
ЛОГИКО-ЛИНГВИСТИЧЕСКАЯ МОДЕЛЬ ИЗВЛЕЧЕНИЯ ФАКТОВ ИЗ СЛАБОСТРУКТУРИРОВАННОЙ ТЕКСТОВОЙ ИНФОРМАЦИИ

Нина Хайрова, Наталья Шаронова

Аннотация: Одним из перспективных направлений информационного поиска является фактографический поиск и разработка фактографических баз данных. Существующие сегодня модели и алгоритмы фактографического поиска в своем большинстве направлены на извлечение фактов из хорошо формализованной информации, в том числе из хорошо формализованной текстовой информации. В работе предлагается модель извлечения фактографической информации из динамически меняющихся слабоформализованных текстовых потоков, не ограниченных определенными предметными областями. Для извлечения некоторого факта используется шаблон «агент-предикат-значение», отображающий отношения, формально выражаемые семантическими падежами участников предложения. В предлагаемой логико-лингвистической модели семантические роли именных групп определяются отношением четко выделенных множеств морфологических, синтаксических и семантически категорий, описываемым с помощью базового аппарата алгебры конечных предикатов. В работе рассмотрена реализация данной модели для извлечения фактографической информации о дате, месте рождения и роде деятельности персоналии из русскоязычных слабоформализованных текстов. Экспериментальная проверка программной имплементации модели показала правильность выделения факта примерно в 94,3% случаев.

Ключевые слова: фактографический поиск, слабоформализованная текстовая информация, лингвистический процессор, алгебра конечных предикатов.

ACM Classification Keywords: H.3.3 .Information Search and Retrieval

Введение

Одним из главных приложений Natural Language Processing (NLP) является информационный поиск, классифицируемый по результатам выдачи на документальный поиск – процесс поиска документа в массивах первичных или вторичных документов, и фактографический поиск – процесс поиска фактов, отвечающих информационному запросу. Факт представляет собой знание в форме утверждения, достоверность которого строго установлена [Барахнин, 1980]. На практике, в сфере информационных технологий, фактографическую информацию обычно трактуют несколько иначе — как конкретные сведения или данные независимо от того, являются ли они фактическими или прогнозируемыми. Главное, что эти сведения сообщают о какой-то предметной области, а не о документах, посвященные этой области.

В свою очередь фактографическую информацию можно разделить на хорошо структурированную и плохо структурированную. К хорошо структурированным сведениям (так называемая параметрическая информация) относятся, прежде всего, сведения количественного характера, а так же качественные сведения, имеющие хорошо регламентированную форму. К плохо структурированной фактографической

информации относятся сведения, представленные различными нерегламентированными словесными конструкциями, представленными на естественном языке [Baeza-Yates, 1999].

Алгоритмы фактографического анализа зависят, в свою очередь, от степени структурированности конкретного документа [Ландэ, 2009]. По степени структурированности данные документа можно разделить, подобно общей классификации степени формализации информации, на табличные данные, отображенные в виде фактов; массивы однородных слабоструктурированных текстовых документов, обычно описывающие конкретную предметную область, и документы произвольного слабоструктурированного типа.

Актуальность исследования

Большое влияние на разработку методов и моделей извлечения фактографической информации оказала серия конференций Message Understanding Conferences, MUC, проходившая с 1987 г. по 1997 г. при поддержке Американского агентства DARPA (Defense Advanced Research Projects Agency), и способствующая упорядочиванию информации по системам фактографического поиска. Но только в последние несколько лет стали появляться реальные системы, включающие элементы такого рода поиска. Например, практически единственная на территории СНГ, находящаяся в тестовом режиме, российская поисковая система pigma.ru. Главным ресурсом извлечения которой, являются слабоструктурированная wikipedia.org, как самая большая база текстов сходной структуры. Системы ask.com и answers.com так же являются системами фактографического поиска, выполняющие поиск в документах, а в качестве ответа на более общие вопросы, использующие ссылки на ресурс wikipedia.org.

