

МИНИСТЕРСТВО ОБРАЗОВАНИЯ И НАУКИ РОССИЙСКОЙ ФЕДЕРАЦИИ
Федеральное государственное бюджетное образовательное учреждение
высшего образования

«КУБАНСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ»
(ФГБОУ ВО «КубГУ»)

Физико-технический факультет

Кафедра теоретической физики и компьютерных технологий

КУРСОВОЙ ПРОЕКТ

МЕТОДЫ МАНИПУЛИРОВАНИЯ ЗНАНИЯМИ

Работу выполнил  Локтев Евгений Сергеевич

Курс 2

Направление 09.03.02 Информационные системы и технологии

Научный руководитель

канд. физ.-мат. наук, доцент  Ю. Г. Никитин

Нормоконтролер инженер  Г. Д. Цой

Краснодар 2017

СОДЕРЖАНИЕ

ВВЕДЕНИЕ	3
1 Основные понятия	4
1.1 Проблемная область искусственного интеллекта	4
1.2 Данные, информация и знания.....	6
2 Основная часть	6
2 1 Инженерия знаний	7
2 2 Получение знаний.....	8
2 3 Обобщение и классификация знаний	12
2 4 Нейронные сети	14
2 5 Экспертные системы	15
2 6 KQML.....	18
ЗАКЛЮЧЕНИЕ.....	18
СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ	19

ВВЕДЕНИЕ

Известно и очевидно, что одной из наивысших и наипочетнейших человеческих квалификаций является работа со знаниями: получение новых и манипулирование имеющимися. Еще совсем недавно считалось, что лишь естественный интеллект способен на такое. Ничто, созданное людьми, не удостоивалось чести считаться весомым элементом Вселенского мыслительного процесса.

Однако сейчас, когда темпы развития технологий растут по экспоненте, определенно стоит пересмотреть некоторые мысли, касающиеся искусственного интеллекта и высказанные в эпоху перфокарт.

Разумеется, я также говорю и о писателях-фантастах, упоминания о которых прошу мне простить – хотя бы за то, что всё больше и больше их фантазий становятся ныне реальностью

Актуальность данного проекта несоизмеримо высока, ибо вычислительные мощности машин неуклонно растут, закон Гордона Мура, гласящий, что число транзисторов на кристалле интегральной схемы, удваивается каждые 2 года, по-прежнему исправно выполняется, а океан данных и знаний давно уж стал столь огромен, что всякий естественный разум однозначно утонет и окажется раздавлен толщей информации. Никогда прежде проблемы работы со знаниями не были так актуальны, как в наши дни.

Целью данного проекта считаю обзор основных методов манипулирования знаниями, рассмотрение некоторых систем искусственного интеллекта, использующих в процессе своей деятельности данные методы, а также представление наработок в данной области, осуществленных зарубежными специалистами и не переведенных на русский язык.

Тема наша большая и объемная, а посему начинать, я считаю, необходимо с самых основ.

1 Основные понятия

Ни в коем случае нельзя приступать к данной работе, не уяснив некоторых крайне важных, фундаментальных понятий, коими мы будем оперировать.

В первую очередь, необходимо дать определение искусственному интеллекту. В 1989 году такое определение предложили А. Барр и Е. Фейгенбаум [1]: «Искусственный интеллект – это область информатики, которая занимается разработкой интеллектуальных компьютерных систем, т.е. систем, обладающих возможностями, которые мы традиционно связываем с человеческим разумом: понимание языка, обучение, способность рассуждать, решать проблемы и т.д.».

Цель искусственного интеллекта – понять функционирование человеческого разума, легко управляющего огромными объемами элементарных частичек информации и решающего три основные проблемы: «сформулировать – запомнить – применить». Задачи искусственного интеллекта имеют две характерные особенности: информация представляется в символической форме и предполагается наличие выбора между некоторыми вариантами, причем порой в условиях неопределенности.

