Санкт-Петербургский политехнический университет Петра Великого

Институт компьютерных наук и технологий

Кафедра компьютерных систем и программных технологий

**Реферат на тему:**

История искусственного интеллекта

**Дисциплина**: Интеллектуальные системы

Выполнил студент гр. 13541/1 Смирнов М.И.

(подпись)

Руководитель Сазанов А.М.

(подпись)

“ ” 2017 г.

Санкт – Петербург

2017

**Содержание**

[Введение 3](#_Toc500327470)

[Глава 1. Искусственный интеллект в современном мире 4](#_Toc500327471)

[1.1. Появление термина “искусственный интеллект” 4](#_Toc500327472)

[1.2. Эвристическое программирование 4](#_Toc500327473)

[1.3. Экспертные системы и их области применения 6](#_Toc500327474)

[1.4. Машинное обучение 8](#_Toc500327475)

[Глава 2. Главные направления разработки в области ИИ 10](#_Toc500327476)

[Глава 3. История развития искусственного интеллекта в России. 12](#_Toc500327477)

[Глава 4. Опасность развития искусственного интеллекта 15](#_Toc500327478)

[Глава 5. Сильный искусственный интеллект 17](#_Toc500327479)

[5.1. Что такое сильный ИИ? 17](#_Toc500327480)

[5.2. Копирование мозга 18](#_Toc500327481)

[Вывод 20](#_Toc500327482)

[Список литературы 21](#_Toc500327483)

# **Введение**

С давних времен человечество было охвачено идеей создать искусственный интеллект, мышление наподобие человеческого мозга, обладающего способностью решать задачи сложного характера. Подтверждением служат некоторые интересные факты. Например, древние египтяне сконструировали движущуюся статую бога Амона. В древнейшей поэме «Илиаде» бог Гефест создавал по подобию человека прекрасные существа-автоматы. В литературных произведениях также немало упоминаний об этой теме: создание существа, обладающего интеллектом, встречается от Галатеи Пигмалиона до Буратино папы Карло. Тем не менее, основоположником идеи искусственного интеллекта признан испанский философ, поэт и математик средних веков – Раймонд Луллий. В ХIII веке он стал первым, кто предпринял попытку сконструировать механическую машину, способную решать задачи различного рода, опираясь на специально разработанную классификацию понятий. Позднее, в XVIII веке, Лейбниц и Декарт развивали эту идею независимо друг от друга и предложили применять универсальные языки для классификации всех наук. Эти научные труды считаются первым теоретическим материалом в области искусственного интеллекта. Однако зарождением научного направления, сфокусированного на создании искусственного интеллекта, признано считать лишь изобретение электронно-вычислительной машины (ЭВМ) в 40-х годах XX века.

# **Глава 1. Искусственный интеллект в современном мире**

* 1. **Появление термина “искусственный интеллект”**

ЭВМ 40-х годов, согласно оценкам науки того времени, обладали высоким уровнем производительности. Этот факт породил первые вопросы в науке искусственного интеллекта, главным из которых стал интерес выяснить, возможно ли изобрести машину, интеллект которой будет равняться или даже превосходить интеллектуальные возможности самого человека.

Спустя десятилетие, в 50-ые годы, научное направление искусственного интеллекта получило следующий этап развития. Ученые активно работали над созданием машин, «разум» которых имитировал мозг человека. К сожалению, ни одна из попыток не увенчалась успехом, поскольку машины оказались совершенно непригодными с аппаратной и программной точек зрения. На научном семинаре в Стэнфордском университете (США) в 1956 г. был впервые предложен широко известный нам сегодня термин - искусственный интеллект.

* 1. **Эвристическое программирование**

В течение следующих десяти лет, в 60-е года, ученые пытались вывести универсальные методы для решения задач разнообразной направленности, подобно тому, как наш мозг ежедневно справляется с задачами широкого класса. Однако и здесь результат не отличался от прежнего – создание универсальных программ, как выяснилось, стало слишком сложной задачей. Суть в том, что программа, обладающая большим набором инструментов для решения широкого класса задач, является практически беспомощной в решении конкретной проблемы. Это подобно ситуации, где человек, собираясь на необитаемый остров, наберет с собой как можно больше вещей первой необходимости, но при столкновении с конкретной проблемой, у него не окажется и половины нужных предметов. Данная проблема подтолкнула к зарождению эвристического программирования.

