**История развития искусственного интеллекта**

Раньше с понятием искусственного интеллекта (ИИ) связывали надежды на создание мыслящей машины, способной соперничать с человеческим мозгом и, возможно, превзойти его. Эти надежды, на долгое время захватившие воображение многих энтузиастов, так и остались несбывшимися. И хотя фантастические литературные прообразы «умных машин» создавались еще за сотни лет до наших дней, лишь с середины тридцатых годов, с момента публикации работ А. Тьюринга, в которых осуждалась реальность создания таких устройств, к проблеме ИИ стали относиться серьезно.

Для того, чтобы ответить на вопрос, какую машину считать «думающей», Тьюринг предложил использовать следующий тест: испытатель через посредника общается с невидимым для него собеседником человеком или машиной. «Интеллектуальной» может считаться та машина, которую испытатель в процессе такого общения не сможет отличить от человека.

Если испытатель при проверке компьютера на «интеллектуальность» будет придерживаться достаточно жестких ограничений в выборе темы и формы диалога, этот тест выдержит любой современный компьютер, оснащенный подходящим программным обеспечением. Можно было бы считать признаком интеллектуальности умение поддерживать беседу, но, как было показано, эта человеческая способность легко моделируется на компьютере. Признаком интеллектуальности может служить способность к обучению. В 1961 г. профессор Д. Мичи, один из ведущих английских специалистов по ИИ, описал механизм, состоящий из 300 спичечных коробков, который мог научиться играть в крестики и нолики. Мичи назвал это устройство MENACE (Matchbox Educable Naughts and Crosses Engine). В названии (угроза) заключается, очевидно, доля иронии, вызванной предубеждениями перед думающими машинами.

До настоящего времени единого и признанного всеми определения ИИ не существует, и это не удивительно. «Достаточно вспомнить, что универсального определения человеческого интеллекта также нет дискуссии о том, что можно считать признаком ИИ, а что нет, напоминают споры средневековых ученых о том, которых интересовало, сколько ангелов смогут разместиться на кончике иглы»1. Сейчас к ИИ принято относить ряд алгоритмов и программных систем, отличительным свойством которых является то, что они могут решать некоторые задачи так, как кто делал бы размышляющий над их решением человек.

**Нейросети**

Идея нейронных сетей родилась в ходе исследований в области искусственного интеллекта, а именно в результате попыток воспроизвести способность нервных биологических систем обучаться и исправлять ошибки, моделируя низкоуровневую структуру мозга. Основной областью исследований по искусственному интеллекту в 60-80е годы были экспертные системы. Такие системы основывались на высокоуровневом моделировании процесса мышления (в частности, на его представлении как манипуляций с символами). Скоро стало ясно, что подобные системы, хотя и могут принести пользу в некоторых областях, не охватывают некоторые ключевые аспекты работы человеческого мозга.

Согласно одной из точек зрения, причина этого состоит в том, что они не в состоянии воспроизвести структуру мозга. Чтобы создать искусственный интеллект, необходимо построить систему с похожей архитектурой.

Мозг состоит из очень большого числа (приблизительно 1010) нейронов, соединенных многочисленными связями (в среднем несколько тысяч связей на один нейрон, однако это число может сильно колебаться). Нейроны — это специальные клетки, способные распространять электрохимические сигналы. Нейрон имеет разветвленную структуру ввода информации (дендриты), ядро и разветвляющийся выход (аксон). Аксоны клетки соединяются с дендритами других клеток с помощью синапсов. При активации нейрон посылает электрохимический сигнал по своему аксону. Через синапсы этот сигнал достигает других нейронов, которые могут в свою очередь активироваться. Нейрон активируется тогда, когда суммарный уровень сигналов, пришедших в его ядро из дендритов, превысит определенный уровень (порог активации).

Интенсивность сигнала, получаемого нейроном (а, следовательно, и возможность его активации), сильно зависит от активности синапсов. Каждый синапс имеет протяженность, и специальные химические вещества передают сигнал вдоль него. Один из самых авторитетных исследователей нейросистем, Дональд Хебб, высказал постулат, что обучение заключается в первую очередь в изменениях силы синоптических связей. Например, в классическом опыте. Павлова каждый раз перед кормлением собаки звонил колокольчик, и собака быстро научилась связывать звонок колокольчика с пищей.

Синоптические связи между участками коры головного мозга, ответственными за слух, и слюнными железами усилились, и при возбуждении коры звуком колокольчика у собаки начиналось слюноотделение.