Каковы же элементы искусственного интеллекта, используемые людьми для решения каких-либо задач? Само собой напрашивается слово «система». Так и есть.

Интеллектуальная система – это информационно-вычислительная система, при решении задач применяющая интеллектуальную поддержку и не требующая оператора.

Интеллектуализированная система же отлична от интеллектуальной лишь в одном пункте – ей необходим оператор.

1.1 Проблемная область искусственного интеллекта

Примечательно, что в данном разделе можно отождествить понятия «проблемы» и «направления развития». Что ж, перечислим их:

- представление знаний;
- манипулирование знаниями;
- общение;
- восприятие;
- обучение;
- поведение.

Как нетрудно догадаться, в рамках данного курсового проекта особое внимание будет уделено второму пункту – манипулированию знаниями.

Для того, чтобы при решении задач умело манипулировать знаниями, искусственный интеллект должен уметь:

- оперировать имеющимися знаниями;
- пополнять знания;
- классифицировать имеющиеся в системе знания;
- обобщать по тем или иным разработанным процедурам знания;
- формировать абстрактные понятия;
- проводить достоверный вывод на основе имеющихся знаний;
- пользоваться моделями рассуждений.

Кстати говоря, манипулирование знаниями и представление оных – два очень тесно связанных друг с другом направления, это продемонстрировано на рисунке 1. Создающаяся в настоящее время теория баз знаний включает исследования, относящиеся к обоим направлениям.

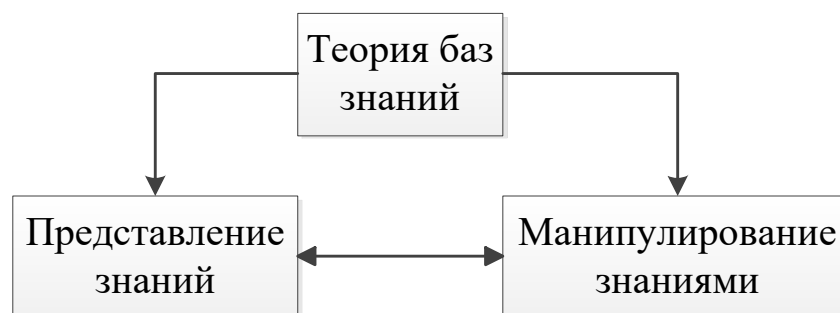


Рисунок 1 – Связь представления знаний и манипулирования знаниями

1.2 Данные, информация и знания

Все мы неоднократно слышали эти слова: данные, информация и знания. Теперь настало время раз и навсегда разграничить эти понятия.

Данные – это некоторые сведения, зафиксированные неким носителем и представленные в форме, которая пригодна для манипулирования.

Информация – это данные, имеющие смысл в конкретном, данном контексте либо для конкретного приемника информации. Следует отметить, субъективность разграничения, то есть то, что для одних приемников является данными, для иных вполне может считаться информацией.

Знания – это информация, непременно проверенная практикой, которая может быть сколь угодно раз использована кем-либо или чем-либо для решения некоторых задач.

Теперь, когда мы отчетливо видим разницу, мы можем свободно оперировать этими понятиями. Однако нужно уяснить еще кое-что. База знаний – это совокупность знаний внутри конкретной предметной области и представленных так, чтобы на их основе возможно было осуществление рассуждений.

2 Основная часть

Теперь, когда мы имеем некоторое понимание фундаментальных основ систем искусственного интеллекта, а также теории знаний, можно приступить к рассмотрению конкретных систем, наук, языков, протоколов и технологий

2.1 Инженерия знаний

Основные трудности в работе со знаниями связаны с проблемой извлечения и структурирования оных. Как раз эти вопросы изучает наука, называемая инженерия знаний (knowledge engineering). Это весьма молодое направление искусственного интеллекта, родившееся, когда практические разработчики столкнулись с проблемами трудности добычи и формализации знаний. В первых трудах по искусственному интеллекту эти факты лишь постулировались, однако довольно скоро все осознали, что необходимы серьезные исследования в этой области.

