

## **Лекция 1.**

### **Раздел 1. История развития систем искусственного интеллекта (СИИ). Основные определения и понятия**

**Тема 1. История искусственного интеллекта, терминология, происхождение и развитие основных направлений.**

**Тема 2. Границы применимости искусственных интеллектуальных систем.**

**Тема 3. Отличие данных от знаний. Проблема представления.**

#### **План лекции.**

1. Введение.
2. История развития искусственного интеллекта. Основные термины и определения.
3. Происхождение и развитие основных направлений (подходов) к ИИ.
4. Границы применимости интеллектуальных систем.
5. Данные и знания. Проблема представления.

#### **Основная часть.**

С конца 40-х годов ученые все большего числа университетских и промышленных исследовательских лабораторий устремились к дерзкой цели; построение компьютеров, действующих таким образом, что по результатам работы их невозможно было бы отличить от человеческого разума. Терпеливо продвигаясь вперед в своем нелегком труде, исследователи, работающие в области искусственного интеллекта (ИИ), обнаружили, что вступили в схватку с весьма запутанными проблемами, далеко выходящими за пределы традиционной информатики. Оказалось, что, прежде всего, необходимо понять механизмы процесса обучения, природу языка и чувственного восприятия. Выяснилось, что для создания машин, имитирующих работу человеческого мозга, требуется разобраться в том, как действуют миллиарды его взаимосвязанных нейронов. И тогда многие исследователи пришли к выводу, что, пожалуй, самая трудная проблема, стоящая перед современной наукой - познание процессов функционирования человеческого разума, а не просто имитация его работы. Что непосредственно затрагивало фундаментальные теоретические проблемы психологической науки. В самом деле, ученым трудно даже прийти к единой точке зрения относительно самого предмета их исследований - интеллекта. Здесь как в притче о слепцах, пытавшихся описывать слона, пытаются придерживаться своего заветного определения. Некоторые считают, что интеллект - умение решать сложные задачи; другие рассматривают его как способность к обучению, обобщению и аналогиям; третьи - как возможность взаимодействия с внешним миром путем общения, восприятия и осознания воспринятого. Тем не менее, многие исследователи ИИ склонны принять тест машинного интеллекта, предложенный в начале 50-х годов выдающимся английским математиком и специалистом по вычислительной технике Аланом Тьюрингом. Компьютер можно считать

разумным, - утверждал Тьюринг, - если он способен заставить нас поверить, что мы имеем дело не с машиной, а с человеком.

### **Происхождение и развитие основных направлений (подходов) к ИИ:**

**Электронный подход.** После второй мировой войны появились устройства, казалось бы, подходящие для достижения заветной цели - моделирования разумного поведения; это были электронные цифровые вычислительные машины. "Электронный мозг", как тогда восторженно называли компьютер, поразил в 1952 г. телезрителей США, точно предсказав результаты президентских выборов за несколько часов до получения окончательных данных. Этот "подвиг" компьютера лишь подтвердил вывод, к которому в то время пришли многие ученые: наступит тот день, когда автоматические вычислители, столь быстро, неумолимо и безошибочно выполняющие автоматические действия, смогут имитировать невычислительные процессы, свойственные человеческому мышлению, в том числе восприятие и обучение, распознавание образов, понимание повседневной речи и письма, принятие решений в неопределенных ситуациях, когда известны не все факты. Таким образом "заочно" формулировался своего рода "социальный заказ" для психологии, стимулируя различные отрасли науки.

Многие изобретатели компьютеров и первые программисты составляли программы для отнюдь не технических занятий, как сочинение музыки, решение головоломок и игры, на первом месте здесь оказались шашки и шахматы. Некоторые романтически настроенные программисты даже заставляли свои машины писать любовные письма.

К концу 50-х годов все эти увлечения выделились в новую более или менее самостоятельную ветвь информатики, получившую название "искусственный интеллект". Исследования в области ИИ, первоначально сосредоточенные в нескольких университетских центрах США - Массачусетском технологическом институте, Технологическом институте Карнеги в Питтсбурге, Стэнфордском университете, - ныне ведутся во многих других университетах и корпорациях США и других стран. В общем исследователей ИИ, работающих над созданием мыслящих машин, можно разделить на две группы. Одних интересует чистая наука и для них компьютер - лишь инструмент, обеспечивающий возможность экспериментальной проверки теорий процессов мышления. Интересы другой группы лежат в области техники: они стремятся расширить сферу применения компьютеров и облегчить пользование ими. Многие представители второй группы мало заботятся о выяснении механизма мышления - они полагают, что для их работы это едва ли более полезно, чем изучение полета птиц и самолетостроения.

В настоящее время, однако, обнаружилось, что как научные, так и технические поиски столкнулись с несоизмеримо более серьезными трудностями, чем представлялось

первым энтузиастам. На первых порах многие пионеры ИИ верили, что через какой-нибудь десяток лет машины обретут высочайшие человеческие таланты. Предполагалось, что, преодолев период "электронного детства" и обучившись в библиотеках всего мира, хитроумные компьютеры, благодаря быстрдействию точности и безотказной памяти постепенно превзойдут своих создателей-людей. Сейчас мало кто говорит об этом, а если и говорит, то отнюдь не считает, что подобные чудеса не за горами.

На протяжении всей своей короткой истории исследователи в области ИИ всегда находились на переднем крае информатики. Многие ныне обычные разработки, в том числе усовершенствованные системы программирования, текстовые редакторы и программы распознавания образов, в значительной мере рассматриваются на работах по ИИ. Короче говоря, теории, новые идеи, и разработки ИИ неизменно привлекают внимание тех, кто стремится расширить области применения и возможности компьютеров, сделать их более "дружелюбными" то есть более похожими на разумных помощников и активных советчиков, чем те педантичные и туповатые электронные рабы, какими они всегда были.

Несмотря на многообещающие перспективы, ни одну из разработанных до сих пор программ ИИ нельзя назвать "разумной" в обычном понимании этого слова. Это объясняется тем, что все они узко специализированы; самые сложные экспертные системы по своим возможностям скорее напоминают дрессированных или механических кукол, нежели человека с его гибким умом и широким кругозором. Даже среди исследователей ИИ теперь многие сомневаются, что большинство подобных изделий принесет существенную пользу. Немало критиков ИИ считают, что такого рода ограничения вообще непреодолимы.