Таким образом, будучи построен из очень большого числа совсем простых элементов (каждый из которых берет взвешенную сумму входных сигналов и в случае, если суммарный вход превышает определенный уровень, передает дальше двоичный сигнал), мозг способен решать чрезвычайно сложные задачи. Определение формального классического нейрона дается следующим образом:

Он получает входные сигналы (исходные данные или выходные сигналы других нейронов сети) через несколько входных каналов. Каждый входной сигнал проходит через соединение, имеющее определенную интенсивность (или вес); этот вес соответствует синоптической активности биологического нейрона. С каждым нейроном связано определенное пороговое значение. Вычисляется взвешенная сумма входов, из нее вычитается пороговое значение и в результате получается величина активации нейрона.

Сигнал активации преобразуется с помощью функции активации (или передаточной функции) и в результате получается выходной сигнал нейрона.

Если при этом использовать ступенчатую функцию активации, то такой нейрон будет работать точно так же, как описанный выше естественный нейрон.

**Нейросети в искусственном интеллекте**

Работы по созданию интеллектуальных систем ведутся в двух направлениях. Сторонники первого направления, составляющие сегодня абсолютное большинство среди специалистов в области искусственного интеллекта, исходят из положения о том, что искусственные системы не обязаны повторять в своей структуре и функционировании структуру и проистекающие в ней процессы, присущие биологическим системам. Важно лишь то, что теми или иными средствами удается добиться тех же результатов в поведении, какие характерны для человека и других биологических систем.

Сторонники второго направления считают, что на чисто информационном уровне этого не удастся сделать. Феномены человеческого поведения, его способность к обучению и адаптации, по мнению этих специалистов, есть следствие именно биологической структуры и особенностей её функционирования.

У сторонников первого информационного направления есть реально действующие макеты и программы, моделирующие те или иные стороны интеллекта. Одна из наиболее ярких работ, представляющих первое направление, это программа «Общий решатель задач» А. Ньюэлла, И. Шоу и Г. Саймона. Развитие информационного направления шло от задачи о рационализации рассуждений путем выяснения общих приемов быстрого выявления ложных и истинных высказываний в заданной системе знаний. Способность рассуждать и находить противоречия в различных системах взаимосвязанных ситуаций, объектов, понятий является важной стороной феномена мышления, выражением способности к дедуктивному мышлению.

Результативность информационного направления бесґспорна в области изучения и воспроизведения дедуктивных мыслительных проявлений. Для некоторых практических задач этого достаточно. Информационное направление наука точная, строгая, вобравшая в себя основные результаты изысканий кибернетики и математическую культуру. Главные проблемы информационного направления ввести в свои модели внутреннюю активность и суметь представить индуктивные процедуры.

Одна из центральных проблем, это «проблема активных знаний, порождающих потребности в деятельности системы из-за тех знаний, которые накопились в памяти системы»1.

У сторонников второго биологического направления результатов пока существенно меньше, чем надежд. Одним из родоначальников биологического направления в кибернетике является У. Мак-Каллок. В нейрофизиологии установлено, что целый ряд функций и свойств у живых организмов реализованы с помощью определенных нейронных структур. На основе воспроизведения таких структур в ряде случаев получены хорошие модели, в особенности это касается некоторых сторон работы зрительного тракта.

Создание нейрокомпьютеров, моделирующих нейронные сети (НС), в настоящее время рассматривается как одно из наиболее перспективных направлений в решении проблем интеллектуализации вновь создаваемых ЭВМ и информационно-аналитических систем нового поколения.

В большей части исследований на эту тему НС представляется как совокупность большого числа сравнительно простых элементов, топология соединений которых зависит от типа сети. Практически все известные подходы к проектированию НС связаны в основном с выбором и анализом некоторых частных структур однородных сетей на формальных нейронах с известными свойствами (сети Хопфилда, Хемминга, Гроссберга, Кохоннена и др.) и некоторых описанных математически режимов их работы. В этом случае термин нейронные сети метафоричен, поскольку он отражает лишь то, что эти сети в некотором смысле подобны живым НС, но не повторяют их во всей сложности. Вследствие такой трактовки нейронные ЭВМ рассматриваются в качестве очередного этапа высоко параллельных супер-ЭВМ с оригинальной идеей распараллеливания алгоритмов решения разных классов задач. Сам термин нейронная ЭВМ нейрокомпьютер, как правило, никак не связан с какими-то ни было свойствами и характеристиками мозга человека и животных. Он связан только с условным наименованием порогового логического элемента как формального нейрона с настраиваемыми или фиксированными весовыми коэффициентами, который реализует простейшую передаточную функцию нейрона-клетки. Исследования в области создания нейроинтеллекта ведутся на различных уровнях: теоретический инструментарий, прототипы для прикладных задач, средства программного обеспечения НС, структуры аппаратных средств. Основными этапами на пути создания мозгоподобного компьютера являются выяснение принципов образования межэлементных связей и мозгоподобных системах адаптивных сетях с большим числом элементов, создание компактного многовходового адаптивного элемента аналога реального нейрона, исследование его функциональных особенностей, разработка и реализация программы обучения мозгоподобного устройства.

