

ISSN 2306-1561

Automation and Control in Technical Systems (ACTS)

2014, No 2, pp. 107-122.

DOI: 10.12731/2306-1561-2014-2-11



Researching some approaches and main problems of native Russian Language understanding

Adamova Larisa Evgen'evna

Russian Federation, Ph. D., Associate professor.

Scientific-Research Institute "MIVAR" (SRI "MIVAR"), 127572, Russian Federation, Moscow, Altufevskoe Highway, 95 B. <http://www.mivar.ru>

larisapers@yandex.ru

Peterson Anastasiya Olegovna

Russian Federation, Student, Department of «Automated Control Systems».

Moscow Automobile & Road construction State Technical University, 125319, Russian Federation, Moscow, Leningradsky prospekt, 64. Tel.: +7 (499) 151-64-12. <http://www.madi.ru>

vasileva5193@gmail.com

Protopopova Dar'ya Aleksandrovna

Russian Federation, Student, Department of «Automated Control Systems».

Moscow Automobile & Road construction State Technical University, 125319, Russian Federation, Moscow, Leningradsky prospekt, 64. Tel.: +7 (499) 151-64-12. <http://www.madi.ru>

darina-1993@mail.ru

Skakunova Elena Anatol'evna

Russian Federation, Student, Department of «Automated Control Systems».

Moscow Automobile & Road construction State Technical University, 125319, Russian Federation, Moscow, Leningradsky prospekt, 64. Tel.: +7 (499) 151-64-12. <http://www.madi.ru>.

skakunova1993@mail.ru

Varlamov Oleg Olegovich

Russian Federation, Doctor of Technical Sciences, Professor, Department of «Applied Mathematics».

Moscow Automobile & Road construction State Technical University, 125319, Russian Federation, Moscow, Leningradsky prospekt, 64. Tel.: +7 (499) 151-64-12. <http://www.madi.ru>

ovar@mivar.ru

Abstract. The problem of native language understanding which is one of the main problems in creating artificial intellect has been discussed in this article. Comparative analysis of the ways in native language understanding has been carried out. In theory practicability of using mivar way to simulate understanding meaning of the text was proved. We confirm new directions in understanding Russian language for example: automatic studying mivar knowledge base, creation mivar concept network, studying virtual persons with help of mivar knowledge base.

Keywords: mivar, mivar's net, mivar technologies, native language, artificial intelligence.

ISSN 2306-1561

Автоматизация и управление в технических системах (АУТС)

2014. – №2. – С. 107-122.

DOI: 10.12731/2306-1561-2014-2-11



УДК 004.9

Исследование подходов и основных проблем понимания естественного Русского Языка

Адамова Лариса Евгеньевна

Российская Федерация, кандидат психологических наук, доцент.

Научно-исследовательский институт «МИВАР» (НИИ «МИВАР»), 127572, Российская Федерация, Москва, Алтуфьевское ш., д. 95 Б. <http://www.mivar.ru>. larisapers@yandex.ru

Петерсон Анастасия Олеговна

Российская Федерация, студент кафедры «Автоматизированные системы управления».

ФГБОУ ВПО «Московский автомобильно-дорожный государственный технический университет (МАДИ)», 125319, Российская Федерация, г. Москва, Ленинградский проспект, д.64, Тел.: +7 (499) 151-64-12, <http://www.madi.ru>. vasileva5193@gmail.com

Протопопова Дарья Александровна

Российская Федерация, студент кафедры «Автоматизированные системы управления».

ФГБОУ ВПО «Московский автомобильно-дорожный государственный технический университет (МАДИ)», 125319, Российская Федерация, г. Москва, Ленинградский проспект, д.64, Тел.: +7 (499) 151-64-12, <http://www.madi.ru>. darina-1993@mail.ru

Скакунова Елена Анатольевна

Российская Федерация, студент кафедры «Автоматизированные системы управления».

ФГБОУ ВПО «Московский автомобильно-дорожный государственный технический университет (МАДИ)», 125319, Российская Федерация, г. Москва, Ленинградский проспект, д.64, Тел.: +7 (499) 151-64-12, <http://www.madi.ru>. skakunova1993@mail.ru

Варламов Олег Олегович

Российская Федерация, доктор технических наук, профессор кафедры «Прикладная математика».

ФГБОУ ВПО «Московский автомобильно-дорожный государственный технический университет (МАДИ)», 125319, Российская Федерация, г. Москва, Ленинградский проспект, д.64, Тел.: +7 (499) 151-64-12, <http://www.madi.ru>. ovar@mivar.ru

Аннотация. Рассмотрена проблема понимания естественного языка, которая является одной из основных проблем в области создания искусственного интеллекта. Проведен сравнительный анализ методов понимания естественного языка. Теоретически обоснована целесообразность применения миварного продукционного подхода к моделированию понимания смысла текстов. Обоснованы новые направления в понимании русского языка: автоматизированное обучение миварной базы знаний; формирование комплексной миварной сети понятий (концептов); обучение виртуальных личностей на основе заполнения миварной базы знаний.

Ключевые слова: мивар, миварные сети, миварные технологии, продукционный подход, естественный язык, искусственный интеллект.

