

Системы искусственного интеллекта

02.03.03 - Математическое обеспечение и администрирование информационных систем, направленность (профиль) - разработка и администрирование информационных систем

09.03.03 - Прикладная информатика, направленность (профиль) - прикладная информатика в экономике

<http://vikchas.ru>

Тема 1. Введение в искусственный интеллект и машинное обучение Лекция 1 «Общее определение искусственного интеллекта»

Часовских Виктор Петрович

д-р техн. наук, профессор кафедры ШИиКМ

ФГБОУ ВО «Уральский государственный экономический университет»

Екатеринбург 2023

Об искусственном интеллекте

Искусственный интеллект — это сложное понятие, для которого не существует однозначного определения.

Понятие искусственного интеллекта (для удобства сокращают как «ИИ») используется специалистами в различных областях: писателями, журналистами, в бизнесе и науке; и разные специалисты вкладывают свой смысл в это понятие.

Мир сегодня разделен надвое в своем отношении к ИИ: одни видят в нем чуть ли не мессию, который спасет наш мир от всех бед и напастей, а другие, наоборот, воспринимают сродни всадникам Апокалипсиса (четыре всадника Апокалипсиса — персонажи из шестой главы Откровения Иоанна Богослова в Библии), которые грозят уничтожить человеческую цивилизацию и превратить нас в рабов машин.

В самом широком смысле искусственным интеллектом называют способность компьютера решать те же интеллектуальные задачи, которые способен решать человек.

Указанное понятие можно конкретизировать на разных уровнях:

- машина, способная воспринимать и понимать мир через сенсоры (например, анализ изображений и звука);
- способная придумывать и создавать новые объекты (например, изображения, видео и тексты);
- способная решать интеллектуальные задачи (например, игра шахматы или го);
- или способная переключаться между задачами и творчески решать сложные интеллектуальные задачи.

При определении искусственного интеллекта надо отталкиваться от понятия «естественный интеллект», то есть способности мозга человека решать интеллектуальные задачи путем приобретения, запоминания и целенаправленного преобразования знаний, применения этих знаний для управления средой.

При этом под интеллектуальной задачей понимается задача, решение которой не было до сих пор известно, то есть отсутствовал алгоритм ее решения.

Исходя из этого, академик АН РФ И.Каляев предлагает следующую формулировку:

искусственный интеллект - это свойство искусственных систем решать интеллектуальные задачи, для которых отсутствует алгоритм решения.

Отсюда следует важный вывод:

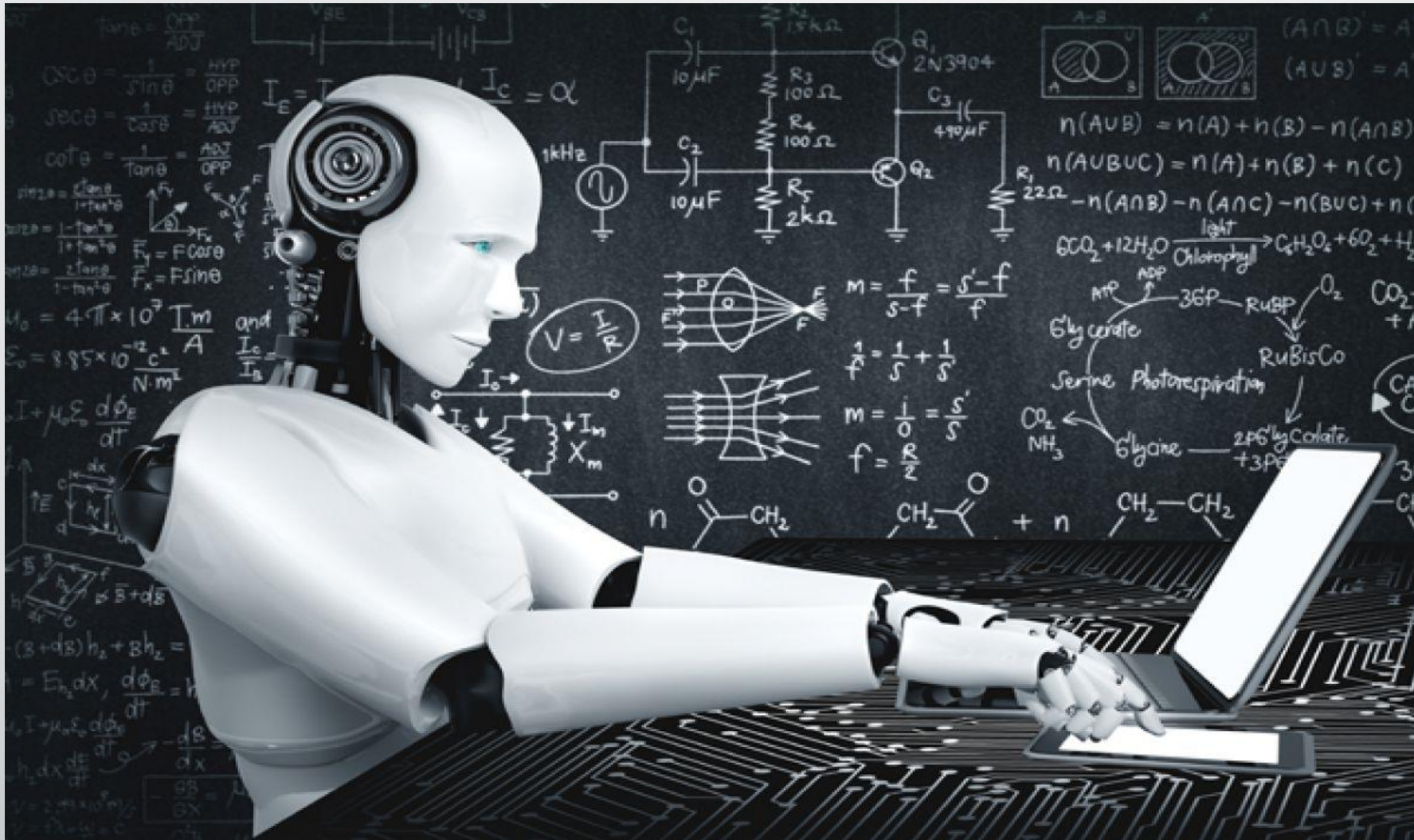
как только задача решена на компьютере, она перестает быть интеллектуальной, поскольку это означает, что для нее уже создан алгоритм решения (как известно, компьютер может работать только по алгоритму).