Эвристика – это не обоснованное правило, которое позволяет уменьшить количество подбора верного варианта в пространстве потенциальных ответов.

Эвристическое программирование – это программирование, занимающееся разработкой плана действий на основе аналогичных или прецедентных случаев. Можно сказать, что в истории искусственного интеллекта 50-60гг. XX века главной задачей стал поиск универсального алгоритма для решения задач разнообразной направленности.

Обычная процедура построения моделей методом эвристического программирования строится следующим образом. Испытуемым пред­лагается решать некоторую задачу, сопровождая свои размышления устными комментариями хода своих рассуждений. Все высказывания испытуемых тщательно протоколируются. Затем протоколы подвергают анализу с целью выявления хода решения, характера применяемых операций, догадок, приемов и т. п. Полученный в ходе анализа материал используется при составлении компьютерной программы - модели данного вида поведения. Таким образом, программа является моделью не испытуемого, а протокола. Такая модель должна выполнять то, что делает испытуемый и так, как это делает он.

Следующий этап процедуры связан с исследованием работы мо­дели при решении задач того же типа, которые предлагались испы­туемым. Если процесс решения отклоняется от зафиксированного в протоколе, программа дорабатывается. Для этого могут быть постав­лены новые эксперименты с испытуемыми, получены новые прото­колы и т. д.

Использование описанной методики связано с рядом затруднений. При самоотчете испытуемый не в состоянии выделить и прокоммен­тировать все без исключения шаги выполненного им решения, осо­бенно если они связаны с догадками, неосознаваемыми заключения­ми и т.п. Соответственно такие шаги не попадают в протокол и ис­следователь должен самостоятельно заполнять разрывы, опираясь на собственные знания и догадки. Поэтому правильнее будет говорить, что окончательная программа моделирует гипотезу исследователя об изучаемом процессе - гипотезу, основанную на протоколе и опреде­ленных данных из области психологии.

Ряд трудностей заключается также в проведении анализа и интер­претации протокольных записей. Здесь необходима типизация прие­мов анализа и разработка мер, ограничивающих произвольность ин­терпретаций.

Развитие эвристического программирования связано с построени­ем разнообразных моделей, таких, как модели поведения при выборе, при определении стратегии размещения ценных бумаг и др. Однако наибольшие успехи теоретического и прикладного характера достиг­нуты при создании систем ИИ, называемых решателями задач.

Разработка программ - решателей задач преследует две основ­ные цели: во-первых, являясь моделью поведения человека в опреде­ленных условиях, решатель обладает некоторой объяснительной си­лой и может быть использован для предсказания действий человека в ходе решения; во-вторых, решатель может быть использован как со­ставная часть в системах автоматизации управления сложными объ­ектами, в частности - роботами.

Обычно решатели строятся для задач, связанных с преобразова­ниями ситуаций. В таких задачах задаются исходная и желаемая си­туации, а также набор операторов или действий, которые могут стро­го определенным образом изменять ситуации. Класс таких задач до­вольно широкий. В него входят, например, задачи формальной логи­ки, планирования целенаправленных действий и др.

* 1. **Экспертные системы и их области применения**

В 70-ые года ученые пересмотрели основной подход в области создания искусственного интеллекта и пришли к выводу, что поиск универсального алгоритма мышления носит безрезультатный и бессмысленный характер. Взамен пришла новая идея – собрать и классифицировать точные знания конкретных специалистов-экспертов в моделированную систему. И тогда в США впервые зародились коммерческие системы, известные как экспертные системы. На смену старого метода пришел новый – представление знаний. В это время были созданы всемирно известные экспертные системы медицины и химии - «MYCIN» и «DENDRAL». Обе системы являются мощным вспомогательным средством врачей при диагностике пациентов. При введении ряда паталогических симптомов болезни в программу «MYCIN» выдаются данные о диагнозе. При помощи «DENDRAL» можно по ряду свойств вещества определить его химический состав. На этом этапе истории ИИ зародилось понятие экспертных систем.