К числу таких скептиков относится и Хьюберт Дрейфус, профессор философии Калифорнийского университета в Беркли. С его точки зрения, истинный разум невозможно отделить от его человеческой основы, заключенной в человеческом организме. "Цифровой компьютер - не человек, - говорит Дрейфус. - У компьютера нет ни тела, ни эмоций, ни потребностей. Он лишен социальной ориентации, которая приобретается жизнью в обществе, а именно она делает поведение разумным. Я не хочу сказать, что компьютеры не могут быть разумными. Но цифровые компьютеры, запрограммированные фактами и правилами из нашей, человеческой, жизни, действительно не могут стать разумными. Поэтому ИИ в том виде, как мы его представляем, невозможен".

**Кибернетический подход.** Попытки построить машины, способные к разумному поведению, в значительной мере вдохновлены идеями профессора Норберта Винера, одной из выдающихся личностей в интеллектуальной истории Америки. Помимо математики он

обладал широкими познаниями в других областях, включая нейропсихологию, медицину, физику и электронику.

Винер был убежден, что наиболее перспективны научные исследования в так называемых пограничных областях, которые нельзя конкретно отнести к той или иной конкретной дисциплины. Они лежат где-то на стыке наук, поэтому к ним обычно не подходят столь строго. "Если затруднения в решении какой-либо проблемы психологии имеют математический характер, пояснял он, - то десять несведущих в математике психологов продвинутся не дальше одного столь же несведущего". Винеру и его сотруднику Джулиану Бигелоу принадлежит разработка принципа "обратной связи", который был успешно применен при разработке нового оружия с радиолокационным наведением. Принцип обратной связи заключается в использовании информации, поступающей из окружающего мира, для изменения поведения машины. В основу разработанных Винером и Бигелоу систем наведения были положены тонкие математические методы; при малейшем изменении отраженных от самолета радиолокационных сигналов они соответственно изменяли наводку орудий, то есть, заметив попытку отклонения самолета от курса, они тотчас рассчитывали его дальнейший путь и направляли орудия так, чтобы траектории снарядов и самолетов пересеклись. В дальнейшем Винер разработал на принципе обратной связи теории как машинного, так и человеческого разума. Он доказывал, что именно благодаря обратной связи все живое приспосабливается к окружающей среде и добивается своих целей. "Все машины, претендующие на "разумность", - писал он, - должны обладать способностью преследовать определенные цели и приспосабливаться, т. е. обучаться". Созданной им науке Винер дает название кибернетика, что в переводе с греческого означает рулевой. Следует отметить, что принцип "обратной связи", введенный Винером был в какой-то степени предугадан Сеченовым в явлении "центрального торможения" в "Рефлексах головного мозга" (1863 г.) и рассматривался как механизм регуляции деятельности нервной системы, и который лег в основу многих моделей произвольного поведения в отечественной психологии.

**Нейронный подход.** К этому времени и другие ученые стали понимать, что создателям вычислительных машин есть чему поучиться у биологии. Среди них был нейрофизиолог и поэт-любитель Уоррен Мак-Каллок<sup>1</sup>, обладавший, как и Винер, философским складом ума и широким кругом интересов. В 1942 г. Мак-Каллок, участвуя в научной конференции в Нью-Йорке, услышал доклад одного из сотрудников Винера о

---

<sup>1</sup> **Уоррен Мак-Каллок (1898-1969)** - американский нейропсихолог, нейрофизиолог, теоретик [искусственных нейронных сетей](#) и один из основателей [кибернетики](#).

механизмах обратной связи в биологии. Высказанные в докладе идеи перекликались с собственными идеями Мак-Каллока относительно работы головного мозга. В течении следующего года Мак-Каллок в соавторстве со своим 18-летним протеже, блестящим математиком Уолтером Питтсом, разработал теорию деятельности головного мозга. Эта теория и являлась той основой, на которой сформировалось широко распространенное мнение, что функции компьютера и мозга в значительной мере сходны. Исходя отчасти из предшествующих исследований нейронов (основных активных клеток, составляющих нервную систему животных), проведенных Мак-Каллоком, они с Питтсом выдвинули гипотезу, что нейроны можно упрощенно рассматривать как устройства, оперирующие двоичными числами. Двоичные числа, состоящие из цифр единица и ноль, - рабочий инструмент одной из систем математической логики. Английский математик XIX в. Джордж Буль, предложивший эту остроумную систему, показал, что логические утверждения можно закодировать в виде единиц и нулей, где единица соответствует истинному высказыванию, а ноль - ложному, после чего этим можно оперировать как обычными числами. В 30-е годы XX в. пионеры информатики, в особенности американский ученый Клод Шеннон, поняли, что двоичные единица и ноль вполне соответствуют двум состояниям электрической цепи (включено-выключено), поэтому двоичная система идеально подходит для электронно-вычислительных устройств. Мак-Каллок и Питтс предложили конструкцию сети из электронных "нейронов" и показали, что подобная сеть может выполнять практически любые вообразимые числовые или логические операции. Далее они предположили, что такая сеть в состоянии также обучаться, распознавать образы, обобщать, т.е. она обладает всеми чертами интеллекта.

Теории Мак-Каллока-Питтса в сочетании с книгами Винера вызвали огромный интерес к разумным машинам. В 40-60-е годы все больше кибернетиков из университетов и частных фирм запирались в лабораториях и мастерских, напряженно работая над теорией функционирования мозга и методично припаявая электронные компоненты моделей нейронов. Из этого кибернетического, или нейромодельного, подхода к машинному разуму скоро сформировался так называемый "восходящий метод" - движение от простых аналогов нервной системы примитивных существ, обладающих малым числом нейронов, к сложнейшей нервной системе человека и даже выше. Конечная цель виделась в создании "адаптивной сети", "самоорганизующейся системы" или "обучающейся машины" - все эти названия разные исследователи использовали для обозначения устройств, способных следить за окружающей обстановкой и с помощью обратной связи изменять свое поведение в полном соответствии с господствовавшей в те времена в психологии школой

бихевиористики<sup>2</sup>, т.е. вести себя так же как живые организмы. Однако отнюдь не во всех случаях возможна аналогия с живыми организмами. Как однажды заметили Уоррен МакКаллок и его сотрудник Майкл Арбиб, "если по весне вам захотелось обзавестись возлюбленной, не стоит брать амебу и ждать пока она эволюционирует".

