

## **Лекция 1. Основные понятия систем искусственного интеллекта.**

### **1.1. Основные понятия.**

**А. Цель.** Интеллектуальная деятельность всегда связана с какой-то целью. **Целью называется конечный результат, на который направлены мыслительные процессы человека.** При проектировании систем ИИ всегда следует помнить о цели, для достижения которой они предназначены.

Целью может быть, например, следующее:

1. Определить кратчайший путь между Москвой и Новгородом.
2. Выбрать вино, больше всего подходящее к определенной рыбе.
3. Научиться завязывать шнурки у ботинок.
4. Найти способ оценки успехов ребенка в арифметике.

#### **Б. Факты и правила.**

Человеческий мозг - это огромное хранилище знаний. Человеку свойственно приобретать новые знания и применять их к возникающим ситуациям. В общем, интеллект можно представить как совокупность фактов и способов их применения для достижения цели. Отчасти цели достигаются с помощью правил использования всех известных фактов. Приведем несколько примеров фактов и правил их использования.

#### **Пример 1.**

Факт 1. Зажженная плита — горячая.

Правило 1. ЕСЛИ положить руку на зажженную плиту, ТО можно обжечься.

#### **Пример 2**

Факт 2. В час пик на улице много машин.

Правило 2. ЕСЛИ попытаться в час пик перейти шоссе, ТО можно попасть под машину.

#### **Пример 3**

Факт 3а. Тихие, темные улицы опасны.

Факт 3б. Пожилые люди обычно не совершают дерзких преступлений.

Факт 3в. Полиция защищает людей от преступников.

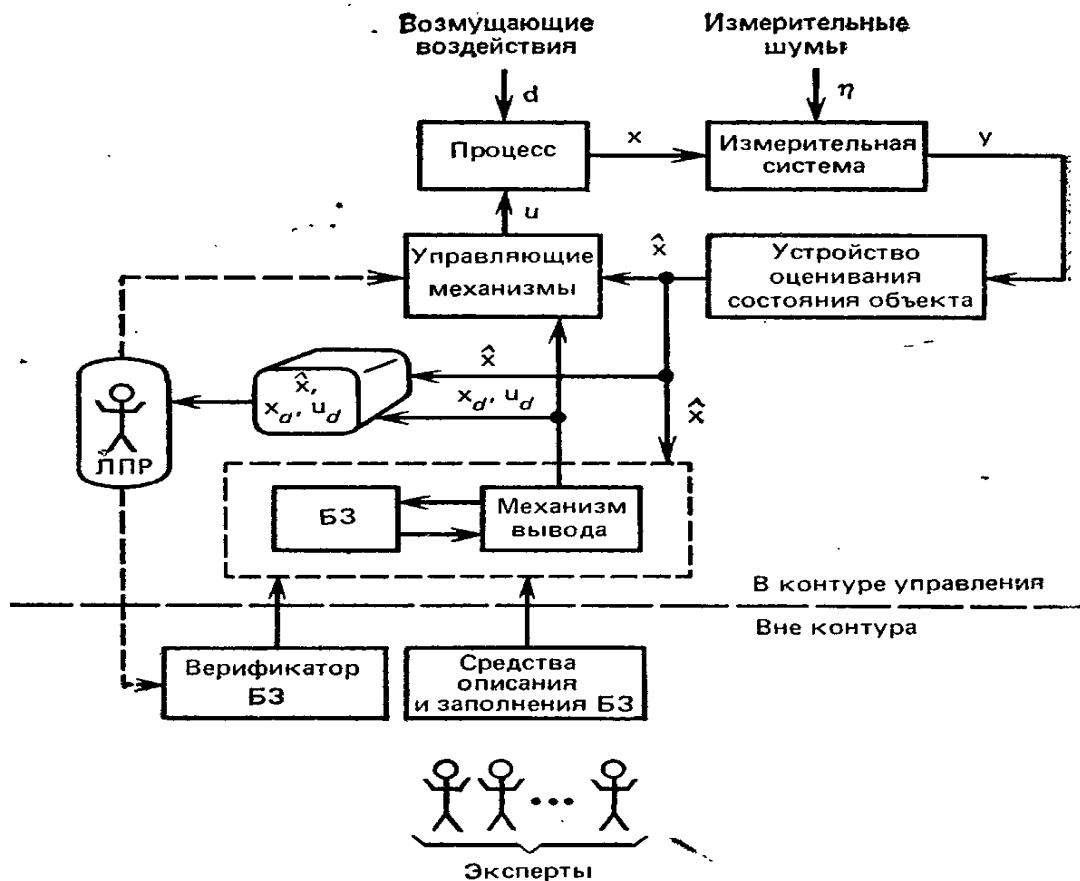


Рис. 1 Структурная схема СИИ АСУ ТП.