Области применения систем, основанных на знаниях, могут быть сгруппированы в несколько основных классов: медицинская диагностика, контроль и управление, диагностика неисправностей в механических и электрических устройствах, обучение.

* Прогнозирование.

Прогнозирующие системы предсказывают возможные результаты или события на основе данных о текущем состоянии объекта. Программная система «Завоевание Уолл-стрита» может проанализировать конъюнктуру рынка и с помощью статистических методов алгоритмов разработать план капиталовложений на перспективу. Она не относится к числу систем, основанных на знаниях, поскольку использует процедуры и алгоритмы традиционного программирования. Хотя пока ещё отсутствуют экспертные системы, которые способны за счёт своей информации о конъюнктуре рынка помочь увеличить капитал, прогнозирующие системы уже сегодня могут предсказывать погоду, урожайность и поток пассажиров. Даже на персональном компьютере, установив простую систему, основанную на знаниях, можно получить местный прогноз погоды.

* Планирование.

Планирующие системы предназначены для достижения конкретных целей при решении задач с большим числом переменных. Дамасская фирма Informat впервые в торговой практике предоставляет в распоряжение покупателей 13 рабочих станций, установленных в холле своего офиса, на которых проводятся бесплатные 15-минутные консультации с целью помочь покупателям выбрать компьютер, в наибольшей степени отвечающий их потребностям и бюджету. Кроме того, компания Boeing применяет экспертные системы для проектирования космических станций, а также для выявления причин отказов самолётных двигателей и ремонта вертолётов. Экспертная система XCON, созданная фирмой DEC, служит для определения или изменения конфигурации компьютерных систем типа VAX и в соответствии с требованиями покупателя. Фирма DEC разрабатывает более мощную систему XSEL, включающую базу знаний системы XCON, с целью оказания помощи покупателям при выборе вычислительных систем с нужной конфигурацией. В отличие от XCON система XSEL является интерактивной.

* Интерпретация.

Интерпретирующие системы обладают способностью получать определенные заключения на основе результатов наблюдения. Система PROSPECTOR, одна из наиболее известных систем интерпретирующего типа, объединяет знания девяти экспертов. Используя сочетания девяти методов экспертизы, системе удалось обнаружить залежи руды стоимостью в миллион долларов, причём наличие этих залежей не предполагал ни один из девяти экспертов. Другая интерпретирующая система - HASP/SIAP. Она определяет местоположение и типы судов в тихом океане по данным акустических систем слежения.

* Контроль и управление.

Системы, основанные на знаниях, могут применяться в качестве интеллектуальных систем контроля и принимать решения, анализируя данные, поступающие от нескольких источников. Такие системы уже работают на атомных электростанциях, управляют воздушным движением и осуществляют медицинский контроль. Они могут быть также полезны при регулировании финансовой деятельности предприятия и оказывать помощь при выработке решений в критических ситуациях.

* Диагностика неисправностей в механических и электрических устройствах.

В этой сфере системы, основанные на знаниях, незаменимы как при ремонте механических и электрических машин (автомобилей, дизельных локомотивов и т.д.), так и при устранении неисправностей и ошибок в аппаратном и программном обеспечении компьютеров.

* Обучение.

Экспертные системы, выполняющие обучение, подвергают диагностике, «отладке» и исправлению (коррекции) поведение обучаемого. Примером является обучение студентов отысканию неисправностей в электрических цепях, обучение военных моряков обращению с двигателем на корабле и обучение студентов-медиков выбору антимикробной терапии. Обучающие системы создают модель того, что обучающийся знает и как он эти знания применяет к решению проблемы. Системы диагностируют и указывают обучающемуся его ошибки, анализируя модель и строя планы исправлений указанных ошибок. Они исправляют поведение обучающихся, выполняя эти планы с помощью непосредственных указаний обучающимся.