Но дело здесь не только во времени. Основной трудностью, с которой столкнулся "восходящий метод" на заре своего существования, была высокая стоимость электронных элементов. Слишком дорогой оказывалась даже модель нервной системы муравья, состоящая из 20 тыс. нейронов, не говоря уже о нервной системе человека, включающей около 100 млрд. нейронов. Даже самые совершенные кибернетические модели содержали лишь несколько сотен нейронов. Столь ограниченные возможности обескуражили многих исследователей того периода.

**Появление персептрона.** Одним из тех, кого ничуть не испугали трудности был Фрэнк Розенблатт<sup>3</sup>, труды которого отвечали самым заметным устремлениям кибернетиков. В середине 1958 г. им была предложена модель электронного устройства, названного им персептроном, которое должно было бы имитировать процессы человеческого мышления. Персептрон должен был передавать сигналы от "глаза", составленного из фотоэлементов, в блоки электромеханических ячеек памяти, которые оценивали относительную величину электрических сигналов. Эти ячейки соединялись между собой случайным образом в соответствии с господствующей тогда теорией, согласно которой мозг воспринимает новую информацию и реагирует на нее через систему случайных связей между нейронами. Два года спустя была продемонстрирована первая действующая машина "Марк-1", которая могла

---

<sup>2</sup> **Бихевиоризм** (англ. *behavior* «поведение») - систематический подход к изучению поведения людей и животных. Он предполагает, что все поведение состоит из рефлексов, реакций на определённые стимулы в среде, а также последствий индивидуальной истории, таких как подкрепление и наказание, совместно с настоящим мотивационным состоянием индивида и контролирующими стимулами. Хотя бихевиористы, как правило, принимают важную роль, которую играет наследственность в предопределении поведения, они, прежде всего, фокусируются на средовых факторах.

Бихевиоризм сочетает элементы философии, методологии и психологической теории. Он появился в конце девятнадцатого века, как реакция на глубинную психологию и другие традиционные формы психологии, которые зачастую не справлялись с экспериментальной проверкой предсказаний.

<sup>3</sup> **Фрэнк Розенблатт (1928-1971)** – известный американский учёный в области психологии, нейрофизиологии и искусственного интеллекта.

научится распознавать некоторые из букв, написанных на карточках, которые подносили к его "глазам", напоминающие кинокамеры. Персептрон Розенблатта оказался наивысшим достижением "восходящего", или нейромодельного метода создания искусственного интеллекта. Чтобы научить персептрон способности строить догадки на основе исходных предпосылок, в нем предусматривалась некая элементарная разновидность автономной работы или "самопрограммирования". При распознавании той или иной буквы одни ее элементы или группы элементов оказываются гораздо более существенными, чем другие. Персептрон мог научиться выделять такие характерные особенности буквы полуавтоматически, своего рода методом проб и ошибок, напоминающим процесс обучения. Однако возможности персептрона были ограниченными: машина не могла надежно распознавать частично закрытые буквы, а также буквы иного размера или рисунка, нежели те, которые использовались на этапе ее обучения.

Ведущие представители так называемого "нисходящего метода" специализировались, в отличие от представителей "восходящего метода", в составлении для цифровых компьютеров общего назначения программ решения задач, требующих от людей значительного интеллекта, например, для игры в шахматы или поиска математических доказательств. К числу защитников "нисходящего метода" относились *Марвин Минский* и Сеймур Пейперт, профессора Массачусетского технологического института. Минский начал свою карьеру исследователя ИИ сторонником "восходящего метода" и в 1951 г. построил обучающуюся сеть на вакуумных электронных лампах. Однако вскоре к моменту создания персептрона он перешел в противоположный лагерь. В соавторстве с южноафриканским математиком Пейпертом, с которым его познакомил Мак-Каллок, он написал книгу "Персептроны", где математически доказывалось, что персептроны, подобные розенблаттовским, принципиально не в состоянии выполнять многие из тех функций, которые предсказывал Розенблатт. Минский утверждал, что, не говоря о роли работающих под диктовку машинисток, подвижных роботов или машин, способных читать, слушать и понимать прочитанное или услышанное, персептроны никогда не обретут даже умения распознавать предмет, частично заслоненный другим. Глядя на торчащий из-за кресла кошачий хвост, подобная машина никогда не сможет понять, что она видит. Нельзя сказать, что появившаяся в 1969 г. эта критическая работа покончила с кибернетикой. Она лишь переместила интерес аспирантов и субсидии правительственных организаций США, традиционно финансирующих исследования по ИИ, на другое направление исследований - "нисходящий метод". Интерес к кибернетике в последнее время возродился, так как сторонники "нисходящего метода" столкнулись со столь же неодолимыми трудностями. Сам Минский публично выразил сожаление, что его

выступление нанесло урон концепции персептронов, заявив, что, согласно его теперешним представлениям, для реального прорыва вперед в создании разумных машин потребуется устройство, во многом похожее на персептрон. Но в основном ИИ стал синонимом нисходящего подхода, который выражался в составлении все более сложных программ для компьютеров, моделирующих сложную деятельность человеческого мозга.

### **Границы применимости интеллектуальных систем**

**Применение ИИ. Нейросети** - это область ИИ, получившая наиболее широкое распространение в различных областях деятельности. *Нейронная сеть* представляет собой совокупность большого числа сравнительно простых элементов - нейронов. В основу искусственных нейронных сетей положены следующие черты живых нейронных сетей, позволяющие им хорошо справляться с интеллектуальными задачами:

- простой обрабатывающий элемент - нейрон;
- большое число нейронов участвует в обработке информации;
- один нейрон связан с большим числом других нейронов;
- изменяющиеся по весу связи между нейронами;
- массированная параллельность обработки информации.