1. Введение

На сегодняшний день важным направлением в области искусственного интеллекта является обработка естественного языка. Понимание естественного языка включает гораздо больше, чем разбор предложений на отдельные части речи и поиск значений слов в словаре [1]. Как известно, естественный язык – это язык используемый для общения людей (в отличие от формальных языков и других типов знаковых систем, также называемых языками в семиотике) и не созданный целенаправленно (в отличие от искусственных языков) [2]. Словарь и грамматические правила естественного языка определяются практикой применения и не всегда бывают формально зафиксированы. В связи с этим возникают проблемы понимания естественного языка при создании искусственного интеллекта [1-24, 26, 27, 30].

2. Проблематика искусственного интеллекта и понимание естественного языка

Сегодня за счет достижений в области искусственного интеллекта создано большое количество научных разработок, которое существенно упрощает жизнь людей. Распознавание речи или отсканированного текста, решение вычислительно сложных задач за короткое время и многое другое - все это стало доступно благодаря развитию искусственного интеллекта. Замена человека-специалиста на системы искусственного интеллекта, в частности на экспертные системы, разумеется, там, где это допустимо, позволяет существенно ускорить и удешевить процесс производства. Системы искусственного интеллекта всегда объективны, и результаты их работы не зависят от моментного настроения и ряда других субъективных факторов, которые присущи человеку [6].

Одной из долгосрочных целей искусственного интеллекта является создание программ, способных понимать человеческий язык и строить фразы на нем. Оно базируется на обширном фоновом знании о предмете беседы и идиомах, используемых в этой области, так же, как и на способности применять общее контекстуальное знание для понимания недомолвок и неясностей, присущих естественной человеческой речи [6].

Задача сбора и организации этого фонового знания, чтобы его можно было применить к осмыслению языка, составляет значительную проблему в автоматизации понимания естественного языка. Для ее решения исследователи разработали множество методов структурирования семантических значений, используемых повсеместно в искусственном интеллекте [6]. Основная часть текущих работ в области понимания естественных языков направлена на поиск формализмов представления, которые должны быть достаточно общими, чтобы применяться в широком круге приложений и уметь адаптироваться к специфичной структуре заданной области. Множество разнообразных методик (большинство из которых являются развитием или модификацией семантических сетей) исследуются с этой целью и используются при разработке программ, способных понимать естественный язык в ограниченных, но достаточно интересных предметных областях [6].

Отметим, что под естественным языком чаще всего понимают английский язык, который значительно отличается от других языков. В нашей работе исследуется проблема понимания естественного русского языка, представленного в виде правильно написанных текстов (без картинок и графических символов). Глобальная цель наших исследований – это создание такой компьютерной системы обработки естественного языка, которая была бы в состоянии приобретать знания самостоятельно, читая текст (самостоятельно вычленив из текста информацию). В настоящее время проблема понимания естественного языка не решена, даже с учетом значительного роста компьютерных мощностей. Существуют отдельные решения в ограниченных предметных областях, но это лишь подчеркивает сложность и важность решения исследуемой проблемы. Скорее всего, проблема понимания естественного языка является столь же сложной и многоплановой, как и создание искусственного интеллекта. Миварные технологии эволюционного накопления и быстрой логико-вычислительной обработки данных позволили создать теоретические основы построения систем логического искусственного интеллекта [5, 12-24, 26-30]. Следовательно, целесообразно исследовать возможности применения миварных технологий для решения проблемы понимания естественного языка, хотя бы на уровне логически взаимосвязанных и обоснованных высказываний на русском языке. Таким образом, исследование новых подходов к пониманию естественного русского языка является актуальным и практически значимым.

3. Языковые неоднозначности и проблемы понимания языка

Обработка естественного языка (Natural Language Processing, NLP) — общее направление [искусственного интеллекта](#) и [математической лингвистики](#) [2]. Оно изучает проблемы [компьютерного](#) анализа и [синтеза естественных языков](#). Применительно к искусственному интеллекту анализ означает понимание языка, а синтез — генерацию грамотного текста. Решение этих проблем будет означать создание более удобной формы взаимодействия компьютера и человека. Одной из ключевых проблем NLP является языковая неоднозначность. Виды языковой неоднозначности [2]:

Синтаксическая (структурная) неоднозначность: во фразе «Юноша коротко попрощался с подругой певицы, которая ему нравилась» не понятно кто понравился юноше, певица или подруга певицы. В этом и заключается синтаксическая неоднозначность.

Смысловая неоднозначность: во фразе «Мама хранит деньги в банке» слово «банке» может означать как финансово-кредитное учреждение «банк», так и стеклянный сосуд - «банка».

Падежная неоднозначность: предлог «по» в предложениях «Он не бежал по дорожке / Он не бежал по болезни» обозначает либо место, либо причину, т.е. представлены совершенно различные отношения.

Референциальная неоднозначность: для системы, не обладающей знаниями о реальном мире, будет затруднительно определить, с каким словом - «торт» или «стол» - соотносится местоимение «он» в косвенном падеже во фразе «Я взял торт со стола и съел его».