Следовательно, большинство из того, что сегодня мы называем ИИ, никакого отношения к интеллекту не имеет, поскольку это просто различные компьютерные программы, работающие в соответствии с уже созданным человеком алгоритмом.

Это орудия труда, такие же, как молоток. Только он усиливает физические возможности человека, а компьютерные программы - интеллектуальные, умственные.

«Поэтому все, что сегодня называется искусственным интеллектом, - говорит академик АН РФ И.Каляев, - было бы корректно обозначить как интеллектуальные компьютерные технологии или псевдоинтеллектуальные компьютерные технологии.

Тем не менее термин «искусственный интеллект» по отношению к компьютерным программам устоялся, принят во всем мире, и мы тоже вынуждены его использовать».



В настоящее время эксперты различают так называемый слабый и сильный искусственный интеллект.

Слабый ИИ предназначается для решения какой-нибудь одной задачи (игра в шахматы, распознавание объектов на картинках, управление автомобилем и т.д. – прим. profiok.com). Это направление в сфере развития ИИ называется логическим. **Сильный искусственный интеллект** в идеале должен со временем научиться решать любые интеллектуальные задачи, то есть стать универсальным аналогом человеческого мозга. Это так называемый нейрокибернетический подход.

На сегодня существует только узкоспециализированный искусственный интеллект: машина, способная решать одну данную интеллектуальную задачу (например, распознавание лиц, игра в шахматы или машинный перевод). При этом неясно, возможно ли в принципе с использованием существующих технологий создать общий искусственный интеллект, то есть машину, которая сможет решать различные сложные интеллектуальные задачи.



СИЛЬНЫЙ ИСКУСТВЕННЫЙ ИНТЕЛЛЕКТ

Возьмем персонального помощника: общий ИИ подразумевает, что машина сможет заменить помощника целиком (будет выполнять различного рода поручения, планировать расписание, отвечать на звонки и так далее, а также сможет самостоятельно обучаться новым задачам). Существующие сегодня технологии гораздо проще: машина может записывать текст по речевому вводу, предлагать короткие ответы на входящие письма вида «Принято в работу. С уважением, (имя)», или устанавливать напоминания по текстовому или речевому описанию, причем за каждую перечисленную функцию отвечает отдельный алгоритм.

С другой стороны, создание алгоритмов выполнения даже этих задач, кажущихся несложными на контрасте с функционалом полноценного персонального помощника, исторически потребовало немало времени.

АН РФ – искусственный интеллект

2020 – 2022 годах проведено обсуждение основных проблем ИИ в форме видео-докладов (более 200 выступлений, 1,5 – 2,5 часа) Академиков АН РФ, Член-кор. АН РФ, профессоров, ведущих специалистов РФ в области ИИ

Заседания Сем-ра 06.04.21, 08.04.21, 20.05.21 (совм. с НСМИИ РАН), 27.05.21 (совм. НЕЙРОФИЛОСОФИЯ) ▶

Заседания Семинара

Образовательный проект НСМИИ РАН «Интеллектуальная информатика, математика, теория систем»
Международный междисциплинарный семинар «АЛГЕБРАИЧЕСКАЯ БИОЛОГИЯ И ТЕОРИЯ СИСТЕМ»
Толоконников Г.К., Алексеев А.Ю., Петухов С.В.
ПЛАТФОРМА ЕДИНОГО КУРСА ИИ И ТЕОРИИ СИСТЕМ
Васов В.Л.
КВАНТОВАЯ ЛОГИКА 2:56:34

Заседание Семинара Секции НСМИИ РАН...

Международный междисциплинарный семинар «АЛГЕБРАИЧЕСКАЯ БИОЛОГИЯ И ТЕОРИЯ СИСТЕМ»
Толоконников Г.К., Алексеев А.Ю., Петухов С.В.
ПЛАТФОРМА ЕДИНОГО КУРСА ИИ И ТЕОРИИ СИСТЕМ 2:38:17

Семинар "Единый курс ИИ и теории систем"...

Международный междисциплинарный семинар «АЛГЕБРАИЧЕСКАЯ БИОЛОГИЯ И ТЕОРИЯ СИСТЕМ»
Толоконников Г.К., Алексеев А.Ю., Журавский Е.Е., Зайцев А.А., Жданов А.А.
«Эволюция и получение эволюционных знаний в вычислительной среде методами ДСМ, семантического вероятностного вывода и автономного адаптивного управления»
Москва, 20 мая 2021 г.
2:37:00

Заседание Семинара 20.05.21, Алексеев А.Ю....

