

ISSN 2306-1561

Automation and Control in Technical Systems (ACTS)

2014, No 4, pp. 100-108.

DOI: 10.12731/2306-1561-2014-4-10



Researching of Method Based on Complex Concepts Network Forming in Mivar Knowledge Base with the Aim of Natural Language Understanding

Protopopova Dariya Alexandrovna

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

State Technical University – MADI, 125319, Russian Federation, Moscow, Leningradsky prospekt, 64.

Tel.: +7 (499) 151-64-12. <http://www.madi.ru>

darina-1993@mail.ru

Abstract. The problem of intelligent systems development, which is one of the major challenges in the field of artificial intelligence is described. The possibility of using mivar technologies in creating intelligent systems has been proved theoretically. The necessity of using concepts for understanding texts on natural language has been explained. Definition of concept in mivar knowledge base has been given accurately. The necessity of complex concepts network forming in mivar knowledge base has been proved theoretically. Forming methodology of concepts network (CNW) in mivar information space has been presented. Mivar information space structure is represented with examples. The structure of the model CNW is examined. Bipartite directed graph is displayed graphically. Concepts of model CNW parts are defined. Relations between concepts are reviewed in details. Concepts relations are classified by their nature. A framework of mivar knowledge base is presented.

Keywords: mivar, mivar method, mivar technologies, mivar nets, mivar knowledge base, concept, concept networks.

ISSN 2306-1561

Автоматизация и управление в технических системах (АУТС)

2014. – №4. – С. 100-108.

DOI: 10.12731/2306-1561-2014-4-10



УДК 004.9

Разработка методики формирования комплексной сети концептов в миварной базе знаний в целях решения задачи понимания текстов на естественном русском языке

Протопопова Дарья Александровна

Российская Федерация, магистрант 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

Аннотация. Рассмотрена проблема создания интеллектуальных систем понимания естественного языка, которая является одной из основных проблем в области искусственного интеллекта. Теоретически обоснована возможность применения миварных технологий для создания интеллектуальных систем. Обоснована необходимость использования концептов для понимания текстов на естественном языке. Дано определение концепта в миварной базе знаний. Теоретически доказана необходимость формирования комплексной сети концептов в миварной базе знаний. Изложена методика формирования сети концептов (ВСО) в миварном информационном пространстве. Представлена структура миварного информационного пространства с примерами. Рассмотрена структура модели ВСО. Графически представлен двудольный ориентированный граф. Даны понятия частей модели ВСО. Подробно рассмотрены связи между концептами. Проклассифицированы связи концептов по их природе. Представлен пример графического представления сети концептов в миварной базе знаний.

Ключевые слова: мивар, миварный подход, миварные сети, миварные технологии, миварные базы знаний, естественный язык, концепт, сети концептов.

1. Введение

В настоящее время, проблема создания интеллектуальных систем и систем искусственного интеллекта остается актуальной и практически значимой. Одним из

наиболее развитых типов интеллектуальных систем являются экспертные системы. Изучая и создавая экспертные системы, важным является представление структуры экспертных систем и их классификация, принципы построения и технология их разработки, представление знаний в экспертных системах и достижение целей. При этом особое внимание уделяется рассмотрению продукционных экспертных систем [1 – 3].

Создание экспертных систем нового поколения позволит автоматизировать решение различных сложных интеллектуальных задач. Более высокий качественный уровень в решении задач предполагает, прежде всего, обеспечение необходимой и достаточной интеллектуальной поддержкой. Интеллектуализация информационно-вычислительных систем (ИВС) имеет в виду использование не только нового поколения инструментальных средств, но и нового поколения математического, алгоритмического и программного обеспечения. Информационно-вычислительные системы с интеллектуальной поддержкой, как правило, применяются для решения сложных задач, где логическая (смысловая) обработка информации превалирует над вычислительной [2]. Одним из примеров таких задач является понимание и синтез текстов на естественном языке (ЕЯ). В настоящее время не создано практически полезных систем понимания текстов на естественных языках, т.к. существовавшие методы математической лингвистики не позволяют учитывать контекст в большом объеме. Использование миварного подхода позволяет разработать теоретические основы построения логического искусственного интеллекта и перейти на новый этап создания систем понимания текстов на естественных языках. Миварные технологии позволяют создавать качественно новые интеллектуальные системы, способные быстро логически обрабатывать большие массивы данных в многомерных эволюционных базах данных и правил [4-30]. Отметим, что для краткости «миварные многомерные активные эволюционные базы данных и правил» принято называть сокращенно: «миварные базы знаний».