Литерация: в диалоге «-Можешь открыть окно? – Мне холодно!» ни просьба, ни ответ не выражены нестандартным способом. В других обстоятельствах на вопрос может быть получен прямой ответ «да»/ «нет», но в данном случае в ответе выражена просьба не открывать окно.

Всякая речь (текст) состоит из слов. Не зная значения отдельных слов, нельзя понять ни сказанного, ни написанного. Наоборот: если известен смысл каждого слова, не трудно разобраться в смысле всего текста. Любое предложение состоит из слов, как сумма из слагаемых: известны все слагаемые – можно найти и сумму. Однако математика - одно, а язык – другое. Поэтому, для полного понимания текста, нужно вложить в него определённую языковую систему. Языковая система в свою очередь состоит из: слов, значений этих слов и связей между ними. При этом, значения слов необходимо сделать краткими, понятными, сохранив при этом их строгую научность. Но в тоже время одно слово может иметь несколько значений.

Тему многозначности слов затрагивал в своей работе австрийский философ и логик Людвиг Витгенштейн [3]. Естественный язык был главным предметом анализа Витгенштейна. Ученый опровергает мнение, что смысл человеческого языка можно вывести из каких-либо объективных основ. В трудах Витгенштейна, как и в теории речи (speech act theory), развитой Остином (Austin) и его последователями (Grice, Searle), значение любого высказывания зависит от человеческого, культурного контекста. Значение слова "сиденье", к примеру, зависит от наличия физического объекта, который можно применить для сидения на нем, а также культурных соглашений об использовании сидений. Когда, например, большой плоский камень можно назвать сиденьем? Почему нелепо так называть королевский трон? Какая разница между человеческим пониманием "сиденья" и пониманием кота или собаки, которые в человеческом смысле сидеть не могут? Атакуя основы смысла, Витгенштейн утверждал, что мы должны рассматривать использование языка посредством выбора и действий в изменчивом культурном контексте [3].

Поэтому, если развивать только языковую систему, без вложения в нее понимания контекста, то полное достижение понимания текста невозможно. В настоящее время в среде ученых сложилось общее мнение, что для понимания текстов необходимо работать с большими контекстами. Известен термин «картина мира», под которым понимают глобальную модель мира, необходимую для обработки текстов на естественном языке. В нашем подходе, «картина мира» – это объединение самых разнообразных контекстов, которое можно накапливать в миварных многомерных базах данных и правил на основе концепции «Вещь-Свойство-Отношение».

4. Методы понимания естественного языка

В области «понимания естественного языка» исследователи занимаются анализом и генерацией текстов, их внутренним представлением, выявлением знаний, необходимых для их понимания, т.е. синтаксических, семантических, прагматических знаний. Главной проблемой на пути к полному пониманию естественного языка остается грамматика и семантика. Также естественный язык слабо поддается формализации. Кроме того, от языков программирования их отличает присутствие многозначности [4].

Для решения проблемы понимания естественного языка применяют методы, приведенные ниже [26, 27].

4.1. Метод машинного перевода

Данный метод был разработан во второй половине XX века. Основой для понимания естественного языка является создание семантической модели представления смысла текстов. Создание такой модели привело к пониманию необходимости создания языка для внутреннего представления знаний [10]. Важным достижением данного метода является понимание того, что анализ текстов на естественном языке состоит из четырех основных этапов: морфологический анализ, синтаксический анализ, семантический анализ, прагматический анализ [10]. Таким образом, достижения в области машинного перевода оказались очень полезными для ИИ, так как они показали, что анализ естественно-языковых конструкций невозможен без создания семантической модели [10].

4.2. Метод использования языка Пролог в среде программирования Велиар-2

Этот метод разработан Киевской лабораторией искусственного интеллекта [11]. В основе лежит принцип вложенности языков, который позволяет постепенно наращивать возможности системы понимания и таким образом пройти под границей зоны комбинаторного взрыва числа вариантов анализируемых альтернатив. Разработчики данного метода используют модель малого русского языка и модель процесса понимания. Основными особенностями использования языка Пролог являются [11, 26, 27]:

- близость к естественному языку;
- инкрементальный рост процедур
- логическая интерпретация;

- систематическое освоение многозначности;
- совершенная система обработки отказов;
- однократность присваивания;
- ориентация на параллельные компьютеры;
- простота грамматических анализаторов;
- термы и списки в качестве структур данных;
- метаинтерпретаторы;
- машинно-независимость
- эффективная реализация;
- ансамбли;
- конструктивное отрицание;
- обобщение в дереве реализации.

4.3. Метод использования Томита-парсера

Данный метод назван в честь японского ученого Масару Томита, автора алгоритма GLR-парсинга (Generalized left-to-right algorithm), на основе которого и был создан Томита-парсер. GLR-парсер (от англ. Generalized Left-to-right Rightmost derivation parser — Обобщенный восходящий магазинный анализатор) — расширенный алгоритм LR-парсера, предназначенный для разбора по недетерминированным и неоднозначным грамматикам. Принципы работы GRL-парсера схожи с алгоритмом LR-парсера, так как он является его производным. Единственная особенность GRL-парсера — способность разрешать недетерминированность и неоднозначность естественных языков [2]. На сегодняшний день, IT-компания «Яндекс» активно использует в своей работе Томита-парсер. Уже четыре сервиса, такие как Яндекс.Работа, Яндекс.Новости, Яндекс.Авто и Яндекс.Почта работают на основе данного парсера. Он анализирует вводимую информацию, выделяет главную сущность и формализует вывод.