Нейросети предпочтительны там, где имеется очень много входных данных, в которых скрыты закономерности. Целесообразно использовать нейросетевые методы в задачах с неполной или «зашумлённой» информацией, а также в таких, где решение можно найти интуитивно. Преимущества нейросети становятся видны тогда, когда довольно часто изменяются «правила игры». Нейросети применяются:

- *Экономика*: для предсказания рынков, оценки риска невозврата кредитов, предсказания банкротств, автоматического рейтингования, оптимизации товарных и денежных потоков, автоматического считывания чеков и форм.
- *Медицина*: обработка медицинских изображений, мониторинг состояния пациентов, диагностика, факторный анализ эффективности лечения, очистка показаний приборов от шумов.
- *Авиация*: обучаемые автопилоты, распознавание сигналов радаров, адаптивное пилотирование сильно поврежденного самолета.
- *Связь*: сжатие видеоинформации, быстрое кодирование-декодирование, оптимизация сотовых сетей и схем маршрутизации пакетов.
- *Интернет*: ассоциативный поиск информации, электронные секретари и агенты пользователя в сети, фильтрация информации в push-системах, рубрикация новостных лент, адресная реклама, адресный маркетинг для электронной



торговли.

- *Политические технологии:* анализ и обобщение социологических опросов, предсказание динамики рейтингов, выявление значимых факторов, объективная кластеризация электората, визуализация социальной динамики населения.
- *Автоматизация производства:* оптимизация режимов производственного процесса, комплексная диагностика качества продукции (ультразвук, оптика, гамма-излучение и т. д.), мониторинг и визуализация многомерной диспетчерской информации, предупреждение аварийных ситуаций, робототехника.

### **Отличие данных от знаний. Проблема представления**

Понятие искусственного интеллекта, задачи, решаемые с помощью методов искусственного интеллекта, и, как следствие, необходимость создания интеллектуальных систем (систем искусственного интеллекта) возникли практически сразу после создания А. Тьюрингом, Фон Нейманом и др. основополагающих принципов построения автоматических дискретных вычислений (ЭВМ, компьютеров).

Появление ЭВМ, работа которых происходит под управлением созданных человеком программ (т.е. с максимально полным привлечением интеллектуальных способностей человека), позволило автоматизировать самые разнообразные процессы обработки данных или, по-другому, самые разнообразные вычислительные процессы.

*Здесь под автоматизацией вычислительных процессов или вычислений понимается выполнение их вычислительным устройством (компьютером) без непосредственного участия человека. При этом важным является то, что*

- а) вычислительный процесс должен представляться в виде последовательности (сколь угодно большой, но конечной длины) элементарных или «рутинных» операций;
- б) формирование последовательности элементарных операций или, по-другому, составление алгоритма решения задачи, осуществляется непосредственно человеком (пользователем ЭВМ);
- в) вычислительное устройство не может само (без участия человека) ни создавать, ни менять алгоритм, если это изменение не предусмотрено самим алгоритмом.

Принято говорить, что вычислительные устройства (в дальнейшем ЭВМ или компьютеры), построенные по классической фон-неймановской схеме (такowymi сейчас являются подавляющее большинство ЭВМ) реализуют так называемые «жесткие» вычисления. Термин «жесткие» вычисления обозначает организацию вычислений по заранее (до начала вычислений) разработанному человеком (пользователем ЭВМ) вполне определенному алгоритму.

$\xrightarrow{F}$  Если обозначить через  $X$  – исходные данные для решения задачи, через  $Z$  – результат решения, то процедуру решения задачи на ЭВМ можно рассматривать как реализацию некоторого отображения исходных данных в конечный результат в соответствии с алгоритмом  $F$  решения задачи (рис.1).

$X \rightarrow Z$

Рис.1.

По сути, в алгоритме  $F$  аккумулированы наши знания о тех или иных законах (математических, физических, химических и т.д.), привлекаемых для решения поставленной задачи (разработки алгоритма ее решения), а также новые приемы решения, специально для этой цели, разработанные человеком, – пользователем ЭВМ.

Таким образом, можно констатировать, что традиционная обработка информации на ЭВМ происходит по схеме «Данные» ----» «Данные» на основе (или с помощью) знаний человека – пользователя ЭВМ. (рис. 2)

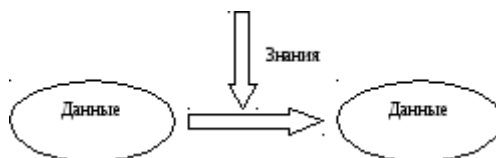


Рис. 2.

Обработка информации на ЭВМ, понимаемая в общепринятом смысле, представляет собой *обработку данных*.

В то же время, характерным признаком интеллектуальных систем является *обработка знаний*.

**Данными** называют информацию фактического характера, описывающую объекты, процессы и явления конкретной предметной области. Как правило, эта информация не требует при своем дальнейшем использовании более глубокого осмысления и анализа. К примеру, в качестве данных могут быть координаты материальной точки  $(x_i, y_i)$

измеренные в процессе ее плоского движения (точнее вращения вокруг начала координат) и соответствующей данным координатам моменты времени (рис. 3).

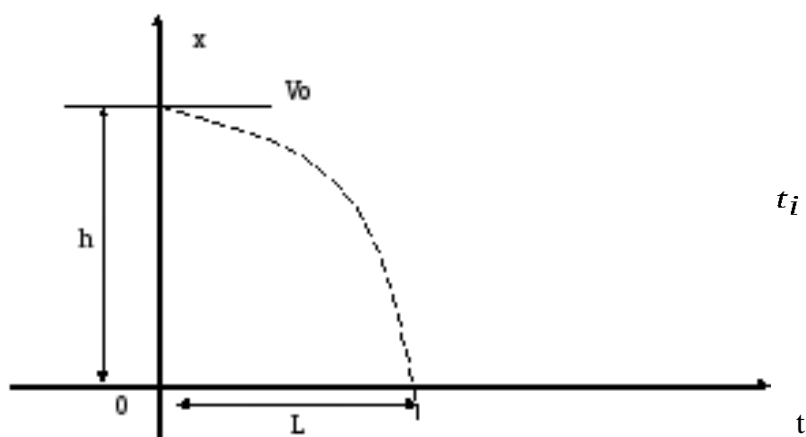


Рис.3.

