

# Базовые информационные технологии: технологии искусственного интеллекта

## 1. Понятие искусственного интеллекта

Термин интеллект происходит от латинского *intellectus* – что означает ум, рассудок, разум; мыслительные способности человека.

Искусственный интеллект (ИИ) – наука и технология создания интеллектуальных машин и систем, особенно интеллектуальных компьютерных программ, направленных на то, чтобы понять человеческий интеллект.

Общая структура интеллектуальной системы состоит из совокупности блоков (база знаний, решатель, естественно-языковой интерфейс пользователя) и связей между ними.

**База знаний** представляет собой совокупность сред, хранящих знания различных типов. Рассмотрим кратко их назначение.

- *База фактов* (данных) хранит конкретные данные, а • база правил — элементарные выражения, называемые в теории искусственного интеллекта *продукциями*.

- *База процедур* содержит прикладные программы, с помощью которых выполняются все необходимые преобразования и вычисления.

- *База закономерностей* включает различные сведения, относящиеся к особенностям той среды, в которой действует система.

- *База метазнаний* (база знаний о себе) содержит описание самой системы и способов ее функционирования: сведения о том, как внутри системы представляются единицы информации различного типа, как взаимодействуют различные компоненты системы, как было получено решение задачи.

- *База целей* содержит целевые структуры, называемые сценариями, позволяющие организовать процессы движения от исходных фактов, правил, процедур к достижению той цели, которая поступила в систему от пользователя, либо была сформулирована самой системой в процессе ее деятельности в проблемной среде.

Управление всеми базами, входящими в базу знаний, и организацию их взаимодействия осуществляет система управления базами знаний. С ее же помощью реализуются связи баз знаний с внешней средой. Таким образом, машина базы знаний осуществляет первую функцию интеллектуальной системы.

Выполнение второй функции обеспечивает часть интеллектуальной системы, называемая **решателем** и состоящая из ряда блоков, управляемых системой управления решателя. *Блок дедуктивного вывода* осуществляет в

решателе дедуктивные рассуждения, с помощью которых из закономерностей из базы закономерностей, фактов из базы фактов и правил из базы правил выводятся новые факты. Кроме этого данный блок реализует эвристические процедуры поиска решений задач, как поиск путей решения задачи по сценариям при заданной конечной цели. Для реализации рассуждений, которые не носят дедуктивного характера, т.е. для поиска по аналогии, по прецеденту и пр., используются *блоки индуктивного и правдоподобного выводов*. Блок планирования используется в задачах планирования решений совместно с блоком дедуктивного вывода. Назначение *блока функциональных преобразований* состоит в решении задач расчетно-логического и алгоритмического типов.

Третья функция — функция общения — реализуется как с помощью компоненты *естественно-языкового интерфейса*, так и с помощью рецепторов и эффекторов, которые осуществляют так называемое невербальное общение и используются в интеллектуальных роботах.

## 2. Методы искусственного интеллекта

В философии не решён вопрос о природе и статусе человеческого интеллекта. Нет и точного критерия достижения компьютерами «разумности», хотя на заре искусственного интеллекта был предложен ряд гипотез, например, **тест Тьюринга** или **гипотеза Ньюэлла — Саймона**.

- Гипотеза Ньюэлла-Саймона, формулировка которой выглядит следующим образом: физическая символическая система имеет необходимые и достаточные средства для того, чтобы производить осмысленные действия. Другими словами, без символических вычислений невозможно выполнять осмысленные действия, а способность выполнять символические вычисления вполне достаточна для того, чтобы быть способным выполнять осмысленные действия. Независимо от того, справедлива ли эта гипотеза, символические вычисления стали реальностью, и полезность этой парадигмы для программирования трудно отрицать.

- Тест Тьюринга – мысленный эксперимент, предложенный в качестве критерия и конструктивного определения интеллектуальности. Тест должен проводиться следующим образом. Судья (человек) переписывается на естественном языке с двумя собеседниками, один из которых – человек, а другой – компьютер. Если судья не может надёжно определить, кто есть кто, считается, что компьютер прошёл тест. Предполагается, что каждый из собеседников стремится, чтобы человеком признали его.