Анализ показывает, что основные усилия разработчиков фактографических информационно-поисковых систем (ФИПС) направлены на хорошо структурированные факты, извлечение которых, с одной стороны, легче автоматизировать, а с другой, к этому типу относится почти вся производственно-экономическая информация, циркулирующая в сфере материального производства и управления. Таким образом, для извлечения фактов, представленных в табличных данных и в массивах однородной структурированной текстовой информации существуют достаточно надежные алгоритмы. Задача же извлечения фактов из произвольных текстов естественного языка до сих пор не имеет сколько-нибудь общего решения [Baeza-Yates, 1999], [Ландэ, 2009].

Проблема создания ФИПС, работающих со слабоформализованной текстовой информацией, еще далека от полного решения. Один из существующих подходов извлечения фактов из подобных текстов заключается в использовании онтологий или тезаурусов предметной области, которые и определяют, что является фактом, в рамках данной онтологии. Но такого рода подход, опять же, ограничивает анализируемые полнотекстовые документы узкой предметной областью онтологии.

Разработка методов и алгоритмов извлечения фактов из динамически меняющихся слабоструктурированных текстовых потоков, не ограниченных определенными предметными областями, требует точного моделирования когнитивной деятельности человека, по пониманию и идентификации фактов, а также наличия мощных средств как синтаксического, так и семантического анализа текстов, учитывающих семантическую эквивалентность и мультязычность.

Общая постановка задачи

Опираясь на определение факта, можно дать определение минимальной смысловой единицы фактографического поиска, представляющую собой триаду: агент-предикат-значение (табл.1). То есть запись фактографической информации должна включать указатель на агента фактографического поиска,

на атрибут или предикат этого объекта и давать конкретное значение этого атрибута. Такое определение позволяет извлекать понятия из слабоструктурированных текстовых источников информации и представлять отношения между ними в структурированном виде. Получаемая структура представляет собой факты, как в виде достаточно простых понятий: ключевых слов, персоналий, организаций, географических названий, так и в более сложном виде, например, имя персоналии с ее должностью и родом деятельности.

Таблица 1. Примеры формального представления фактической информации

<i>Агент</i>	<i>Атрибут</i>	<i>Значение</i>
пациент Иванов	резус-фактор	положительный
системный блок	вес	5 кг
Афродита	место рождения	Кипр

В общем случае выделение фактов из слабоструктурированной текстовой информации включает следующие этапы:

- 1) Entity Extraction – извлечение слов или словосочетаний, важных для описания смысла текста (списки терминов предметной области, персоналий, организаций, географических названий и т.д);
- 2) Feature Association Extraction – исследование связей между извлеченными понятиями;
- 3) Event and Fact Extraction – извлечение сущностей, распознавание фактов и действий.

Для реализации первого этапа выделения понятий используется стандартный лингвистический процессор [Хайрова, 2010], включающий графемную, морфологическую, синтаксическую и контекстную этапы обработки, с добавлением специализированных методов и алгоритмов обработки документов. Так как очень часто в задачах по извлечению фактографической информации нужно найти в тексте упоминания лиц, компаний, правительственных организаций и местоположений, и другие подобные типы сущностей, то для их выделения используются специальные формализмы графемного анализа. На этапе морфологического анализа используется декларативный и алгоритмический методы. Каждый неправильный глагол английского языка представлен в базе данных во всех его формах, то есть глагол *write* иметь формы *write-writes-wrote-written-writing*, формы правильных глаголов определяются алгоритмически. Русское словоизменение определяется алгоритмически, с использованием словарей окончаний.

А так как для построения триады фактографической информации необходимо выделить сущности, представленные в текстах под разными именами, то особое значение приобретает этап разрешения кореферентности (*coreference resolution*) синтаксического анализа, для определения синонимов, интересующих сущностей. На этом этапе такие местоимения как «он», «она», «они», «ним» и т.д. должны быть ассоциированы со своими антецедентами, соотнесены их с именуемой сущностью данной предметной области.