Инженерия знаний – направление разработок и исследований в области искусственного интеллекта, цель которого – разработка методов, моделей и систем для получения, структурирования и формализации знаний с целью проектирования баз знаний.

Направления исследований данной науки представлены на рис. 2

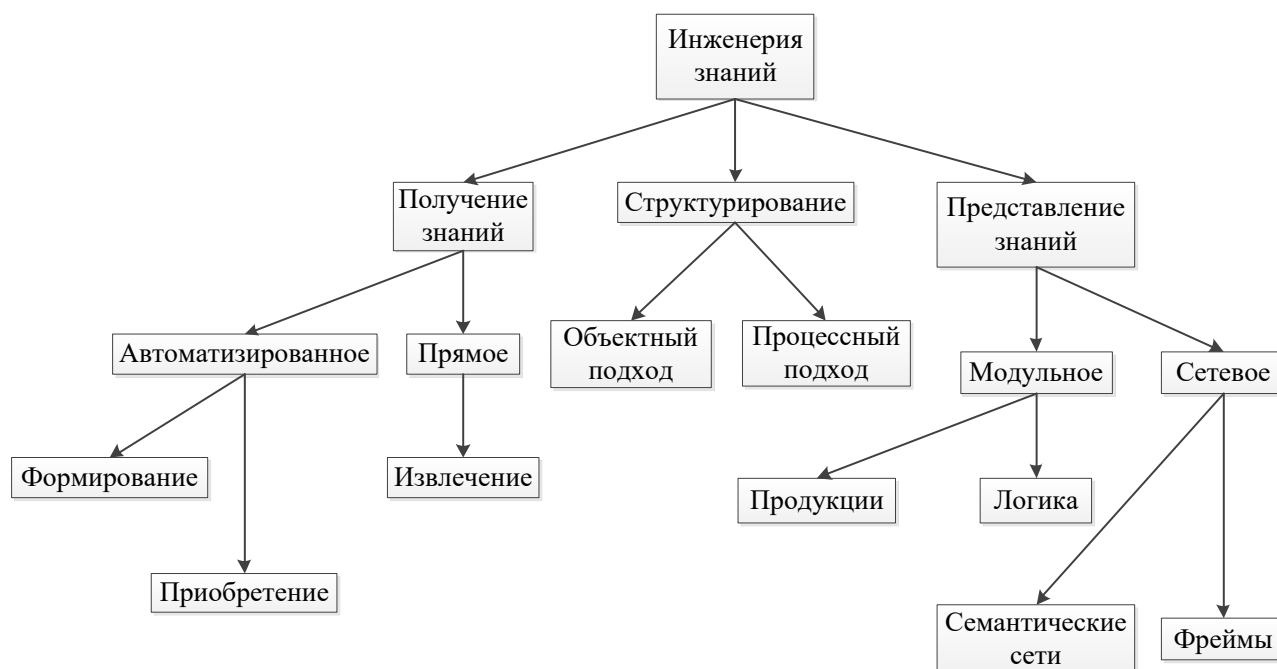


Рисунок 2 - Структура инженерии знаний

Инженерия знаний – довольно обширная, хоть и совсем молодая, наука.

2 2 Получение знаний

Для того, чтобы знаниями манипулировать, прежде всего их необходимо заполучить.

Приобретение знаний – это выявление знаний из источников и преобразование их в нужную форму.

Знания могут быть объективизированными (различные книги, документы) и субъективизированными (экспертные или эмпирические знания).

Объективизированные знания непременно должны быть приведены к форме, доступной для потребителя. Экспертные знания имеются у специалистов и не зафиксированы во внешних по отношению к ним хранилищах. Эмпирические знания, которые, как и экспертные, являются субъективными, получаются путём наблюдения за окружающей средой.

Ввод в базу знаний объективизированных знаний не представляет никаких проблем, а вот выявление и ввод субъективных и, в частности, экспертных знаний, довольно трудны. Представим схему приобретения знаний, она изображена на рисунке 3.