Поэтому несмотря на наличие множества подходов как к пониманию задач ИИ, так и созданию интеллектуальных информационных систем можно выделить два основных подхода к разработке ИИ:

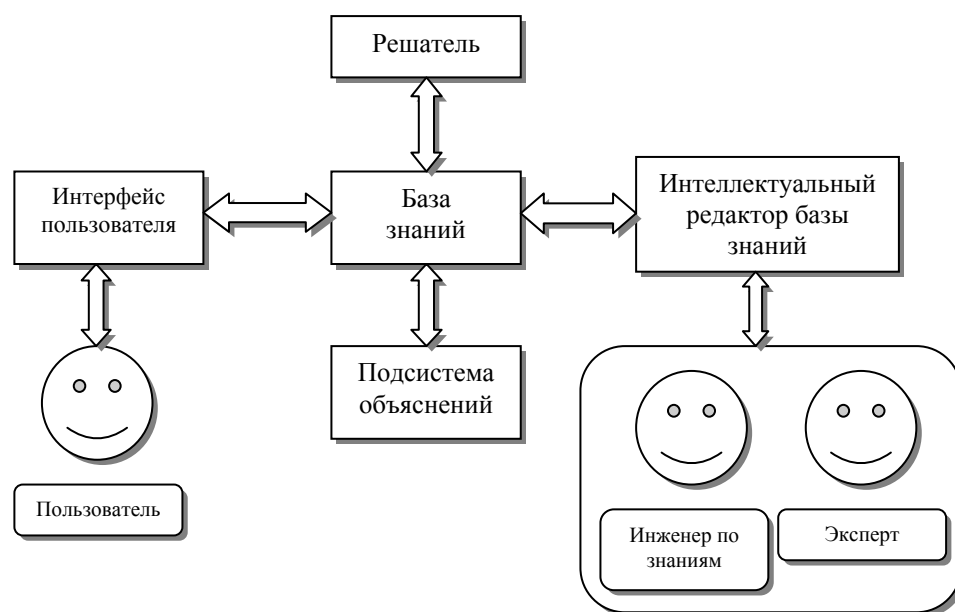
- нисходящий (англ. Top-Down AI), семиотический — создание экспертных систем, баз знаний и систем логического вывода, имитирующие высокоуровневые психические процессы: мышление, рассуждение, речь, эмоции, творчество и т. д.;

- восходящий (англ. Bottom-Up AI), биологический — изучение нейронных сетей и эволюционных вычислений, моделирующих интеллектуальное поведение на основе более мелких «неинтеллектуальных» элементов.

**3. Первый подход** реализуется в следующих методах и системах:

**3.1. Экспертная система (ЭС)** – компьютерная программа, которая действуя по определенным правилам, обрабатывает большое количество информации, и в результате выдает заключение или рекомендацию на ее основе. Структура экспертной системы может быть выражена следующей схемой (рис. 1.).

Рис. 1. Обобщенная структура экспертной системы



Главным элементом экспертной системы является база знаний (БЗ), состоящая из правил анализа информации от пользователя по конкретной проблеме.

Решатель, называемый также блоком логического вывода, представляет собой программу, моделирующую ход рассуждений эксперта на основании знаний, содержащихся в БЗ.

Подсистема объяснений – программа, позволяющая пользователю получать ответы на вопросы: «Как была получена та или иная рекомендация?» и «Почему система приняла то или иное решение?» Ответ на вопрос «как» – это трассировка всего процесса получения решения с

указанием использованных фрагментов БЗ, т.е. всех шагов цепи умозаключений. Ответ на вопрос «почему» – ссылка на умозаключение, непосредственно предшествовавшее полученному решению, т.е. отход на один шаг назад.

ЭС создается при помощи инженеров по знаниям (аналитиков), которые разрабатывают ядро ЭС и, зная организацию базы знаний, заполняют ее при помощи эксперта по специальности.

