

# Об одной проблеме Смейла

Н. Б. Новиков

На основе анализа истории научных открытий, а также данных когнитивной психологии и теории искусственного интеллекта установлено, что основными факторами, исключаящими алгоритмический характер творческой деятельности в науке, являются вероятностная природа индуктивного вывода, метод проб и ошибок (метод последовательного перебора), фактор случая в научном открытии и теорема Геделя о неполноте. Это позволяет решить 18-ю проблему С. Смейла: каковы пределы интеллекта, как искусственного, так и человека?

**Ключевые слова:** творческая деятельность, искусственный интеллект, индуктивный вывод, фактор случая в научном открытии, теорема Геделя о неполноте.

## 1. Введение

В 1997 году выдающийся американский математик Стивен Смейл сформулировал восемнадцать нерешенных математических проблем. Излишне говорить, что данные проблемы отличаются высокой степенью сложности. Достаточно вспомнить теорему А. Пуанкаре (вторую проблему С. Смейла) об эквивалентности топологических свойств односвязного трехмерного пространства и трехмерной сферы, не поддававшуюся решению на протяжении 100 лет и доказанную Г. Перельманом в 2002 году.

В списке проблем С. Смейла под номером 18 фигурирует задача, касающаяся пределов искусственного и человеческого интеллектов. В лекции «Математические проблемы следующего столетия» [1], прочитанной в июне 1997 года по случаю 60-летия В. И. Арнольда в Филдсовском институте (США, Торонто), С. Смейл ставит перед математиками вопрос: «Каковы пределы интеллекта, как искусственного, так и человека?» Далее С. Смейл поясняет свой вопрос: «Пенроуз

в [20] пытается привести некоторые ограничения для искусственного интеллекта. Фигурирующий в его доказательстве интересный вопрос — это разрешимость множества Мандельброта (рассмотренная в [7]) и выводы из теоремы Геделя о неполноте. Однако необходимо более широкое изучение, которое включало бы более глубокие модели разума, а также компьютера, и проясняющее, что общего между искусственным и человеческим интеллектом, и чем они отличаются. Я бы начал исследования в том направлении, где вместе с теорией действительных чисел, приближениями, теорией вероятностей и геометрией значительную роль играют обучение, решение задач и теория игр» [1, с. 297]. Ссылаясь на работу Р. Пенроуза [20], С. Смейл имеет в виду его монографию «Новый ум короля» (1991).

## 2. Достижения в области искусственного интеллекта

Несмотря на то, что компьютерная наука (теория искусственного интеллекта) развивается чуть более 50-ти лет, она достигла заметных, а в ряде случаев впечатляющих успехов в деле моделирования человеческого познания. По справедливому замечанию Л. Н. Ясницкого, методы искусственного интеллекта позволили создать эффективные компьютерные программы в самых разнообразных сферах человеческой деятельности: медицине, биологии, зоологии, социологии, культурологи, политологии, экономике, бизнесе, криминалистике и т. д. [2]. Современные компьютеры успешно борются за звание чемпиона мира по шахматам, создают музыкальные произведения, доказывают математические теоремы, распознают образы и сцены, анализируют и интерпретируют предложения, выраженные на естественном языке, хранят огромные массивы информации и способны предоставить по первому требованию определенные фрагменты этой информации. Различные вычислительные устройства, созданные по образу и подобию человеческого мозга, успешно справляются с управлением сложными техническими объектами, диагностикой заболеваний человека, неисправностей сложных технических систем, предсказывают погоду и курсы валют, результаты голосований, выявляют хакеров и потенциальных банкротов, помогают абитуриентам правильно выбрать специальность. Сегодня искусственный интеллект — это обширная область исследований и разработок, предназначенных

для работы в трудно формализуемых областях деятельности человека. Для задач, решаемых методами искусственного интеллекта, характерно наличие большого числа степеней свободы (высокой степени неопределенности). В отличие от жестко детерминированных компьютерных программ системы искусственного интеллекта сами ищут пути решения поставленной задачи. При этом они могут менять свои параметры и структуру, совершенствоваться и развиваться, накапливать знания, адаптировать их для той или иной проблемной ситуации.