Рисунок 3 - Схема приобретения знаний

Под посредником обычно понимается человек, обладающий специфическими знаниями и называющийся когнитологом или инженером по знаниям. Посредник обязательно должен обладать знаниями о конкретной формализуемой модели.

Также в роли посредника может выступать специальная программа, однако когнитолог имеет ряд преимуществ перед ней, в частности универсальность по отношению к моделям знаний и предметным областям. Но с другой стороны, программа дает возможность специалисту-предметнику напрямую проектировать модель. Таким образом, можно сформулировать еще одно определение приобретения знаний.

Приобретение знаний – это процесс получения знаний от эксперта или каких-либо других источников и передача их в некоторую модель. В англоязычной литературе тут и там встречаются три термина: acquisition (приобретение), elicitation (извлечение) и discovery (формирование). Считаю важным разграничить эти понятия меж собой.

Если процесс получения данных от источников осуществляется без применения компьютерных средств поддержки этого процесса через непосредственный контакт когнитолога и источника знаний, то речь идет об извлечении знаний. Если же использование специальных программных средств поддержки имеет место быть, то уместно говорить о приобретении знаний. И третье: если процесс получения знаний от источников осуществляется с использованием программ обучения при наличии репрезентативной (т.е. достаточно представительной) выборки примеров принятия решений, то уместно говорить о формировании знаний.

Наиболее широкую трактовку имеет термин «приобретение», уделим ему больше внимания. Приобретение знаний может характеризоваться следующими аспектами:

- фазы приобретения знаний;
- модели приобретения знаний;
- номенклатура приобретения знаний;

- уровни знаний;
- средства отладки знаний.

Первые два аспекта носят скорее методический и технологический характер, третий и четвертый – теоретический, а последний же аспект носит специфический характер, определяемый инженерными особенностями разработки той или иной системы. Рассмотрим подробнее фазы и модели приобретения знаний.

Несмотря на существование различных точек зрения на число фаз в процессе приобретения знаний, большинство исследователей выделяют три фазы, отражающие изменение функций участников проектирования систем на данном этапе.

Предварительная фаза – это фаза непосредственно извлечения знаний из источника знаний на домашних этапах разработки (идентификация проблемы, получение знаний, структурирование, формализация)

Начальная фаза – это фаза приобретения знаний, которая осуществляется на этапе реализации прототипа интеллектуальной или интеллектуализированной системы и обеспечивает наполнение системы знаниями об области экспертизы

Фаза наполнения – это фаза, выполняемая в основном на этапах реализации и тестирования и связанная с решением таких задач, как:

- обнаружение неполноты, неточности или противоречивости знаний;
- извлечение новых знаний, устраняющих обнаруженные «нефакторы»;
- преобразование новых знаний в понятный вид.

Иными словами, в широком смысле под приобретением знаний понимаются все вышеперечисленные фазы, а в узком – лишь фаза накопления, когда происходит непосредственная передача знаний в действующий прототип интеллектуальной или интеллектуализированной системы, находящийся на этапах реализации и тестирования.

Рассмотрим самую сложную фазу – фазу извлечения знаний, процессы которой в отличие от процессов иных фаз не поддаются формализации и осуществляются вручную, оставаясь больше искусством, нежели наукой. При

всем этом разработчикам таких систем (инженерам по знаниям) приходится практически самостоятельно разрабатывать методы извлечения знаний в каждом конкретном случае, сталкиваясь со следующими трудностями:

-неудачный способ извлечения знаний, не совпадающий со структурой знаний в данной области;

-значительное упрощение «картины мира» эксперта;

-неумение наладить контакт с экспертом;

-терминологические проблемы;

-отсутствие целостной системы знаний;

-неадекватная модель представления знаний;

-неувязки организационного, финансового и др. характера.