Интеллектуальный редактор БЗ – программа, предоставляющая инженеру по знаниям возможность создавать БЗ в диалоговом режиме.

Интерфейс пользователя – комплекс программных средств, реализующих диалог пользователя с ЭС как для ввода информации, так и для получения результатов работы ЭС.

Задачи, решаемые при помощи экспертных систем, чаще всего относятся к одной из следующих областей:

- Интерпретация данных – это одна из традиционных задач для экспертных систем. Под интерпретацией понимается определение смысла данных, результаты которого должны быть согласованными и корректными. Примеры существующих ЭС: SIAP (обнаружение и идентификация различных типов океанских судов), АВТАНТЕСТ, МИКРОЛЮШЕР (определение основных свойств личности по результатам психодиагностического тестирования).

- Диагностика – это обнаружение неисправности в некоторой системе. Трактовка неисправности как отклонения от нормы позволяет с единых теоретических позиций рассматривать и неисправность оборудования в технических системах, и заболевания живых организмов, и всевозможные природные аномалии. Примеры существующих ЭС: ANGY (диагностика и терапия сужения коронарных сосудов), CRIB (диагностика ошибок в аппаратуре и математическом обеспечении компьютера).

- Мониторинг – это непрерывная интерпретация данных в реальном масштабе времени и сигнализация о выходе тех или иных параметров за допустимые пределы. Главные проблемы – «пропуск тревожной ситуации» и инверсная задача «ложного» срабатывания. Сложность этих проблем состоит в размытости симптомов тревожных ситуаций и необходимости учета временного контекста. Примеры существующих ЭС: СПРИНТ (контроль за работой электростанций), REACTOR (помощь диспетчерам атомного реактора), FALCON (контроль аварийных датчиков на химическом заводе).

- Проектирование состоит в подготовке спецификаций на создание «объектов» с заранее определенными свойствами. Под спецификацией понимается весь набор необходимых документов – чертеж, пояснительная записка и т.д. Примеры существующих ЭС: XCON (проектирование конфигураций ЭВМ), CADHELP (проектирование БИС), SYN (синтез электрических цепей).

- Прогнозирование – это логический вывод вероятных следствий из заданных ситуаций. В прогнозирующей системе обычно используется параметрическая динамическая модель, в которой значения параметров «подгоняются» под заданную ситуацию. Выводимые из этой модели следствия составляют основу для прогнозов с вероятностными оценками. Примеры существующих ЭС: WILLARD (предсказание погоды), PLANT (оценки будущего урожая), ECON (экономические прогнозы).

- Планирование – нахождение планов действий, относящихся к объектам, способным выполнять некоторые функции. В таких ЭС используются модели поведения реальных объектов с тем, чтобы логически вывести последствия планируемой деятельности. Примеры существующих ЭС: STRIPS (планирование поведения робота), ISIS (планирование промышленных заказов), MOLGFN (планирование эксперимента).

- Обучение – процесс диагностирования ошибки при изучении какой-либо дисциплины с помощью компьютера и подсказывают правильные решения. Они аккумулируют знания о гипотетическом «ученике» и его характерных ошибках, а затем в ходе работы способны диагностировать слабости в знаниях обучаемых и находить соответствующие средства их ликвидации. Кроме того, они планируют процесс общения с учеником в зависимости от успехов ученика с целью передачи знаний. Примеры существующих ЭС: PROUST (обучение языку программирования Pascal).

3.2. *CBR-системы (рассуждение по аналогии)* представляют собой реализацию методологии искусственного интеллекта, применяемую при построении компьютеризированных консультационных систем, которые базируются на накопленном опыте. В отличие от классических экспертных систем, действующих на основе логических правил, CBR-системы хранят успешные решения ряда реальных проблем, называемых примерами или прецедентами, и при появлении новой проблемы находят по определенному алгоритму (зачастую при помощи машины логического вывода, с количественной оценкой) наиболее подходящие (похожие) прецеденты, после чего предлагает соответственно модифицированную комбинацию их решений. Если новая проблема оказывается таким образом успешно решенной, это решение заносится в базу прецедентов для повышения эффективности работы системы в будущем. Главный недостаток CBR-систем состоит в том, что они не создают моделей или правил, обобщающих накопленный опыт.

