

УДК 004.896

## Создание искусственного интеллекта: мифы и реальность

***Ненов И.А.**, студент,  
кафедра: «Радиоэлектронные системы и устройства»  
Россия, 105005, г. Москва, МГТУ им. Н.Э. Баумана*

*Научный руководитель: Гаврилина Е.А., к.ф.н., доцент  
Россия, 105005, г. Москва, МГТУ им. Н.Э. Баумана  
[gavrilina@bmstu.ru](mailto:gavrilina@bmstu.ru)*

*Аналитическая машина не претендует на создание чего-либо нового. Ее способности  
непревосходят наших знаний о том, как приказать ей что-либо исполнить...*

Ада Байрон (AdaByron), графиня Лавлейс

На протяжении многих тысячелетий человек пытается определить, какие процессы происходят у него в голове. Так и в сфере искусственного интеллекта (ИИ) ученым предстоит решить еще более сложную задачу. Ведь в данной области специалистам предстоит не только понять сущность интеллекта, но и создать интеллектуальные сущности. Примечательно, что единого определения искусственного интеллекта не существует. В различных научных работах, посвященных ему, имеются различные толкования данного явления. Они могут охватывать не только мыслительные процессы, но и формулировки относительно поведения индивидуума. Отчасти привлекательность ИИ в том и состоит, что он является оригинальным и мощным орудием для исследования именно этих проблем.

Искусственный интеллект можно определить как научную дисциплину, которая занимается моделированием разумного поведения. Это определение имеет один существенный недостаток – понятие интеллекта трудно объяснить: является ли он чем-то единым, или же этот термин объединяет набор разрозненных способностей. Возможно, ли вообще достичь разумности посредством компьютерной техники, или же сущность интеллекта требует богатства чувств и опыта, присущего лишь биологическим существам? На эти вопросы ответа пока не найдено, но все они помогли сформировать задачи и методологию, составляющие основу современного искусственного интеллекта.

Создание искусственного интеллекта позволяет расширить возможности компьютерных наук, а не определить их границы. Таким образом, искусственный интеллект можно определить как область компьютерной науки, занимающуюся автоматизацией разумного поведения, опираясь на теоретические и прикладные принципы. Можно не связывать термин искусственного интеллекта напрямую с пониманием человеческого интеллекта: просто инженеры и ученые, работающие над созданием ИИ, могут использовать для решения конкретных проблем методы, не свойственные человеческому мышлению.

Вопрос парадигмы искусственного интеллекта настолько сложный и комплексный, что им занимаются сразу несколько наук: не только компьютерная наука, но также философия, нейро–бионауки, футурология и множество других. Искусственный интеллект как направление научных исследований изучает природу и суть интеллектуальной творческой деятельности человека, ищет возможности воспроизвести в искусственных системах отражательную способность человеческого сознания. Но при этом непосредственно суть искусственного интеллекта понимается как кибернетическая система, которая перерабатывает информацию, поступающую из внешней среды, чтобы на ее основании принимать решения. Очень интересный и важный момент: слово «интеллект» в этом понятии метафорично, поскольку ИИ-системы пока не воспроизводят процессы, происходящие в мозгу человека. На сегодня общепринято, что искусственный интеллект – это система, которая должна формировать решения, удовлетворяющие предъявляемым требованиям. Это научное направление, в рамках которого ставятся и решаются задачи аппаратного или программного моделирования тех видов человеческой деятельности, которые традиционно считаются интеллектуальными. Выполнять функции, например творческие, которые традиционно считаются прерогативой человека.

Одна из первых работ о машинном разуме, «Вычислительные машины и интеллект» была написана еще в 1950 году британским математиком Аланом Тьюрингом. Он рассмотрел вопрос о том, можно ли заставить машину действительно думать. Тест Тьюринга сравнивает способности предположительно разумной машины со способностями человека – лучшим и единственным стандартом разумного поведения. В тесте, который Тьюринг назвал «имитационной игрой», машину и ее человеческого соперника – «следователя» помещают в разные комнаты, отделенные от помещения в котором находится «имитатор». Следователь не может видеть их или говорить с ними напрямую – он общается с ними исключительно при помощи текстового устройства. Задача следователя – отличить компьютер от человека только на основе их ответов на вопросы, задаваемые через это текстовое устройство. Если следователь не может отличить