4.4. Миварные технологии

Миварный подход включает технологии накопления данных и обработки информации в едином миварном информационном пространстве. Для понимания языка надо собрать и поддерживать в актуальном состоянии огромную базу данных фактов и такое же большое количество правил, которые позволяют выявлять нюансы смысла разных понятий в различных ситуациях [5, 19-22]. Миварный подход основывается на представлении всей информации в виде "тройки": "вещь-свойство-отношение". Тогда все процедуры, подпрограммы, агенты, сервисы и т.п. могут быть представлены в виде неких отношений, которые хранятся в базе данных, а при необходимости запускаются на основе миварной сети. Создание миварной сети — это одновременное параллельное выявление миварных правил и занесение их в базу данных. Миварные сети постепенно связываются из них в полноценную миварную сеть большого объема [5, 9, 12-24]. Миварный подход позволяет для одной предметной области хранить несколько разных процедур, решающих одинаковые задачи. Более того, если вычислительные ресурсы

позволяют, то все эти процедуры могут запускаться одновременно, а потом на основе конкуренции из представленных ими результатов выбирается наиболее подходящий для каждого конкретного случая [5, 9]. Миварные технологии позволяют создавать большие описания предметных областей и использовать их в качестве контекста при обработке текста. Миварные сети, обладающие линейной вычислительной сложностью логического вывода, позволяют очень быстро обрабатывать все контексты.

5. Продукционный миварный подход к пониманию языка

Однако же с точки зрения миварного подхода задача понимания русского языка решается без применения предикатов и формальной грамматики [5, 19-22]. Задача понимания естественного языка базируется на применении логики и использования больших баз правил реального использования русского языка. Как известно, миварный подход применяется на логическом уровне исследований в области искусственного интеллекта. На этом уровне все знания можно представить в виде причинно-следственных зависимостей, т.е. в виде продукций «Если ..., то...». Таким образом, предложено использовать продукционный подход для моделирования понимания естественного языка. Фактически, миварные сети позволяют развить продукционный подход (причинно-следственные зависимости "Если ..., то...") и создать автоматическую обучаемую логически рассуждающую систему [5, 12-24].

Продукционная модель знания – модель, основанная на правилах, позволяет представить знание в виде предложений типа «Если ... (условие), то ... (действие)» [7]. Продукционная модель обладает тем недостатком, что при накоплении достаточно большого числа (порядка нескольких сотен) продукций они начинают противоречить друг другу [7]. В общем случае продукционную модель можно представить в следующем виде [7]:

$i = \langle S; L; A \rightarrow B; Q \rangle$, где:

S — описание класса ситуаций; L – условие, при котором продукция активизируется; $A \rightarrow B$ – ядро продукции; Q – постусловие продукционного правила. В настоящее время, продукции – это наиболее проработанная и распространенная модель представления знаний, в частности в экспертных системах. Полнота базы знаний определяет возможности системы по удовлетворению потребностей пользователей. Логический вывод в продукционных системах основан на построении прямой и обратной цепочек заключений, образуемых в результате последовательного просмотра левых и правых частей соответствующих правил, вплоть до получения окончательного заключения [8].

Джордж Ф. Люгер считал, что [3]: «Действительно, гибкость, позволяющая биологическому разуму реагировать практически на бесконечное количество различных ситуаций приемлемым, если даже и не оптимальным образом - отличительная черта разумного поведения. Справедливо замечание, что управляющая логика, используемая в большинстве традиционных компьютерных программ, не проявляет великой гибкости или силы воображения, но неверно, что все программы должны писаться подобным

образом. Большая часть работ в сфере ИИ за последние 25 лет была направлена на разработку таких языков программирования и моделей, призванных устранить упомянутый недостаток, как продукционные системы, объектные системы, сетевые представления и другие модели...». В миварах продукционный подход получил значительное развитие, он стал более мощным. Поэтому, ограничения, указанные Люгером, на данный момент преодолены. Миварные сети работают с сервисно-ориентированными архитектурами, имеют сложные механизмы запуска процедур и обладают линейно вычислительной сложностью логического вывода. Миварные сети работают с большими базами данных, поэтому создано новое поколение продукционных систем. Этот подход, в целях понимания контекста, позволяет реализовывать быструю логическую обработку больших текстов.

6. Новые задачи для развития миварного понимания естественного языка

Миварные технологии позволяют создавать качественно новые интеллектуальные системы, способные быстро логически обрабатывать большие массивы данных в многомерных эволюционных базах данных и правил. Для понимания текстов на естественных языках необходима обработка контекстов [5, 19-22]. Контексты представляют собой большие массивы дополнительной информации, которая может быть представлена в виде универсального миварного информационного пространства. Таким образом, теоретически обоснована целесообразность применения миварного продукционного подхода к моделированию понимания смысла текстов. В миварных технологиях уже получены разнообразные новые научные результаты [12-24]. Большую часть этих достижений можно непосредственно применять к решению задачи понимания естественного языка [19-22].