Обычно данные представляются в виде таблиц, диаграмм, графиков. Так, данные о движении точки (рис. 3) можно представить в виде следующей таблицы (таблица 1)

<b>x</b>	<b>t</b>
$x_1$	$t_1$
$x_2$	$t_3$
....	...
$x_N$	$t_N$

**Знания** являются более сложной категорией информации по сравнению с данными. Знания описывают не только отдельные факты, но и взаимосвязи между ними. Поэтому, знания иногда структурированными данными. Знания могут быть получены:

- А) на основе обработки экспериментальных данных (данных эксперимента);
- Б) в результате мысленной деятельности человека.

Интеллектуальные системы позволяют производить автоматическую (без участия человека) обработку данных в условиях существенной априорной неполноты знаний о том, как нужно вести эту обработку для получения требуемого результата. Очевидно, что для этого интеллектуальные системы должны быть способны сами генерировать (получать) недостающие знания путем:

- А) логического (дедуктивного) вывода;
- Б) обучения;

- В) поиска;
- Г) обработки экспериментальных данных.

В первых трех случаях обработка информации происходит по схеме «знания» --- «знания», «знания» --- «знания» --- «данные».

С помощью существующих на настоящий момент времени методов искусственного интеллекта (нечеткой логики, нейронных сетей и генетических алгоритмов) в последнем случае – по схеме «данные» --- «знания» с помощью нейросетевых методов аппроксимации и интерпретации данных (рис. 4).

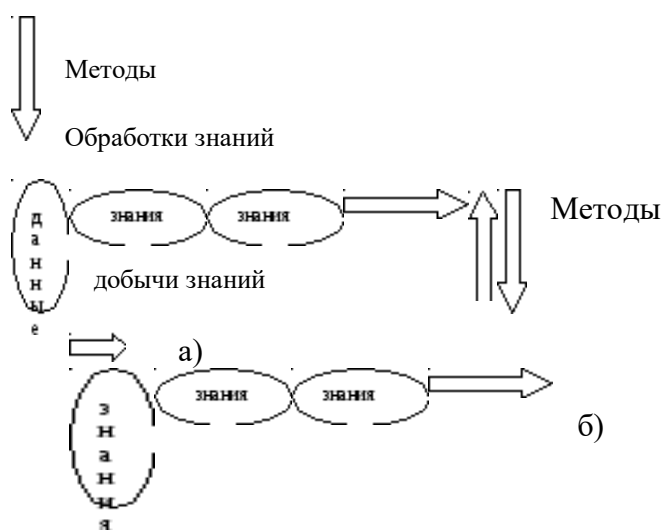


Рис. 4.

Так, если аппроксимировать приведенные в табл.1 данные о движении материальной точки, брошенной с высоты  $h$  под углом к горизонту, то  $\alpha = 0$  получим следующую функциональную зависимость высоты  $x$  от времени  $t$  движения (1)

$$x = h - \frac{gt^2}{2}$$

Где  $g=9.8$  – гравитационная постоянная.

Выражение (1) можно трактовать как запись (на языке математических формул) наших знаний о законе движения материальной точки под действием силы земного притяжения.

Сопоставляя табл. 1 и формулу (1), можно отметить, что запись информации о движении материальной точки с помощью знаний дает (по сравнению с совокупностью данных о ее движении) более целостную и системную картину движения.

*Общая характеристика задач, решаемых методами ИИ.* Рассмотрим главные или, что более правильно, сущностные отличия задач, решаемых на ЭВМ с помощью методов искусственного интеллекта, от обычных задач, решаемых традиционными методами и способами.

**Степень использования человеческого интеллекта.** Как известно, традиционно решаемые на ЭВМ задачи требуют максимально полного использования интеллекта (способностей, знаний) при выборе метода и составлении алгоритма решения задачи.

При этом на ЭВМ возлагается лишь задача правильного выполнения (или реализации) разработанного человеком алгоритма.

Напротив, решение задач с привлечением методов искусственного интеллекта (или задач, решаемых системами искусственного интеллекта) основывается не только на использовании знаний человека, но и дополнительных знаниях, полученных самой ЭВМ.

Методы и структурные решения, лежащие в основе получения (вывода) знаний, являются предметом рассмотрения сравнительно молодой науки (ей не более 50 лет), называемой искусственным интеллектом.

*Полнота априорной информации.* Традиционно решаемые на ЭВМ задачи (разумеется речь идет не о простых задачах, а о достаточно сложных) требуют для своего успешного решения большого объема априорной информации о закономерностях поведения исследуемого объекта или процесса. Например, если рассматривается движение летательного аппарата в атмосфере, то должны быть точно известны:

- а) физические законы, определяющие силы, действующие на летательный аппарат;
- б) полученные на их основе математические состояния (математическая модель объекта), определяющие реакцию летательного аппарата (изменение его высоты, скорости полета и т.п.) на эти силы и на управляющие воздействия со стороны системы управления летательным аппаратом.

На практике это весьма сложно обеспечить, учитывая существенную нестационарность условий полета (внешних и внутренних). Действительно, для современных летательных аппаратов характерен большой диапазон изменения характеристик атмосферы, возможность возникновения нештатных (или критических) ситуаций как в атмосфере (грозы, смерчи, турбулентные потоки и т.п.), так и на борту летательного аппарата (отказы оборудования, неправильные действия летчиков). Заранее все это при разработке алгоритма управления сложно предусмотреть. «Жесткие» алгоритмы управления современными летательными аппаратами не обеспечивают требуемой эффективности (в том числе боевой) их применения.

Интеллектуальные или «мягкие» алгоритмы управления, основанные на применении методов искусственного интеллекта («мягких» вычислений), существенно снижают требования к объему необходимой априорной информации за счет ее доопределения интеллектуальной системой непосредственно в процессе функционирования (в режиме on-line).

*«Продвинутость» задач.* На сегодняшний день практически все ЭВМ имеют фон-неймановскую архитектуру, основанную на функциональных принципах построения дискретных вычислений, изложенных в работах Ч. Бэббиджа, Поста, А. Тьюринга. Это накладывает определенные ограничения на класс задач, решаемых на ЭВМ. В частности, решение задач должно допускать возможность описания его с помощью некоторого алгоритма.