Правило 3а. ЕСЛИ на тихой, темной улице встретится пожилой человек, ТО можно не очень беспокоиться.

Правило 3б. ЕСЛИ на тихой, темной улице вы видите полицейского, ТО можно чувствовать себя в безопасности.

### В. Упрощение.

В мозгу существует сложная система, руководящая выбором правильной реакции на конкретную ситуацию. Такой выбор называется упрощением. Механизм упрощения блокирует мысли, не имеющие отношения к решаемой в данный момент задаче.

Когда человек сталкивается с какой-то ситуацией, механизм упрощения заставляет его мозг сосредоточиться только на фактах и правилах, нужных для достижения поставленной цели.

### Г. Механизм вывода.

Достигая цель, человек не только приходит к решению поставленной перед ним задачи, но одновременно приобретает новые знания. Рассмотрим такой пример:

1. Иван и Мария — родители Юры.
2. Иван и Мария — родители Анны.

Цель заключается в том, чтобы определить, кем приходятся друг другу Юра и Анна. Механизм упрощения заставляет человека обратиться к хранящемуся в его мозгу правилу: ЕСЛИ у девочки и мальчика одни и те же родители, ТО мальчик и девочка — брат и сестра. Цель мгновенно достигнута.

Ответ на вопрос о степени родства Юры и Анны получен из известного ранее

правила. Кроме того, в процессе достижения цели получен новый факт: Юра и Анна — брат и сестра.

**Часть интеллекта, которая помогает извлекать новые факты, называется механизмом вывода.**

Именно механизм вывода позволяет человеку учиться на опыте, так как он дает возможность генерировать новые факты из уже существующих, применяя имеющиеся знания к новой ситуации.

#### **Д. База знаний.**

Факты формулируются в виде вопросов, ответы на которые помогают человеку принять окончательное решение. Факты и правила хранятся в компьютере в так называемой **базе знаний**.

#### **Е. Экспертная система.**

Конкретные сферы человеческой деятельности, в которых могут применяться системы ИИ, называются **предметными областями**.

Примерами предметных областей могут служить оценка эффективности обучения и выбор маршрута автобуса.

Невозможно создание единой системы ИИ, охватывающей все предметные области. Для такой системы необходимо бесконечное число фактов и правил. Даже если бы такая система была создана, понадобилось бы длительное время на наполнение ее знаниям

**Система ИИ, созданная для решения задач в конкретной проблемной области, называется экспертной системой.** Источником знаний для наполнения экспертных систем служат люди - эксперты в соответствующей предметной области.

Когда человек сталкивается с проблемой, не все соображения для него равноценны, например, если надо попасть на работу вовремя, он не будет слишком заботиться о том, чтобы сидеть в автобусе. Можно сказать, что человек «взвешивает» различные соображения. Числа, хотя бы приближенно оценивающие тех или иных фактов, поэтому называются **весами** или **весовыми факторами** фактов.

Получив сумму весовых факторов для положительных или отрицательных ответов, можно узнать, как велики шансы на положительное или отрицательное решение задачи.

Итоговое число, оценивающее как положительные, так и отрицательные шансы, называется общим весовым фактором. Общий весовой фактор- это некоторая количественная оценка.

Весовые факторы выбираются не случайно, они представляют собой знания, полученные в результате исследования проблемной области. Работа всех экспертных систем основана строго на экспертной информации, полученной в конкретной проблемной области.

#### **Ж. Получение данных.**

После того как определены общие факты, необходимые для достижения цели, надо получить конкретные данные и присвоить значения переменным.

Прежде всего, факты следует представить в форме вопросов, ответив на которые можно получить необходимую информацию.

! Программа в режиме диалога выводит на дисплей вопросы, ответы на

которые будут использоваться для оценки цели.

Проинициализированные переменные становятся частью базы данных.

### **1.2. Прямая и обратная цепочки рассуждений.**

Из сказанного выше следует, что разработка программы для И.И. состоит из:

1. Определения целей.
2. Определения фактов, имеющих отношение к этим целям.
3. Получения данных, соответствующих фактам, характерным для заданной ситуации или объекта.
4. Оценки данных, используя правила и механизм вывода.

Процесс достижения целей описанным способом называется **прямой цепочкой рассуждений**, т. е. **цепочкой от данных к логическому заключению**. Он позволяет логически переходить от одного шага к другому.