Однако, в настоящее время выявлены и обоснованы три новые актуальные задачи. Прежде всего, учитывая важность, сложность и трудоемкость «ручного» заполнения огромных миварных баз данных структурированной информацией, необходимо решить проблему и выполнять автоматизированное обучение миварной базы знаний.

Миварный подход к пониманию естественного русского языка основывается на том, что можно выделить три уровня работы со словами:

- 1) словоформы,
- 2) слова (начальные формы слов),
- 3) понятия-концепты.

Концепты – это слова с конкретным значением, например, описанным в толковых словарях русского языка. Многие слова обладают многозначностью, но концепты позволяют однозначно определять значение слова (в смысле толкования на основе словаря значений). При таком подходе возникает важная, новая и актуальная задача: формирование комплексной миварной сети понятий (концептов).

Наконец, третьей актуальной задачей является обучение виртуальных личностей на основе заполнения миварной базы знаний. Отметим, что еще в самых первых работах по миварам [5, 12] было предложено разрешить пользователям создавать свои

индивидуальные информационные пространства. Дальнейшее развитие науки и техники подтвердило необходимость этого и сделало актуальным персонализацию миварных баз данных и правил. Отметим, что такой подход не отменяет централизованное создание общеупотребимых энциклопедий как всеобщей «картины мира» – миварного инфопространства, а лишь дополняет и позволяет каждому человеку – пользователю создавать и свои «индивидуальные картины мира».

7. Виртуальные личности и миварные технологии понимания языка

Каждый человек воспринимает окружающий мир в рамках развития своего сознания. Эти рамки формируются под информационным воздействием в ходе воспитания и обучения человека, находящегося в определенной социальной среде обитания. В ситуации с ИИ – наоборот. Его мир бесконечен, и в нем нет никаких разделений. Понимание контекста человеком рассматривается как результат интерпретации отношений.

Полное понимание высказывания наступает тогда, когда субъекту становится ясно, о какой ситуации идет речь и как представляет себе ситуацию говорящий. Человек не обладает знанием законов языка, на котором будет говорить, с рождения. В определённый период своего развития он усваивает нормы и правила языка, на котором общаются окружающие его люди. Однако, в жизни не всегда все как в правилах, люди сталкиваются с тем, что одной общей онтологии не достаточно [13, 28]. В естественном языке не может быть жестких правил, так как, в зависимости от многих критериев, люди смотрят на вещи по-разному и вкладывают в смысл одного и того же слова разный смысл. Например, человек, живущий в городе, считает, что «домашнее животное» - это кот, собака, черепаха, попугай, рыбки и т.п., а человек, живущий на ферме, считает, что «домашнее животное» – это, в первую очередь, коровы, свиньи, козы и т.п. Следовательно, роль контекста в жизни человека очень велика. А в случае с ИИ важность возрастает в несколько раз. Учитывая этот факт, необходимо понимать, что создание ИИ подразумевает создание виртуальной личности (ВЛ). ВЛ – есть нечто, призванное подменить собой личность реальную, способное «думать» на том же уровне. Обучение ВЛ является одной из проблем создания ИИ.

Для решения этой проблемы необходимо решить следующие задачи:

1. Сформировать у ВЛ представления о зрительных образах с помощью картинок.
2. Сформировать у ВЛ умения получать информацию об окружающем мире с помощью базы знаний.
3. Обучать ВЛ использовать получаемую информацию в познавательной и коммуникативной деятельности.

ВЛ имеет свою базу знаний, на основе которой происходит понимание контекста. Понимание контекста ВЛ базируется на основных принципах и методах коррекционного и восстановительного обучения [25, стр. 248]. Обучение происходит на основе преодоления нарушений различных видов гнозиса и праксиса, рассмотренных в книге

Т.Г. Визель [25, стр. 307- 341]. В миварных технологиях могут быть реализованы несколько способов обучения ВЛ [25, 28].

1) Смысловое обыгрывание (рисунок 1)

Для этого используется метод [25, стр. 290], направленный на обобщение категориального отношения к цвету: преодоление цветового барьера. При этом используются понятия на основе стереотипных образов (пример: при описании красный цвет, ВЛ предъявляется картинка красного помидора, рябины и т.д.)

