

ТУБ 18 ИНТЕЛЛЕКТУАЛЬНЫЕ СИСТЕМЫ

18. Понятие и структура интеллектуальной системы

Развитием систем распознавания явились *интеллектуальные системы* (ИС). Они воплотили в себе наукоемкие технологии с высоким уровнем автоматизации подготовки информации для принятия решений, а также автоматизации процессов нахождения вариантов решений, опирающихся на полученные данные.

ИС можно рассматривать как технические и программные системы, ориентированные на решение большого класса задач, называемых неформализованными. Система становится интеллектуальной, если в ней данные заменяются на знания, а алгоритмы функционирования заменяются на методы искусственного интеллекта.

Отличительными особенностями ИС являются:

- интерфейс с пользователем на естественном языке с применением бизнес-понятий, характерных для прикладной области пользователя;
- способность объяснять свои действия и помогать пользователю в работе с системой;
- представление модели объекта в виде базы знаний в сочетании с возможностью работы с неполной или неточной информацией;
- способность автоматического обнаружения закономерностей бизнеса в ранее накопленных фактах и включения их в базу знаний.

ИС особенно эффективны в применении к слабо структурированным задачам, в которых пока отсутствует строгая формализация, и для решения которых применяются эвристические процедуры, позволяющие в большинстве случаев получить решение. Отчасти этим объясняется необычайно широкий диапазон применения ИС.

На рисунке 1 приведена структура интеллектуальной системы.

Блок «интерфейс пользователя» предназначен для связи компьютера с пользователем, для которого предпочтительным языком «общения» является естественный язык или близкий к нему. Понятие «естественный язык» фактически является искусственным языком, полученным из естественного языка путем удаления из него неоднозначностей. Такой язык в теории автоматизированных систем носит название «искусственный язык» или «информационный язык». Основу ИС составляют блоки: «База данных» (БД), «База правил», «Машина логического вывода». БД хранит исходные данные. В базе правил фиксируются знания и опыт эксперта. Машина логического вывода выводит результат, взаимодействуя с БД и базой правил. Все три блока должны быть описаны математически. Правила записываются через интерфейс эксперта.

По запросу пользователя компьютер с помощью блока «Объяснение» может дать ответ на вопрос КАК получен результат и ПОЧЕМУ компьютер

Одной из главных проблем обработки знаний, а также создания базы знаний и ИС в целом является проблема представления знаний. Это задача представления взаимосвязей в конкретной предметной области в форме, понятной системе искусственного интеллекта. Представление знаний – это их формализация и структурирование, с помощью которых отражаются характерные признаки знаний: внутренняя интерпретируемость, структурированность, связность, семантическая метрика и активность. Кроме того, представление знаний – это процесс описания знаний человека о проблемной области посредством выражений на формальном языке, называемом языком представления знаний.

При проектировании модели представления знаний следует учитывать такие факторы, как однородность представления и простота понимания. Однородность представления приводит к упрощению механизма управления логическим выводом и знаниями. Простота понимания предполагает доступность понимания представления знаний и экспертом, и пользователем системы. В противном случае затрудняется приобретение знаний и их оценка.

Способ представления знаний определяет, каким образом знания описываются в памяти компьютера, а также каковы возможности базы знаний. Для того чтобы ЭВМ имела возможность манипулирования знаниями о проблемной области, они должны быть представлены в виде модели.

Модель представления знаний – это способ и результат формального описания знаний в базе знаний. Она должна быть понятной пользователю и обеспечивать однородность представления знаний, за счет чего упрощаются управление знаниями и логический вывод.

Популярными моделями представления знаний являются искусственная нейронная и семантическая сети. Поскольку искусственные нейронные сети были рассмотрены ранее, остановимся на семантических сетях.

Семантические сети позволяют выделять смысл текста в виде понятий и связей между ними, образующих граф. Понятия семантической сети записываются в вершинах графа, а отношения между понятиями – это дуги графа. Количество типов отношений в семантической сети определяется ее разработчиком исходя из конкретных целей. Часто используются иерархические семантические сети, в которых отношения образуют древовидную структуру. Также отношения в сетях могут быть разных типов: функциональные, количественные, пространственные, временные, логические и др.