Таким образом, актуальной является задача разработки методики формирования комплексной сети концептов в миварной базе знаний.

2. Обзор задач и работ по теме исследования. Необходимость формирования комплексной сети концептов в миварной базе знаний

Миварный подход включает технологии накопления данных и обработки информации в едином миварном информационном пространстве [4 – 30]. Ранее в работах [4 – 5] была обоснована целесообразность применения миварного подхода к решению задачи создания систем понимания текстов на естественных языках. В качестве примера был выбран русский язык.

Для понимания языка необходимо 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. Как известно [4], разработанные ранее в математической лингвистике методы и модели не позволяли хранить и обрабатывать большие контексты. Это было важное ограничение, обусловленное отсутствием моделей многомерного хранения в базах данных и быстрых

алгоритмов логического вывода. Это ограничение позволяет снять использование миварных технологий накопления информации и линейной сложности логической обработки [4 – 30].

Применение миварного подхода к пониманию естественного русского языка основывается на том, что можно выделить три уровня работы со словами [4 – 5]:

1) *словоформы* – слова в некоторой грамматической форме (в частном случае – в единственно имеющейся у слова форме).

Пример: «сады», «садами», «белый», «белую», «пишет», «вчера».

2) *слова* – начальные формы слов

Пример: «природа», «математика», «родители», «кот».

3) *понятия-концепты* – это слова в конкретном их значении.

Пример: слово «край» имеет 2 концепта такие, как:

- «край - это предельная линия, которая ограничивает поверхность плоского предмета, грань; часть поверхности, близкая к её предельной линии, кромка; самая удалённая от центра часть, окраина; верхний обрез стенки сосуда»;
- «край» - страна, область, местность.

Однако, присутствие многозначности в русском языке сильно затрудняет правильную автоматическую обработку естественного языка. Поэтому очень важно производить обработку языка на уровне контекста. Это возможно осуществить, если «понять», по смыслу других – однозначных слов, в каком значении используется многозначное слово в данном тексте. Использование концептов при обработке естественного языка позволяет однозначно определять значение слова (в смысле толкования на основе словаря значений). Поэтому решение проблемы правильной обработки естественного языка подразумевает формирование комплексной сети концептов в миварной базе знаний.

В настоящее время недостаточно проработаны вопросы формирования комплексной сети концептов в миварной базе знаний для понимания естественного русского языка. Следовательно, необходимо решить задачу разработки методики формирования комплексной сети концептов в миварной базе знаний.

3. Метод решения поставленной задачи. Методика формирования сети концептов в миварном информационном пространстве

Эволюционные базы данных и правил (знаний) с изменяемой структурой образуют миварное информационное пространство унифицированного представления данных и правил, которое, в свою очередь, базируется на "тройке" "вещь-свойство-отношение" [4 – 30]. Методика создания данной "тройки" для решения задачи понимания текстов основывается на правилах русского языка. Можно привести следующий пример, которые показывает основные направления такого создания миварной базы знаний. В процессе обучения русскому языку человек выявляет вещи (субъекты, объекты), которые соотносит через элементарные отношения с другими объектами. Также он понимает, что все вещи имеют какие-то свойства. Следовательно, можно привести такую

аналогию, что в процессе обучения русскому языку в голове у человека появляется многомерное пространство представления данных и правил.

