Эта часть работы выложена в ознакомительных целях. Если вы хотите получить работу полностью, то приобретите ее воспользовавшись формой заказа на странице с готовой работой:
https://stuservis.ru/magisterskaya-rabota/77573

Тип работы: Магистерская работа

Предмет: Программирование

ВВЕДЕНИЕ 3
1 ПРИНЦИПЫ ОРГАНИЗАЦИИ ЗНАНИЙ 4
1.1 Понятие искусственного интеллекта. Направления и подходы искусственного интеллекта 4
1.2 Интеллектуальный анализ данных 9
1.3 Технологические подходы к обработке знаний 11
1.4 Концепция Semantic Web. Уровни представления Semantic Web 13
2 МАТЕМАТИЧЕСКИЕ МОДЕЛИ ФОРМАЛИЗАЦИИ ЗНАНИЙ 19
2.1 Методы представления и извлечений знаний 19
2.2 Методы представления и хранения семантической информации 24
2.3 Веб-технологии поиска и обработки семантической информации 35
3 РАЗРАБОТКА ПРОГРАММНОЙ РЕАЛИЗАЦИИ 46
3.1 Выбор языка и технологии программирования 46
3.2 Описание основных компонентов приложения 50
3.3 Порядок запуска и работы приложения 52
ЗАКЛЮЧЕНИЕ 57
СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ 58
ПРИЛОЖЕНИЕ 61

ВВЕДЕНИЕ

Развитие технологии World Wide Web является одним из ключевых направлений компьютерных и сетевых технологий. Можно говорить о существенно определяющем характере влияния технологии Web на состояние информационного общества.
Развитие информационной инфраструктуры связано с ростом масштаба компьютерных сетей и объема передаваемых в них данных. В этих условиях механическое увеличение мощности инфраструктуры сообразно решаемым задачам может дать эффект не во всех случаях. В этой связи на первый план выходят вопросы, связанные как с поиском новых типов сетевых решений, так и с возможностью изменения качественной структуры распространяемого в них контента. Так, введение интеллектуальных механизмов распространения, хранения и обмена данными во многом повысит эффективность применения уже существующих решений.
Так, среди них можно выделить нейросетевые алгоритмы маршрутизации.
Целью работы является анализ существующих методов формализации информации, применяемых в условиях сети Интернет. Для ее реализации необходимо решить следующие задачи:
• получить представление о интеллектуальном анализе данных;
• проанализировать методы представления данных и знаний;
• проанализировать методику представления семантической информации и запросов к ней в условиях сети Интернет;
• выполнить программную реализацию семантического анализа информации с возможностью ввода данных и просмотра результатов.

1 ПРИНЦИПЫ ОРГАНИЗАЦИИ ЗНАНИЙ

1.1 Понятие искусственного интеллекта. Направления и подходы искусственного интеллекта

Термин «искусственный интеллект» (artificial intelligence) предложен в 1956 году на семинаре с аналогичным названием, который состоялся в США и был посвящен решению логических задач.
Искусственный интеллект можно определить как научную дисциплину, в задачи которой входит исследование разумного поведения. Это определение, однако, не конкретизирует понятие разумности.

Формализация последнего представляет собой довольно сложную задачу. Она в существенной степени определяется исходными данными и спецификой предметной области. В этой связи можно говорить о моделировании паттернов принятия решений, свойственных человеку.
Искусственный интеллект (ИИ) - совокупность научных дисциплин, изучающих методы решения задач интеллектуального (творческого) характера с использованием ЭВМ.
Искусственный интеллект - одно из направлений информатики, целью которого является разработка аппаратно-программных средств, позволяющих пользователю-непрограммисту ставить и решать свои, традиционно считающиеся интеллектуальными задачи, общаясь с ЭВМ на ограниченном подмножестве естественного языка.
Конструктивность искусственного интеллекта состоит в потенциальной возможности исследования проблем для теорий интеллекта. Модели, заимствованные из этих теорий, можно математически формализовать, после чего оценить их возможности с применением ЭВМ.
Можно выделить следующие направления исследований искусственного интеллекта (рисунок 1.1).