К достоинствам семантических сетей можно отнести следующее:

- универсальность, достигаемая за счет выбора соответствующего набора отношений;
- наглядность системы знаний, представленной графически;
- близость структуры сети, представляющей систему знаний, семантической структуре фраз на естественном языке;
- соответствие современным представлениям об организации долговременной памяти человека.

Недостатки семантических сетей:

- сетевая модель не всегда дает ясное представление о структуре предметной области, поэтому формирование и модификация такой модели могут быть затруднительными;
- сетевые модели представляют собой структуры, для обработки которых необходим специальный аппарат формального вывода;
- проблема поиска решения в семантической сети сводится к задаче поиска ее фрагмента, отражающего поставленный запрос. Это может обуславливать сложность поиска решения в семантических сетях;
- представление, использование и модификация знаний при описании систем реального уровня сложности оказывается трудоемкой процедурой, особенно при наличии множественных отношений между ее понятиями.

Семантические сети используются в системах понимания естественного языка, в вопросно-ответных системах, в других различных предметно-ориентированных системах. Чаще других востребованы сети смешанного типа, т. е. содержащие, в зависимости от области применения, самые разные типы отношений.

18.3 Технологии искусственного интеллекта

В технологиях искусственного интеллекта можно выделить два основных направления их исследования: общетеоретическое, связанное с разработкой моделей и методов реализации отдельных аспектов получения и преобразования знаний; прикладное, нацеленное на разработку комплексных технологий конечных пользователей (рисунок 2).

В рамках первого направления разрабатываются следующие технологии:

1. *Технологии извлечения и представления знаний.* В первом случае основной задачей является разработка методов: формального описания поисковых образов и их признаков, формализации, распознавания и сравнения образов, получения знаний от экспертов. Во втором – решаются задачи, связанные с формализацией знаний для их представления в памяти ИС. Решение этих двух задач позволяет разработчикам комплексных технологий получить ответы на три важных вопроса: какие знания необходимо представлять в ИС, кто или что является источником этих знаний, какие методы и модели требуются для представления этих знаний в ИС.

2. *Технологии манипулирования знаниями.* Решение интеллектуальных задач предполагает не только представление знаний в ИС, но и их обработку, т.е. необходимо научить ИС оперировать ими. Поэтому здесь изучаются вопросы пополнения знаний на основе их неполных описаний, классификации знаний в ИС, разрабатываются процедуры и методы обобщения знаний, достоверного вывода и др.

3. *Технология общения.* Переход к ИС знаменует новую технологию общения конечных пользователей с ЭВМ и требует решения таких проблем, как понимание связных текстов на ограниченном и неограниченном