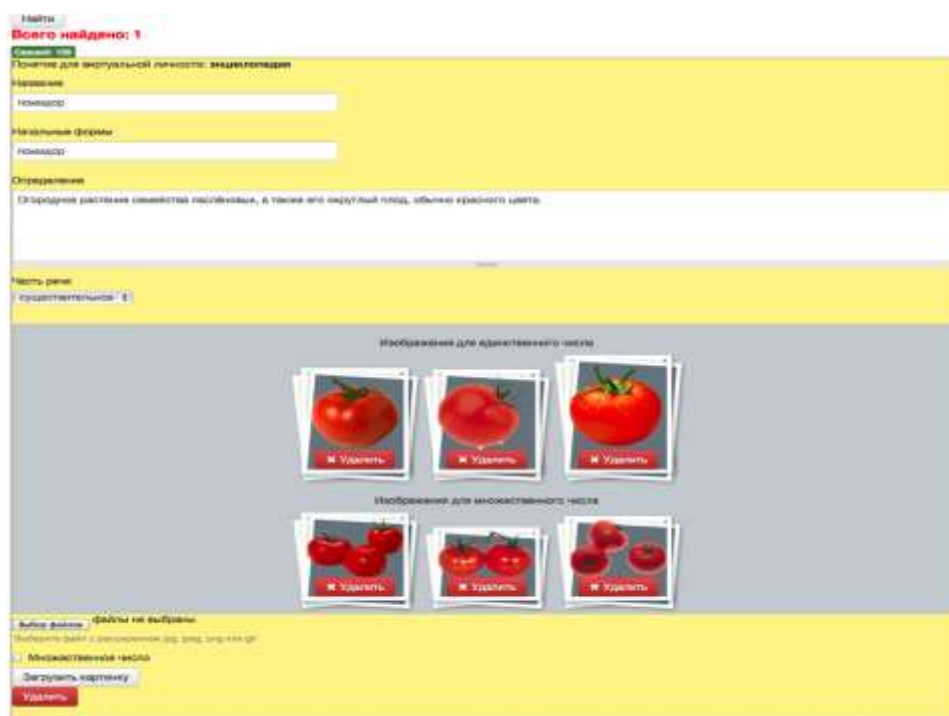


Рисунок 1 – Смысловое обыгрывание

2) Оперирование объектами в пространстве (рисунок 2)

Для реализации этого метода [25, стр. 291] необходимо использовать два объекта. Один объект неподвижен, другой меняет свое местоположение относительно первого: преодоление схематических представлений о пространственном соотношении объектов в действительности. Сюда входят: предлоги и слова с пространственным значением (в, на, под, над, из, через, перед, возле, около).

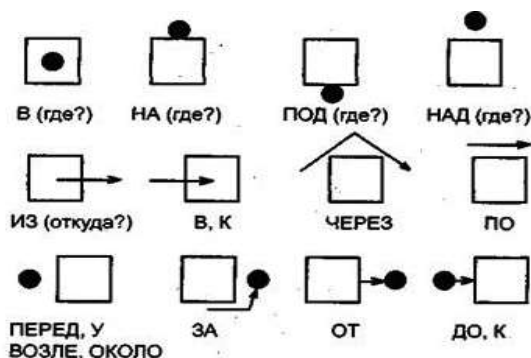


Рисунок 2 – Оперирование объектами в пространстве

3) Преодоление семантического барьера

Использование соответствующих морфологических признаков [25, стр. 284]: приставки, предлоги, суффиксы. (подойти - отойти, подъехать – отъехать, кулачок-кулачище)

4) Восприятие геометрических форм (рисунок 3).

Выработка дифференциального восприятия округлых и угловатых форм на материале геометрических фигур [25, стр. 293]. Схематизация понятий формы и размера.

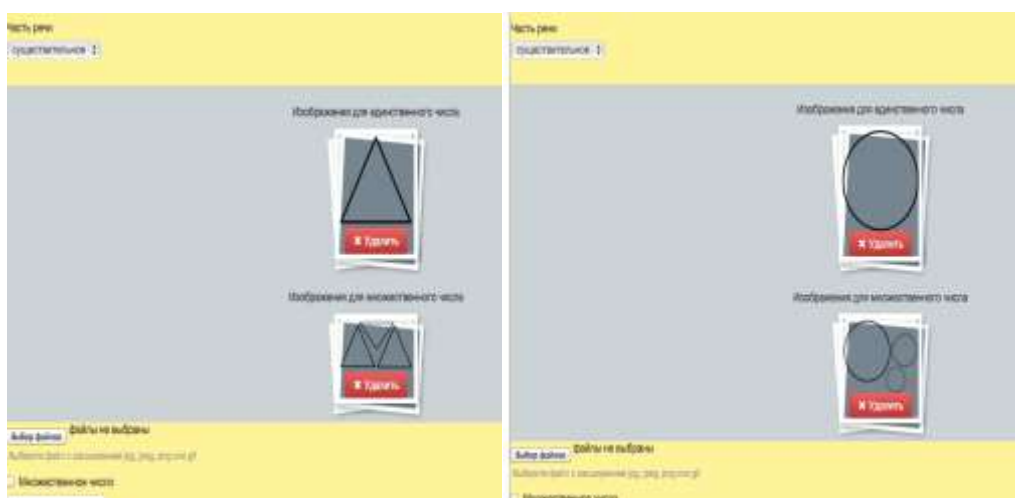


Рисунок 3 – Восприятие округлых и угловатых форм

5) Обучение соответствию целого и его частей (рисунок 4)

Метод основан на отношении между некоторой совокупностью предметов и отдельными предметами, образующими эту совокупность [25, стр. 293]: выделение в окружающем мире отдельных объектов, а в объектах – составляющие их элементы. (пример: Тело человека состоит из: головы, шеи, туловища, рук и ног. Голова в свою очередь состоит из лица, затылка, макушки, уха и волос. А на лице уже расположены нос, глаза и т.д.