3.3. **Байесовская сеть доверия** – это вероятностная модель, представляющая собой множество переменных и их вероятностных зависимостей. Например, байесовская сеть может быть использована для вычисления вероятности того, чем болен пациент по наличию или отсутствию ряда симптомов, основываясь на данных о зависимости между

симптомами и болезнями. Существуют эффективные методы, которые используются для вычислений и обучения байесовских сетей. Байесовские сети используются для моделирования в биоинформатике (генетические сети, структура белков), медицине, классификации документов, обработке изображений, обработке данных и системах принятия решений.

**3.4. Поведенческий** (модульный) метод построения систем ИИ, при котором система разбивается на несколько сравнительно автономных программ поведения, которые запускаются в зависимости от изменений внешней среды

**4. Второй подход** реализуется в следующих основных методах:

**4.1. Нейронные сети:** коннекционистские модели нервной системы, демонстрирующие, в частности, высокие способности к распознаванию образов.

Нейронная сеть (НС) – это распределенный параллельный процессор, состоящий из элементарных единиц обработки информации, накапливающих экспериментальные знания и предоставляющих их для последующей обработки. Она представляет собой действующую модель нервной системы и сходна с мозгом с двух точек зрения:

Знания поступают в нейронную сеть из окружающей среды и используются в процессе обучения.

Для накопления знаний применяются связи между нейронами, называемые синаптическими весами.

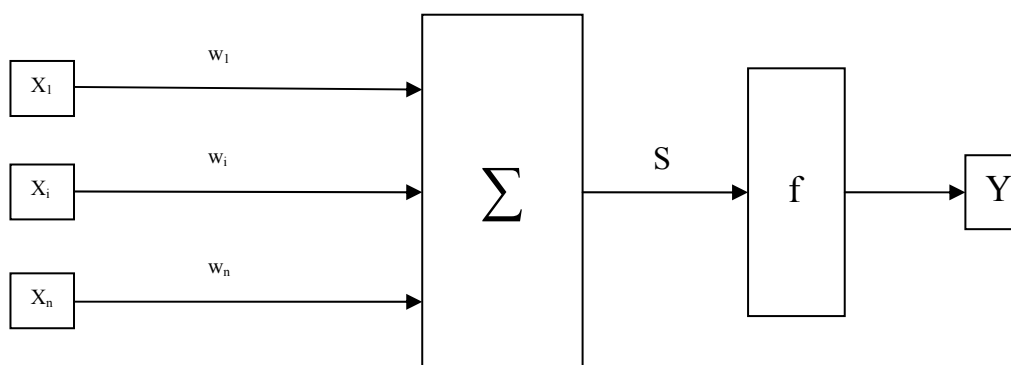


Рис. 2. Структурная схема нейрона

Сигналы  $X_i$ , поступающие на вход нейрона, умножаются на соответствующие весовые коэффициенты  $w_i$ , после чего суммируются.