В миварной базе знаний используется модель «вещь-свойство-отношение», в которой (в первом варианте) все «вещи» являются существительными, «свойства» - прилагательными и наречиями, а «отношения» – глаголами [4]. Для остальных частей речи тоже сформировано отображение в модель ВСО, но оно выходит за рамки нашей работы (первого варианта системы «МИВАР-Текст») и требует дополнительного научного исследования. Соответственно, под концептами в миварной сети понимаются части модели «вещь-свойство-отношение».

4. Результаты решения поставленной задачи. Формирование сети концептов в миварных базах знаний

Под сетью концептов принято понимать: множество концептов (вершин графа) и множество связей (дуг графа) между концептами, хранящееся в единой базе знаний вместе с самими концептами и образующее некоторый граф в миварном информационном пространстве.

Для более полного понимания смысла данного понятия, необходимо его графическое представление в виде миварной сети [4 – 30], которая представляет собой двудольный ориентированный граф (рисунок 1). Многомерность графа заключается в том, что все вершины графа находятся в многомерном пространстве и имеют соответствующие координаты. Отметим, что у вершин графа могут быть специально выделенные атрибуты - «свойства», но сам граф остается, в простейшем случае, двудольным.

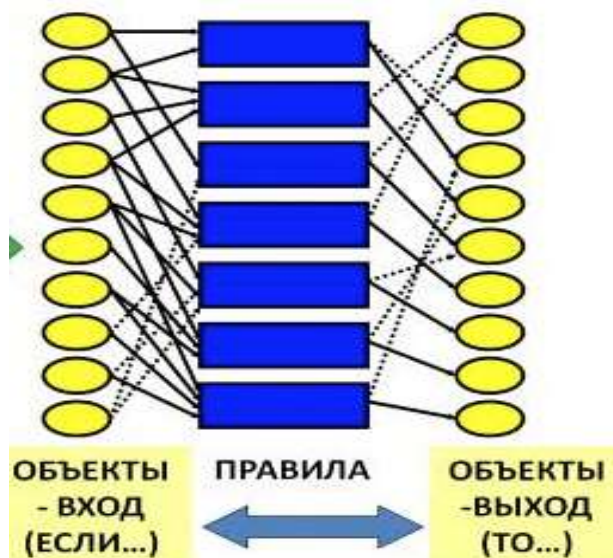


Рисунок 1 – Графическое представление двудольного ориентированного графа

Теория графов обладает устоявшейся терминологией, поэтому определение понятия «граф» звучит так: граф - это совокупность непустого множества вершин и

наборов пар вершин (связей между вершинами). Как известно из данной теории, объекты представляются как вершины, или узлы графа, а связи – как дуги, или ребра.

В миварной базе знаний концепты являются вершинами или узлами графа, а связи между концептами - дугами и ребрами. Так как миварная база знаний основана на модели ВСО («вещь-свойство-отношение»), то:

- «вещь» в данном графе принято обозначать в форме овала (вершина графа);
- «отношение» - прямоугольника (связи вершин – дуги графа);
- «свойство» - треугольника (атрибуты вершин графа).

Пример графического представление сети концептов в миварной базе знаний представлен на рисунке 2. За основу примера взято краткое содержание детской сказки А.Н. Толстого про зайца:

«В лесу жил заяц. Он хвастался перед другими животными. Ворона узнала про зайца. Она хотела наказать зайца. Заяц спас ворону. Ворона простила зайца.»



Рисунок 2 – Графическое представление сети концептов в миварной базе знаний

Необходимо отметить, что сети концептов основываются на множестве связей между концептами. Связи – отношения между концептами, в свою очередь имеют разную природу, поэтому их можно классифицировать:

1. «СИС (словосочетание, имеющее смысл)»

Связь между концептами, которая существует в реальном мире.

Пример: «кресло директора», «рубить дрова», «мама работает».