машину от человека, тогда, утверждает Тьюринг, машину можно считать разумной. Изолированность следователя от машины и другого, живого человека, исключает предвзятое отношение – на решение следователя не будет влиять ни вид машины, ни ее электронный голос. Чтобы обмануть эту стратегию, компьютер должен знать, когда ему следует выдать ошибочное число, чтобы показаться человеком. Чтобы обнаружить человеческое поведение на основе эмоциональной природы, следователь может попросить обоих субъектов высказаться по поводу стихотворения или картины. Компьютер в таком случае должен знать об эмоциональном складе человеческих существ. Однако тест Тьюринга, несмотря на свою интуитивную притягательность, уязвим для многих оправданных нападков. Например: китайская комната – модель системы, разработанная Джоном Серлем, в которой критикуется возможность моделирования человеческого понимания естественного языка в искусственных системах. В какой-то степени, китайская комната является критикой теста Тьюринга. Аналогичную позицию занимает и Роджер Пенроуз, который в своей книге «Новый ум короля», аргументирует невозможность получения процесса мышления на основе формальных систем. Иногда же, напротив, тест Тьюринга обвиняют в попытках втиснуть машинный интеллект в форму интеллекта человеческого. Но одно из наиболее слабых мест – пристрастие в пользу чисто символических задач. Тест не затрагивает способностей, требующих, например, ловкости рук, хотя подобные аспекты являются важными составляющими человеческого интеллекта.

Все же, тест Тьюринга является важной составляющей в тестировании и аттестации современных интеллектуальных программ. Тьюринг затронул проблему осуществимости построения интеллектуальной программы на базе цифрового компьютера – определить, умеет ли машина мыслить. Тест Тьюринга был признан эталоном, до последнего времени ученым не удавалось создать программу, которая бы успешно преодолела тест.

Трудно сказать, когда же, собственно, начались исследования в области искусственного интеллекта. Джордж Буль (1815– 1864 гг.) высказал множество идей, касающихся математических методов исследования мыслительных процессов, и ряд выдвинутых им положений до сих пор сохраняют свою актуальность. Но компьютера у Буля не было, придется согласиться с тем, что не он является основоположником исследований в области искусственного интеллекта.

Известно, что историки по обе стороны Атлантики не могут прийти к общему мнению относительно того, кто создал первую программируемую вычислительную машину. Подобным же образом среди них нет единства взглядов и по вопросу о том, с чего нача-

лись исследования в области искусственного интеллекта. Можно считать, что история создания ИИ начинается с момента создания первых ЭВМ в 40-х г. Тогда следующим этапом в истории создания искусственного интеллекта являются 50-е годы, когда исследователи пытались строить разумные машины, имитируя мозг. Эти попытки оказались безуспешными по причине полной непригодности, как аппаратных, так и программных средств. Американские же историки ведут отсчет от проведенной в 1956 году в Дартмуте конференции, где был впервые предложен термин искусственный интеллект (англ. *artificial intelligence*). Хотя, как указывает председатель Петербургского отделения Российской ассоциации искусственного интеллекта Т. А. Гаврилова, в английском языке словосочетание «*artificial intelligence*» не имеет той слегка фантастической антропоморфной окраски, которую оно приобрело в довольно неудачном русском переводе. Словом «интеллект» (латин. *intellectus* — понимание, познание) обозначают относительно устойчивую структуру умственных способностей. Английское слово «*intelligence*» означает «умение рассуждать разумно», а вовсе не «интеллект», для которого есть английский аналог «*intellec*».

Существуют различные толкования этого понятия, в частности его, отождествляют с мыслительными операциями, со стилем и стратегиями решения проблемных ситуаций, со способностью к учению, к познанию, с индивидуальными особенностями ориентирования в ситуации, когнитивным стилем, с биопсихической адаптацией к существующим обстоятельствам жизни.

60-е года в истории искусственный интеллект отметились попытками отыскать общие методы решения широкого класса задач, моделируя сложный процесс мышления. Чем шире класс задач, которые может решать одна программа, тем беднее оказываются ее возможности при решении конкретной проблемы. В целом, 50-60 г.г. в истории искусственного интеллекта можно отметить как время поиска универсального алгоритма мышления.