27 мая 2021 г., четверг, 14.00 - 16.30
Хреников А.Ю., доктор физико-математических наук, профессор, профессор математики факультета информатики и компьютерных наук, Санкт-Петербургский государственный университет
Квантовая модель бессознательно-сознательного взаимодействия и эмоциональной окраски сознательного опыта 2:26:40

Сем. "Алг.биол. и теор.систем" и...

Конференция 24.06.21 НСМИИ РАН при участии Семинара АБиТС "Искусственный интеллект: стратегии образования" ▶

ВОСПРОИЗВЕСТИ ВСЕ

Конференция 24.06.21 НСМИИ РАН при участии Семинара АБиТС "Искусственный интеллект: стратегии образования"

ЛЕКТОРСКИЙ
Владислав Александрович
академик РАН
Председатель Научного Совета по искусственному интеллекту при Президиуме РАН (НСМИИ РАН)
2:15

выступление академика Лекторского В.А.

СИГОВ
Александр Сергеевич
академик РАН
Президент МНРЭА
2:44

выступление академика Сигова А.С.

МАКАРОВ
Валерий Леонидович
академик РАН, доктор физико-математических наук, профессор, научный руководитель Центрального экономико-математического института РАН, заместитель председателя НСМИИ РАН, научный руководитель направления Искусственные общества НСМИИ РАН, г. Москва
8:15

выступление академика Макарова В.Л.

СЕМЁНОВ
Алексей Пылович
академик РАН и академик РАН, доктор физико-математических наук, профессор, заведующий кафедрой математической логики и теории алгоритмов МГУ им. М.В. Ломоносова, директор Института кибернетики и образовательной информатики им. А.И. Берга Федерального исследовательского центра «Информатика и управление» РАН
7:23

приветствие академика Семенова А.Л.

ГОНЧАРОВ
Сергей Савастьянович
академик РАН, доктор физико-математических наук, профессор, директор Института математики Сибирского отделения РАН, г. Новосибирск
6:30

выступление академика Гончарова С.С.

ЧЕРНИГОВСКАЯ
Татьяна Владимировна
доктор биологических наук, доктор филологических наук, профессор, член бюро НСМИИ РАН, директор Института когнитивных исследований заведующая кафедрой проблем конвергенции естественных и гуманитарных наук Санкт-Петербургского государственного университета
12:14

выступление проф. Черниговской Т.В.

Общая история искусственного интеллекта

Рассмотрим краткую историю создания «умных» алгоритмов и того, как менялось понятие искусственного интеллекта.