Для того, чтобы разобраться в природе извлечения знаний, предлагается выделить три основных аспекта этой процедуры:

Психологический, в рамках которого рассматривают три слоя психологических проблем, возникающих при извлечении знаний:

-контактный

-процедурный

-когнитивный

Лингвистический, в рамках которого рассматривают слои важных для инженерии знаний проблем (проблема «общего кода», понятийная структура)

Гносеологический, в рамках которого рассматривают отдельные закономерности процессов отражения действительности в сознании человека (описание и обобщение фактов, установление связей, построение модели).

Классификация методов извлечения знаний представлена на рисунке 4

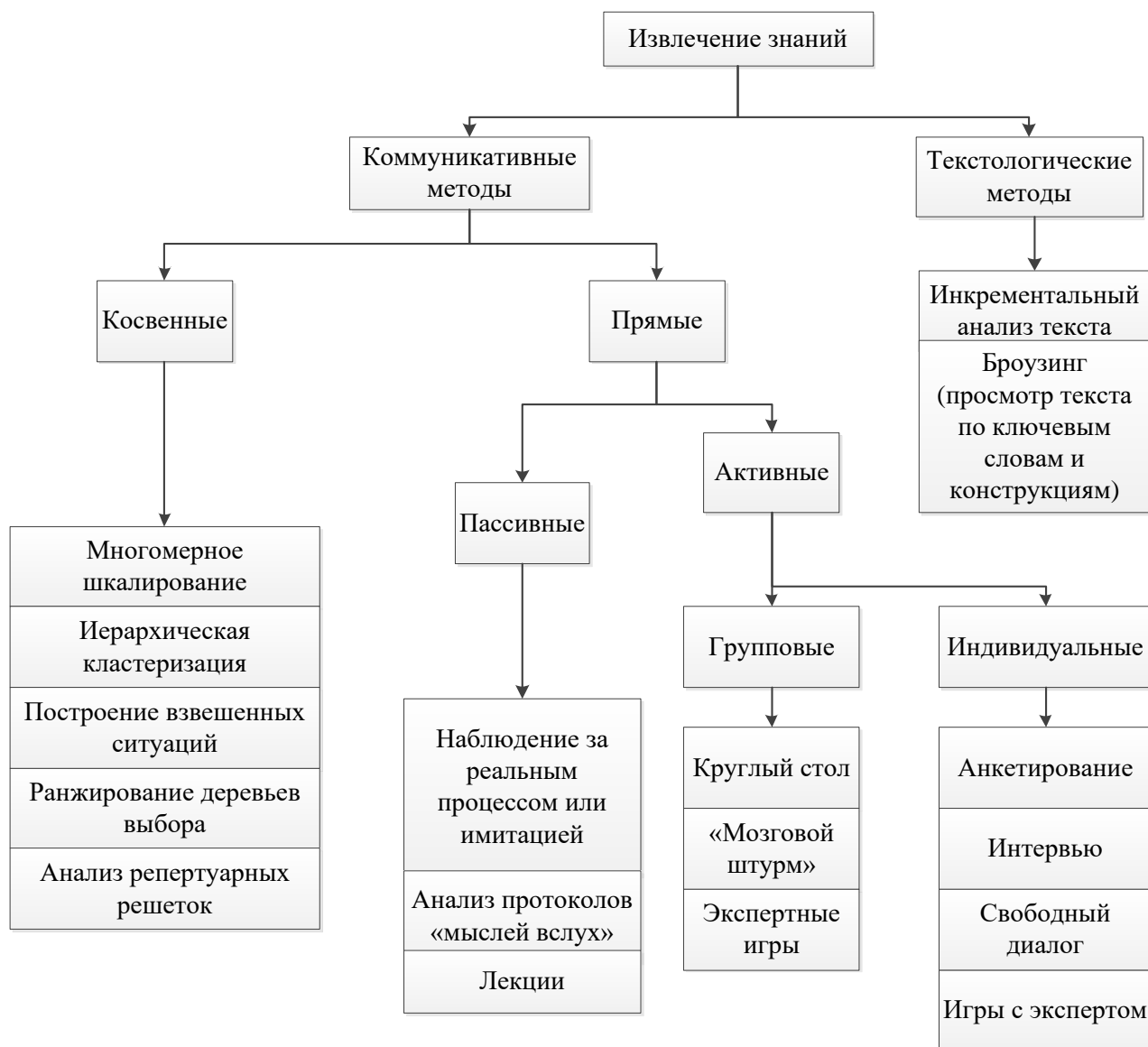


Рисунок 4 - Классификация методов извлечения знаний

Теперь рассмотрим приобретение знаний в широком смысле (т.е. учитываются все три фазы приобретения), что в общем случае предполагает выполнение следующей последовательности задач:

- определение необходимости модификации знаний системы;
- извлечение знаний в случае необходимости такой модификации;
- преобразование новых знаний в «понятную» системе форму;
- модификация знаний системы и переход к первой задаче.

2 3 Обобщение и классификация знаний

В основе любого научного исследований лежит способность естественного интеллекта к обобщению. Ведь именно в результате обобщения и рождаются новые знания, т.е. знания, не следующие непосредственно из прежде известных. В моделирующих мышление системах обобщение понимают как процесс получения знаний, способный объяснять имеющиеся знания и классифицировать или предсказывать новые.