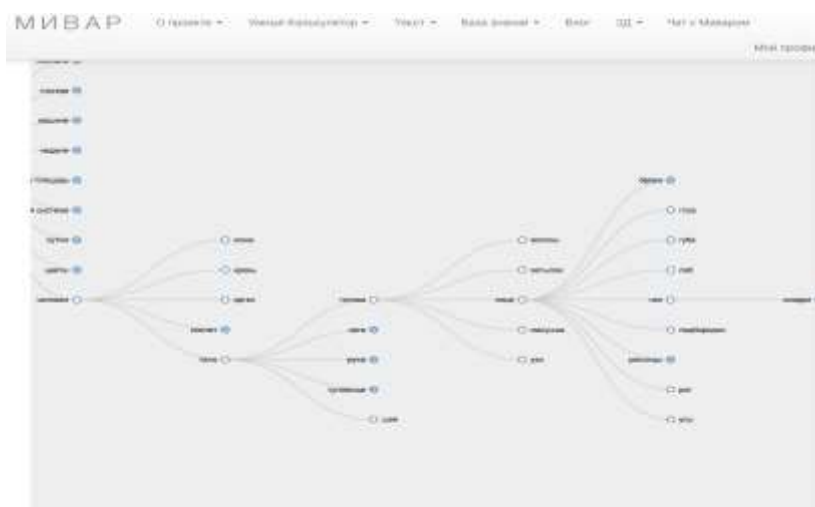


Рисунок 4 – Тип отношений: «часть-целое»

7. Заключение

Теоретически обоснована возможность применения миварных технологий как развития продукционного подхода к пониманию естественного языка на основе моделирования и логической обработки сверхбольших объемов данных и правил, т.е. контекстов. Произведен сравнительный анализ методов понимания естественного языка, и показаны преимущества применения миварных технологий.

Миварный подход включает технологии накопления данных и обработки информации в едином миварном информационном пространстве. Для понимания языка надо собрать и поддерживать в актуальном состоянии огромную базу данных фактов и такое же большое количество правил, которые позволяют выявлять нюансы смысла разных понятий в различных ситуациях.

Миварные сети позволяют развить продукционный подход (причинно-следственные зависимости "Если..., то...") и создать автоматическую обучаемую логически рассуждающую систему. Этот подход, в целях понимания текста, позволяет реализовывать большие контексты и быструю логическую обработку. Практические исследования показали необходимость персонализации контекстов через создание различных виртуальных личностей. Миварные технологии доказали на практике свои преимущества в области понимания естественного языка: обработка в реальном времени сверхбольших объемов данных и правил. Для дальнейшего развития необходимо решить следующие научные задачи.

1. Разработать методику автоматизированного обучения миварной базы знаний.
2. Разработать методику формирования комплексной сети понятий в миварной базе знаний.
3. Разработать методику обучения виртуальных личностей на основе заполнения миварной базы знаний с использованием тонкого клиента.

Учитывая достижения миварных технологий в накоплении и обработке информации, решение трех указанных задач позволит перейти на новый качественный уровень автоматизированной обработки текстов (АОТ) на основе логической обработки больших массивов данных и учета контекста.

Список информационных источников

- [1] Хромин А.А. Проблема искусственного интеллекта. [Электронный ресурс] — Режим доступа. — URL:<http://www.structuralist.narod.ru/articles/ai.htm> (дата обращения: 20.05.2014).
- [2] Википедия. [Электронный ресурс] — Режим доступа. — <http://ru.wikipedia.org/>
- [3] Люгер Дж.Ф. Искусственный интеллект: стратегии и методы решения сложных проблем, 4-е издание: пер. с англ. — М.: Издательский дом "Вильямс", 2005.
- [4] Лорьер Ж.-Л. Системы искусственного интеллекта: Пер. с франц. — М.: Мир, 1991. — 568 с.
- [5] Варламов О.О. Эволюционные базы данных и знаний для адаптивного синтеза интеллектуальных систем. Миварное информационное пространство. - М.: Радио и связь, 2002. - 288 С.