Существенный прорыв в практических приложениях искусственного интеллекта произошел в 70-х гг., когда на смену поискам универсального алгоритма мышления пришла идея моделировать конкретные знания специалистов — экспертов. В США появились первые коммерческие системы и экспертные системы, основанные на знаниях. Пришел новый подход к решению задач искусственного интеллекта как представление знаний. Созданы «MYCIN» и «DENDRAL» — ставшие уже классическими экспертные системы для медицины и химии. Обе эти системы в определенном смысле можно назвать

диагностическими, поскольку в первом случае «MYCIN» по ряду симптомов (признаков патологии организма) определяется болезнь (ставится диагноз), во втором – по ряду свойств определяется химическое соединение. В принципе, этот этап в истории искусственного интеллекта можно назвать рождением экспертных систем.

Следующий значимый период в истории создания искусственного интеллекта – это 80-е года. На этом отрезке ИИ пережил второе рождение. Стала развиваться область машинного обучения. Проблематика машинного обучения касается процесса самостоятельного получения знаний интеллектуальной системой в процессе её работы. К области машинного обучения относится большой класс задач на распознавание образов. Например, это распознавание символов, рукописного текста, речи, анализ текстов. Многие задачи успешно решаются с помощью биологического моделирования. Особо стоит упомянуть компьютерное зрение, которое связано ещё и с робототехникой. Интеллектуальность требуется роботам, чтобы манипулировать объектами, выполнять навигацию с проблемами локализации и планировать движение. Примером интеллектуальной робототехники могут служить игрушки – роботы AIBO, QRIO.

Последний этап, развиваемый с начала 1990-х годов, называемый агентно-ориентированным, основан на использовании интеллектуальных агентов. Согласно этому методу, агент – это вычислительная часть, способен достигать поставленных перед интеллектуальной машиной целей. Сама такая машина будет интеллектуальным агентом, воспринимающим окружающий его мир с помощью датчиков, и способной воздействовать на объекты в окружающей среде с помощью исполнительных механизмов.

Вданный момент времени результаты науки искусственного интеллекта, достигнутые ранее, применяются в различных областях науки, промышленности, в бизнесе и повседневной жизни. Но и наука не стоит на месте. В современном мире существует четыре основных подхода к изучению построения систем искусственного интеллекта. Рассмотрим основные подходы:

- логический

В своей основе содержит так называемую Булеву алгебру. Большинство систем искусственного интеллекта, построенных по логическому принципу, представляют собой определенную машину доказательства теорем: исходная информация содержится в виде аксиом, а логические выводы формулируются по правилам отношений между этими аксиомами. В каждой такой машине есть блок генерирования цели, причем система

вывода доказывает эту цель как теорему. Эта система больше известна под названием экспертной системы.

- структурный

В качестве основы системы искусственного интеллекта используется моделирование структуры мозга человека. Среди первых подобных попыток необходимо отметить перцептрон Розенблатта. Основная структурная моделируемая единица – нейрон. Со временем возникли новые модели, которые в настоящее время известны, как нейронные сети.

- эволюционный

При построении систем искусственного интеллекта, уделяется основное внимание построению начальной модели, а также тем правилам, по которым эта модель может эволюционировать. Классическим примером эволюционного алгоритма является генетический алгоритм.

- имитационный

Одно из базовых понятий данного подхода – это объект, поведение которого имитируется, то есть, так называемый «черный ящик». Таким образом, моделируется способность человека копировать действия других. В основном используется в кибернетике.

Все эти подходы решают два основных направления развития искусственного интеллекта:

- решение проблем, связанных с приближением специализированных систем искусственного интеллекта к возможностям человека, и их интеграции, которая реализована природой человека;
- создание искусственного разума, представляющего интеграцию уже созданных систем искусственного интеллекта в единую систему, способную решать проблемы человечества.