Схема обобщения знаний представлена на рисунке 5.

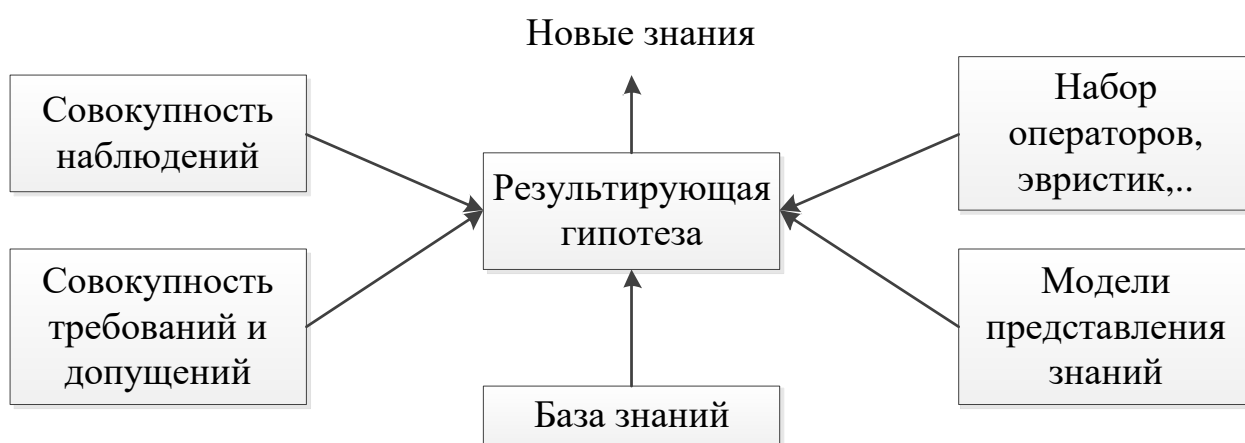


Рисунок 5 - Схема обобщения знаний

Классификация знаний – это одна из простейших регулярных структур-взаимосвязей однотипных понятий. Её значение, в первую очередь, определяется тем, что она задает на множестве рассматриваемых понятий однородную структуру (семантическую сеть), которая носит глобальный характер, однако лишь в рамках конкретной предметной области.

Как логическое средство целостного описания некоторой части реального мира, классификация играет фундаментальную роль, так как предшествует этапу анализа более тонких и частных связей между понятиями предметной области, которые и должны быть выявлены при решении конкретных прикладных проблем. Классификация представляет системным аналитикам и разработчикам функциональных задач интеллектуализированной информационной системы целостную совокупность понятий, которые выполняют роль естественных

координат для описания функциональных задач предметной области и тем самым позволяют ограничиться рассмотрением только допустимых классов сущностей без потери информации. Кроме того, увеличение степени абстракции, достигаемое при переходах меж уровней классификационной схемы, позволяет серьезно повысить выразительность спецификации предметной области, обеспечивая более ясное и, кстати, более сжатое представление информации.