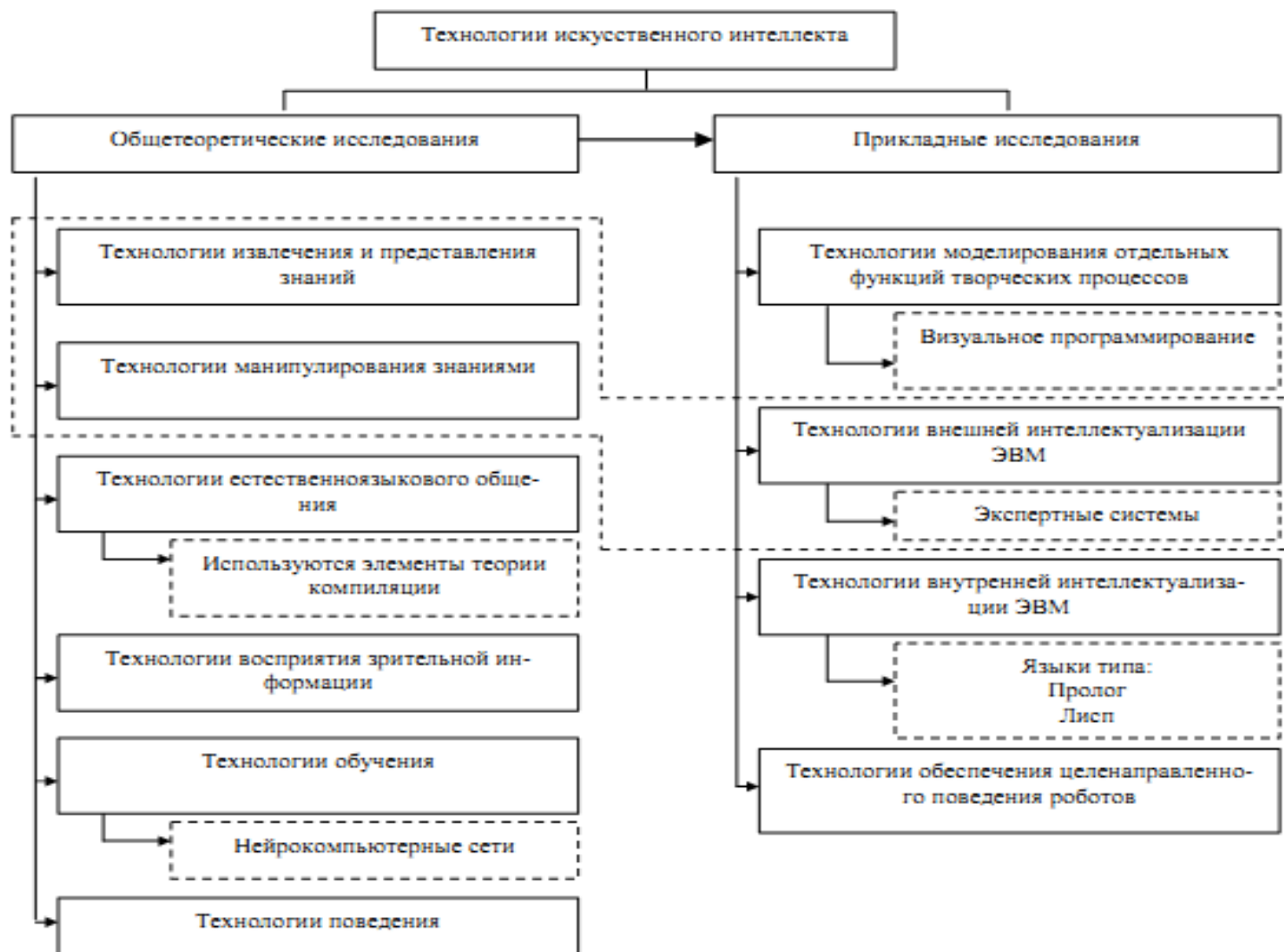


Рисунок 2 – Технологии искусственного интеллекта

естественном языке, понимание речи и ее синтез, разработка коммуникативных моделей «пользователь-ЭВМ», формирование объяснений и т.п. Главная цель данных исследований – обеспечение комфортных условий для общения человека и ИС.

4. *Технологии восприятия.* Разработка этих технологий предполагает создание методов: анализа трехмерных сцен, представления информации о зрительных образах в базе знаний ИС, трансформации зрительных сцен в текстовые описания и обратно, а также разработку процедур когнитивной графики и др.

5. *Технологии обучения.* Отличительной особенностью ИС должна стать их способность решать задачи, в явном виде не представленные в базе знаний, что требует наделения ИС способностью к обучению. Для этих целей необходимо: создать методы формирования условий задачи по описанию проблемной ситуации или по наблюдению за этой ситуацией; обеспечить переход от известного решения частных задач к решению общей задачи; наделить ИС способностью декомпозировать исходную задачу на более мелкие, решение которых известно; разработать нормативные и декларативные модели самого процесса обучения.