Большинство экспертных систем включают знания, по содержанию которых их можно отнести одновременно к нескольким типам. Например, обучающая система может также обладать знаниями, позволяющими выполнять диагностику и планирование. Она определяет способности обучаемого по основным направлениям курса, а затем с учетом полученных данных составляет учебный план. Управляющая система может применяться для целей контроля, диагностики, прогнозирования и планирования.

* 1. **Машинное обучение**

Следующий этап наступает в 80-е года, в которые научное течение об искусственном интеллекте получило совершенно новый виток в истории развития. На искусственный интеллект посмотрели под новым углом. Учеными был признан значимый потенциал использования искусственного интеллекта в области исследования и развития производства. В это время стали появляться первые коммерческие программы. Также зародилось машинное обучение, благодаря которому процесс перенесения базы данных экспертной системы в машинную программу стал гораздо проще и быстрее. Появление программ, которые самостоятельно повышают и расширяют диапазон эвристических правил, послужило наиболее важным этапом в области искусственного интеллекта.

Таким образом, в начале 80-ых годов ученые во всем мире начали производить обработку данных, вести важнейшие международные и национальные исследовательские работы, в фокусе которых находились интеллектуальные вычислительные системы пятого поколения.

Различают два типа обучения:

1. Обучение по прецедентам, или индуктивное обучение, основано на выявлении эмпирических закономерностей в данных.
2. Дедуктивное обучение предполагает формализацию знаний экспертов и их перенос в компьютер в виде базы знаний.

Дедуктивное обучение принято относить к области экспертных систем, поэтому термины машинное обучение и обучение по прецедентам можно считать синонимами.

Многие методы индуктивного обучения разрабатывались как альтернатива классическим статистическим подходам. Многие методы тесно связаны с извлечением информации (Information extraction), интеллектуальным анализом данных (Data mining).

Общая постановка задачи обучения по прецедентам выглядит следующим образом: имеется множество объектов (ситуаций) и множество возможных ответов (откликов, реакций). Существует некоторая зависимость между ответами и объектами, но она неизвестна. Известна только конечная совокупность прецедентов — пар «объект, ответ», называемая обучающей выборкой. На основе этих данных требуется восстановить неявную зависимость, то есть построить алгоритм, способный для любого возможного входного объекта выдать достаточно точный классифицирующий ответ. Эта зависимость не обязательно выражается аналитически, и здесь нейросети реализуют принцип эмпирически формируемого решения. Важной особенностью при этом является способность обучаемой системы к обобщению, то есть к адекватному отклику на данные, выходящие за пределы имеющейся обучающей выборки. Для измерения точности ответов вводится оценочный функционал качества.

Данная постановка является обобщением классических задач аппроксимации функций. В классических задачах аппроксимации объектами являются действительные числа или векторы. В реальных прикладных задачах входные данные об объектах могут быть неполными, неточными, нечисловыми, разнородными. Эти особенности приводят к большому разнообразию методов машинного обучения.

# **Глава 2. Главные направления разработки в области ИИ**

В работах в области ИИ можно выделить 2 главных перспективных направления:

1. Нейрокибернетика. Основная идея: любое мыслящее устройство должно каким-то образом воспроизводить структуру человеческого мозга. Эта наука ориентирована на аппаратное моделирование. Мозг человека состоит из порядка 1021 нейронов, связанных между собой. Есть успешные попытки моделирования множества нейронов – нейронные сетию. Первые успехи были впечатляющими. Был сделан первый объект – персептрон – некоторая матрица нейронов, которая могла распознавать 2 состояния (например падает свет или нет). Она могла, например, распознавать образ. Появилась новая проблема – нужно делать большие матрицы из0за большого количество информации. Из-за этого об этом направлении на 10-15 лет забыли. В последнее время нейрокибернетика снова начала развиваться из-за скачка в развитии ЭВМ. Появились нейрокомпьютеры, транспьютеры.