2. «Общее - частное»

Связь между концептами, имеющими единую природу, в которой частные концепты содержат в своем составе смысл(понятие) главного концепта.

Пример: общий концепт – «фрукт»

частный концепт – «яблоко», «банан», «апельсин».

3. *«Часть - целое»*

Связь между концептами, в которой частные концепты являются частью главного концепта.

Пример: концепт-целое – «яблоко»

концепт-частное – «семена», «кожура», «мякоть».

4. *«Свойство»*

Связь между концептами, в которой смысл частного концепта отражает свойства главного концепта.

Пример: главный концепт – «фрукт»

частный концепт – «вкус», «форма», «цвет», «масса», «запах».

Таким образом, зная вид отношения между концептами, то есть вид связи, можно вводить различные типы связей для усложнения используемого контекста и построения более адекватных моделей для понимания текстов на естественном русском языке.

5. Заключение

Теоретически обоснована возможность применения миварных технологий для создания интеллектуальных систем, включая и системы понимания текстов на естественном русском языке. Информационно-вычислительные системы с интеллектуальной поддержкой, как правило, применяемые для решения сложных задач, где необходима логическая обработка информации, также подразумевают использование миварных технологий.

Миварные технологии накопления данных и обработки информации в едином миварном информационном пространстве позволяют эволюционно наращивать базу знаний. Поэтому необходимо четко и грамотно представлять данные, хранящиеся в миварных базах знаний.

Комплексная сеть концептов в миварной базе знаний позволяет совершать большое количество операций над данными за минимальное время. Поэтому формирование комплексной сети концептов необходимо для обработки в реальном времени сверхбольших объемов данных и правил.

Решение такой задачи, как формирование комплексной сети концептов в миварной базе знаний, позволит перейти на новый качественный уровень автоматизированной обработки текстов на основе логической обработки больших массивов данных и учета контекста.

Список информационных источников

- [1] Лорьер Ж.-Л. Системы искусственного интеллекта: Пер. с франц.-М.: Мир, 1991.- 568с.
- [2] Гушин А.Н., Радченко И.А. Экспертные системы: учебное пособие // Балт. гос. техн. ун-т.- СПб., 2007 . С. 3-5.

- [3] Шамис А.Л. Поведение, восприятие, мышление: проблемы создания искусственного интеллекта. — М.: Едиториал УРСС, 2005.
- [4] Варламов О.О., Адамова Л.Е., Петерсон А.О., Протопопова Д.А., Скакунова Е.А. Исследование подходов и основных проблем понимания естественного русского языка // Автоматизация и управление в технических системах. – 2014. – № 2; URL: auts.esrae.ru/10-196 (дата обращения: 15.09.2014).
- [5] Varlamov O.O., Adamova L.E.E., Eliseev D.V., Mayboroda Yu.I., Antonov P.D., Sergushin G.S., Chibirova M.O. Mivar Technologies in Mathematical Modeling of Natural Language, Images and Human Speech Understanding // International Journal of Advanced Studies. 2013. Т. 3. № 3. С. 17-23.
- [6] 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(4); С. 30-45. [Электронный ресурс]. URL: auts.esrae.ru/4-80 (дата обращения: 26.05.2013).
- [7] Варламов О.О. Обзор 25 лет развития миварного подхода к разработке интеллектуальных систем и создания искусственного интеллекта // Труды НИИР. 2011. № 1. С. 34-44.
- [8] Подкосова Я.Г., Варламов О.О., Остроух А.В., Краснянский М.Н. Анализ перспектив использования технологий виртуальной реальности в дистанционном обучении // Вопросы современной науки и практики. Университет им. В.И. Вернадского. 2011. № 2. С. 104-111.
- [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] 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.
- [11] Владимиров А.Н., Варламов О.О., Носов А.В., Потапова Т.С. Программный комплекс "УДАВ": практическая реализация активного обучаемого логического вывода с линейной вычислительной сложностью на основе миварной сети правил // Труды НИИР. 2010. Т. 1. С. 108-116.
- [12] Варламов О.О. Разработка адаптивного механизма логического вывода на эволюционной интерактивной сети гиперправил с мультиактивизаторами, управляемой потоком данных // Искусственный интеллект. 2002. № 3. С. 363-370.
- [13] Варламов О.О. Основы многомерного информационного развивающегося (миварного) пространства представления данных и правил // Информационные технологии, 2003. № 5. С. 42-47.
- [14] Варламов О.О. Разработка метода распараллеливания потокового множественного доступа к общей базе данных в условиях недопущения взаимного искажения данных // Информационные технологии. 2003. №1. С. 20-28.
- [15] Варламов О.О. Параллельная обработка потоков информации на основе виртуальных потоковых баз данных // Известия высших учебных заведений. Электроника. 2003. № 5. С. 82-89.
- [16] Варламов О.О. Системный анализ и синтез моделей данных и методы обработки информации для создания самоорганизующихся комплексов оперативной диагностики // Искусственный интеллект. 2003. № 3. С. 299-305.