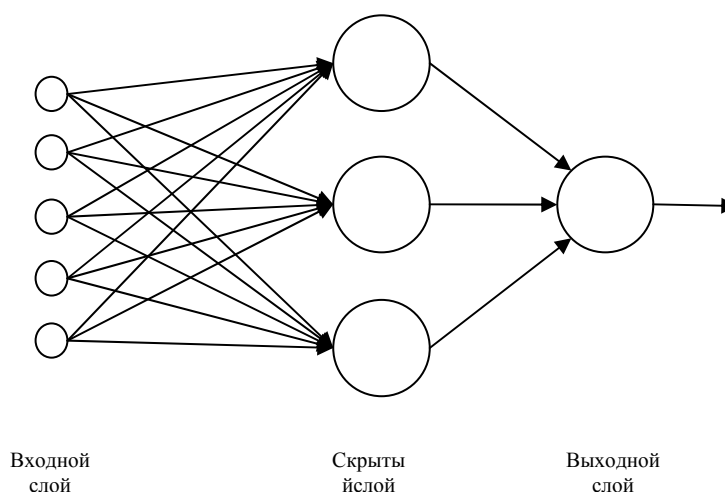


Рис. 3. Структура слоистой нейронной сети

Результат суммирования поступает на нелинейный преобразователь, реализующий некоторую нелинейную функцию, называемую функцией активации или передаточной функцией нейрона: результат ее действия поступает на выход нейрона. Существует множество способов построения нейронных сетей из отдельных нейронов. Наиболее распространенной является слоистая архитектура НС, обобщенная схема которой представлена на рис. 3. Сигналы входного слоя поступают на входы нейронов первого слоя, после прохождения которого распространяются дальше, пока не достигнут выходов нейронной сети. Для обучения нейронных сетей такого типа используется, как правило, т.н. алгоритм обратного распространения ошибки, позволяющий рассчитать изменения весовых коэффициентов, необходимые для того, чтобы согласовать выходные значения НС с выборкой образцов. Использование нейронных сетей обеспечивает следующие полезные свойства систем:

- **Нелинейность.** Это качество нейронной сети особенно важно в том случае, если сам физический механизм, отвечающий за формирование входного сигнала, сам является нелинейным (например, человеческая речь).
- **Адаптивность.** Нейронные сети обладают способностью адаптировать свои синаптические веса к изменениям окружающей среды. Для работы в нестационарной среде могут быть созданы нейронные сети, изменяющие синаптические веса в реальном времени.
- **Контекстная информация.** Знания представляются в самой структуре нейронной сети. Каждый нейрон сети потенциально может быть подвержен влиянию всех остальных ее нейронов.
- **Отказоустойчивость.** Аппаратно реализованные нейронные сети потенциально отказоустойчивы. Это значит, что при неблагоприятных условиях их производительность падает незначительно. Например, если

поврежден какой-то нейрон или его связи, извлечение запомненной информации затрудняется. Однако, принимая в расчет распределенный характер хранения информации в нейронной сети, можно утверждать, что только серьезные повреждения структуры нейронной сети существенно повлияют на ее работоспособность.

Представим некоторые проблемы, решаемые применением нейронных сетей:

- **Классификация образов.** Задача состоит в указании принадлежности входного образа, представленного набором признаков, одному или нескольким предварительно определенным классам. К известным приложениям относятся распознавание букв, распознавание речи, классификация сигнала электрокардиограммы и т.п.

- **Кластеризация/категоризация.** Кластеризация основана на подобию образов: НС размещает близкие образы в один кластер. Известны случаи применения кластеризации для извлечения знаний, сжатия данных и исследования свойств данных.

- **Аппроксимация функций.** Задача аппроксимации состоит в нахождении оценки некоторой искаженной шумом функции, генерирующей обучающую выборку.

- **Предсказание/прогноз.** Задача прогнозирования состоит в предсказании некоторого значения для заданного момента времени на основании ряда значений, соответствующим другим моментам времени.

- **Оптимизация.** Задачей оптимизации является нахождение решения, которое удовлетворяет системе ограничений и максимизирует или минимизирует целевую функцию.

- **Ассоциативная память.** Данная сфера применения НС состоит в организации памяти, адресуемой по содержанию, позволяющей извлекать содержимое по частичному или искаженному образцу.

Дальнейшее повышение производительности компьютеров все в большей степени связывают с НС, в частности, с нейрокомпьютерами, основу которых также составляют аппаратно реализованные нейронные сети.

**4.2 Нечеткие системы:** методики для рассуждения в условиях неопределенности.

Направление базируется на принципах нечеткой логики и теории нечетких множеств – раздела математики, являющегося обобщением классической логики и теории множеств. Данные понятия были впервые предложены американский ученым Лотфи Заде в 1965 г. Основной причиной появления новой теории стало наличие нечетких и приближенных рассуждений при описании человеком процессов, систем, объектов. В общем случае механизм логического вывода в рамках нечеткой логики включает в себя четыре этапа: введение нечеткости (фазификация), нечеткий вывод,



композиция и приведение к четкости или дефазификация (см. рис. 4). Алгоритмы нечеткого вывода различаются главным образом видом используемых правил, логических операций и разновидностью метода дефазификации.