Это направление стабильно держится на первом месте. Продолжается совершенствование алгоритмов обучения и классификации в масштабе реального времени, обработки естественных языков, распознавания изображений, речи, сигналов, а также создание моделей интеллектуального интерфейса, подстраивающегося под пользователя. Среди основных прикладных задач, решаемых с помощью нейронных сетей, - финансовое прогнозирование, раскопка данных, диагностика систем, контроль за деятельностью сетей, шифрование данных. В последние годы идет усиленный поиск эффективных методов синхронизации работы нейронных сетей на параллельных устройствах.

Можно выделить 3 способа реализации нейросетей:

а) аппаратные (платы и т.д.);

б) программный (нейросеть моделируется в памяти компьютера);

в) гибридный – среднее между 1 и 2. В настоящий момент нейрокибернетика развивается в направлении нейрокомпьютеров и клавная задача – распознание образов.

1. Кибернетика черного ящика. Не имеет значение как устроено мыслящее устройство, главное, чтобы на заданные входные воздействия оно реагировало также как человеческий мозг.

Когда попытались заняться исследованием алгоритмов человеческого мозга поняли, что это очень сложно. Науке не известно как мыслит человек, как он принимает решение.

В основном решение проблемы происходит с помощью модели лабиринтного поиска (поиск по дереву решений). Но это дерево может быть очень большим. Возникает проблема его хранения и поиска нужного пути, иногда некоторые ветки могут быть неизвестными. Применяются эвристики. Эвристика – правило, теоретически не обоснованное, но позволяющее сократить перебор в дереве решений. На основе эвристик появилось эвристическое программирование – разработка стратегий действий на основе заранее заданных известных эвристик. Появилось понятие экспертная система (ЭС) или система, основанная на знаниях. В 70-80 гг. Это направление начало развиваться, в 90-х достигло своего пика. Экспертные системы использовались в медицине и диагностике. Это направление продолжает развиваться в сторону ЭС.

# **Глава 3. История развития искусственного интеллекта в России.**

Коллежский советник Семён Николаевич Корсаков (1787—1853) ставил задачу усиления возможностей разума посредством разработки научных методов и устройств, перекликающуюся с современной концепцией искусственного интеллекта, как усилителя естественного. В 1832 году С. Н. Корсаков опубликовал описание пяти изобретённых им механических устройств, так называемых «интеллектуальных машин», для частичной механизации умственной деятельности в задачах поиска, сравнения и классификации. В конструкции своих машин Корсаков впервые в истории информатики применил перфорированные карты, игравшие у него своего рода роль баз знаний, а сами машины по существу являлись предтечами экспертных систем.

В СССР работы в области искусственного интеллекта начались в 1960-х годах. В Московском университете и Академии наук был выполнен ряд пионерских исследований, возглавленных Вениамином Пушкиным и Д. А. Поспеловым.

В 1964 году была опубликована работа ленинградского логика Сергея Маслова «Обратный метод установления выводимости в классическом исчислении предикатов», в которой впервые предлагался метод автоматического поиска доказательства теорем в исчислении предикатов.

В 1966 году В. Ф. Турчиным был разработан язык рекурсивных функций Рефал.

До 1970-х годов в СССР все исследования ИИ велись в рамках кибернетики. По мнению Д. А. Поспелова, науки «информатика» и «кибернетика» были в это время смешаны, по причине ряда академических споров. Только в конце 1970-х в СССР начинают говорить о научном направлении «искусственный интеллект» как разделе информатики. При этом родилась и сама информатика, подчинив себе прародительницу «кибернетику». В конце 1970-х создаётся толковый словарь по искусственному интеллекту, трёхтомный справочник по искусственному интеллекту и энциклопедический словарь по информатике, в котором разделы «Кибернетика» и «Искусственный интеллект» входят наряду с другими разделами в состав информатики. Термин «информатика» в 1980-е годы получает широкое распространение, а термин «кибернетика» постепенно исчезает из обращения, сохранившись лишь в названиях тех институтов, которые возникли в эпоху «кибернетического бума» конца 1950-х — начала 1960-х годов. Такой взгляд на искусственный интеллект, кибернетику и информатику разделяется не всеми. Это связано с тем, что на Западе границы данных наук несколько отличаются.