В свою очередь, это означает, что на ЭВМ могут быть реализованы только алгоритмические процедуры, допускающие представление в виде совокупности базовых (элементарных) операций (в современных ЭВМ - это сложение и сдвиг).

Однако с помощью алгоритмов и алгоритмических процедур в классическом понимании можно автоматизировать решение только так называемых «рутинных» задач, не связанных с получением качественно новой информации (новых знаний), а связанных с организацией вычислительной процедуры их решения при условии, что априорно имеется вся необходимая для этого информация.

Решение же более содержательных по смыслу (более интеллектуальных задач) только с помощью алгоритмических процедур невозможно. Например, невозможно с помощью алгоритмов описать процессы, реализующие

- а) решение задач, начиная от словесной постановки и кончая получением результата решения;
- б) перевода текстов с одного языка на другой;
- в) игру в шахматы, карты и т.п.
- г) диагностики болезней;
- д) доказательства математических теорем и др.

Для этого нужно располагать качественно новым математическим аппаратом и вычислительными машинами, позволяющими моделировать процесс мышления человека.

Рассмотрим характерные особенности данного процесса.

1. Деятельность человека всегда целесообразна, т.е. связана с достижениями некоторой цели. Это означает, что мыслительные процессы человека направлены на достижение цели (цель заставляет человека думать).

2. Человеческий мозг хранит огромное количество фактов и правил их использования. Для достижения определенной цели надо только обратиться к нужным фактам и правилам.

3. Принятие решений всегда осуществляется на основе специального механизма упрощения, позволяющего отбрасывать ненужные (малосущественные) факты и правила. Не имеющие отношения к решаемой в данный момент задаче и, наоборот, выделять главные, наиболее значимые факты и правила, нужные для достижения цели.

4. Достигая цели, человек не только приходит к решению поставленной перед ним задачи, но и одновременно приобретает новые знания. Та часть интеллекта, которая позволяет ему делать соответствующие заключения (выводы) на основании правил, сформулированных человеком, а также генерировать новые факты из уже существующих, называется механизмом *логического вывода*.

Так, типовая схема решения математической задачи часто выглядит следующим образом. Выбираются неизвестные величины, подлежащие определению. На основании анализа условий (ограничений), содержащихся в исходной формулировке задачи, составляется система уравнений, связывающих указанные неизвестные. Далее, применяя какой-либо из стандартных методов решения полученных уравнений, находим искомое решение задачи. Заметим, то, решив один раз конкретную задачу по описанной схеме, мы решим (и гораздо быстрее) другую подобную (и даже более сложную) задачу, отличающуюся значениями исходных данных, числом неизвестных, формой представления условий и т.д.

Поскольку система ИИ принимает решения аналогично тому, как это делает человек, то она должна включать в себя следующие *ключевые элементы*: цель, факты и данные; правила, механизмы вывода и упрощения. Все эти компоненты системы ИИ показаны на рис. 5. - выделена база знаний, которая содержит всю располагаемую информацию о внешнем мире (моделях решаемых задач). Условно она может быть разделена на три части (или области), называемые базой целей, базой правил и базой данных. Первая область содержит информацию о целях, для достижения которых предназначена система ИИ. Вторая область включает в себя сведения, которые отражают закономерности, характерные для решаемого класса задач. Это правила, механизмы упрощения и вывода, которые позволяют не только выводить новые факты, не зафиксированные ранее в базе данных, но и приобретать новые знания в ходе функционирования системы или на этапе ее обучения. В третьей области содержатся в некотором упорядоченном виде качественные данные, необходимые для решения данной задачи. В силу той особой роли, которую играет база

знаний в процессе формирования решений, системы ИИ называют системами, основанными на знаниях.

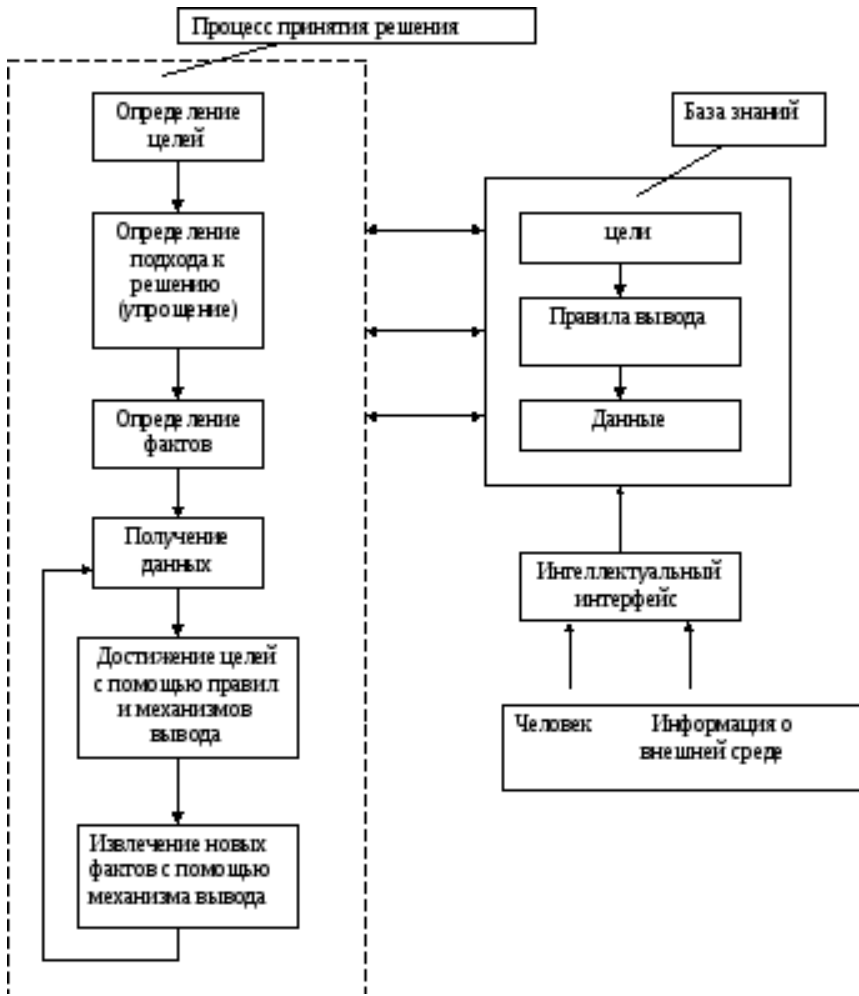


Рис. 5



## Семиотический треугольник Фреге

**Характеристики знаний.** Основной проблемой, решаемой во всех системах ИИ, является проблема *представления знаний*. Информация представляется в компьютере в процедурной и декларативной форме. В процедурной форме представлены программы, в декларативной – данные. В системах искусственного интеллекта возникла концепция новой формы представления информации – знания, которая объединила в себе черты как процедурной, так и декларативной информации. Перечислим основные характеристики знаний:

1. *Внутренняя интерпретируемость.* Каждая информационная единица должна иметь уникальное имя, по которому система находит ее, а также отвечает на запросы, в которых это имя упомянуто. Данные в памяти лишены имен и могут идентифицироваться только программой, извлекающей их из памяти. При переходе к знаниям в память вводится информация о некоторой *протоструктуре информационных единиц и словари имен данных*. Каждая единица информации будет экземпляром протоструктуры. СУБД обеспечивают реализацию внутренней интерпретируемости информации, хранимой в базе данных.

2. *Структурированность.* Информационные единицы должны соответствовать «принципу матрешки», то есть рекурсивной вложенности одних информационных единиц в другие. Другими словами, должна существовать возможность произвольного установления между отдельными информационными единицами отношений типа «часть – целое», «род – вид» или «элемент – класс».

3. *Связность.* Между информационными единицами должна быть предусмотрена возможность установления связей различного типа, характеризующих отношения между информационными единицами. Семантика отношений может носить как декларативный характер, например, в отношениях «одновременно» и «причина – следствие», так и процедурный характер, например, в отношении «аргумент – функция». Все отношения можно разделить на четыре категории: *отношения структуризации* (задают иерархию информационных единиц), *функциональные отношения* (несут процедурную информацию, позволяющую вычислять одни информационные единицы через другие), *каузальные отношения* (задают причинно-следственные связи) и *семантические отношения* (все остальные отношения).

4. *Семантическая метрика.* Между информационными единицами задают *отношения релевантности*, которые характеризуют ситуационную близость информационных единиц, то есть силу ассоциативной связи между информационными единицами (например, «покупка» или «регулирование движения на перекрестке»).

Отношение релевантности позволяет находить знания, близкие к найденным ранее знаниям.

5. *Активность.* Выполнение программ в информационных системах должно инициализироваться не командами, а состоянием информационной базы, например, появлением в базе фактов или описаний событий или установление связей между информационными единицами.

Перечисленные характеристики определяют разницу между *данными* и *знаниями*, при этом *базы данных* перерастают в *базы знаний*.

**Модели представления знаний.** В интеллектуальных системах используются четыре основных типа моделей знаний:

1. *Логические модели.* В основе моделей такого типа лежит *формальная система*, задаваемая четверкой вида  $M = \langle T, S, A, B \rangle$ . Множество  $T$  есть множество *базовых элементов*, например, слов из некоторого словаря, или деталей из некоторого набора. Для множества  $T$  существует некоторый способ определения принадлежности или непринадлежности произвольного элемента к данному множеству. Процедура такой проверки может быть любой, но она должна давать ответ на вопрос, является ли  $x$  элементом множества  $T$  за конечное число шагов. Обозначим эту процедуру  $P(T)$ .

Множество  $S$  есть множество *синтаксических правил*. С их помощью из элементов  $T$  образуют синтаксически правильные совокупности. Например, из слов словаря строятся синтаксически правильные фразы, а из деталей собираются конструкции. Существует некоторая процедура  $P(S)$ , с помощью которой за конечное число шагов можно получить ответ на вопрос, является ли совокупность  $X$  синтаксически правильной.

Во множестве синтаксически правильных совокупностей выделяется некоторое подмножество  $A$ . Элементы  $A$  называются *аксиомами*. Как и для других составляющих формальной системы, должна существовать процедура  $P(A)$ , с помощью которой для любой синтаксически правильной совокупности можно получить ответ на вопрос о принадлежности ее к множеству  $A$ .

Множество  $B$  есть множество *правил вывода*. Применяя их к элементам  $A$ , можно получать новые синтаксически правильные совокупности, к которым снова можно применять правила из  $B$ . Так формируется *множество выводимых* в данной формальной системе *совокупностей*. Если имеется процедура  $P(B)$ , с помощью которой можно определить для любой синтаксически правильной совокупности, является ли она выводимой, то соответствующая формальная система называется *разрешимой*. Это показывает, что именно правила вывода являются наиболее сложной составляющей формальной системы.

Для знаний, входящих в базу знаний, можно считать, что множество  $A$  образуют все информационные единицы, которые введены в базу знаний извне, а с помощью правил вывода из них выводятся новые *производные знания*. Другими словами, формальная система представляет собой генератор порождения новых знаний, образующих множество *выводимых* в данной системе знаний. Это свойство логических моделей позволяет хранить в базе лишь те знания, которые образуют множество  $A$ , а все остальные знания получать из них по правилам вывода.

2. *Сетевые модели.* Сетевые модели формально можно описать в виде  $H = \langle I, C_1, C_2, C_n, G \rangle$ . Здесь  $I$  есть множество информационных единиц;  $C_1, C_2, C_n$  - множество типов связей между ними. Отображение  $G$  задает связи из заданного набора типов связей между информационными единицами, входящими в  $I$ .

В зависимости от типов связей, используемых в модели, различают *классифицирующие сети, функциональные сети и сценарии, нейронные сети*. В классифицирующих сетях используются отношения структуризации. Такие сети позволяют в базах вводить иерархические отношения между информационными единицами. Функциональные сети характеризуются наличием функциональных отношений. Их часто называют *вычислительными моделями*, так как они позволяют описывать процедуры «вычислений» одних информационных единиц через другие. Нейронные сети можно отнести к классу функциональных сетей, однако нечёткие продукционные нейронные сети представляют собой гибридную модель, соединяющую в себе черты логической, продукционной и сетевой моделей. В сценариях используются каузальные отношения, а также отношения типа «средство – результат». Если в сетевой модели допускаются связи различного типа, то ее называют *семантической сетью*.