Особую популярность и широкое практическое применение приобрели экспертные системы — сложные программные комплексы, аккумулирующие знания специалистов в конкретных предметных областях и тиражирующие эти знания для консультации менее квалифицированных пользователей. Наряду с формализуемыми знаниями, которые можно четко изложить в виде законов, формул, моделей, алгоритмов, разработчики экспертных систем пытаются вложить в них плохо формализуемые знания, которые трудно описать в словесной форме. При этом специалистов не останавливает тот факт, что эти знания часто являются субъективными и приблизительными, сформированными на подсознательном уровне в результате обобщения большого личного опыта. Работа с такими знаниями оказывается весьма трудоемкой, требующей значительных затрат времени и сил.

В попытках создать обучающиеся компьютеры исследователи разработали различные варианты искусственных нейронных сетей (ИНС), как однослойных, так и многослойных. Начало этому направлению положили работы У. Мак-Каллока и В. Питтса, которые предложили конструкцию сети из математических нейронов и показали, что такая сеть в принципе может выполнять числовые и логические операции. Они высказали идею о том, что сеть из математических нейронов в состоянии обучаться, распознавать образы, обобщать, то есть она обладает свойствами человеческого интеллекта. Идея У. Мак-Каллока и В. Питтса была реализована в 1958 году Ф. Розенблаттом сначала в виде компьютерной программы для ЭВМ ИВМ-794, а затем, спустя два года, в виде электронного устройства, моделирующего человеческий глаз. Это устройство, имеющее в качестве элементной базы модельные нейроны Мак-Каллока–Питтса и названное персептроном, удалось обучить решению сложнейшей интеллектуальной задачи — распознаванию букв латинского алфавита. Эффективный алгоритм обучения многослойных персептронов, от-

крывший путь для их широкого практического применения, стал известен в 1986 году благодаря работе Д. Румельхарта, Дж. Хинтона и Р. Вильямса. С появлением этого метода, представляющего собой модификацию классического метода градиентного спуска, начался период широкого практического применения нейросетевых технологий для решения самых разнообразных задач. Конечно, в реальных (биологических) нейронах обучающий сигнал не распространяется обратно по тому же аксону, по которому он двигался в противоположном направлении. Но со временем нейрокомпьютер вполне возможно наделить множеством волокон, одни из которых будут передавать сигнал от анализаторов к ассоциативным зонам переработки информации, а другие — от указанных зон к анализаторам (органам чувств).

В 1960-е годы нейроинформатика обогатилась замечательной идеей Д. Холланда об использовании генетических алгоритмов в компьютерных (машинных) программах решения различных научных и технических задач по аналогии с тем, как те же алгоритмы, то есть эволюционные методы решения задач, использует сама природа. Если до Д. Холланда ученые в основном создавали компьютерные программы, включающие эволюционные механизмы мутаций и отбора, то Д. Холланд включил в эти программы механизм рекомбинации генетического материала при спаривании особей. К началу 1990-х годов идея генетических алгоритмов приобрела черты концепции, суть которой Д. Холланд изложил в развернутой статье [3]. В данной статье Д. Холланд объясняет, почему необходимо перенести алгоритмы биологической эволюции в область научного и технического творчества: «Живые организмы — прирожденные решатели задач. Универсальность их способностей не идет ни в какое сравнение даже с самыми лучшими компьютерными программами. Это особенно досадно для программистов-теоретиков, затрачивающих месяцы, а иногда и годы на создание какого-либо алгоритма, в то время как организмы легко справляются с задачей благодаря, казалось бы, слепому механизму эволюции и естественного отбора. Ученые-прагматики рассматривают замечательные возможности эволюции как нечто, чему следовало бы скорее подражать, нежели завидовать» [3, с. 32]. В настоящее время генетические алгоритмы, имитирующие природные оптимизационные процессы, происходящие при эволюции живых организмов, являются новым перспективным разделом информатики.