Классификационные системы, используемые в большинстве прикладных наук, биологии, геологии, медицине и т.д., представляют собой не что иное, как иерархию обобщений.

Итак, классификация – это выделение на основе существенных признаков из некоторого множества понятий универсального класса всех входящих в него подмножеств и установление между выделенными подмножествами отношения порядка. Признаки, на основе коих производится выделение данных подмножеств, называются классификационными.

В классификационных системах класс сходных сущностей называют классификационным таксоном, а способ членения этих самых сущностей на отдельные части, позволяющий установить их сходства – мерономией. Таким образом, таксон – это объем некоторого класса, а мерономия – содержание понятия, связываемого с данным классом.

Упорядоченную совокупность признаков, характеризующих некоторый таксон с точки зрения внутренней структуры входящих в него сущностей, называется архетипом. Архетип, собственно говоря, это некоторая внутренняя структура, которую можно найти во всех сущностях соответствующего таксона.

Очевидно, сколь тесно связаны меж собой таксономия и мерономия.

2 4 Нейронные сети

Одним из наиболее активно развиваемых сегодня направлений в искусственном интеллекте являются нейронные сети – модель, представляющая

собой в некотором смысле гибрид критериальных, вероятностных и логических методов. Развитие теории началось более полувека назад и идёт параллельно исследованиям функционирования головного мозга.

Итак, для чего же нужны нейронные сети? Нейронные сети применяются для решения комплексных задач, требующих сложных вычислений, аналогичных совершаемым человеческим мозгом. В числе основных применений нейронных сетей классификация, предсказание и распознавание.

Ввиду специфики принципов работы нейронных систем, для освещения которых, кстати, потребовался бы не один том печатного текста, с определенным кругом задач их использовать логично, рационально и весьма разумно, для иных же они непригодны.

2.5 Экспертные системы

Итак, мы добрались до одного из самых интересных направлений искусственного интеллекта. Прежде, мы давали определение базам знаний, и именно в теории экспертных систем оно чаще всего встречается. Но что же такое экспертная система?

Экспертная система – это комплекс программного обеспечения, помогающий оператору принимать обоснованные решения. Экспертные системы работают с информацией, полученной от экспертов – людей, являющихся специалистами в области, в которой, собственно говоря, конкретная экспертная система работает.

Появление таких систем было весьма логичной вехой развития в истории человечества. Очевидно, что науки ныне стали очень большими и неповоротливыми, и едва ли какая голова сможет вместить в себя хоть одну, но целиком. Да, мы разбиваем науки на составляющие и готовим узких специалистов, но тем самым обостряем то условие, что все важные открытия делаются на стыке областей. Человечеству было необходимо направление,

исследующее и разрабатывающее программы и устройства, которые используют знания и процедуры вывода для решения задач, трудных для людей-экспертов.

Об экспертных системах можно рассказывать очень и очень долго, в рамках данного проекта мы заострим внимание на том, как данные системы манипулируют знаниями. Для примера, на рисунке 6 изображена схема решения задачи экспертной системой в режиме консультации.

Итак, экспертная система состоит из следующих основных компонентов: решателя, рабочей памяти (базы данных), базы знаний, компонентов приобретения знаний, объяснительного и диалогового.

База данных подобна оперативной памяти компьютера, она предназначена для хранения исходных и промежуточных данных решаемой в данный момент задачи.

База знаний же аналогична постоянным запоминающим устройства, она хранит долгосрочные данные, описывающие рассматриваемую область и правила, описывающие целесообразные преобразования данных.

Решатель, комбинируя знания из баз данных и знаний, формирует последовательность правил, приводящих к решению задачи.