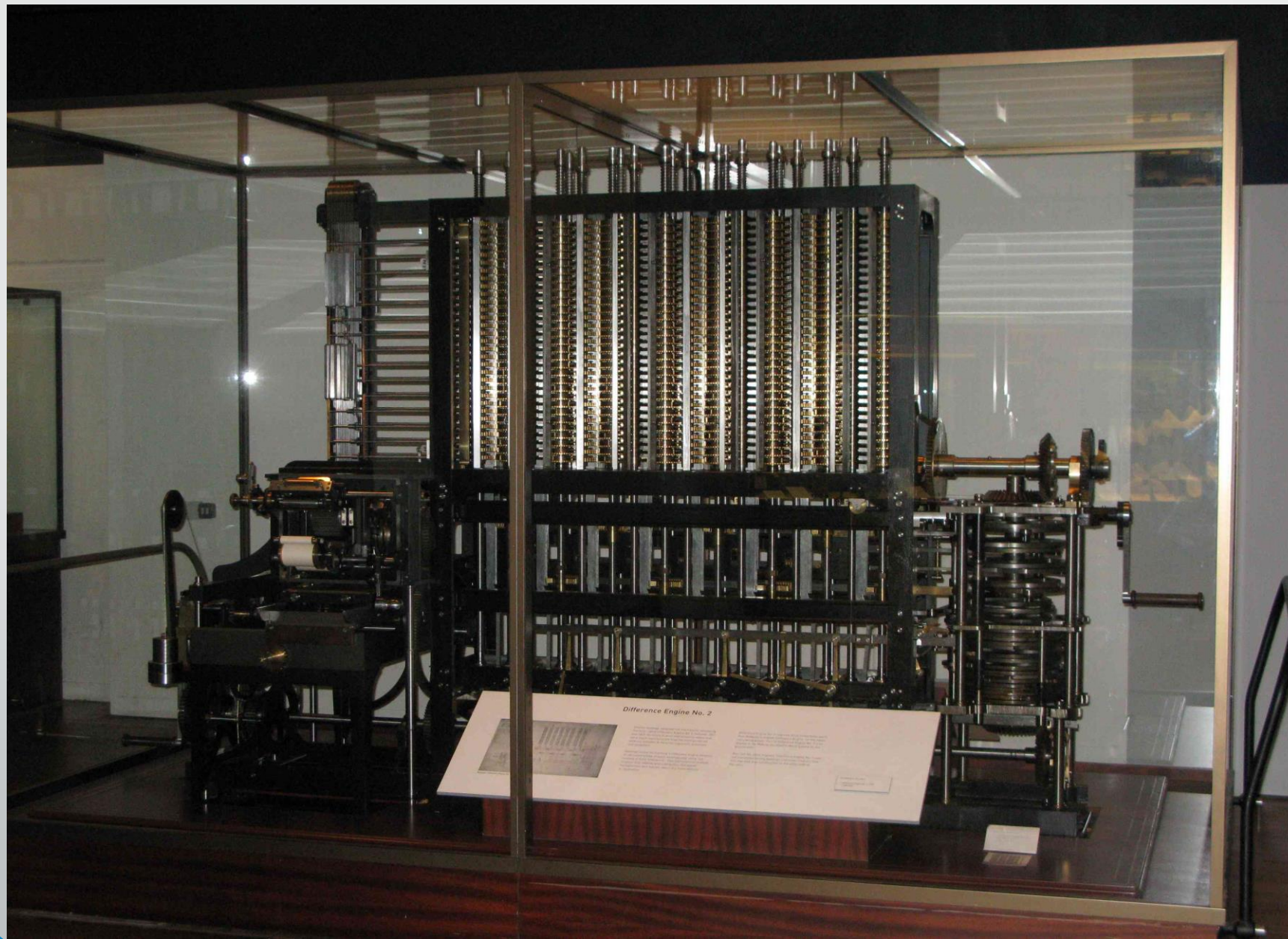
Первая половина 19 века: Ада Лавлейс

Ада Лавлейс, которая создала (в первой половине 19 века) описание вычислительной машины Бэббиджа и написала первую в мире программу для этой машины. Она считается первым программистом в истории человечества. Машины, например машины Бэббиджа, могут выполнять операции в соответствии с инструкциями, считала она, но они не могут самостоятельно выдвигать идеи или иметь намерения. «Аналитическая машина не претендует на создание чего-то своего, — писала она в своих “Примечаниях”, — она может выполнить любую команду, которую мы сумеем задать. Она может провести анализ, но от нее никак нельзя ожидать вывода каких-либо аналитических соотношений или установления законов». Столетие спустя один из создателей первых компьютеров — Алан Тьюринг — назвал это утверждение «Возражением леди Лавлейс». <...> Тьюринг задался вопросом: если машина может изменить свою собственную программу на основе обрабатываемой ею информации, не форма ли — это обучения? Не может ли это привести к созданию искусственного интеллекта?

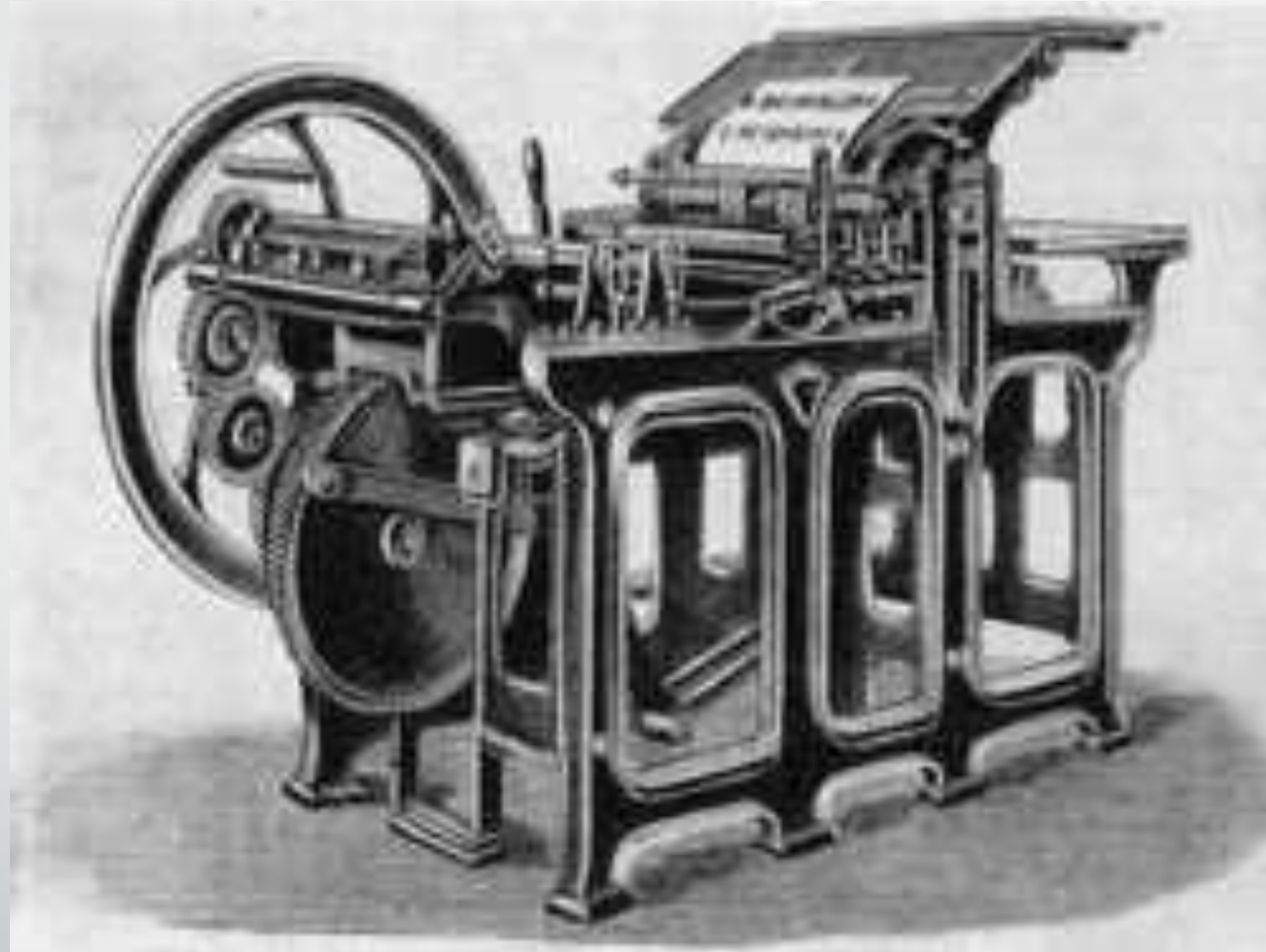
Язык программирования Ада.