- [6] Портал искусственного интеллекта. [Электронный ресурс] — Режим доступа. — <http://www.aiportal.ru/>.
- [7] Белоус Е.С., Кудинов В.А., Желнин М. Э. Современные модели представления знаний в обучающих системах // Ученые записки. Электронный научный журнал курского государственного университета – 2010, № 1, С. 9-14. (<http://elibrary.ru/item.asp?id=13084124>)
- [8] Тузовский А.Ф., Чириков С.В., Ямпольский В.З. Системы управления знаниями (методы и технологии). - Томск: Изд-во НТЛ, 2005.
- [9] Владимиров А.Н., Варламов О.О., Носов А.В., Потапова Т.С. Применение многопроцессорного вычислительного кластера НИИР для распараллеливания алгоритмов в научно-технических и вычислительных задачах // Труды Научно-исследовательского института радио. 2009. № 3. С. 120-123.
- [10] Искусственный интеллект, справочник в 3-х томах. // Под ред. Захарова В. Н. и Хорошевского В. Ф. - М.: Радио и связь, 1990,
- [11] Казаков А.И., Обухов Л.И. Система понимания естественного языка. [Электронный ресурс] – Режим доступа. – URL: http://neurotechnica.info/?page_id=4 (дата обращения: 25.05.2014).
- [12] Варламов О.О. Системный анализ и синтез моделей данных и методы обработки информации в самоорганизующихся комплексах оперативной диагностики // автореферат диссертации на соискание ученой степени доктора технических наук. - Москва, 2003.
- [13] Варламов О.О. Основы многомерного информационного развивающегося (миварного) пространства представления данных и правил // Информационные технологии. 2003. № 5. С. 42.
- [14] Варламов О.О. Системный анализ и синтез моделей данных и методы обработки информации для создания самоорганизующихся комплексов оперативной диагностики // Искусственный интеллект. 2003. № 3. С. 299.
- [15] Варламов О.О. Разработка метода распараллеливания потокового множественного доступа к общей базе данных в условиях недопущения взаимного искажения данных // Информационные технологии. 2003. № 1. С. 20.
- [16] Варламов О.О. Системы обработки информации и взаимодействие групп мобильных роботов на основе миварного информационного пространства // Искусственный интеллект. 2004. № 4. С. 695.
- [17] Варламов О.О., Владимиров А.Н., Бадалов А.Ю., Чванин О.Н. Развитие миварного метода логико-вычислительной обработки информации для АСУ, тренажеров, экспертных систем реального времени и архитектур, ориентированных на сервисы // Труды Научно-исследовательского института радио. 2010. № 3. С. 18-26.
- [18] Варламов О.О. Обзор 25 лет развития миварного подхода к разработке интеллектуальных систем и созданию искусственного интеллекта // Труды Научно-исследовательского института радио. 2011. № 1. С. 34-44.
- [19] Варламов О.О., Сергушин Г.С., Елисеев Д.В., Адамова Л.Е., Майборода Ю.И., Антонов П.Д., Чибирова М.О. О миварном подходе к моделированию процессов понимания компьютерами смысла текстов, речи и образов. Новые возможности расширения границ автоматизации умственной деятельности человека // Автоматизация и управление в технических системах. – 2013. – № 2. – С.38-51.
- [20] Varlamov O.O., Adamova L.E.E., Eliseev D.V., Mayboroda Yu.I., Antonov P.D., Sergushin G.S., Chibirova M.O. Mivar Thechnologies in Mathematical Modeling of Natural Language, Images and Human Speech Understanding // International Journal of Advanced Studies. 2013. T. 3. № 3. С. 17-23.

- [21] Варламов О.О., Адамова Л.Е., Елисеев Д.В., Майборода Ю.И., Антонов П.Д., Сергушин Г.С., Чибирова М.О. Комплексное моделирование процессов понимания компьютерами смысла текстов, речи и образов на основе миварных технологий // Искусственный интеллект. – 2013. – № 4. – С. 15-27.
- [22] Чибирова М.О. Анализ подходов к построению систем поддержки принятия решений: онтологии и мивары // Автоматизация и управление в технических системах. 2014. № 1.2 (9). С. 44-60.
- [23] Варламов О.О. Практическая реализация линейной вычислительной сложности логического вывода на правилах "ЕСЛИ-ТО" в миварных сетях и обработка более трех миллионов правил // Автоматизация и управление в технических системах. – 2013. – № 1. – С. 60-97.
- [24] Максимова А.Ю., Варламов О.О. Миварная экспертная система для распознавания образов на основе нечеткой классификации и моделирования различных предметных областей с автоматизированным расширением контекста // Известия Южного федерального университета. Технические науки. 2011. № 12. С. 77-87.
- [25] Визель Т.Г. Основы нейропсихологии: учеб. для студентов вузов. - М.: АСТАстрель Транзиткнига, 2005. - 384 с.
- [26] Остроух А.В. Основы построения систем искусственного интеллекта для промышленных и строительных предприятий: монография / А.В. Остроух. – М.: ООО «Техполиграфцентр», 2008. - 280 с. – ISBN 978-5-94385-033-2.
- [27] Остроух А.В. Системы искусственного интеллекта в промышленности, робототехнике и транспортном комплексе: монография / А.В. Остроух - Красноярск: Научно-инновационный центр, 2013. – 326 с. – ISBN 978-5-906314-10-9.
- [28] Васюгова С.А. Исследование перспектив и проблем интеграции человека с компьютером: искусственный интеллект, робототехника, технологическая сингулярность и виртуальная реальность / С.А. Васюгова, А.В. Остроух, М.Н. Краснянский, А. Самаратунга // Перспективы науки. – Тамбов: «ТМБПринт», 2011. - № 4(19). - С. 109 – 112.
- [29] Варламов О.О. Анализ возможностей миварного подхода для систем искусственного интеллекта и современной робототехники / О.О. Варламов, А.В. Остроух, М.Н. Краснянский, Т.Л. Давыдова // Вестник ТГТУ. – 2011. – Т.17. – № 3. - С.687-694.
- [30] Остроух А.В. Ввод и обработка цифровой информации: учебник для нач. проф. образования / А.В. Остроух. – М.: Издательский центр «Академия», 2012. – 288 с. – ISBN 978-5-7695-9457-1.