В 1974 году при Комитете по системному анализу при президиуме АН СССР был создан Научный совет по проблеме "Искусственный интеллект", его возглавил Г. С. Поспелов, его заместителями были избраны Д. А. Поспелов и Л. И. Микулич.

По инициативе Совета было организовано пять комплексных научных проектов, которые были возглавлены ведущими специалистами в данной области. Проекты объединяли исследования в различных коллективах страны: "Диалог" (работы по пониманию естественного языка, руководители А. П. Ершов, А. С. Нариньяни), "Ситуация" (ситуационное управление, Д. А. Поспелов), "Банк" (банки данных, Л. Т. Кузин), "Конструктор" (поисковое конструирование, А. И. Половинкин), "Интеллект робота" (Д. Е. Охоцимский).

В 1980-1990 гг. проводятся активные исследования в области представления знаний, разрабатываются языки представления знаний, экспертные системы (более 300).

В 1988 г. создается АИИ - Ассоциация искусственного интеллекта. Ее членами являются более 300 исследователей. Президентом Ассоциации единогласно избирается Д. А. Поспелов. Крупнейшие центры - в Москве, Петербурге, Переславле-Залесском, Новосибирске.

Российские примеры компаний, работающих в сфере ИИ:

Поскольку сфера искуственного интеллекта является сращиванием математических наук и программирования, у России, имеющей солидную базу и школы в этих направлениях, неплохие шансы на получение статуса глобального игрока при достаточном внимании к этой сфере со стороны в первую очередь профильных государственных ведомств в виде программ и, разумеется, крупных частных игроков.

Среди разработок и компаний можно назвать и инициативу сервиса онлайн-заказа такси «Яндекс.Такси» с технологией интеллектуального распределения заказов с переходом на технологии интеллектуального распределения заказов, с учетом дорожной ситуации и специальных пожеланий пользователей. ПАО «КамАЗ» разрабатывает систему полуавтономного управления автомобилем, которая будет строиться на комплексе технологий искусственного интеллекта — компьютерного зрения, машинного обучения, речевых технологий. В сотрудничестве с российской Cognitive Technologies автопроизводитель готовится к выпуску предпромышленной версии системы помощи водителю ADAS (Advanced Driver Assistance System) первого уровня. С другой стороны, разработки группы ЦРТ в области речевых технологий вошли в технологический стэк Microsoft. Также примером удачного использования машинного обучения в области популярных ныне фильтров для соцсетей является нашумевшее приложение Prisma, разработанное в недрах Mail.Ru Group независимыми разработчиками. Пионер в сфере использования чат-ботов мессенджер Telegram стал локомотивом развития индустрии в России с таким ярким российским стартапом, как Chatfuel, в который вложились крупнейшие зарубежные венчурные компании. А робот «Вера», созданный петербургской компанией Stafory, проводит собеседование с потенциальными кандидатами на открытые вакансии, делая за полчаса ту работу, которую три-четыре человека делают неделю.

«Сбербанк» в конце 2016 года анонсировал запуск робота-юриста. Годом ранее банк запустил систему искусственного интеллекта Iron Lady, которая занимается обзвоном должников.

# **Глава 4. Опасность развития искусственного интеллекта**

Развитие искусственного интеллекта однажды приведёт к превосходству его над умственными способностями человека. Однако не станет ли это опасным для человечества? Изучить ситуацию можно в результате более точного определения понятия ИИ, взяв за основу для сравнения естественный интеллект. Может ли в одном человеке сочетаться ум и интеллект одновременно? Или же умный человек не может быть интеллектуалом и наоборот?