Можно также отметить существенный прогресс в разработке роботов, первоначально предназначенных для автоматизации физи-

ческого труда человека, но со временем приобретающих признаки интеллектуальности. Например, в одном из соревнований (2002) перед роботами стояла задача войти в комплекс зданий, где проводилась конференция, найти путь к бюро регистрации, зарегистрироваться на конференции и произнести речь. В соревновании Robocup, инициатором проведения которого в 1995 году стал японский специалист Китано со своими коллегами, провозглашена цель — к 2050 году разработать команду полностью автономных роботов-гуманоидов, которые смогут одержать победу над командой чемпионов мира по футболу среди людей. Игры происходят в лигах для роботов, имитирующих людей, колесных роботов различных размеров и четырехногих роботов Aibo компании Sony. В 2002 году это соревнование привлекло команды почти из 30 различных стран и на нем присутствовало больше 100 тысяч зрителей [4].

### **3. Актуальные проблемы искусственного интеллекта**

#### **3.1. Сенсорное восприятие**

Человек познает мир при помощи органов чувств. Чувственное восприятие, являющееся начальным звеном в цепи событий, связанных со сбором и обработкой информации (входных сигналов), основано на работе сенсорной системы индивида, которая включает пять видов чувствительности: зрение, слух, обоняние, осязание и вкус. Соответственно, восприятие инициируется внешними сигналами (стимулами), состоящими из света, звука, молекулярных соединений и давления. Эти стимулы, обнаруживаемые органами чувств, преобразуются (конвертируются) в сообщения, понимаемые мозгом. Восприятие явлений происходит одновременно с их интерпретацией (репрезентацией), осуществляемой ансамблями нейронов. Эта интерпретация основана на прошлых знаниях, которые, в свою очередь, корректируются и дополняются новыми сенсорными сигналами. Имеющиеся у нас ожидания (предварительные установки) относительно того, что мы можем увидеть или услышать, непосредственно влияют на наше восприятие, позволяя устранять неоднозначности в интерпретации сенсорной информации. С другой стороны, сенсорные сигналы периодически изменяют наши ожидания, задают рамки классификации и категоризации объектов внешнего мира. Здесь многочисленны

петли обратной связи, когда информационные потоки циркулируют вверх и вниз по сложной разветвленной иерархии. Человеческий мозг насквозь пронизан обратными связями. Например, обмен между новой корой мозга (неокортексом) и всеми остальными частями мозга, включая таламус (главный подкорковый центр), построен таким образом, что количество обратных связей превышает количество исходящих почти в десять раз! По словам Дж.Хокинса, «это значит, что на каждое волокно, подающее информацию в неокортекс, приходится десять волокон, отправляющих обратную информацию к органам чувств» [5, с. 33]. Следовательно, принцип обратной связи, обеспечивающий взаимодействие различных сенсорных структур с центрами хранения и обработки информации, должен стать одним из основных принципов работы искусственного интеллекта.

Другой важной особенностью восприятия и обработки входных сигналов в человеческом мозге является принцип формирования инвариантных структур. Когда мы видим, чувствуем или слышим, кора головного мозга получает впечатления, которые сохраняет в инвариантной форме. Способность к инвариантному опознанию возникает у живых организмов на относительно ранних стадиях онтогенеза. Согласно Д. Марру (1987), на этапе обучения зрительная система по двумерным проекциям предъявляемых объектов строит в мозге наблюдателя представления трехмерных объектов, которые запоминаются как шаблоны. В дальнейшем эти шаблоны, основанные на выделении ключевых признаков объектов и пренебрежении второстепенными признаками, используются для идентификации новых образов. Несмотря на то, что в силу варьирования угла наблюдения, масштаба и освещенности объекта сетчатка глаз получает множество различающихся между собой двумерных изображений, мозг выделяет инвариантные характеристики объекта и успешно распознает его. Для современного компьютера такая задача является очень сложной или вообще невыполнимой. Вычислительная машина не в состоянии принять решение о принадлежности предъявляемых проекций тому или иному объекту, то есть распознать изображение, если оно было подвергнуто перемещению, ротации, изменению масштаба или любому другому возможному превращению. Следовательно, важной проблемой является разработка таких систем искусственного интеллекта, которые обладали бы способностью к формированию инвариантных (абстрактных) структур при восприятии многочисленных и постоянно меняющихся входных сигналов.