Центральной задачей получения фактографической информации является второй этап обработки, представляющий извлечение информации об отношениях между сущностями. При этом, для извлечения некоторого факта необходимо определить некий шаблон, отображающий семантические (или понятийные) связи в предложении. Для задания таких смысловых отношений предлагается использовать грамматику семантических падежей. Для чего необходимо разработать строгую модель, связывающую информацию, содержащуюся в определении семантических ролей с элементами поверхностной структуры предложений естественного языка. Такой подход рассматривается в рамках падежной грамматики и основывается на понятии глубинных падежей, введенных Ч. Филлмором, выделившим

пропозицию, или основной смысл предложения, как предикат, выражаемый в поверхностной структуре чаще глаголом, связанным с помощью определенных глубинных падежей с участниками данной ситуации, или партиципантами [Филлмор, 1981].

Для того, чтобы получить возможность использования глубинных падежей в задачах автоматической обработки смысла, необходимо формально определить глубинный семантический падеж через поверхностную структуру предложения данного языка. Семантические падежи в различных естественных языках имеют разные формы формального выражения. Например, в русском и украинском языках, семантическая информация партиципантов кодируется, в основном, грамматическими поверхностными падежами, тогда как в английском — она передается либо сочетанием с предлогом, либо, порядком слов в предложении. Но сегодня, в связи с отсутствием четко сформулированных критериев выделения семантических ролей, отсутствуют формальные модели, содержащие полные наборы семантических ролей, с достаточной степенью точности выражаемых элементами поверхностной структуры предложения, что приводит к низкому уровню использования данного подхода при автоматической обработке текстов.

Описание математической модели

Введем на универсуме U , включающем все возможные понятия и объекты анализа сложной языковой системы [Хайрова, 2012], множество грамматико-семантических характеристик лексем предложения $M = \{m_1, \dots, m_n\}$, где n — количество характеристик системы. Используя формальный аппарат алгебры конечных предикатов [Бондаренко, 2007], формально определим отношения между морфологическими и семантическими характеристиками существительного, формально выражающими семантические падежи партиципантов предложения русского языка. Для этого представим отношения между характеристиками в виде $m_i * m_j$, где $m_i, m_j \in M$, а знак $*$ — обозначает, что данные характеристики соответствуют существительному, относящемуся к определенному семантическому падежу. На множестве M введем систему предикатов S так, чтобы любой предикат $P(q_m) \in S$, обращался в 1 на множестве лексем с грамматико-семантической информацией, соответствующей определяемому семантическому падежу и был равен 0 в противном случае, сопоставляя множество предикатов S множеству семантико-грамматических характеристик приписанных лексеме.

Количество и состав семантических ролей, и предметных переменных, выделяемых при описании языка, могут существенно различаться в зависимости от задач описания, языка и его степени детализации [Филлмор, 1981]. Для формального определения семантических падежей русского языка выделим достаточно четко сформированное множество семантико-грамматических признаков, с помощью несократимого набора трех переменных: X -признак одушевленности (со значениями x^o — предметная переменная, характеризующая семантический признак живого, x^h — предметная переменная характеризующая семантический признак неживого); Y — элемент семантического значения существительного (y^m — механизм, y^c — имя собственное, y^i — инструмент, y^t — часть тела, y^r — плоскость/точка, y^o — объемное пространство, y^s — определенное время, y^n — период, y^u — пункт назначения); Z — грамматический падеж существительного ($z^a, z^p, z^d, z^s, z^t, z^n$ — предметные переменные, описывающие свойства существительных обладать тем или иным грамматическим падежом). Область изменения введенных переменных формально задается следующим образом:

$$\begin{aligned} x^o \vee x^h &= 1, \\ z^a \vee z^p \vee z^d \vee z^s \vee z^t \vee z^n &= 1, \\ y^m \vee y^c \vee y^i \vee y^t \vee y^o \vee y^s \vee y^n \vee y^u &= 1, \end{aligned} \quad (1)$$