День рождения Лавлейс 10 декабря. День программиста 10 декабря.

Вычислительная машина Беббиджа



Принтер Беббиджа



1940-е: тест Тьюринга

С момента разработки в 40-х годах прошлого века первых цифровых электронно-вычислительных машин возник вопрос о границах возможностей компьютеров и их способностях достигнуть уровня развития человека.

В 1950 году английский ученый Алан Тьюринг описал идею теста Тьюринга — формального способа определить, может ли машина выполнять мыслительную деятельность по аналогии с человеком. Суть теста в следующем: человек с помощью текстовых сообщений задает вопросы компьютеру и другому человеку, не зная, кто из них, кто. Задача человека — определить, какие ответы приходят от компьютера, а какие — от человека.

В XX веке было предпринято много попыток пройти тест Тьюринга, например программа Элиза (Джозеф Вейценбаум, 1966 год), однако ни одна из них к цели значительно не приблизилась. Сегодня практическим проведением теста Тьюринга занимаются различные организации по всему миру, конкретные правила проведения теста варьируются в зависимости от площадки (длительность теста, сколько раз компьютеру нужно «обмануть» человека и т. д.).

В 2014 году тест, организованный на базе Лондонского королевского общества, был пройден программой «Женя Густман» (Владимир Веселов, Санкт-Петербург), имитирующим ребенка (которому допускается не знать каких-то вещей), который обманул жюри в 33% случаев (при пороге прохождения теста в 30% случаев). Тест Тьюринга нередко подвергается критике, в частности за то, что он помогает определить, смогла ли машина имитировать человека, но не оценивает ее настоящие способности решать интеллектуальные задачи.

1943 год - модель нейрона, предложенная Уорреном Мак-Каллоком и Уолтером Питтсом.

Мак-Каллок и Питтс в своей статье предприняли попытку описать работу нейронных механизмов мозга, используя аппарат математической логики. Их модель была чисто теоретической попыткой представить мозг как совокупность связанных в сеть логических элементов. К этому подталкивало то, что живой нейрон похож, а при определенных ограничивающих предположениях и становится логическим пороговым элементом, имеющим фиксированное число двоичных входов и двоичный выход.

Свойства нейронов и нейронной сети определялись следующими аксиомами:

Возбуждение нейрона соответствует принципу "все или ничего".
Время делится на дискретные моменты - такты.

Возбуждение нейрона в какой-то момент времени происходит, если в предшествующий момент времени произошли возбуждения определенного фиксированного числа синапсов. Это число не зависит ни от предыдущей активности, ни от расположения синапсов на нейроне.

(Синапс (от греч. synapsis соединение) - область контакта (связи) нервных клеток (нейронов) друг с другом и с клетками исполнительных органов. Термин был введен в 1897 г. английским физиологом Чарльзом Шеррингтоном. Большой Энциклопедический словарь).

Возбуждение по связи от одного нейрона к любому другому происходит без задержки (за один такт).

Синапсы могут быть как возбуждающими, так и тормозящими.

Входной сигнал, прошедший через тормозящий синапс, абсолютно исключает возбуждение данного нейрона в рассматриваемый момент времени.

С течением времени структура сети не изменяется.

Удовлетворяющие приведенной аксиоматике формальные нейроны и нейронные сети Мак-Каллока и Питтса обладают следующими свойствами.

Во-первых, доказано, что эти формальные нейроны могут реализовать **любую двоичную логическую функцию**.

Во-вторых, доказано, что из этих формальных нейронов можно построить сеть, реализующую **любое высказывание логики высказываний**.

Что из этого следует? Пересказывая фон Неймана, можно сказать, что если функционирование системы заключается в формировании реакций на ее выходе в зависимости от входов, то с помощью формальной нейронной сети Мак-Каллока и Питтса в этой системе возможна реализация любого такого функционирования, которое можно точно и однозначно описать конечным числом слов.

Работы этих ученых показали, например, что любая вычислимая функция может быть вычислена с помощью некоторой сети из соединенных нейронов и что все логические связки (**“И” “ИЛИ”, “НЕ” и т.д.) могут быть реализованы с помощью простых сетевых структур.**

Кроме того, Мак-Каллок и Питтс выдвинули предположение, что сети, структурированные соответствующим образом, **способны к обучению.**

1950-е: персептрон Розентблатта

В 50-х годах XX века американский нейрофизиолог Фрэнк Розенблатт разработал **персептрон** — математическую модель, будто бы моделирующую восприятие информации человеческим мозгом.

Позднее выяснилось, что человеческий мозг устроен значительно сложнее, и персептрон имеет с ним мало общего. Персептрон Розенблатта был реализован несколькими годами позднее на компьютере IBM и был способен распознавать рукописные буквы латинского алфавита, за что заслужил много внимания со стороны бизнеса и государства и породил надежды на будущие разработки.

Однако продвинуться далеко в совершенствовании персептрона не удалось, и в последующие годы получить гранты на исследования в области искусственного интеллекта оказалось сложно — период, часто называемый «**зимой искусственного интеллекта**».