### 3.2. Память

Объем памяти современных компьютеров постоянно увеличивается, но структура их памяти коренным образом отличается от способов хранения и извлечения информации человеческим мозгом. Можно перечислить основные особенности нашей памяти, описанные Дж. Хокинсом [5]:

- неокортекс запоминает последовательности элементов, а не отдельные элементы окружающего мира;
- неокортекс вспоминает последовательности автоассоциативно;
- неокортекс запоминает последовательности в инвариантной форме;
- неокортекс сохраняет последовательности иерархически.

Примером запоминания последовательности элементов является изучение алфавита. Все мы знаем алфавит, но если попробовать произнести его в обратном порядке, то легко убедиться, что это не такая уж простая задача. В нашей памяти алфавит представлен как последовательность сигналов. Именно поэтому он не поддается ментальному извлечению или извлечению в произвольном порядке. Запоминание песни — еще один пример временной последовательности. Зная песню, нельзя припомнить ее слова в обратном порядке или охватить своим вниманием песню целиком. Мы вспоминаем ее только в том порядке, в каком заучили. Точно так же обстоит дело с множеством другой информации. В противоположность нашему мозгу память компьютера не хранит последовательности сигналов. Как указывает Дж. Хокинс, с помощью разных программных доработок можно этого достичь, но память компьютера не способна на автоматическое выполнение такого задания.

Наша память имеет автоассоциативную природу. Нейроны, специализирующиеся на запечатлении входных сигналов, постоянно ассоциируют их друг с другом. Сравнение (сопоставление) различных элементов и фрагментов информации определяет нашу способность замечать сходство между ними, что, в конечном счете, позволяет проводить аналогию между фактами и идеями. В процессе решения сложных задач мозг по ассоциации извлекает из памяти знания, применявшиеся ранее при решении аналогичных задач, и использует их в новой ситуации. Даже слова хранятся в неокортексе в связанном (ассоциированном) состоянии. Это обстоятельство хорошо известно психолингвистам, которые ввели понятие логогена — элемента мо-

дели распознавания слов (блока памяти, который подобно «узлу» связывает все аспекты слова). Т. Н. Ушакова в [6] отмечает: «Совокупность множества межсловесных связей образует так называемую вербальную сеть (паутину). Вербальная сеть — психофизиологическое образование, вырабатываемое в детстве (при усвоении новых языков — в любом возрасте) и затем стабильно существующее в мозге в течение его жизни. Все известные человеку слова включаются в форме логогенов в ее структуру. Объединяясь множественными „межлогогенными“, межсловесными связями, логогены становятся образующими элементами, узлами вербальной сети» [6, с. 207]. Выделение логогена имени (слова) вызывает активацию соответствующего поля в вербальной сети. В активное состояние приходят связанные с найденным именем слова и вербальные клише. На основе диффузной активации выделяются адекватные случаю конкретные словесные логогены. Поиск всего набора слов, подходящих для выражения актуальной, словесно еще не оформленной интенции, происходит как движение (блуждание) по путям вербальной сети (сети нейронов, хранящих весь известный нам лексикон). Благодаря автоассоциативной природе нашей памяти мы способны по одному признаку определить (вспомнить) объект, а по объекту — всю совокупность признаков. Другими словами, наш мозг способен дополнять образы, воспроизводить полную картину на основе неполной или искаженной входной информации.