Рисунок 1.1. Направления исследования искусственного интеллекта
К сфере, связанных с моделированием результатов интеллектуальной деятельности можно отнести широкий круг задач. Так, к ним можно отнести задачи, связанные с поиском оптимального решения, распознаванием образов, логические задачи и т. п. При этом найденные алгоритмы решения задач должны работать на существующей аппаратной базе ЭВМ. Данное направление можно отнести к программно-прагматическому [1,2].
Моделирование билогических систем можно рассматривать как бионическое направление. Оно специализируется на исследовании структур и процессов, которые характерны для поведенческой деятельности биологических систем. Так, одним из его составляющих элементов является моделирование нейронных сетей и систем. На его базе наука нейроинформатика, играющая важную роль в развитии нейрокомпьютерных систем [4,5].
В настоящее время традиционным (классическим) принято считать программное прагматическое направление, при котором не ставится вопрос об адекватности используемых структур и методов тем, которыми пользуется в аналогичных случаях человек, а рассматривается лишь конечный результат решения задачи. В рамках этого направления сначала велись поиски моделей и алгоритма человеческого мышления. Ни одна из существующих наук (философия, психология, лингвистика) не смогла предложить такого алгоритма. Тогда специалисты в области искусственного интеллекта модель лабиринтного поиска. Этот подход представляет задачу как некоторый граф, отражающий пространство состояний, и в этом графе проводится поиск оптимального пути от входящих данных к результирующим;
Эвристика – правило, теоретически не обоснованное, но позволяющее сократить количество переборов в пространстве поиска; использование методов математической логики. На основе метода резолюций, позволившего автоматически доказывать теоремы при наличии исходных аксиом, в 1973 г. был создан язык Пролог.
Можно выделить следующие методы, применяемые в моделях искусственного интеллекта [2,3]:
• Символьный подход.
• Логический подход.
• Агентно-ориентированный подход.
• Гибридный подход.
Символьный подход к ИИ представляет собой круг методов интеллектуальной обработки информации, которые допускают ее формализацию в классе символических представлений. Последние, в свою очередь, описывают предметную и семантическую составляющую исходной информации. Это, в свою очередь, позволяет установить между отдельными символами связи логического характера. Таким образом, символические вычисления дают возможность получения результата на основе манипулирования правилами формализованной предметной области. Это, в свою очередь, может приводить к созданию новых правил преобразования данных, основанных на композиции старых.
В качестве примера применения символьного подхода можно привести решение задач, связанных с применением методов символьной математики. Одним из первых языков, поддерживающих такие возможности, был язык Lisp.
К недостаткам символьного подхода можно отнести достаточно ограниченные возможности интерпретации результатов его применения. Это, в свою очередь, обусловлено сложностью формализации предметной области.

В основе логического подхода лежит попытка формализации мышления на уровне сознания и рассуждений. В качестве теоретического аппарата в этом случае используются методы математической логики и теории предикатов.
Логический подход может быть в полной мере реализован в языке логического программирования Пролог. Написанные на нем программы можно рассматривать как наборы фактов и предикатов. В последних задаются правила и порядок выполнения логического вывода. При этом конкретный алгоритм вывода не задается.
Достоинствами рассмотренного подхода являются относительная прозрачность процесса логического вывода и легкость его естественно-языковой интерпретации. Также само поведение процесса логического вывода однозначно определяется конфигурацией исходных данных.
К недостаткам логического подхода следует отнести сложность формализации предметной области в условиях неопределенности. Наличие последней может быть связано с нечеткостью и недостаточностью знаний, отсутствием конечной цели, интервальностью или вероятностным характером данных и т.п. Также процесс принятия решения в условиях применения логических систем достаточно сложно распараллеливать.
Агентно-ориентированный подход является одним из наиболее перспективных подходов, применяемых в настоящее время. Он опирается на использовании интеллектуальных (рациональных) агентов. Концепция была предложена Шохемом (англ. Yoav Shoham) в 1990 г.
В основе концепции агентно-ориентированного подхода лежат концепции понятия агента и его поведения, которое может носить ментальный характер. Структура взаимодействия агента и среды приведена на рисунке 1.2.