1956 год: искусственный интеллект

На 2-х месячном семинаре в Дартмуте летом 1956 года, при участии Джона Маккарти, Марвина Минского, Клода Шеннона, Натаниэля Рочестера, Тренчарда Мура, Артура Самюэла, Рея Соломонова, Оливера Селфриджа Аллен Ньюэлл и Герберт Саймон рассмотрев различные логические программы и язык обработки списков приняли соглашение о новом названии, предложенное Маккарти, — **искусственный интеллект**.

Из анализа докладов семинара можно понять, с чем связана необходимость преобразовать **искусственный интеллект** в отдельную область знаний.

Почему нельзя было бы публиковать все работы, выполненные в рамках искусственного интеллекта, под флагом теории управления, или исследования операций, или теории решений, которые в конечном итоге имеют цели, аналогичные искусственному интеллекту?

Или почему искусственный интеллект не рассматривается как область математики? Ответом на эти вопросы, во-первых, является то, что искусственный интеллект с самого начала впитал идею моделирования таких человеческих качеств, как творчество, самосовершенствование и использование естественного языка.

Эти задачи не рассматриваются ни в одной из указанных областей.

Во-вторых, еще одним ответом является методология.

Искусственный интеллект - это единственная из перечисленных выше областей, которая, безусловно, является одним из направлений компьютерных наук (хотя в исследовании операций также придается большое значение компьютерному моделированию), кроме того, искусственный интеллект - это единственная область, в которой предпринимаются попытки создания машин, действующих автономно в сложной, изменяющейся среде.

1969 – 1979 год: системы, основанные на знаниях

Основной подход к решению задач, сформированный в течение первого десятилетия исследований в области искусственного интеллекта, представлял собой механизм поиска общего назначения, с помощью которого предпринимались попытки связать в единую цепочку элементарные этапы проведения рассуждений для формирования полных решений.

Все относящиеся к делу теоретические знания, требуемые для решения указанных проблем, были преобразованы из наиболее общей формы (из “исходных принципов”) в эффективные специальные формы (в “рецепты поваренной книги”).

Первая зима ИИ (1974-1980). Сдерживающие факторы - отсутствие методов и алгоритмов, адекватных сложности поставленных задач - исключение ИИ из числа национальных приоритетов, значительное сокращение финансирования нейронных сетей

Возобновление работ по ИИ (1981 – 1986) –ЭВМ 5-го поколения (поддержка диалога, перевод языков, интерпретация изображений, построение причинно-следственных связей, Япония), экспертные систем, генетические алгоритмы

1980-е: экспертные системы

В начале 80-х годов получили широкое распространение **экспертные системы** — программы, имитирующие знания и умения человека.

Для создания таких программ проводился подробный опрос экспертов с целью выяснить, каким именно образом они решают задачи, например, по каким правилам психолог отвечает пациенту или врач определяет диагноз. Полученные правила затем реализовывались в виде компьютерной программы.

Однако и эти системы быстро достигли своего потолка, положив начало второй «зиме искусственного интеллекта». Тем не менее, экспертные системы по-прежнему используются в некоторых областях, особенно в тех, в которых особое внимание уделяется надежности, например в медицине.

Вторая зима (1987-1993) ИИ и ИНС. Сдерживающие факторы - ограниченные вычислительные возможности - ограниченные возможности методов машинного обучения и анализа данных - высокие требования к разработчикам, значительная трудоемкость создания технологий - ограниченный объем данных для обучения и настройки систем, слабая обобщающая способность; - значительное сокращение финансирования.

Возобновление работ по ИИ (1994 г. – н.в.)

1990-е: Deep Blue

В конце 90-х годов XX века и начале XXI века драйверами развития искусственного интеллекта стали увеличивающиеся вычислительные мощности, активное применение математических методов, а также упор на решение конкретных задач.

Искусственный интеллект начал использоваться в различных областях: в логистике, экономике, медицине. Одним из известных прорывов стала разработка суперкомпьютера Deep Blue, обыгравшего в 1997 году действующего чемпиона мира по шахматам Гарри Каспарова.

В начале XXI века одним из основных драйверов разработки алгоритмов искусственного интеллекта стали данные, а акценты в сфере искусственного интеллекта сместились в сторону таких областей, как машинное обучение, анализ больших данных и др. В отличие от экспертных систем, формализующих человеческий опыт в виде программы, машинное обучение извлекает знания из большой базы данных, благодаря чему получается более точный алгоритм, корректно работающий в более широком круге случаев.

2010-е: Watson и DeepMind

В 2011 году компания IBM представила систему [Watson](#), способную находить ответы на вопросы, заданные на естественном языке. Для демонстрации навыков система «сыграла» в игру «Jeopardy!» (американский аналог русскоязычной «Своей игры») и обыграла обоих соперников. В основе Watson — алгоритмы информационного поиска, обработки естественного языка, машинного обучения и построения логических цепочек.

В 2015 году британская компания DeepMind (один из крупнейших сегодняшних лидеров в области искусственного интеллекта) представила AlphaGo — систему, [обыгравшую](#) чемпиона мира по игре в Го, Ли Седоля. Го намного сложнее шахмат из-за большего количества возможных позиций, что создает сложности в применении алгоритмов поиска ходов, активно используемых в DeepBlue. Основу AlphaGo составляют алгоритмы машинного обучения: программа обучалась на данных 160 тысяч партий профессионалов.

2020: GPT-3, AlphaFold

В 2020 году американская компания OpenAI представила систему GPT-3, способную генерировать англоязычные тексты, неотличимые от написанных человеком, с сохранением темы по всему объему текста, грамматической корректностью и т. д.