Современные компьютеры на это не способны. Ученые уже несколько десятков лет пытаются реализовать автоассоциативные воспоминания в искусственных нейронных сетях, но полученные результаты пока не очень впечатляют. Вычислительные машины осуществляют поиск знаний в памяти последовательно, а не ассоциативно, на что тратят колоссальное время. Как отмечает Дж. Люгер в [7], «люди быстрее справляются с задачами, когда получают больше информации, в то время как компьютеры, наоборот, замедляют работу. Это замедление происходит за счет увеличения времени последовательного поиска в базе знаний» [7, с. 52]. Одним из способов повышения эффективности памяти компьютеров может оказаться применение так называемых хаотических алгоритмов. Ряд исследователей связывают с такими алгоритмами большие надежды, в частности, рассчитывают наделить искусственный интеллект ассоциативным поиском информации. А. Дмитриев в [8] пишет: «Голубая мечта пользователей — возможность поиска мелодии, видеосюжета

или нужных фотографий не по их атрибутам (названию директории и файла, дате создания и т. д.), а по содержанию или ассоциации, чтобы, например, по фрагменту мелодии можно было найти и воспроизвести музыкальное произведение. Оказывается, такой ассоциативный поиск можно осуществить с помощью технологий на основе детерминированного хаоса» [8].

Четвертый принцип работы нашей памяти, отличающий мозг от вычислительных машин, — иерархический способ сохранения информации. Любая иерархическая система характеризуется тем, что одни элементы расположены выше, а другие — ниже. Все функциональные зоны головного мозга «обитают» в одной и той же ткани коры. Но одна зона оказывается «выше» или «ниже» другой в зависимости от того, как они связаны и взаимодействуют друг с другом. Первичные сенсорные зоны мозга, в которые непосредственно поступает информация об окружающем мире, являются низшими функциональными зонами. Эти области занимаются обработкой первичной информации на самом простом, базовом уровне. Например, зрительная информация поступает в кору головного мозга через первичную зрительную зону (зону V1). Зона V1 отвечает за восприятие мелких контурных сегментов, простых составляющих движения, основных цветов, сигналов о контрастности. Зона V1 посылает информацию в другие зоны, называемые вторичными зонами (зоны V2), которые из отдельных деталей формируют целостные образы. В свою очередь, вторичные зоны, обработав сигналы, передают их в ассоциативные зоны, где происходит сравнение (ассоциирование) разных целостных образов. Аналогичная иерархическая структура существует не только в зрительной коре, но и в других отделах мозга (слуховой, тактильной, двигательной и т. д.). Информационные сигналы в коре головного мозга передаются в двух направлениях: от зон низшего порядка к высшим и обратно, в нисходящем направлении, причем обратные информационные потоки, идущие сверху вниз, имеют большую информационную насыщенность. Очередная задача, стоящая перед компьютерной наукой, — разработать искусственные нейронные сети (ИНС), наделенные иерархической организацией памяти. Безусловно, это будет шагом на пути создания самообучающихся компьютеров, ведь самообучение предполагает последовательное восхождение от конкретного (детализированного, фрагментарного) знания к абстрактным (обобщенным, целостным) репрезентациям.

### 3.3. Понимание речи

В начале 1950-х годов многие думали, что компьютеры сильно помогут при переводе с языка на язык. Считалось достаточным просто загрузить в компьютер словарные эквиваленты, ввести один язык и получить на выходе другой. Однако даже если делается перевод один к одному в контексте синтаксической информации, результаты получаются довольно странные. Р. Солсо в [9] замечает: «Например, когда пассаж из Библии „Стремится дух, да плоть слаба“ перевели на русский, а затем обратно на английский, то получилось — „Вино было приятным, но мясо протухло“» [9, с. 480]. Современные программы по распознаванию речи работают успешно лишь в очень ограниченных ситуациях, когда количество слов, которые человек может произнести в каждое отдельное мгновение, строго ограничено. А вот людям распознавание речи дается без труда, потому что неокортекс не только воспринимает отдельные слова, но и предугадывает содержание целых предложений, а также рамки общего контекста. В процессе распознавания устной речи мы прогнозируем идеи, фразы, отдельные слова. Мало того, кора головного мозга выполняет всю эту работу автоматически. Необходимость прогнозирования связана с тем, что словесные сообщения, которыми обмениваются люди, зависят от контекста и ряда других вещей, не допускающих однозначной интерпретации речи. Многие современные психолингвисты подчеркивают вариабельность речевого сигнала, его принципиальное несовершенство, в результате чего слушающий не столько извлекает информацию из акустического речевого сигнала, сколько реконструирует сообщение, основываясь на части акустических признаков и используя эвристические процедуры, базирующиеся на максимально широких основаниях — от знания собеседника до общей картины мира. Некоторые ситуации восприятия и понимания речи включают этап вероятностной идентификации слов в рамках столь же вероятностно определенных синтаксических структур. В речи людей то, что не говорится, столь же важно для эффективного сообщения, как и то, что говорится. Чтобы понять скрытое содержание текста, необходимо использовать различные виды умозаключений.