3. *Продукционные модели.* В моделях этого типа используются некоторые элементы логических и сетевых моделей. Из логических моделей заимствована идея правил вывода, которые здесь называются *продукциями*, а из сетевых моделей – описание знаний в виде семантической сети. В результате применения правил вывода к фрагментам сетевого описания происходит трансформация семантической сети за счет смены ее фрагментов, наращивания сети и исключения из нее ненужных фрагментов. Таким образом, в продукционных моделях процедурная информация явно выделена и описывается иными средствами, чем декларативная информация. Вместо логического вывода, характерного для логических моделей, в продукционных моделях появляется *вывод на знаниях*. **Продукционная модель** или модель, основанная на правилах, позволяет представить знания в виде предложений типа «Если (условие), то (действие)». Под *условием* понимается некоторое предложение - образец, по которому осуществляется поиск в базе знаний, а под

*действием* - действия, выполняемые при успешном исходе поиска (они могут быть промежуточными, выступающими как условия, и терминальными или целевыми, завершающими работу системы). При использовании продукционной модели база знаний состоит из набора правил, Программа, управляющая перебором правил, называется *машиной вывода*. Чаще всего вывод бывает прямой (от данных к поиску цели) или обратный (от цели для ее подтверждения – к данным).

*Данные* - это исходные факты, на основании которых запускается машина вывода. Если в памяти системы хранится некоторый набор продукций, то они образуют систему продукций. В системе продукций должны быть заданы специальные процедуры управления продукциями, с помощью которых происходит актуализация продукций и выполнение той или иной продукции из числа актуализированных.

4. *Фреймовые модели*. В отличие от моделей других типов во фреймовых моделях фиксируется жесткая структура информационных единиц, которая называется *протофреймом*. В общем виде она выглядит следующим образом:

(*Имя фрейма*:

*Имя слота 1 (значение слота 1)*

*Имя слота 2 (значение слота 2)*

. . . . .

*Имя слота K (значение слота K)).*

Значением *слота* может быть все, что угодно: числа, математические соотношения, тексты на естественном языке, программы, правила вывода, ссылки на другие слоты данного фрейма или других фреймов. В качестве значения слота может выступать набор слотов более низкого уровня, что позволяет реализовать во фреймовых представлениях «принцип матрешки».

При конкретизации фрейма ему и слотам присваиваются имена, и происходит заполнение слотов. Таким образом, из протофреймов получают *фреймы-экземпляры*. Переход от исходного протофрейма к фрейму-экземпляру может быть многошаговым, за счет постепенного уточнения значений слотов. Связи между фреймами задаются значениями специального слота с именем «Связь». Некоторые специалисты по ИС не выделяют фреймовые модели в отдельный класс, так как в ней объединены все основные особенности моделей остальных типов.

**Семиотический треугольник Фреге.** Знак является единством формы, представляющей, замещающей некоторый предмет, и информации о нем. Форма (означающее), значение (сигнификат, отражение свойств и признаков денотата в сознании, представление о предмете) и денотат (объекты действительности, которые могут быть

обозначены данным знаком) образуют структуру знака, так называемый семантический (семиотический) треугольник, или, по имени немецкого ученого Готлоба Фреге (1848-1925) – *треугольник Фреге*.

Готлоб Фреге, один из основателей математической логики, так писал о знаке: «Употребляя знак, мы хотим сказать что-то не о знаке, но главным, как правило, является его значение».

*Знак*: в лингвистике, например, фонетическое слово или написанное слово; в математике – математический символ; иное название, принятое в философии и математической логике, – имя.

Тип вещи, предмета, явления действительности называют *денотатом знака*.

Конкретную вещь называют *референтом знака*.

Поскольку человек обозначает вещи не напрямую, а через свое восприятие вещи, то в эту вершину треугольника часто помещают *смысл*, представление о вещи, сигнификат знака. Понятие также иногда называют *концепт* или *десигнат* знака.



Рис. 6 Треугольник Фреге

*Семиотика* – наука о знаковых системах. Основоположник семиотики – американский ученый Чарлз Сандерс Пирс (1839 – 1914). Основные труды Пирса были опубликованы после его смерти, и как наука семиотика начала развиваться только в 30-х гг. XX века. Классификация Пирса является самой известной из существующих классификаций знаков. Знаки, по Пирсу, делятся на три вида:

1. *Символы* (условные, или конвенциональные знаки – денотат (означаемое) связан с формой как бы по соглашению, договору, негласно заключенному между пользующимися этими знаками; форма его выражения ни в каком отношении не сходна с

его денотатом: формы слов книга, человек, письмо, храбрость ничуть не похожи на человека, письмо и т. д.).

2. *Индексы* (указательные, или дейктические знаки – форма и денотат находятся в отношениях смежности, «соприкасаются» друг с другом в пространственном или временном отношениях: знак направления в виде стрелки, слова «там», «сейчас», «здесь» и т. д.).

3. *Иконы* (изобразительные знаки – форма и денотат похожи друг на друга в том или ином отношении: различные пиктограммы (знаки в виде картинок), звукоподражательные слова (жужжать) и т. д.).

Важными свойствами любого знака являются:

- наличие означающего (экспонент знака), т.е. нечто, доступное зрению или слуху (органам чувств), какие-либо материальные признаки;
- наличие обозначаемого (содержания знака – понятия, образа);
- наличие условной связи (не природной, естественной и неизбежной, а основанной на договоренности);
- соотнесенность с другим знаком – если нет противопоставленности другому знаку – знак не существует. Знаковая система может состоять из знака, имеющего в качестве означающего какой-либо материальный сигнал, и знака, у которого означающее – отсутствие такого сигнала. Такой знак называется нулевым (например, нулевое окончание, пробел и т. д.)

Не каждая группа знаков может называться системой. Основным свойством системы является наличие у нее структуры. *Структура* подразумевает наличие отношений между всеми ее элементами. *Система*<sup>4</sup> - это структура в действии, т.е. структура, способная функционировать.

---

<sup>4</sup> **Система** - множество элементов, находящихся в отношениях и связях друг с другом, которое образует определённую целостность, единство.