Семантический падеж существительного предложения определяются через предикат P , связывающий элементы семантического значения существительного x и y с его грамматическими значениями z :

$$P(x, y, z) \rightarrow P(x) \bullet P(y) \bullet P(z), \quad (2)$$

где \bullet — операция конъюнкции. Так как возможность согласования морфосемантической информации не зависит от того, к какой конкретно словоформе она относится, на декартовом квадрате множества $S * S$ можно задать предикат $\gamma(x_n, y_n, z_n)$, принимающий значение 1, если морфосемантическая информация словоформы n формирует некоторый семантический падеж лексемы, и значение 0 в противном случае. Таким образом, отношения морфосемантических признаков существительных предложения, выражающих семантические падежи, требуемые валентностью глагола, можно задать формулой:

$$P(x_n) * P(y_n) * P(z_n) = \gamma_k(x_n, y_n, z_n) \bullet P(x_n) \bullet P(y_n) \bullet P(z_n), \quad (3)$$

Практически никогда подмножество согласующейся морфосемантической информации, выражающей семантические падежи, не совпадает с декартовым произведением на множестве морфологических и семантических признаков. Те морфосемантические признаки, которые в своем согласовании не формируют семантический падеж существительного должны исключаться из формулы (2) множителями $\gamma_k(x_n, y_n, z_n)$, $k \in [1; m]$, где m — количество, принятых к рассмотрению в системе семантических падежей. В соответствии с формулой (3) семантический падеж *агенса* определяется множеством возможных связей морфосемантической информации существительного:

$$P_1(x_n, y_n, z_n) = (x_n^o z_n^m \vee z_n^m x_n^h y_n^m \vee z_n^m x_n^o y_n^c) (y_n^m \vee y_n^c \vee y_n^h \vee y_n^u \vee y_n^t \vee y_n^o \vee y_n^b \vee y_n^u) (x_n^o \vee x_n^h) (z_n^m \vee z_n^p \vee z_n^d \vee z_n^b \vee z_n^t \vee z_n^n) \quad (4)$$

Множество возможных связей грамматической и семантической информации существительного семантического падежа *со-агенса* задается предикатом $P_2(x_n, y_n, z_n)$:

$$P_2(x_n, y_n, z_n) = (z_n^t x_n^o y_n^c) (y_n^m \vee y_n^c \vee y_n^h \vee y_n^u \vee y_n^t \vee y_n^o \vee y_n^b \vee y_n^u \vee y_n^u) (x_n^o \vee x_n^h) (z_n^m \vee z_n^p \vee z_n^d \vee z_n^b \vee z_n^t \vee z_n^n) \quad (5)$$

Множество возможных связей грамматической и семантической информации существительного семантического падежа *темпоралис* задается предикатом $P_3(x_n, y_n, z_n)$:

$$P_3(x_n, y_n, z_n) = (z_n^b x_n^h y_n^b \vee z_n^p x_n^h y_n^n) (y_n^m \vee y_n^c \vee y_n^h \vee y_n^u \vee y_n^t \vee y_n^o \vee y_n^b \vee y_n^u \vee y_n^u) (x_n^o \vee x_n^h) (z_n^m \vee z_n^p \vee z_n^d \vee z_n^b \vee z_n^t \vee z_n^n) \quad (6)$$

Множество возможных связей грамматической и семантической информации существительного семантического падежа *локатив* задается предикатом $P_4(x_n, y_n, z_n)$:

$$P_4(x_n, y_n, z_n) = (z_n^p x_n^h y_n^t \vee z_n^p x_n^h y_n^m \vee z_n^p x_n^h y_n^u \vee z_n^p x_n^h y_n^o) (y_n^m \vee y_n^c \vee y_n^h \vee y_n^u \vee y_n^t \vee y_n^o \vee y_n^b \vee y_n^u \vee y_n^u) (x_n^o \vee x_n^h) (z_n^m \vee z_n^p \vee z_n^d \vee z_n^b \vee z_n^t \vee z_n^n) \quad (7)$$