Такие вопросы возникают в связи с приближением эры ИИ, о возможной опасности которого человечество должно знать заранее и своевременно принять меры для обеспечения своей безопасности. В первую очередь опасность ИИ будет связана с его самостоятельностью и неконтролируемым принятием решений. На данный момент уже выделены средства для изучения этой проблемы. В институте OpenAI изучаются перспективы развития ИИ. На теперешней стадии развития систем ИИ опасность его применения может быть связана со следующим факторами:

* ошибками ПО (программного обеспечения). Любому программному обеспечению может угрожать такая опасность;
* самостоятельной активностью ИИ, вредной для человека. Опасность от ИИ может исходить после изобретения умного компьютера. Поэтому необходимо определиться со степенью ума для компьютера, которая способна быть допустимой, а также чрезмерной, представляющей опасность для человека. Эти свойства должны быть точно измерены из-за сближения умственных способностей человека и компьютера, которые неизбежны. Информационные системы, существующие сегодня, представляют собой человеко-машины, которые могут работать благодаря интеллекту пользователя либо эксперта по компьютерам.

Например, какая опасность будет исходить от интеллектуальной бухгалтерской системы, которой может выдаваться неверная информация? Опасность может возникнуть, когда у такой системы появятся элементы личности, например, собственная заинтересованность, не имеющая ничего общего с человеческой. Решением такой проблемы может быть запрет на создание систем, отличающихся возможностью эволюции.

Кроме того, опасность может быть связана с содержанием в ИИ логических ошибок. Он может использоваться для решения достаточно сложных задач, список которых сразу неизвестен. Поэтому должны быть предусмотрены специальные меры для подтверждения правильности решение, которое будет получено. Скорее всего, возникнет необходимость в разработке всевозможных способов контроля таких систем, например, специальных программных средств, которыми будет автоматически проверяться правильность решения и при этом участие человека не потребуется.

# **Глава 5. Сильный искусственный интеллект**

* 1. **Что такое сильный ИИ?**

Говоря о серьезных угрозах и мощных потенциальных возможностях в будущем, как правило, мы думаем о сильном искусственном интеллекте. Что значит сильный ИИ? Имеется в виду, что созданный человеком искусственный интеллект станет способным самостоятельно мыслить и оценивать себя, как индивидуальность, но при этом процесс мышления будет необязательно похожим на человеческий. По общей договоренности считается, что машина с подобным искусственным интеллектом должна уметь следующие возможности:

* Принимать решения, использовать определенные стратегии, уметь решать головоломки и проблемы, возникающие в условиях неопределенности;
* Представлять знания, в том числе общие представления об окружающей действительности;
* Планировать действия;
* Обладать общими представлениями и знаниями о реальном мире;
* Уметь не только обучаться на основе данных ПО, но и самообучаться, фильтруя информацию из множества источников;
* Свободно общаться и поддерживать любые темы, включая темы с субъективными оценочными явлениями;
* Уметь объединить и применить все выше перечисленные возможности для достижения конкретной цели.

Ключевой фактор заключается в самосознании искусственного интеллекта. Т.е. машина должна осознавать себя как нечто обособленное от внешнего мира, уметь оценивать себя и свои возможности. Это подобно тому, как мы спрашиваем себя – кто мы и что мы здесь делаем? К сожалению, развитие искусственного интеллекта задерживается, так как наука на данном этапе развития не может ответить на вопрос, какие из областей человеческого мозга отвечают за наше самосознание, а, следовательно, не имея знаний о принципе действия нашего сознания, мы не можем смодулировать его в некой другой системе.

* 1. **Копирование мозга**

Наиболее простым и логичным кажется просто воспроизвести искусственный интеллект, скопировав работу нашего мозга. И лишь только с изобретением работоспособной копии хотя бы части мозга человека или животного можно будет говорить о по-настоящему значимом прорыве в создании сильного искусственного интеллекта. Однако на сегодняшний день результаты научных работ в этой области оставляют желать лучшего. Наилучший результат пока что у команды ученых Blue Brain Project, работающих над моделированием неокортекса головного мозга.