Сейчас существуют проекты, на которые необходимо обратить внимание. Огромный проект, стартовавший в 2005 году по созданию искусственного мозга под названием «BlueGene» – (Голубой мозг). Сверхмощный компьютер. Основная цель проекта заключалась в создании модели структуры и функционирования мозговой активности различных животных с целью дальнейшего моделирования человеческого неокортекса. Разработкой проекта занимаются ученые-исследователи, представители Федеральной политехнической школы (г.Лозанна). Они сумели создать модель-схему расположения синапсов в головном мозге крыс. Все полученные результаты могут помочь



проанализировать деятельность головного мозга на локальном уровне, но не в масштабах всей его работы. Вполне возможно, что исследователи смогут ответить – придет ли на смену человеческому разуму искусственный и будет ли он более высокоразвитым?

В Лозанне планируется к запуску еще один аналогичный проект – «HumanBrainProject» – (Человеческий мозг.). В его рамках к 2023 году ученые из 13 стран собираются создать самый крупный в мире компьютерный мозг, в котором будет работать столько же нейронов, сколько и в человеческом мозге – сто миллиардов. По мнению директора проекта профессора Маркрама, создание компьютерной модели мозга просто необходимо, ведь благодаря этому можно унифицировать ход исследований и проводить эксперименты, совершенствуя и исправляя ее.

Еще одним проектом, который стартовал в 2010 году, является проект компании DARPA совместно с SRI International. Суть его заключается в разработках прорывного искусственного интеллекта, который будет способен обрабатывать и передавать данные, копируя механизмы работы человеческого мозга. Искусственный интеллект не только будет анализировать сложную информацию из различных источников сразу, но и сможет динамически переписывать себя для оптимизации взаимодействия с окружающей средой. И при этом система компактна и потребляет минимум энергии. Результатом трехлетней работы стал когнитивный чип, в котором процесс вычислений осуществляется по средствам нейронных связей, память представлена в виде синапсов, и связи в виде аксонов. Новые чипы не содержат биологические элементы, сообщили в компании, но имеют цифровые кремниевые схемы.

Еще в конце 90-х годов японцы публикуют новость: впервые в мире ведутся работы по созданию биокомпьютера, принцип действия которого основан на биологических датчиках. Вполне вероятно, что в будущем с помощью биотехнологий будут управлять химическим заводом, регулировать биологические процессы внутри человеческого организма, производить гормоны и лекарственные вещества. Сохранение важной информации в случае ядерной катастрофы – серьезная проблема, так как электронные носители уязвимы в случае воздействия сильного электромагнитного импульса. Альтернативным способом хранения данных могут стать бактерии.

В 2012 году в Японии был разработан первый прототип искусственного мозга. Искусственный интеллект может обрабатывать огромное количество информации, но еще не наделен способностью мыслить. Разработчики пока с этим не спешат. По мнению исследователей, роботы ближайшего будущего во многом будут похожи на людей: они смогут ходить на двух ногах, смогут различать лица, поддерживать беседу, выполняют

просьбы, однако по своей сути – это всего лишь машины, подобные человеку. Все их действия подчинены заранее подготовленному алгоритму, а потому – примитивны. И только в том случае, если удастся реализовать технологию бимолекулярного вычисления, машины смогут мыслить и получат способность к творчеству.

Ученые института Вайзмана в Иерусалиме создали биокомпьютер величиной с каплю воды. Роль "аппаратной" части выполняют энзимы, "программной" – молекулы ДНК. Нанокомпьютер состоит из триллиона клеток. Главным свойством биокомпьютеров является то, что каждая их клетка – миниатюрная химическая лаборатория. Если биоорганизм запрограммирован, то он просто производит нужные вещества. Наш мозг – это тоже своего рода биокомпьютер.

Существуют эффективные компьютерные системы, предназначенные для мониторинга и управления различными механизмами. Робот Asimo (Advanced Step in Innovative Mobility), разработанный в компании Honda оснащен 16 гибкими «суставами». Задача сохранения равновесия изделия и управления его передвижениями возложена на 4-процессорный компьютер. Поскольку спина у робота негнущаяся, его руки пришлось несколько удлинить, чтобы машина могла подбирать предметы с пола. Движения Asimo при перемещении в пространстве удивительно похожи на движения человека, но мыслить он не в состоянии.