Для английского языка в дополнение к семантическим признакам вместо морфологических категорий выбираются синтаксические признаки употребления предлогов, т.к. именно предлоги после глаголов выражают валентность глагола английского языка. Признаки могут быть как уникальными для определенных глаголов, так и общими, как, например, признак направления движения может быть выражен глаголами *go, run, drive, ride, transport, ship* и т.д. Признаки могут быть описаны, так же как и для объектов с помощью семантических и грамматических характеристик. Например, для семантического падежа *Локатива*, если именная форма имеет признак плоскости / точки, употребляется английский

предлог *on*, а если объемного пространства, то предлог *in*. Падеж *темпоралис* определяется в английском языке предлогами *in* и *on*. Предлог *in* выражает год, в котором произошло событие, предлог *on* выражает точную дату, например, *He was born on February, 7 1902*. Падеж *локатива* определяется предлогом *in*, при выражении места, где произошло событие. Падеж *фактитив*, не требует предлога после глагола, что и является главным требованием (наличие после определенного слова предлога показывает, что глагол не требует падежа *фактитив*).

Структурное описание модели

Для извлечения из неструктурированной текстовой знаний или фактов необходимо определить семантические падежи, определяющие данные факты, глаголы языка, выражающие предикат в триаде факта «агент – предикат – значение» и предикаты морфосемантических (или семантико-синтаксических) признаков существительных предложения разработанной модели (3), выражающие семантические падежи объекта триады факта.

Для определения фактов о дате, месте рождения и виде деятельности персоналии в виде: «агент-предикат-значение» используем, кроме семантического падежа *агент*, три семантических падежа, выражающие информацию, соответствующую требованиям: *темпоралис* - временная характеристика события, позволяющая определить дату рождения человека; *локатив* - падеж, характеризующий местонахождение, положение или состояние объекта, позволяя определить место рождения человека; и *фактитив* - падеж, определяющий сферы и продукты деятельности человека. Нами был выбран именно *фактитив*, а не *объектив*, так как именно *фактитив* дает значение результата действия.

В процессе реализации модели был определен набор глаголов, требующих определенного семантического падежа участников предложения. Для этого с использованием толковых и переводных словарей были проанализированы около 130 биографий ученых и деятелей и выделены наиболее распространенные в данном типе слабоструктурированной текстовой информации глаголы английского и русского языка, выражающие предикат требуемого действия (табл. 2). Количество и список глаголов может меняться в зависимости от цели и объектов (фактов) поиска.

Таблица 2. Семантические падежи и наиболее распространенные предикаты, соответствующие фактам даты, места рождения и деятельности некоторой персоналии.

Падеж	Английский предикат	Русский предикат	Пример
Темпоралис	born, came up, saw the light, came into the world	родиться	He was born on February, 7 1902.
Локатив	born, came up, saw the light, came into the world	родиться	She came into the world in Frankfurt.
Фактитив	write, produce, paint, draw, pencil, design, project, invent, discover, engineer, compose, create, construct, publish, establish, investigate, research, explore, contribute, make, describe, work и др.	писать, изготавливать, рисовать, проектировать, открывать, изобретать, разрабатывать, создавать, строить, публиковать, учреждать, исследовать, вносить вклад, сделать, описывать, работать	There he wrote his first poem. In 1843 he completed the drama NazarStodolya.

Для определения соответствия информации определенному факту были заданы требуемые семантические падежи для каждого глагола. Например, русский глагол «*написать*», аналогично английскому глаголу «*write*» формирует семантические падежи участников представленные в табл. 3.

Таблица 3. Формальное определение требуемых семантических падежей глагола «написать» («write»)

towritetosmb	адресат
towritesmth	объектив (работа, рукопись, издание)
towritewithsmth	инструменталис (инструмент письма)
tobewrittenbysmb	агент (автор)
whenwritein	темпоралис (время написания)
ftowritein	локатив (место написания)

Осуществление детального анализа текстов биографий позволило определить дополнительные условия выражения семантических падежей, определяющий дату, место и род занятий персоналии:

- семантический падеж *локатив* выражается именем собственным (обычно графически выражаемым большой буквой), так как нас интересует населенный пункт, а не местоположение, как например *in mansion*;
- семантический падеж *фактитив* допускает несколько вариантов выражения: значащее слово написано с большой буквы, слово или словосочетание взято в скобки, или оно является существительным.