В том же 2020 году компания DeepMind представила AlphaFold, решающую задачу прогноза третичной структуры белков, которая являлась одной из сложнейших и важнейших задач в современной биологии и решение которой не могли найти около 50 лет.

Наибольшие научные результаты в «естественном интеллекте» достигнуты физиологами и математиками.

Открытие и описание **нейронов** сделал в 1836 году немецко-швейцарский **физиолог** Габриэль Густав Валентин.

В 1943 году американские **нейрофизиолог** Уоррен Маккаллок и **математик** Уолтер Питтс предложили первую модель **искусственного нейрона** и основанную на нём **модель нейронной сети.**

Текущее состояние дел в ИИ и ИНС на лето 2022 года можно охарактеризовать следующим образом.

Во-первых, с точки зрения адаптивности все существующие системы ИИ на основе любых известных подходов предполагают функционирование в ограниченном наборе задач и условий и не способны самообучаться функционированию в условиях существенно новых. То есть они являются программируемыми, хотя сложность их программирования заметно снизилась, а способность к обучению возросла.

Во-вторых, с точки зрения автономности систем ИИ все существующие системы не являются автономными и **не могут** полноценно функционировать без живого **оператора**, отвечающего за запуск и остановку, техобслуживание, целеполагание и определение режимов работы в зависимости от тех или иных условий или задач. То есть они остаются управляемыми. Особенно это очевидно в областях, характеризующихся высокими рисками или высокой априорной неопределенностью (даже робот-пылесос без помощи человека не проживет и пары дней в нормальной квартире).

В-третьих, с точки зрения интегративности (*сделать отдельные программные компоненты, совместимыми с другими компонентами*) современные системы ИИ являются не системами, обладающими интеллектом как таковым (даже ограниченным), а

системами компьютерного зрения,
обработки естественного языка,
анализа данных (машинного обучения),
обработки символьной информации (рассуждений на основе знаний)

и т.д., то есть интегративными не являются.

Техническая составляющая ИИ



1962- цифровая система контроля или управления некоторым объектом

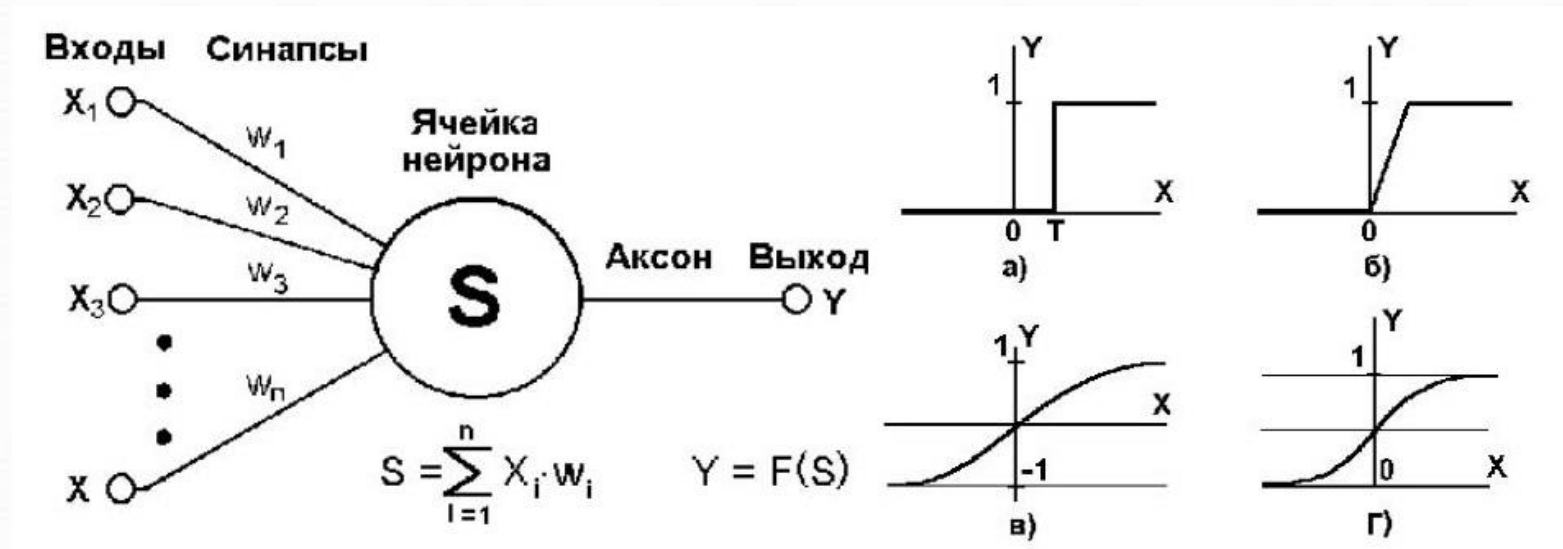
1970 - система аппаратного и программного обеспечения на базе микропроцессора или микроконтроллера, предназначенная для выполнения специальных функций в более крупной системе.

1980 Сетевые встроенные системы - появились стандартные технологии объединения Ethernet.

2006 + Киберфизические системы - комплексы, состоящие из природных объектов, искусственных подсистем и контроллеров.