Процесс, в котором **заключение используется для поиска подтверждающих его данных**, называется **обратной цепочкой рассуждений**.

Рассмотрим пример, иллюстрирующий обратную цепочку рассуждений.

Совершено преступление: в квартире был обнаружен труп с пулевыми ранами.

Полиция начала расследование. Первое, чем заинтересовались полицейские, - кто кроме потерпевшего, имел ключ от квартиры? Они узнали, что у убитого был приятель, который часто пользовался его квартирой. Расследование показало, что друзья недавно поссорились.

Опросив свидетелей, полиция сделала вывод, что приятель потерпевшего и является вероятным убийцей (прямая цепочка рассуждений);

Но для завершения дела были необходимы неопровержимые улики. Осмотрев квартиру подозреваемого, они ничего не обнаружили. Но в соседнем переулке в мусорном баке один из полицейских нашел ружье с отпечатками пальцев подозреваемого, а баллистические эксперименты подтвердили, что человека убили из этого ружья. Преступление было раскрыто. В данном случае, получая новые данные и проверяя, согласуются ли они с изначальным заключением, полиция идентифицировала убийцу.

Здесь заключение использовалось для поиска подтверждающих его данных, т.е. была проведена обратная цепочка рассуждений.

Заключение - это подозреваемый, а данные - это оружие.

### **1.3. Агенты и среды.**

Искусственный интеллект занимается созданием интеллектуальных сущностей, объектов, которые принято называть **агентами** или **носителями**.

Агент воспринимает внешнюю среду с помощью датчиков  $x_1, x_2, \dots, x_m$  и воздействует на нее посредством исполнительных органов  $z_1, z_2, \dots, z_n$ , подобно тому, как человек воспринимает внешнюю среду или просто среду с помощью органов чувств и воздействует на нее с помощью таких частей тела, как руки, ноги и т.п. В понятия датчиков и исполнительных органов закладывают самый широкий смысл. Например, датчиком может быть некий аналог уха, воспринимающий речевые сообщения, а исполнительным органом — органы речи, позволяющие передавать сообщения на каком-либо языке. Обычно воздействие агента на среду называют **реакцией**, а восприятие агентом среды — **восприятием**.

Если каждый исполнительный орган  $z_j$  сопоставить с одноименной выходной переменной  $z_j$ , принимающей множество значений  $\gamma_j$ , и каждое такое значение назвать **микрореакцией**, то реакция будет представлять собой набор значений  $\gamma_1, \gamma_2, \dots, \gamma_n$ .

Аналогично, если каждый датчик  $x_i$  сопоставить с одноименной входной переменной  $x_i$ , принимающей множество значений  $\alpha_i$ , называемых **микровосприятиями**, то восприятие будет представлять собой набор значений  $\alpha_1, \alpha_2, \dots, \alpha_m$ .

**Поведение агента состоит в переработке восприятий в реакции.** Эта переработка осуществляется агентом с помощью специального **решателя**, функционирующего на основе заложенных в него знаний.

Не существует какой-либо общепринятой классификации агентов. В зависимости от сложности решаемых задач выделим следующие четыре типа агентов:

**комбинационные; последовательностные; целенаправленные; целевыбирающие.**

А) Поведение комбинационного агента внешне выглядит достаточно простым. В определенный момент времени  $t$  агент получает с датчиков  $x_1, \dots, x_m$  восприятие  $\alpha_1, \alpha_2, \dots, \alpha_m$ , характеризующее состояние среды.

На основании только этого восприятия и неизменяемых в процессе всего существования агента знаний, хранящихся в его памяти, он в этот же момент времени с помощью исполнительных органов  $z_1, z_2, \dots, z_n$  формирует реакцию  $\gamma_1, \gamma_2, \dots, \gamma_n$ . Конечно, при практической реализации агента на формирование реакции по данному восприятию требуется время, но теоретически считается, что все происходит мгновенно в момент времени  $t$ , и этот момент времени нас может даже не интересовать. Существенно лишь то, что комбинационный агент не порождает новые знания. Каждый раз, когда надо вырабатывать очередную реакцию по вновь поступившему восприятию, он использует одни и те же знания, хранящиеся в его памяти.

Б) Агентов, которые используют запомненную в предыдущие моменты времени информацию, называют **последовательностными**.

В) Поведение **целенаправленного** агента принципиально отличается от комбинационного и последовательностного, поскольку их поведение основано на восприятиях в настоящий или предыдущий момент времени и использовании правил, учитывающих только эти восприятия или производные от них состояния.