Программная имплементация модели

Программная имплементация модели представляет собой веб-приложение, анализирующее текст или список анализируемых текстовых файлов. При нажатии кнопки *checkname* (рис. 1) программа выделяет первые 2 слова с большой буквы в первом предложении и выводит их на экран, как возможные, имя и фамилия лица, о котором идет речь. Так как правила написания биографий, представляющих слабоформализованные тексты на естественном языке, практически, одинаковы, было определено, что первые 2 слова с большой буквы, с вероятностью 99,5, определяют имя и фамилию персоналии, о которой идет речь в данной библиографии. В случае, если выделенные данные не верны, пользователь может сам вписать имя, это понадобится в разборе биографии *ПаблоПикасо*, имеющего полное имя *Pablo Diego José Francis code Paula Juan Nepomuceno María de los Remedios Cipriano de la Santísima Trinidad Ruiz y Picasso*. Результат работы программы представлены в виде диалогового окна,

Извлеченная системой фактографическая информация представляется пользователю форме диалогового окна (см. рис.1). Программа отображает извлеченную фактографическую информацию в виде факта и первичных предложений, из которых данный факт был извлечен. В поле *birthday* находится информация даты рождения - *7 February 1812* и предложения, из которых была извлечена данная информация. Аналогичным образом в полях *origin* и *activity* отображается фактографическая информация о месте рождения и роде деятельности. Если в полях *birthday* и *origin* существует лишь один верный ответ, то поле *activity* может содержать несколько истинных ответов. Информация о деятельности записывается последовательно, каждый факт с нового абзаца. Факты деятельности располагаются в порядке значимости, определенной системой.

Экспериментальная проверка, проведенная на 47 полнотекстовых библиографических текстах электронного фонда ХГНБ, показали правильность определения даты рождения – 97,9% случаев, правильность определения места рождения – 95,7% , деятельность персоналии – 89,4%.

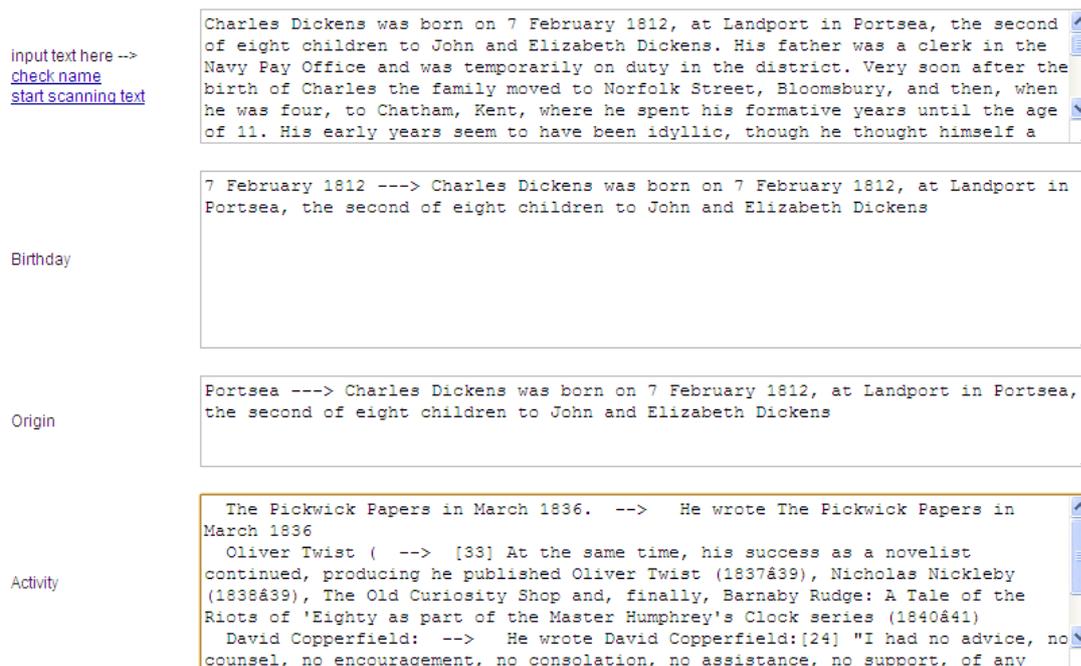


Рис. 1. Окно результата извлеченной фактографической информации рассматриваемой модели.