Blue Brain Project показали результат в моделировании одной колонки неокортекса молодой крысы. Это в разы проще, чем у человека. Стоит уточнить, что одна колонка неокортекса в реальности представляет собой всего лишь пару кубических миллиметров живого мозга, но при этом обладает передачей информации в сотни гигабайт за секунду. Теперь становится ясно, что на данном этапе исследований не может быть и речи о главной цели - создать настоящее сознание. Пока что необходимо успешно имитировать хотя бы маленькую часть мозга.

Кроме того, даже при успешном копировании части мозга, которое мы вряд ли застанем в ближайшие пару десятков лет, не означает создание искусственного интеллекта. Например, повторив внешний дизайн и форму компьютера, мы не получим работающую машину, поскольку внутри нет программного обеспечения.

Приведем несколько примечательных исследований в области ИИ:

1. В 2013 году был запущен интересный проект под названием «Национальный месяц создания романов». Дариус Казими создал программу, способную генерировать текст. В ее рамках было использовано несколько чат-ботов, запущенных в виртуальном пространстве, в котором они могли взаимодействовать между собой. Спустя три года, в 2016 году, ученые из японского университета Хакодате выпустили программу, написавшую роман «День, когда компьютер пишет роман». Роман принял участие в литературном конкурсе и вышел в финал, обогнав 1450 работ, авторами которых являлись живые люди.

2. Созданная программа «Eugene Goostman» олицетворяет собой мальчика из Одессы. В 2012 году был проведен первый тест, в ходе которого «мальчик» обманул судей в 20,2% случаев. В тестах Университета Рединга в 2014 году показатель повысился до 33%. Чуть позже, в 2015 году, новая программа «Соня Гусева» улучшила показатель до 47%. Важно заметить, что тестирование было ограничено по времени до 5 минут на эксперта. Ограничение по времени делает результаты тестирования не столь однозначными.

# **Вывод**

Сделать программу, способную обмануть собеседника, создать у него полное ощущение, что он общается с реальным человеком, не так уж и сложно. Совсем скоро мы познакомимся с виртуальными собеседниками, которые смогут не только разговаривать с нами, как настоящие люди, но и даже флиртовать, шутить и вести себя ничуть не «искусственнее» окружающих людей. Но не стоит себя обманывать, это не искусственный разум, даже не близко.

Искусственному интеллекту ещё долгое время необходимо развиваться и важно не забрасывать это перспективное направление. Мы все еще очень далеки от создания настоящего сильного искусственного интеллекта, способного мыслить также как люди, иметь эмоции и чувства. Некоторые ученые придерживаются мнения, что это вовсе невозможно воплотить в жизнь. Но нельзя не признать, что сделаны серьезные шаги в этом направлении, начиная с 40-ых годов.

Я думаю, что ИИ всё равно не может полностью заменить человека. И есть множество профессий, где человек должен брать на себя большую ответственность за каждое свое действие, где из-за неправильного, неверного решения могут погибнуть люди. Там необходимо присутствие человека. По крайней мере (из-за человеческого фактора) ИИ должен помочь человеку в принятии решения, но не заменить его.

Возможно, из-за этической стороны вопроса, ученые не смогут сделать тот самый широкий шаг на путь к настоящему ИИ. И государство прикроет всякое дальнейшее исследование в этой области. Во всяком случае, любое развитие в истории происходит скачкообразно. И если сейчас идет достаточно бурный рост в этой сфере, то скоро может всё приостановиться на определенный срок. Остается только ждать.

# **Список литературы**

1. Д.В. Смолин. Введение в искусственный интеллект. Конспект лекций. 2007 г.
2. Портал искусственного интеллекта [Электронный ресурс]. – Режим доступа: http://www.aiportal.ru/ (Дата обращения: 02.12.17).
3. Life. Наука. [Электронный ресурс]. – Режим доступа: https://life.ru/ (Дата обращение: 02.12.17).
4. Geektimes. Портал об индустрии высоких технологий. [Электронный ресурс]. – Режим доступа: https://geektimes.ru/ (Дата обращения: 02.12.17).
5. Prorobot. Роботы и робототехника. [Электронный ресурс]. – Режим доступа: http://www.prorobot.ru/ (Дата обращения: 03.12.17).