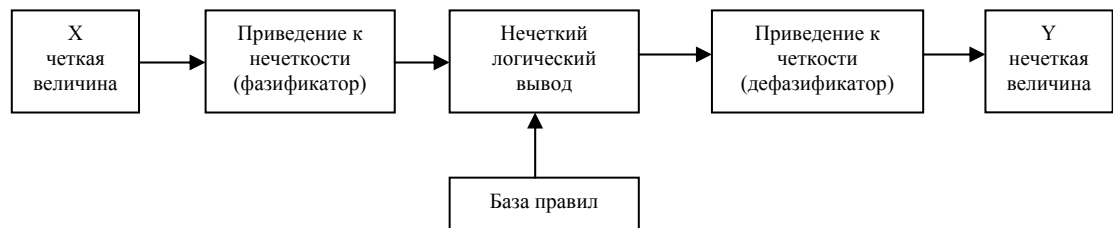


Рис. 4. Механизм нечеткого логического вывода

Нечеткая логика оказала сильное влияние на другие парадигмы искусственного интеллекта. Объединение ее принципов с методами иных направлений породило такие новые направления, как:

- Нечеткие нейронные сети
- Адаптивные нечеткие системы
- Нечеткие запросы
- Нечеткие ассоциативные правила
- Нечеткие когнитивные карты
- Нечеткая кластеризация

Альтернативные методы искусственного интеллекта дополняют методологию нечеткой логики и используются в различных комбинациях для создания гибридных интеллектуальных систем.

**4.3. Эволюционные вычисления:** модели, использующие понятие естественного отбора, обеспечивающего отсеивание наименее оптимальных согласно заданному критерию решений. В этой группе методов выделяют генетические алгоритмы и т.н. муравьиный алгоритм.

Генетический алгоритм, составляющий основу эволюционных вычислений, – это эвристический алгоритм поиска, используемый для решения задач оптимизации и моделирования путем последовательного подбора, комбинирования и вариации искомых параметров с использованием механизмов, напоминающих биологическую эволюцию. Сущность алгоритма состоит в следующем. Задача кодируется таким образом, чтобы ее решение могло быть представлено в виде вектора (такой вектор называется хромосомой). Случайным образом создается некоторое количество начальных векторов (начальная популяция). Они оцениваются с использованием т.н. функции приспособленности, в результате чего каждому вектору присваивается определенное значение (приспособленность), которое

определяет вероятность выживания организма, представленного данным вектором. После этого с использованием полученных значений приспособленности выбираются вектора (селекция), допущенный к скрещиванию. К этим векторам применяются т.н. генетические операторы (в большинстве случаев – скрещивание и мутация), создавая таким образом следующее поколение. Особи следующего поколения также оцениваются, затем производится селекция, применяются генетические операторы и т.д. Так моделируется эволюционный процесс, продолжающийся несколько жизненных циклов (поколений), пока не будет выполнен критерий останова алгоритма.

Генетические алгоритмы служат, главным образом, для поиска решений в очень больших, сложных пространствах поиска, и применяются для решения следующих задач:

- Оптимизация функций
- Разнообразные задачи на графах (задача коммивояжера, раскраска, нахождение паросочетаний)
- Настройка и обучение нейронной сети
- Задачи компоновки
- Составление расписаний
- Игровые стратегии
- Аппроксимация функций
- Искусственная жизнь
- Биоинформатика

Сверхзадачей искусственного интеллекта является построение компьютерной интеллектуальной системы, которая обладала бы уровнем эффективности решений неформализованных задач, сравнимым с человеческим или превосходящим его.

### **Задание для самостоятельной работы**

1. Какие существуют разновидности интеллектуальных систем?
2. Каковы особенности и основные свойства гибридных экспертных систем?