Выводы

Результатом данного исследования является разработка логико-лингвистической модели извлечения фактов из слабоструктурированных текстов, не ограниченных определенными предметными областями. Для извлечения факта использован шаблон «агент-предикат-значение», отображающий семантические (смысловые) отношения в предложении. Для задания таких смысловых отношений используется грамматика семантических падежей. В предлагаемой логико-лингвистической модели семантические падежи именных групп определяются предикатами, выражающими отношения четко выделенных множеств морфологических, синтаксических и семантически категорий. Множество возможных связей грамматической и семантической информации участников предложения описано средствами аппарата алгебры конечных предикатов. В работе рассмотрена реализация модели для извлечения фактографической информации о дате, месте рождения и роде деятельности персоналии из русскоязычных слабоформализованных текстов. Экспериментальная проверка программной имплементации модели показала правильность выделения факта примерно в 94,3% случаев.

Литература:

- [Baeza-Yates, 1999] Baeza-Yates R., Ribeiro-Neto B. Modern Information Retrieval. — Addison-Wesley, 1999. — 340 p.
- [Баряхнин, 1980] Баряхнин В.Б., Федотов А.М. Проблемы разработки технологии фактографического поиска – М.: Институт вычислительных технологий СО РАН, 1980. – 150 с.
- [Бондаренко, 2007] Бондаренко М.Ф., Шабанов-Кушнарченко Ю.П. Теория интеллекта. Харьков. Изд-во СМИТ. 2007. 576 с.
- [Ландэ, 2009] Ландэ Д.В., Снарский А.А., Безсуднов И.В. Интернетика: Навигация в сложных сетях: модели и алгоритмы - М.: Либроком (Editorial URSS), 2009. - 264 с.
- [Ушаков, 1939] Толковый словарь русского языка: В 4 т. / Под ред. Д. Н. Ушакова. Т. 1. М., 1935; Т. 2. М., 1938; Т. 3. М., 1939; Т. 4. М., 1940.

-
- [Филлмор, 1981] Филлмор Ч. Дело о падеже открывается вновь // Новое в зарубежной лингвистике. – М.: Изд. иностр.лит., 1981, вып. 10. – С. 496-530
- [Хайрова, 2010] Хайрова Н. Ф., Тарловский В. А. Использование семантико-ориентированного лингвистического процессора для добывания новых знаний из потока документов корпоративной информационной системы / Вісник Національного технічного університету «ХПІ». Збірник наукових праць. Мематичний випуск «Системний аналіз, управління та інформаційні технології». — Х.: НТУ «ХПІ». — 2010. — № 67. — С. 132-138.
- [Хайрова, 2012] Хайрова Н. Ф. Використання логіко-алгебраїчної моделі семантичних відмінків для семантичного аналізу речення/ Н. Ф. Хайрова // Збірник наукових праць Військового інституту Київського національного університету імені Тараса Шевченка. – К.: ВІКНУ, 2012.- Вип. № 38. – С. 239 – 245.
-

Сведения об авторах

Хайрова Нина – доцент кафедри інтелектуальних комп'ютерних систем Національного технічного університету «Харьковский политехнический институт», ул. Фрунзе, 21, Харьков, 61002, Украина; e-mail: nina_khajrova@yahoo.com

Научные интересы: искусственный интеллект, обработка знаний, автоматическая обработка текстов

Шаронова Наталья – профессор, заведующий кафедрой интеллектуальных компьютерных систем Національного технічного університету «Харьковский политехнический институт», ул. Фрунзе, 21, Харьков, 61002, Украина; e-mail: nvsharonova@mail.ru

Научные интересы: искусственный интеллект, математическое моделирование, автоматизированные библиотечные системы