Одним из наиболее существенных путей расширения функционального диапазона НС, а также повышения их эффективности для традиционных задач является более целенаправленное использование в моделях механизмов и принципов организации мозга. Обоснованием этого служит достаточно экономная реализация функций в мозге, пока не доступная для самых совершенных супер-ЭВМ. В мозге, как и в любой сложной системе, процесс функционирования представляет собой совокупный результат работы его элементов и способов их взаимодействия. Оба эти фактора находят свое отражение в системной работе мозга.

В настоящее время становится очевидным, что успех разработки нейрокомпьютеров и интеллектуализации ЭВМ нового поколения в значительной степени определяется успехом работы над созданием нового класса базовых элементов с использованием данных о работе мозга. В первую очередь, это касается усложнения архитектуры, простанственно-временного распределения процессов в самом базовом элементе и расширении его функциональных возможностей. Поэтому актуальна необходимость в новом взгляде на перераспределение основных функций обработки информации между самими базовыми элементами нейрокомпьютера и сетевыми ресурсами в сторону увеличения логической нагрузки на базовые элементы.

Это связано с тем, что только в самое последнее время, на основе данных практической нейрофизиологии появилась возможность выделить из огромного числа процессов в мозге небольшое их количество наиболее значимых для переработки информации и выполнения сложных функций принятия конечных решений. Минимально необходимый набор структур, обеспечивающих эти процессы, значительно сузился и вследствие установленных ограничений существующих ЭВМ, которые не могут быть преодолены в настоящее время без использования свойств работы мозга. Кроме того, широко практикуемые однородные структуры искусственных НС на формальных нейронах не используют в полной мере возможностей реальных нейронов: их разнотипность, свойства распределенной и параллельной работы, многоуровневую иерархическую структурированность и соподчиненность в организации базовых структур головного мозга.

Из огромного числа данных о деятельности мозга, по-видимому, наиболее близко к решению проблемы интеллектуализации разрабатываемых ЭВМ относятся факты о механизмах и принципах элементной и сетевой организации процессов и функций в коре больших полушарий (КБП). Это определяется ее функциональной значимостью и уровнем современных данных о конкретных механизмах ее работы. Известно, что КБП является основным субстратом выполнения высших функций, определяющим уровень интеллекта личности.

В настоящее время накоплен и в значительной мере систематизирован экспериментальный и теоретический материал об элементарной организации корковых функций.

Все это дает основания предполагать, что данные о работе высшего отдела мозга могут иметь существенное значение и для идеологии создания нейрокомпьютеров, и для конструктивных решений отдельных их блоков.

В плане общего подхода к моделированию нейрокомпьютера существенно то, что по мере накопления фактов о морфологии, цитохимии и нейрофизиологии появляется все больше путей для перехода от вероятностных к детерминированным сетям корковой деятельности, основанных на данных об архитектурных принципах организации КБП. На основе этих данных все четче прослеживается связь особенностей функций КБП с конкретной спецификой ее элементов и связей. Это позволяет уже на исходной стадии моделирования решать принципиальный вопрос о соотношении функциональных нагрузок отдельного элемента и сети в целом, определяющим саму стратегию разработки нейрокомпьютера.

На практике этот выбор связан, прежде всего, с определением набора функций и свойств базового элемента и зависит как от уровня технической базы, так и от конструкторского решения их реализации. Обоснованием пересмотра концепции базового элемента нового типа являются данные практической нейрофизиологии, выявившие необходимый минимальный набор базовых свойств реальных нейронов, обеспечивающий реализацию основных информационных функций мозга у животных и человека. В соответствии с этими данными. В задаче создания новых поколений интеллектуальных вычислительных систем и задаче развития робототехники путь интеллектуализации за счет введения квазибиологических автоматов, в конечном счете, окажется технически и экономически более целесообразным направлением по сравнению с введением элементов интеллекта на основе информационно-логических методов.

Для того, чтобы интегрировать эти навязанные биологией наблюдения в логичные теоретические рамки, необходимо также найти некартезианское концептуальное пространство для размышлений о жизни и разуме, пространство в рамках строгой непрерывности. К счастью такое пространство существует в концепции Аристотеля.