В 2013 году исследователи из Иллинойского университета в Чикаго протестировали IQ самой совершенной в мире системы искусственного интеллекта. Цель данного исследования заключалась в том, чтобы привлечь внимание к "проблемным моментам" и испытаниям развития системы ИИ. Исследователи предложили "умной" системе пройти через «Weschler Preschool and Primary Scale of Intelligence Test» – стандартный в США тест для определения интеллекта дошкольников. Результаты оказались довольно странными: в то время как по баллам система показала уровень интеллекта 4-летнего ребенка, очки в разных частях теста были распределены крайне неравномерно. "Мы по-прежнему безумно далеки от создания искусственной системы, наделенной здравым смыслом, которая могла бы соревноваться в интеллекте с хотя бы восьмилетним ребенком", – отметил Роберт Слоан, глава кафедры компьютерных наук в Иллинойском университете. Каким же образом заложить в программу здравый смысл, никто пока что не знает.

Один из наиболее известных на сегодняшний момент футурологов Рэй Курцвейл предсказывает, что к 2029 году компьютер сможет пройти тест Тьюринга, что, по его мнению, будет доказывать у него наличия разума.

Некоторые из самых известных систем искусственного интеллекта:



- компьютер «DeepBlue» – победил чемпиона мира по шахматам. Матч Каспаров против суперЭВМ не принёс удовлетворения ни компьютерщикам, ни шахматистам, и система не была признана.
- экспертные системы «Watson» – перспективная разработка IBM, способная воспринимать человеческую речь и производить вероятностный поиск, с применением большого количества алгоритмов. Для демонстрации работы «Watson» принял участие в американской игре «Jeopardy», где системе удалось выиграть.
- «MYCIN» – одна из ранних экспертных систем, которая могла диагностировать небольшой набор заболеваний, причем часто так же точно, как и доктора.
- «ViaVoice» – системы распознавания речи, способны обслуживать потребителей.
- роботы в ежегодном турнире «RoboCup» соревнуются в упрощённой форме футбола.

Банки применяют системы искусственного интеллекта (СИИ) в страховой деятельности, при игре на бирже и управлении собственностью. Методы распознавания образов (включая, как более сложные и специализированные, так и нейронные сети) широко используют при оптическом и акустическом распознавании, медицинской диагностике, спам-фильтрах, в системах ПВО для обеспечения ряда задач национальной безопасности.

Разработчики компьютерных игр применяют искусственный интеллект в той или иной степени проработанности. Это образует понятие – игровой искусственный интеллект. Стандартными задачами искусственного интеллекта в играх являются имитация поведения боевой единицы, нахождение путей в двумерном или трёхмерном пространстве, расчёт верной экономической стратегии.

Пытаться предсказать конкретно, как именно благожелательный искусственный интеллект поможет человечеству, или недружественный искусственный интеллект повредит – это рискованное интеллектуальное предприятие. Доктор математических наук Винер, основатель кибернетики, написал множество книг, посвященных машинам, их использованию, их будущему и настоящему. Сотрудничая с физиологами и психологами, он доказывает некоторое сходство работы человеческого мозга и электронного. Однако, как отмечает доктор Винер в своей книге «Кибернетика», даже обладая определенным набором рецепторов, эффекторов искусственный интеллект является лишь имитацией

простейших условных рефлексов живых существ. Все вышесказанное полностью относится и к человекоподобным роботам—это весьма примитивные модели человека. Однако, в будущем, доктор Винер считает возможным создание не только разумной машины, но и машины «умнее своего создателя». Не исключает он и бунта машин. Он полагает, что не только робот, но и любая стратегическая машина способна вызвать катастрофу.

В 2014 году на базе Кембриджского университета будет создан "Центр изучения глобальных рисков", одной из целей которого является определение принципов разработки "безопасного" искусственного интеллекта. Глава создаваемого центра профессор философии Хью Прайс опубликовал статью " Искусственный интеллект: сможем ли мы удержать его в коробке?", в которой призвал серьезно рассмотреть возможные угрозы. "Мы думаем, что в качестве первого правильного шага стоило бы прекратить относиться к искусственному интеллекту как к предмету научной фантастики и начать думать о нем как о части реальности, с которой мы или наши потомки могут рано или поздно столкнуться". Основной причиной опасности ультра—умных машин создатели проекта видят в принципиальном отличии предполагаемого интеллекта от человека. Все ценности — такие, как «любовь, счастье и даже выживание», важны для нас, потому что у нас есть определенная эволюционная история — нет ни одной причины полагать, что машины будут разделять их с нами, что повлечет за собой катастрофические последствия.