Неспособность компьютера правильно интерпретировать речь объясняется тем, что он действует на основе заложенной в него программы, а не учится на собственном опыте. Ребенок осваивает лексикон, в котором каждое слово обозначает предмет, действие или харак-

теристику предмета, путем обучения. В свою очередь, взрослые оказывают постоянное воздействие на этот процесс обучения, поощряя правильные речевые действия ребенка и корректируя ошибки. При восприятии некоторой ситуации (предмета, действия) и одновременно звучащего слова, произносимого взрослыми, мозг ребенка связывает (ассоциирует) ситуацию со словом. Многократное повторение этих взаимосвязанных событий приводит к формированию стереотипов, то есть к пониманию значения тех или иных слов, которые сохраняются в памяти. Здесь возможны обобщения, описываемые по схеме: это слово часто звучало одновременно при предъявлении конкретного предмета или при выполнении действия, следовательно, данное слово имеет отношение к тому или иному предмету или действию. Подобные обобщения — важный элемент речевого онтогенеза. В конечном счете, они способствуют выработке обобщенной ассоциации (абстрактного стереотипа): каждая вещь или действие имеют имя. Со слов Т.Н. Ушаковой, «ассоциативные процессы, связывающие впечатления от воспринимаемых объектов с восприятием звучания называющих их слов, играют важную роль и в речевом онтогенезе. Наблюдения показывают, что они вступают в действие в очень раннем возрасте: свидетельством этого служат широко практикующиеся воспитателями разного рода словесные клише, приспособленные для выработки и поощрения межсловесных ассоциаций у детей (гуси, гуси — га-га-га; обезьяна — чи-чи-чи и др.). На основе объединения локально существующих словесных структур (часто это структуры, соответствующие отдельным именам) образуется „вербальная сеть“. Тогда имена теряют изолированность и превращаются тем самым в слова языка, его лексическую составляющую. Образуется стабильно сохраняющаяся область индивидуального знания, вербальная память» [6, с. 203]. Освоение слов (существительных, прилагательных и глаголов), а также первых правил грамматики, перенимаемых путем имитации, позволяет строить сначала простые фразы, а затем и сложные предложения.

Чтобы понимать речь, то есть смысл предложений, передающих сущность тех или иных ситуаций, компьютер должен, прежде всего, иметь опыт столкновения с этими ситуациями. Поэтому Дж. Хокинс отмечает: «Чтобы полностью понимать человеческую речь, машина должна многое „пережить“ и „научиться“ тому же, что и люди. Возможно, нам понадобятся долгие годы, чтобы создать разумную машину, которая понимает язык так же хорошо, как вы и я» [5, с. 213].

Об этом же говорят С. Рассел и П. Норвиг в [4]: «Задача перевода является сложной, поскольку, вообще говоря, для ее решения требуется глубокое понимание текста, а для этого, в свою очередь, необходимо глубокое понимание ситуации, о которой идет речь. Это утверждение является справедливым применительно даже к очень простым текстам, в частности, даже к „текстам“, состоящим из одного слова» [4, с. 1025]. Резюмируя, укажем, что искусственному интеллекту, способному понимать естественный язык, нужны как минимум: семантические и синтаксические правила; база знаний о мире и о социальном контексте; какие-нибудь методы обработки неоднозначностей, имеющих в обычно употребляемом языке.