6. *Технологии поведения.* Взаимодействие ИС со средой требует разработки специальных поведенческих процедур, которые бы позволили им адекватно реагировать на те или иные изменения в среде. Такое взаимодействие предполагает создание моделей целесообразного, нормативного и ситуативного поведения, а также разработку методов многоуровневого планирования и коррекции планов в динамических ситуациях.

Второе направление, по существу, связано с использованием результатов общетеоретических исследований на практике. Практические исследования в области технологий искусственного интеллекта ведутся в следующих аспектах:

1. *Технологии моделирования на ЭВМ отдельных функций, реализуемых человеком.* Это – игровые задачи, автоматическое доказательство теорем, автоматический синтез программ, анализ и синтез элементов творчества человека.

2. *Технологии внешней интеллектуализации ЭВМ.* Исследования в данном направлении привели к созданию: интеллектуальных информационно-поисковых систем, интеллектуальных пакетов прикладных программ и расчетно-логических систем, систем поддержки принятия решений, экспертных систем, а также созданию комплексного диалогового интерфейса.

3. *Технологии внутренней интеллектуализации ЭВМ.* Они направлены на разработку ЭВМ новой архитектуры, ориентированных на обработку символьной информации. К ним относятся программы на языках ЛИСП и ПРОЛОГ, лингвистические процессоры, машины баз данных и машины баз знаний.

4. *Технологии обеспечения целенаправленного поведения роботов.* Здесь создаются ЭВМ для специальных механических и энергетических систем. Работы в данном направлении особенно важны, т.к. существенно расширяют возможности человеческой деятельности в опасных условиях, где присутствие человека невозможно или нецелесообразно.

8.3 Классификация интеллектуальных систем

Рассмотрим наиболее популярные виды интеллектуальных систем с точки зрения их особенностей и прикладного применения.

Системы на естественном языке специфичны и предназначены преимущественно для таких целей, как машинный перевод, генерация документов, автоматическое аннотирование и реферирование.

Экспертные системы предполагают высокую степень формализации процессов. Разновидностью экспертных систем можно считать расчетно-логические системы, оперирующие с функциями вместо правил.

Интеллектуальные пакеты в настоящее время стали оболочками экспертных систем. Их еще называют системами без базы правил.

Искусственные нейронные сети представляют собой фактически разновидность систем автоматического управления, использующие свойства нейрона.

Генетические алгоритмы относятся к разновидности эволюционных эвристических методов.

Мультиагентные системы преследуют цель согласования теорий баз данных и баз знаний.

ИС чаще всего оперируют с дискретными величинами, но все больше необходима связь дискретных и непрерывных величин, вследствие чего появился класс гибридных моделей. Существует множество вариантов реализации гибридных систем. На роль гибридных систем претендуют искусственные нейронные сети, системы на основе генетических алгоритмов, экспертные системы реального времени.

8.4 Знания в интеллектуальной системе

Знания – это совокупность сведений о сущностях реального мира, их свойствах и отношениях между ними в определенной предметной области. Другими словами, знания – это выявленные закономерности предметной области, позволяющие решать задачи в этой области. С точки зрения искусственного интеллекта знания можно определить как формализованную информацию, на которую ссылаются в процессе логического вывода.

В этом случае, под проблемной областью понимается область человеческих знаний, в терминах которой формулируются задачи и в рамках которой они решаются. Проблемная область представляется описанием части

реального мира, которое в силу своей приближенности рассматривается как ее информационная модель.

Проблемная область – это содержательное описание в терминах проблемной области проблемы совместно с комплексом условий, факторов и обстоятельств, вызвавших ее возникновение.

Наиболее общее определение трактует знание как всю совокупность данных, необходимую для решения задачи. В этом определении подчеркивается, что данные в привычном понимании также являются знаниями. Однако знания в информационном плане не ограничиваются рамками данных. В полном объеме информация, содержащаяся в знаниях, должна включать сведения о: системе понятий предметной области, в которой решаются задачи; системе понятий формальных моделей, на основе которых решаются задачи; соответствии систем понятий, упомянутых выше; методах решения задачи; текущем состоянии предметной области.