**Целенаправленный же агент прежде, чем принять решение, на основании известной ему цели (в нашем примере места назначения и времени, к которому он туда должен прибыть) заранее планирует свои реакции.** Иными словами, на основании имеющихся у него правил агент заранее до того, как он начнет действовать, пытается построить план, гарантирующий ему достижение цели, или обнаруживает, что такого плана не существует. В случае обнаружения недостижимости цели он может запросить дополнительные правила и продолжить или повторить процесс поиска. План является последовательностью пар восприятие - реакция (или только реакций), называемых также действиями и ведущих к цели. Если план найден, то целенаправленный агент его выполняет и достигает цели.

Таким образом, решатель целенаправленного агента использует не раз и навсегда данное ему множество правил, предписывающих, какие реакции выдавать в ответ на восприятия, а всякий раз для каждой вновь возникающей цели порождает план достижения именно этой цели. Исходными для работы такого решателя могут быть также правила, описывающие не реакции агента на конкретные восприятия, а некие общие законы его поведения в среде, законы поведения самой среды и законы порождения планов достижения целей.

Г). **Целевыбирающий агент**, помимо возможности построения планов достижения целей, так же, как это делает целенаправленный агент, способен на большее.

Во-первых, при наличии одной цели **он может выбирать из множества всех конкурирующих планов достижения цели наилучший**, иногда и без полного построения всех планов.

Во-вторых, при наличии нескольких конкурирующих целей, достижение каждой из которых заранее нельзя оценить с полной уверенностью, **он способен определить степень успеха достижения каждой цели в зависимости от ее важности**.

В-третьих, на основании предшествующего опыта, **он может обучаться и корректировать или пополнять свои знания**.

**Агент всегда функционирует в некоторой среде**. От свойств конкретной среды зависит выбор типа агентов и всего, что ему необходимо для успешного функционирования в этой среде.

Рассмотрим в общих чертах свойства сред в виде взаимоисключающих пар.

А). Существуют **дискретные** и **непрерывные** среды.

Дискретные среды таковы, что число различных восприятий и реакций, которые требуются агенту при функционировании в среде, конечно.

Непрерывные среды могут породить бесконечное число восприятий, реакций или того и другого. Примером дискретной среды является, например, среда шахмат, а непрерывной — среда агента-водителя, если для его функционирования требуется восприятие значения, например, скорости со сколь угодно высокой точностью. Если же все параметры среды воспринимаются агентом (как это обычно бывает на практике) с определенной точностью и в заданных пределах, например, скорость с точностью до 1 км/ч в пределах от 1 до 200 км/ч, то такая среда с точки зрения агента также может считаться дискретной.

Б). **Различают детерминированные и недетерминированные среды**.

В детерминированных средах по любому восприятию агент формирует строго одну реакцию. Недетерминированные же среды таковы, что вследствие каких-либо причин, например недоступности всех необходимых восприятий, агент не в состоянии сформировать единственную реакцию.

Кроме того различают статические и динамические среды.

В). **Среда является статической**, если за время, протекающее между получением агентом любого восприятия и выработкой им реакции, в среде ничего не изменяется. В противном случае среда называется **динамической**. При функционировании агента в статической среде необязательно, чтобы он наблюдал за ней, пока занимается выработкой реакции. Но даже если среда является динамической, на практике чаще всего считается, что для агента неважно, какие

изменения в ней происходят, пока он вырабатывает реакцию. Агент игнорирует эти изменения, считая динамическую среду статической.

**Предмет искусственного интеллекта - наука о создании агентов, а под СИИ понимается сообщество агентов, способных решать интеллектуальные задачи в средах.**

Предполагается, что создание агента любого типа осуществляется человеком, и он всегда способен решать задачи суперагента любого уровня. Процесс создания агента сам по себе неформален и качество описания, выражающееся в степени адекватности поведения получаемого агента задуманному, зависит от учета создателем всех необходимых аспектов его будущего поведения. Иначе говоря, создатель агента должен включить в его описание все правила, необходимые для задуманного поведения. **Совокупность всех таких правил называют базой знаний агента.**

Эту базу знаний можно представить в некотором формальном языке, в частности, языке логики.

В ответ на свое восприятие агент с помощью логических рассуждений на основе знаний, хранящихся в базе знаний, способен вырабатывать реакции. Механизм рассуждений зависит от типа агента и от языка представления базы знаний.

#### **Логические рассуждения.**

Рассуждением или умозаключением обычно называют ряд мыслей, изложенных в логически последовательной форме.

Агент должен уметь находить интересующие его состояния среды (**целевые состояния**), если он что-либо знает о других ее состояниях. **Определение целевых состояний осуществляется с помощью поиска или рассуждений в пространстве состояний.**