Многие ученые высказываются, что самая главная угроза выживанию человеческой цивилизации — это увеличение вычислительных мощностей компьютеров. Рано или поздно этот процесс приведет к созданию "единого искусственного интеллекта", который будет способен писать программы самостоятельно и без вмешательства человека производить роботов, свое "потомство". Безусловно, искусственный интеллект и гипотетические последствия его выхода из—под контроля человека являются одной из любимых тем современной философии. Однако большинство действующих исследователей в этой области на данный момент воздерживаются от каких—либо громких прогнозов о возможности серьезных прорывов в ближайшем будущем — проблемы сегодня скорее заключаются не в необходимых мощностях, а в отсутствии новых подходов.

Тем не менее, международная неправительственная организация HumanRightsWatch призвала правительства всех стран отказаться от создания полностью автономных вооружений, способных нарушить права мирного населения во время военных действий. По оценке организации, боевые роботы не смогут отличить мирных

жителей от военных и нарушат множество положений Женевских конвенций («законы войны») и требования Декларации Мартенса.

Что же может произойти после реального появления искусственного интеллекта? При текущей скорости прогресса первые дни после этого события могут стать самым опасным временем в истории человечества. Наступит эра господства слабого и ограниченного искусственного интеллекта, очень опасной комбинации, которая сможет нанести непоправимый ущерб существующей цивилизации. Существует множество гипотетических направлений появления искусственного интеллекта. Под сингулярностью зачастую ошибочно понимают искусственный интеллект, который просто умнее человека. Однако, это не так. Большая часть из созданных образцов искусственного интеллекта уже превышают человеческие возможности. Однако современный искусственный интеллект недостаточно сложный и продвинутый, чтобы причинить значительный ущерб инфраструктуре человечества. Проблемы появятся, когда в случае сингулярности искусственный интеллект начнет многократно улучшать самого себя. По словам математика Ирвина Джона Гуда, событие сингулярного перехода будет подобно интеллектуальному взрыву, который действительно, как бомба, поразит нас. Контроль человека над машинами навсегда будет отодвинут на второй план. Технологическая сингулярность, представляет собой гипотетический момент появления рекурсивного улучшения искусственного интеллекта – искусственный сверхразум. Он достигает развития могучего, подобного интеллекту человека либо искусственный сверхразум значительно превосходит человеческий. Особое беспокойство вызывает возрастание использования слабого искусственного интеллекта в качестве экспертной системы, соответствующей человеческому интеллекту или даже превосходящей его, в конкретной узконаправленной области. Следствием этого может стать независимая и бесконтрольная работа этих систем.

Центр по изучению экзистенциальных рисков (CSER) г. Кембриджа провел исследования, которые доказали, что современные технологии могут уничтожить человеческую цивилизацию уже в ближайшем будущем. Трудно точно узнать как, когда и где произойдет первая настоящая катастрофа, вызванная искусственным интеллектом. Но можно с уверенностью утверждать, что у нас в запасе есть еще несколько десятилетий. Наша инфраструктура прочна и еще не достаточно интегрирована, чтобы с ней могло случиться что-то действительно страшное. Но наш полностью цифровой и более взаимосвязанный мир уязвим перед такими угрозами.

### Список литературы

1. Гэри Антес. Формула жизни, Computerworld Россия, №4 (501), 2006, - 36с.
2. Андрей Плахов – О проблеме создания искусственного интеллекта /Научно-популярный журнал «Мембрана», [Электронный ресурс].  
<http://www.membrana.ru/articles/readers/2002/12/16/210400.htm> (дата обращения 28.10.2013г.) .
3. Сотник С. Л. – Экспертные системы. Базовые понятия. Методика построения. Статистический подход. Режим доступа: <http://ai.obrazec.ru/aiexpert.htm> (дата обращения 25.10.2013г.).
4. Алан Тьюринг – Могут ли машины мыслить? Портал «Гуманитарное образование». Режим доступа: <http://humanities.edu.ru/db/msg/27602> (дата обращения 20.10.2013г.).