Рисунок 1.2. Модель взаимодействия агента с окружающей средой
Динамика агента определяется свойствами и влиянием среды, в которой он находится. Конечным итогом функционирования агента является решение некоторой задачи или достижение цели в условиях среды, в которой он находится.
Интеллект в контексте агента можно рассматривать как элемент планирования его поведенческой стратегии. Связь между агентом и окружающей средой можно реализовать через систему датчиков, считывающих реакцию среды, и с помощью исполнительных механизмов, которые могут влиять на ее состояние.
В рамках агентно-ориентированной методики приоритет отдается изучению методов и алгоритмов, которые обеспечивают агенту минимизировать риски при выполнении его задачи в условиях внешней среды.
Гибридный подход предполагает интеграцию разнородных моделей искусственного интеллекта. В нем утверждается, что объединение возможностей подходов на основе различных парадигм может дать синергетический эффект, которых приведет к максимизации когнитивных и вычислительных возможностей системы.
Гибридный подход можно проиллюстрировать на следующем примере. Например, в некоторой системе правила умозаключений на экспертном уровне могут генерироваться нейронными сетями. Порождающие правила, в свою очередь, можно синтезировать на основе механизма статистического вывода. Таким образом, можно ожидать, что интеллектуальные системы на основе гибридного подхода могут обладать большими когнитивными возможностями по сравнению с традиционно применяемыми решениями.

1.2 Интеллектуальный анализ данных

Под интеллектуальным анализом данных (ИАД) следует понимать анализ предварительно обработанных данных. Он может носить достаточно многообразный характер. При этом могут использоваться различные математические методы, относящиеся машинному обучению. К ним можно отнести методы, основанные на статистике, генетические алгоритмы, деревья решений и т. д. Другую группу методов образуют нейросетевые модели, различнве подходы на базе искусственного интеллекта и классификации данных.
Целью применение анализа такого уровня сложности является извлечение новой информации из данных, отражающей глубинные связи между их элементами [7,8].
Методы ИАД позволяют решить многие задачи, с которыми сталкивается аналитик.
В общем случае процесс ИАД состоит из трёх стадий [6,7]:
1) выявление закономерностей (свободный поиск);
2) использование выявленных закономерностей для предсказания неизвестных значений (прогностическое моделирование);

3) анализ исключений, предназначенный для выявления и толкования аномалий в найденных закономерностях.
Иногда в явном виде выделяют промежуточную стадию проверки достоверности найденных закономерностей между их нахождением и использованием (стадия валидации).
Все методы ИАД подразделяются на две большие группы по принципу работы с исходными обучающими данными. В первом случае исходные данные могут храниться в явном детализированном виде и непосредственно использоваться для прогностического моделирования и/или анализа исключений; это так называемые методы рассуждений на основе анализа прецедентов. Главной проблемой этой группы методов является затрудненность их использования на больших объемах данных, хотя именно при анализе больших хранилищ данных методы ИАД приносят наибольшую пользу.
Во втором случае информация вначале извлекается из первичных данных и преобразуется в некоторые формальные конструкции (их вид зависит от конкретного метода). Согласно предыдущей классификации, этот этап выполняется на стадии свободного поиска, которая у методов первой группы в принципе отсутствует. Таким образом, для прогностического моделирования и анализа исключений используются результаты этой стадии, которые гораздо более компактны, чем сами массивы исходных данных. При этом полученные конструкции могут быть либо "прозрачными" (интерпретируемыми), либо "черными ящиками" (нетрактуемыми). Прежде чем использовать технологию Data Mining, необходимо тщательно проанализировать ее проблемы, ограничения и критические вопросы, с ней связанные, а также понять, чего эта технология не может.
Интеллектуальный анализ данных (ИАД) является междисциплинарной областью науки, возникшей и развивающейся на базе достижений прикладной статистики, распознавания образов, искусственного интеллекта, теории баз данных и других областей современной науки. Схема междисциплинарных связей в DataMining представлена на рисунке 1.3 [8].