- [17] Варламов О.О. Системы обработки информации и взаимодействие групп мобильных роботов на основе миварного информационного пространства // Искусственный интеллект. 2004. № 4. С. 695-700.
- [18] Варламов О.О. Создание интеллектуальных систем на основе взаимодействия миварного информационного пространства и сервисно-ориентированной архитектуры // Искусственный интеллект. 2005. № 3. С. 13-17.
- [19] Варламов О.О. Анализ взаимосвязей GRID и САС ИВК, SOA и миварного подхода // Искусственный интеллект. 2005. № 4. С. 4-11.
- [20] Варламов О.О. О необходимости перехода от теории искусственного интеллекта к разработке теории активного отражения // Известия Южного федерального университета. Технические науки. 2007. Т. 77. № 2. С. 89-95.
- [21] Варламов О.О. Миварные технологии: переход от продукции к двудольным миварным сетям и практическая реализация автоматического конструктора алгоритмов, управляемого потоком входных данных и обрабатывающего более трех миллионов продукционных правил // Искусственный интеллект. 2012. № 4. С. 11-33.
- [22] 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.
- [23] Варламов О.О., Чибирова М.О., Сергушин Г.С., Елисеев Д.В. "Облачная" реализация миварного универсального решателя задач на основе адаптивного активного логического вывода с линейной сложностью относительно правил "Если-То-Иначе" // Автоматизация и управление в технических системах. – 2013. – № 2. С. 7-23.
- [24] Сергушин Г.С., Варламов О.О., Чибирова М.О., Елисеев Д.В., Муравьева Е.А. Исследование возможностей информационного моделирования сложных систем управления технологическими процессами на основе миварных технологий // Автоматизация и управление в технических системах. – 2013. – № 2. С. 46-60.
- [25] 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.
- [26] Чибирова М.О., Сергушин Г.С., Варламов О.О., Елисеев Д.В., Хадиев А.М. и др. Реализация общедоступного миварного универсального решателя задач на основе адаптивного активного логического вывода с линейной сложностью и облачных технологий // Искусственный интеллект. – 2013. – № 3. – С. 512-523.
- [27] Белоусова А.И., Варламов О.О., Остроух А.В., Краснянский М.Н. Подход к формированию многоуровневой модели мультиагентной системы с использованием миваров // Перспективы науки. 2011. № 20. С. 57-61.
- [28] 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.
- [29] Варламов О.О. Миварный подход к разработке интеллектуальных систем и проект создания мультипредметной активной миварной интернет-энциклопедии // Известия Кабардино-Балкарского научного центра РАН. 2011. № 1. С. 55-64.
- [30] Варламов О.О. Эволюционные базы данных и знаний. Миварное информационное пространство // Известия Южного федерального университета. Технические науки. 2007. Т. 77. № 2. С. 77-81.