Компонента приобретения знаний автоматизирует процесс наполнения экспертной системы знаниями, который осуществляется пользователем-экспертом.

Объяснительная компонента объясняет, каким образом система получила или не получила решение исходной задачи, что она при этом использовала.

Диалоговая компонента направлена на дружелюбное общение со всеми категориями пользователей на всех этапах работы экспертной системы.



Рисунок 6 – Решение задачи экспертной системой в режиме консультации

Следует также указать, какие типы задач выполняет экспертная система.

Интерпретация символов или сигналов – это составление смыслового описания по входным данным.

Предсказание – это определение последствий наблюдаемых ситуаций.

Диагностика – определение состояния неисправностей по известным признакам.

Конструирование – разработка объекта с заданными свойствами при соблюдении установленных ограничений.

Планирование – определение последовательности действий, приводящих к желаемому состоянию объекта.

Слежение – наблюдение за изменяющимся состоянием объекта и сравнение его показателей с установленными или желаемыми.

Управление – воздействие на объект для достижения желаемого поведения.

2.6 KQML

То, как различные виды интеллектуальных и интеллектуализированных систем устроены и как манипулируют знаниями – это, несомненно, важно и интересно, однако давайте вернемся к фундаментальным основам, знание которых необходимо, независимо от того, какая разновидность искусственного интеллекта избрана.

В числе таких основ находится язык KQML - Knowledge Query and Manipulation Language, что переводится с английского как «язык управления знаниями и запросов к знаниям».

Данный язык и одноименный протокол для обмена сообщениями между программными агентами и системами был создан в рамках одного из проектов небезызвестного агентства DARPA. Первоначально, проектировался интерфейс к системам, основанным на знаниях, однако позже было решено, что главной целью языка и протокола KQML должно быть взаимодействие между агентами.

ЗАКЛЮЧЕНИЕ

В рамках выполнения данного курсового проекта был проведен обзор основных аспектов манипулирования знаниями, рассмотрены некоторые системы, работающие со знаниями, а также рассказано о языке KQML и

одноименном протоколе. Выполнен перевод информации, обнаруженной в англоязычных источниках.

Очевидно, сколь юна и слабо развита данная тема, и как разнятся понятия и классификации у разных авторов. Отсутствие общепринятых норм в первую очередь указывает на острую актуальность изучения и проведения исследований в рамках данной темы.

СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ

1. Barr A. Handbook of Artificial Intelligence / E. Feigenbaum – Los Altos, CA: William Kaufman, 1989. – 442 с

2. Костенко К. И. Формализмы представления знаний и модели интеллектуальных систем: [Учебное пособие] / К. И. Костенко. – Краснодар, 2015. – 300с.
3. Гаскаров Д. В. Интеллектуальные информационные системы / Д. В. Гаскаров. – М.: Высш. Шк., 2003. – 431 с.
4. Осовский С. Нейронные сети для обработки информации / Пер. с польского И. Д. Руднинского. – М.: Финансы и статистика, 2004. – 344 с.
5. Введение в искусственный интеллект: конспект лекций / Д. В. Смолин, Е. Н. Глебова, Е. В. Третьяков, А. А. Логунов. – М.: ФИЗМАТЛИТ, 2004. – 208 с.
6. Частиков А. П. Разработка экспертных систем. Среда CLIPS / А. П. Частиков, Т. А. Гаврилова, Д. Л. Белов. – СПб.: БХВ-Петербург, 2003. – 608 с.
7. Системы искусственного интеллекта. Практический курс / В. А. Чулюков, И. Ф. Астахова, А. С. Потапов и др. – М.: БИНОМ. Лаборатория знаний, 2008. – 292 с.
8. Девятков В. В. Системы искусственного интеллекта: Учебное пособие для вузов / В. В. Девятков, Г. В. Росс, Е. И. Пупырев. – М.: Изд-во МГТУ им. Н.Э.Баумана, 2001. – 352 с.