2011 + Киберфизические системы на основе сквозных технологий цифровой экономики РФ (4-ая промышленная революция)

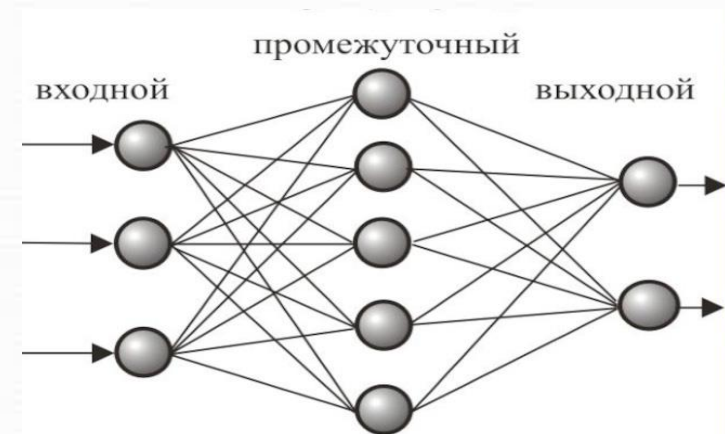
Нейросетевой подход: модель нейрона и сеть



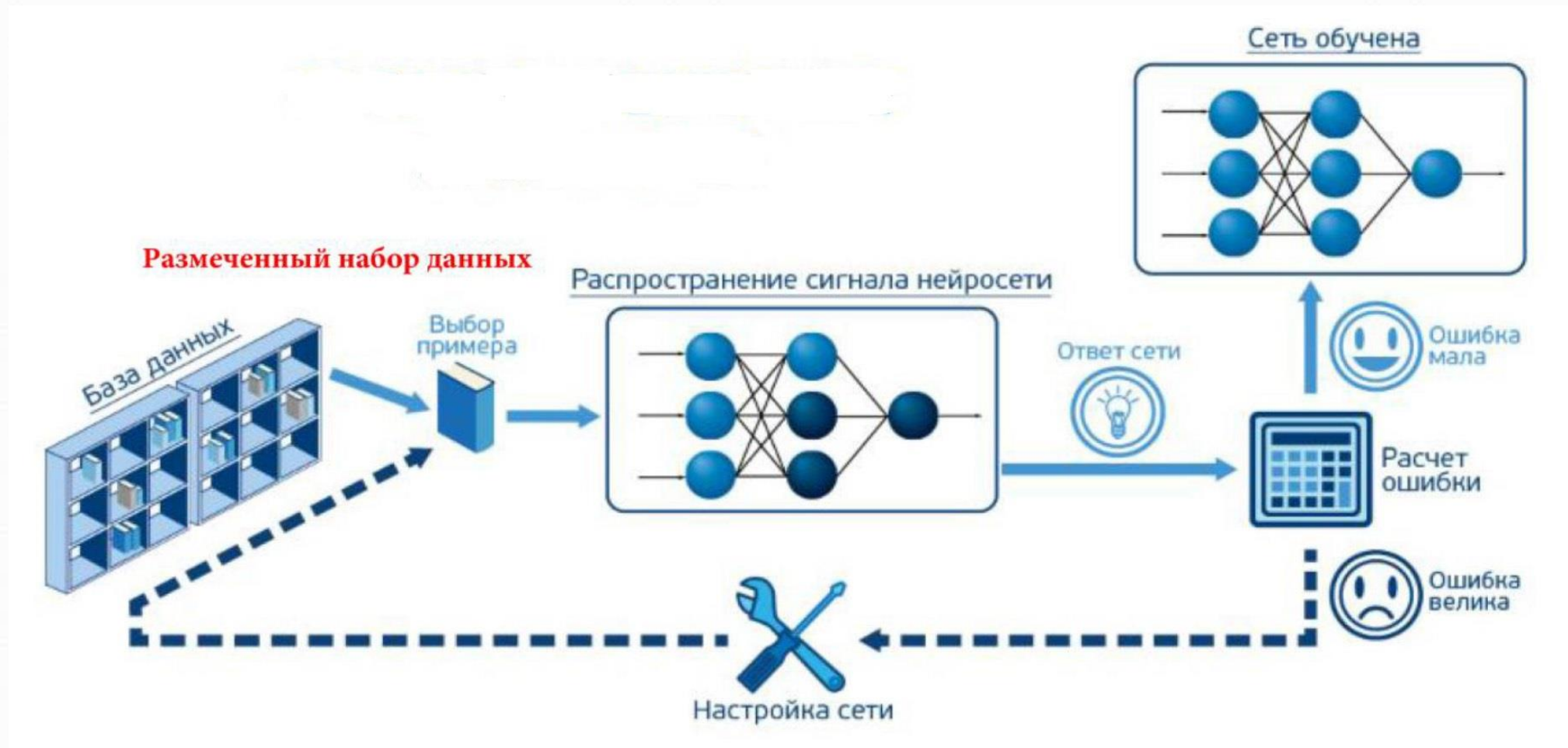
Математическая модель нейрона

Различные виды функции активации

- нейроны
- связи и их веса
(сумма взвешенных входов - активация нейрона)
- функция активации нейрона
- последовательность слоев



Процесс обучения нейронной сети



Направления исследований в области ИИ и будущее

Исследования в области ИИ развиваются более 60 лет по двум основным направлениям: **логическому** и **нейрокибернетическому**.

Логический подход направлен на создание прикладного (**слабого**) ИИ, т.е. компьютерных программ, предназначенных для решения какой-либо одной «интеллектуальной» задачи или их небольшого множества.

Нейрокибернетический подход направлен на создание универсального (**сильного**) ИИ, т.е. аналога человеческого мозга, способного решать любые «интеллектуальные» задачи.

Все достижения ИИ зависят от ЭВМ – супер ЭВМ и квантовые компьютеры