Несмотря на сложности определения считается общепризнанным, что знания имеют ряд свойств, позволяющих отличить их от данных. Это – внутренняя интерпретируемость, внутренняя структурированность, внешняя взаимосвязь структурных единиц, шкалирование, погружение в пространство с семантической метрикой, активность. Если данные обладают этими свойствами, можно говорить о перерастании их в знания.

Внутренняя интерпретируемость означает наличие в памяти ЭВМ сведений не только о значении, но и о наименовании информационной единицы. Это свойство присуще некоторым моделям представления данных, например, реляционной.

Внутренняя структурированность отражает вложенность одних информационных единиц в другие или в самих себя.

Внешняя взаимосвязь единиц определяет, с какой информационной единицей имеет связь данная информационная единица и какова эта связь.

Шкалирование означает использование шкал, предназначенных для фиксации соотношения различных величин. Прежде всего шкалирование необходимо для фиксации соотношений качественной информации.

Погружение в пространство с семантической метрикой используется для задания меры близости информационных единиц.

Активность знаний выражается в возможности вызова той или иной процедуры в зависимости от структуры, сложившейся между информационными единицами.

Перечисленные особенности информационных единиц определяют ту грань, за которой данные превращаются в знания, а базы данных перерастают в базу знаний. Совокупность средств, обеспечивающих работу со знаниями, образуют систему управления базой знаний.

Базы данных фиксируют экстенциональную семантику заданной проблемной области, состояние конкретных объектов, конкретные значения параметров для определенных моментов времени и временных интервалов.

База знаний определяет интенциональную семантику моделей и содержит описание абстрактных сущностей: объектов, отношений, процессов.

Таким образом, при использовании знаний происходит переход к формуле: *знания + вывод = система*.

Работа со знаниями лежит в основе работы с интеллектуальными системами. В свою очередь обработка знаний включает в себя:

- *извлечение знаний из источников*, кроме человека;
- *приобретение знаний от профессионалов* (экспертов);
- *представление знаний*, т.е. их формализация, позволяющая в дальнейшем использовать знания для проведения логического вывода на ЭВМ;
- *манипулирование знаниями*, включающее пополнение, классификацию, обобщение знаний и вывод на знаниях;
- *объяснение на знаниях*, позволяющее дать ответ, как и почему сделан тот или иной вывод.

Различают два типа знаний: экстенциональные и интенциональные.

Экстенциональные знания – это набор количественных и качественных характеристик различных конкретных объектов. Они представляются перечислениями объектов предметной области, экземпляров объектов, свойств объектов. Иными словами, экстенциональные знания – это данные, хранящиеся в базах данных. Их еще называют предметными или фактографическими знаниями.

Интенциональные знания – это совокупность основных терминов, применяемых в проблемной области, и правил над ними, позволяющих получать новые знания. Интенциональные знания описывают абстрактные объекты, события, отношения.

Интенциональные знания подразделяются на декларативные, процедуральные и метазнания.

Декларативные знания отражают понятия проблемной области и связи между ними. Они не содержат в явном виде описания каких-либо процедур. Эти знания еще называются понятийными или концептуальными.

Процедуральные знания описывают процедуры, т.е. указывают операции над понятиями, позволяющие получать новые понятия. В отличие от декларативных знаний они содержат в явном виде описания процедур. Примером процедуральных знаний является программа, хранящаяся в памяти ЭВМ. Иногда процедуральные знания называются алгоритмическими.

Метазнания – это знания об организации всех остальных типов знаний. Иначе они называются специальными. Метазнания содержат признаки декларативных и процедуральных знаний.

Поверхностные – знания о видимых взаимосвязях между отдельными событиями и фактами в предметной области.

Глубинные – абстракции, аналогии, схемы, отображающие структуру и природу процессов, протекающих в предметной области.

Концепция баз знаний является результатом естественного научно-технического развития концепции БД в направлении увеличения независимости содержащейся информации от программ и их целостности. База знаний – это целостная и независимая модель проблемной области.