**Все еще картезианская**

Будет очень полезно на некоторое время сконцентрироваться на одном специфическом спорном вопросе, по которому расходятся ортодоксальная и биологическая наука о мышлении, а именно выражение соотношения которое существует между нейробиологическими/биохимическими свойствами живых организмов с одной стороны и мышлением с другой стороны. Различие в этом вопросе может быть объяснено тем фактом, что два вида науки о мышлении сформированы в радикально различающихся философских концепциях. В целом биологическая наука о мышлении наиболее органично ложится в рамки общей аристотелевой концепции, в то же время у ортодоксальной науки о мышлении наблюдаются картезианские корни.

Любой, кто даже не проявлял особого интереса к философии разума, знает, что Декарт считал мыслимое и физическое двумя различными, но взаимодействующими онтологическими реальностями. Однако другой вклад, сделанный Декартом в изучение разума, менее широко известен. Этим вторым вкладом является форма психологического обоснования – дуализм обоснования – который одновременно поддерживает следующие тезисы: (1) для объяснения физического явления, некто нуждается в привлечении только специфических физических сущностей и состояний и специфических физических законов; (2) для объяснения психологических явлений, некто нуждается в привлечении только специфических мыслительных сущностей и состояний и специфических законов мышления. Дуализм обоснования прекрасно согласуется с идеей, что мыслительные события являются предельной формой физических явлений. Для физической онтологии, предлагаем ли мы физический или психологический стиль обоснования зависит описания, к которому, с текущими заданными конкретными целями обоснования, склоняются взятые интересующие нас явления.

Важно то, что Декарт мыслил органическое тело мыслителя как еще один физический объект в физическом мире. Учитывая дуализм обоснования, данная идея приводила его к тому, что нейробиологическое/биохимическое обоснование событий в теле мыслителя неуместно в психологическом обосновании событий в разуме мыслителя, в том смысле, что психологическое обоснование может быть проведено в отсутствии любого, сколько бы то ни было детализированного нейробиологического/биохимического знания о теле мыслящего объекта. Данное обосновательное отделение разума от физического носителя приводил в результате к прерывистости в обосновании в данном контексте между жизнью и разумом. Научное обоснование процессов, которые рассматривались картезианством как органические, относящиеся к телесной жизни (такие процессы как пищеварение, размножение и рост) немедленно попадало в область биологических объяснений, которая толковалась как принижаемая физической наукой. Научное обоснование физических процессов, с другой стороны, нуждается в изложении на языке, совершенно отличающемся от биологического языка, в языке специфическом для психологии. Это равнозначно отклонению положений строгой непрерывности. Другими словами, дуализм обоснования несовместим с биологической наукой о мышлении.

В настоящее время функционалисты в философии разума придерживаются взглядов, что определяющее свойство типа ментального состояния является причинная роль того, что состояние играет в посредничестве между (1) сенсорными входами, (2) другими типами ментальных состояний и (3) моторикой поведения. Строго говоря, функционализм не проводит связей к природе основы, на которой реализованы ментальные состояния, так как сущность в некотором частном ментальном состоянии уже является, как уже говорилось, сущностью в некотором специфическом функциональном состоянии, и совершенно эквивалентные функциональные состояния могут быть, в принципе, реализованы биохимически на углеродной основе, в виде кремниевого мозга или в виде картезианского разума самого по себе. Здесь, в принципе, функционализм входит составной частью в дуализм субстанций. Данный принципиальный факт может показаться незначительным, если считать, что функционализм обычно является рабочей лошадкой теории распознавания, согласно которой любая сущность данного типа ментального состояния является единственной и сходна с некоторым физическим состоянием в физической системе. Но, «дополнение к требованиям распознавания не рассматривает нейробиологические и биохимические детали тела биологического мыслящего объекта относящегося к процессу психологического обоснования»1. Согласно функционалистам, процесс психологического обоснования может проводиться в превосходной изоляции от этих частных деталей. Подобная позиция уже рассматривалась ранее: функционализм является формой картезианского дуализма обоснования.

Так как же база дуализма обоснования в функционализме уместна в понимании ортодоксальной науке о мышлении? Ответом является, что ортодоксальная наука о мышлении построена на функционализме. В самом деле, вычислительные состояния (тип состояний к которым прибегают ортодоксальный искусственный интеллект и ортодоксальная наука о мышлении) прекрасные примеры функционально определенных состояний. При этом нет никакого противоречия в том, что одно из классических положений функционализма было выражено в теории путем использования машины Тьюринга. Как только принимается функционалисткие основы ортодоксальной науки о мышлении и вместе с ней общее картезианское рассмотрение отношений между живым телом и разумом которые порождаются данными основами, можно увидеть почему ортодоксальная наука о мышлении связана с идеей, что может быть описано без понимания или существенных ссылок на нейробиологический или биохимический базис данного процесса мышления. Другими словами, можно видеть почему ортодоксальная наука о мышлении принимает положения совершенно неприемлемые биологической наукой о мышлении.