Рисунок 1.3. Междисциплинарные связи Data Mining
ИАД представляет собой спектр технологий и методик выявления статистических взаимосвязей, закономерностей, правил в наборах данных. На сегодняшний день технология ИАД содержит множество различных подходов к обнаружению знаний. Каждый из них имеет свои преимущества и недостатки. При этом выбор конкретного подхода определяется спецификой предметной области и организацией данных.

1.3 Технологические подходы к обработке знаний

Важной стороной технологического аспекта развития искусственного интеллекта является развитие интеллектуальных систем. Интеллектуальные системы настроены на решение специализированных задач, принадлежащих определенной предметной области.
Для ИС характерны следующие признаки [1-3]:
• развитые диалоговые и коммуникационные возможности;
• умение решать слабо формализованные задачи;
• возможность самоорганизации и самообучения;
• адаптивность и гибкость.
Каждому из перечисленных признаков условно соответствует свой класс ИС. Различные системы могут обладать одним или несколькими признаками интеллектуальности с различной степенью проявления.
Характерным признаком интеллектуальных систем является наличие знаний, необходимых для решения задач конкретной предметной области. При этом возникает естественный вопрос, что такое знания и чем они отличаются от обычных данных, обрабатываемых ЭВМ [9-11].
Типичным представителем систем искусственного интеллекта являются экспертные системы. Экспертную систему (ЭС) от других автоматизированных систем на этапе ее использования отличают высокая интеллектуальность, специализация и ориентация на решение задач в определенной области.
В ЭС выполняется процесс решения определенного класса задач, которые носят интеллектуальный характер. Качество результатов при этом должно быть сопоставимо с результатами, получаемыми квалифицированным экспертом.
В составе ЭС можно выделить следующие компоненты:
• базу знаний, которая предназначена для хранения информации, описывающей предметную область, и правил, описывающих целесообразные преобразования данных этой области;
• интерпретатор, который формирует последовательность правил, обеспечивающих решение задачи;

• компоненту извлечения знаний, отвечающая за формализацию процесса наполнения ЭС знаниями, получаемыми от эксперта.
Пользовательский интерфейс обеспечивает взаимодейств

СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ

1 Джонс М. Т. Программирование искусственного интеллекта в приложениях. Пер. с англ. Осипов А. И. — М.: ДМК Пресс, 2006. — 312 с.
2 Макконелл Дж. Основы современных алгоритмов: Учеб. пособие – М.: Техносфера, 2004. — 366 с.
3 Рассел С. Искусственный интеллект. Современный подход – М.: Вильямс, 2007. — 1410 с.
4 Д. Рутковская, М. Пилиньский, Л. Рутковский. Нейронные сети, генетические алгоритмы и нечёткие системы – М.: Горячая линия - Телеком, 2006. — 452 с.
5 В. В. Емельянов, В. М. Курейчик, В. В. Курейчик. Теория и практика эволюционного моделирования – М.: ФИЗМАТЛИТ, 2003. — 431 с.
6 Барсегян А. А.Технологии анализа данных: Data Mining, Visual Mining, Text Mining, OLAP: учеб. Пособие / А. А. Барсегян, М. С. Куприянов, В. В. Степаненко, И. И. Холод. – СПб.: БХВ-Петербург, 2007.
7 Барсегян А. А. Анализ данных и процессов: учеб. пособие. – СПб.: БХВ-Петербург, 2009.
8 Чубукова И. А. Data Mining.: учеб. пособие. – М.: БИНОМ. Лаборатория знаний, 2010.
9 Гущин А.Н., Радченко И.А. Экспертные системы: учебное пособие / А.Н. Гущин, И.А. Радченко; Балт. гос. техн. ун-т. — СПб., 2007
10 Поспелов Д. А. Моделирование рассуждений. Опыт анализа мыслительных актов.— М.: Радио и связь, 1989.—184 с.
11 Д. А. Гольдин, А. М. Чесноков, Экспертные системы и продукционные правила в интеллектуальных комплексах распределения ресурсов автономных систем электроснабжения // УБС, 2008, выпуск 20, С. 95–106
12 О. М. Васильев, Д. П. Ветров, Д. А. Кропотов, Представление и обнаружение знаний в экспертных системах для задач распознавания образов, Ж. вычисл. матем. и матем. физ., 2007, том 47, номер 8, С. 1428–1454.
13 Васильев, В.В. Практикум по Web-технологиям. Гриф УМО вузов России / В.В. Васильев. - М.: Форум, 2015. - 275 с.
14 Дронов, Владимир PHP, MySQL и Dreamweaver MX 2004. Разработка интерактивных Web-сайтов / Владимир Дронов. - М.: "БХВ-Петербург", 2016. - 448 с.
15 Евсеев, Д. А. Web-дизайн в примерах и задачах / Д.А. Евсеев, В.В. Трофимов. - М.: КноРус, 2015. - 272 с.
16 Якобсон, Й. Концепция разработки Web-сайтов. Как успешно разработать Web-сайт с применением мультимедиа-технологий / Й. Якобсон. - М.: НТ Пресс, 2006. - 496 с.
17 Semantic Web Challenge. - http://challenge.semanticweb.org/.
18 Рогушина Ю.В., Гришанова І.Ю. Средства интеллектуализации поиска информационных ресурсов в сети Интернет // Тез. VI Междунар. конф. "Интеллектуальный анализ информации ИАИ-2007", 2007. – С. 322–331.
19 Rogushina J., Gladun A. Ontological Approach to Domain Knowledge Representation for Informational Retrieval in Multiagent Systems // International Jornal "Information Theories & Applications", 2006. – V.13, N.4. – P. 354–362.
20 Искусственный интеллект. Междисциплинарный подход: моногр. ; ИИнтелл - М., 2013. - 448 с.
21 Калиниченко Л. А., Рывкин В. М. Машины баз данных и знаний; Главная редакция физико-математической литературы издательства "Наука" - М., 2011. - 296 с.
22 Кейтер Дж. Компьютеры - синтезаторы речи; Мир - М., 2012. - 238 с.
23 Коваль С. А. Лингвистические проблемы компьютерной морфологии; Издательство Санкт-Петербургского университета - М., 2016. - 152 с.
24 Ковальски Р. Логика в решении проблем: моногр. ; Главная редакция физико-математической литературы издательства "Наука" - М., 2013. - 280 с.
25 Крейг Джон Введение в робототехнику. Механика и управление; Институт компьютерных исследований - М., 2013. - 564 с.
26 Курцвейл Рэй Эволюция разума; Эксмо - М., 2015. - 352 с.
27 Макаров И. М., Лохин В. М., Манько С. В., Романов М. П. Искусственный интеллект и интеллектуальные системы управления; Наука - М., 2012. - 336 с.
28 Ж. Бланшет, М. Саммерфилд «QT 4: программирование GUI на C++», КУДИЦ-Пресс, 2008
29 М. Шлее «Qt 5.3. Профессиональное программирование на C++», БХВ-Петербург, 2015

30 Мэтиз, Э. Изучаем PYTHON. Программирование игр, визуализация данных, веб-приложения / Э. Мэтиз. СПб.: Питер, 2017. – 496 с.
31 Лутц, М. Программирование на Python, I том / М. Лутц. СПб.: Символ-плюс, 2015. – 992 с.
32 Саммерфилд, М. Программирование на Python 3. Подробное руководство / М. Саммерфилд. СПб.: Символ-плюс, 2015. – 608 с.

Эта часть работы выложена в ознакомительных целях. Если вы хотите получить работу полностью, то приобретите ее воспользовавшись формой заказа на странице с готовой работой:
https://stuservis.ru/magisterskaya-rabota/77573