последней буквы основы при письменном представлении слов и их форм в этой модели был принят знак пробела, и это не приводило к появлению беглой буквы в основе слова. В устном же варианте при склонении этих слов неизбежна беглость звука основы («-шуй», «-шуйго» и т.д.), и поэтому приходится в роли последнего звука основы принять беглый согласный «й». Если письменную модель скорректировать по устной, то первая примет более логичный вид, поскольку теперь все без исключения последние буквы основы окажутся согласными.

Выводы

Анализируя полученные результаты сравнения, где рассматриваются две конкретные мозгоподобные структуры, а именно устная и письменная модели склонения полных непряжательных имен прилагательных, заключаем, что в дальнейшем планируется устранение представленных недостатков, что позволит решить стоящую перед нами задачу — объединение моделей и проверка их универсальности. Последующие задачи, которые будут исследоваться и разрабатываться, также охватят морфологическую структуру естественного языка,

которая позволит рассмотреть реляционную сеть потенциальных полных непряжательных имен прилагательных, представленных в устной и письменной формах как отдельные части, близкие к совершенной сети.

Список литературы: 1. Бондаренко, М.Ф. Теория интеллекта [Текст]/ М.Ф. Бондаренко, Ю.П. Шабанов-Кушнарченко // Теория интеллекта: Учебник. — Х.: ООО «Компания СМИТ», 2006. — 576 с. 2. Мальцев, А.И. Алгебраические системы [Текст]/ А.И. Мальцев. — М.: Наука, 1970. — 392 с. 3. Бондаренко, М.Ф. Алгебра предикатов и предикатных операций [Текст] / М.Ф. Бондаренко, З.В. Дударь, Н.Т. Процай, В.В. Черкашин, В.А. Чикина, Ю.П. Шабанов-Кушнарченко // Радиоэлектроника и информатика. — 2000. — № 4. — С. 15–23. 4. Гильберт, Д. Основы теоретической логики [Текст] / Д. Гильберт, В. Аккерман. — М.: ИЛ, 1947. — 302 с. 5. Бондаренко, М.Ф. О мозгоподобных структурах академика Виктора Михайловича Глушкова [Текст] / М.Ф. Бондаренко, Н.Е. Русакова, Ю.П. Шабанов-Кушнарченко // Бионика интеллекта науч.-техн. журнал. — 2011. — № 2 (76). — С. 3–9. 6. Дударь, З.В. Математические модели флективной обработки словоформ и их использование в системах автоматической обработки текста русского языка [Текст]: Дис. ... канд. техн. наук. — Х.: ХИРЭ, 1984. — 215 с. 7. Лецинский, В.А. Модели бинарных логических сетей и их применение в искусственном интеллекте [Текст]: Дис. ... канд. техн. наук. — Х.: ХНУРЭ, 2006. — 157 с. 8. Мельникова, Р.В. Алгебраические модели морфологии и их применение в логических сетях. Дис. ... канд. техн. наук. — Х.: ХНУРЭ, 2005 — 152 с. 9. Бондаренко, М.Ф. Модели языка [Текст] / М.Ф. Бондаренко, В.А. Чикина, Ю.П. Шабанов-Кушнарченко // Бионика интеллекта: науч.-техн. журнал. — 2004. — № 1(61). — С. 27–36. 10. Русакова, Н.Е. Модели устной речи [Текст] / Н.Е. Русакова // Бионика интеллекта: науч.-техн. журнал. — 2010. — №1(72). — С. 94–97.

Поступила в редколлегию 22.12.2011

УДК 519.7

Аналіз флексії письмової та усної моделей повних непряжательних прикметників російської мови / І.В. Каменева // Біоніка інтелекту : наук.-техн. журнал. — 2012. — № 1 (78). — С. 116–118.

У статті розглянуто розробки попередників, які займалися морфологічними структурами російської мови, проведено порівняльний аналіз ефективності реляційної мережі відмінювання повних непряжательних прикметників письмової та звукової моделей російської мови, що дало змогу виявити деякі недоліки, щоб врахувати їх для удосконалення майбутніх розробок.

Л. 2. Бібліогр.: 10 найм.

UDC 519.7

The adjectives inflection of Russian language/I. Kamenieva // Bionics of Intelligense: Sci. Mag. — 2012. — № 1 (78). — P. 116–118.

The article discussed the precursors's development who were involved in morphological structures of the Russian language. The comparative analysis of the effectiveness of the relational network adjectives of the Russian language allowed to identify some shortcomings for improve future developments.

Fig. 2. Ref.: 10 items.