### 3.4. Выполнение сложных двигательных актов

Человек не смог бы приобретать знания, если бы не обладал телом. Чтобы получать информацию о структуре объектов, окружающих его, нужно физически воздействовать на них, причем формы воздействия столь многочисленны, что их невозможно перечислить. Одним из факторов, обусловивших рост знаний человека, послужило изготовление орудий. Для выполнения десятков разнообразных двигательных операций, связанных с изготовлением орудий, требовалась рука, освобожденная от функции ходьбы. С появлением экспериментальной науки, давшей мощный толчок общественному развитию, роль человеческой руки многократно возросла. Ведь постановка любого эксперимента связана с «ручным» изготовлении различных приборов и технических средств, позволяющих проникать в тайны природы. Ярким примером может служить постановка экспериментов, позволивших исследовать необычные электронные свойства графена, за что отечественные физики Константин Новоселов и Александр Гейм были удостоены в 2010 году Нобелевской премии. А. Гейм в своей Нобелевской лекции [10] пишет о том, какой невероятно тонкой работы пальцев требовало изготовление кристаллов графита для изучения их электронных свойств: «Для того чтобы пинцетом перенести такой кристалл (толщиной  $\sim 20$  нм) со скотча, а затем сделать четыре близко расположенных контакта с помощью всего лишь серебряной пасты и зубочистки, требуется высочайший уровень экспериментаторского мастерства. В наше время немногие экспериментаторы имеют пальцы достаточно ловкие для того, чтобы изготовить такие образцы» [10, с. 1290].

Работа рук, о которой мы говорим, должна стать принадлежностью искусственного интеллекта, если он хочет приблизиться к совершенству человеческого разума (разум без руки — беспомощное дитя). В настоящее время тонкость и изящность движений руки недоступны машинам (роботам). Когда специалисты пытаются запрограммировать движение руки робота так, чтобы он мог поймать летящий мяч, им приходится решать сложные математические уравнения. Сначала они пытаются вычислить траекторию полета мяча и определить его пространственное расположение в момент контакта с рукой. Затем они занимаются вычислением, определяя все стобы руки робота, то есть положения его руки в различные моменты времени. Но человеческий мозг решает задачу перехвата мяча иначе. Он просто обращается к памяти, в которой хранятся моторные команды для поимки мяча. Эти моторные команды появляются в памяти после того, как вы методом проб и ошибок уже обучились ловить летящий мяч. Мозг не решает никаких сложных уравнений, он просто автоматически активизирует воспоминания о последовательности движений, необходимых для перехвата того или иного предмета. Дж. Хокинс в [5] описывает эту ситуацию следующим образом: «Воспоминание о том, как нужно ловить мяч, не было запрограммировано в вашем мозге. Вы обучились этому в процессе жизни, соответственно информация не вычисляется, а сохраняется в нейронах коры вашего головного мозга» [5, с. 72]. Память о моторных командах, присущая человеческому мозгу, подчиняется тем же принципам, что и память о зрительных, слуховых и тактильных сигналах: 1) запоминание последовательности элементов, 2) автоассоциативное извлечение последовательностей элементов, 3) запоминание последовательностей элементов в инвариантной форме, 4) иерархическая организация памяти. Так же, как и в случае формирования следов памяти на визуальные и слуховые сигналы, обратная связь играет существенную роль в формировании информации о моторных командах в нашем мозге. Отметим, что для совершенного владения рукой и другими конечностями искусственный интеллект нуждается не в программировании, а в самообучении. С легкой руки теоретиков искусственного интеллекта это самообучение на примерах получило название «индуктивного обучения». А. А. Жданов, анализируя ходьбу современных роботов в [11], пишет: «... Даже в самых современных роботах не удастся смоделировать ходьбу человека ни в отношении ее скорости, ни в отношении ее свойств. И более того, известные эффектно выглядящие антро-

поморфные шагающие роботы шагают совсем не по тем принципам, по которым шагает человек, и потому их прогресс в умении ходить отнюдь нельзя сравнивать с развитием процесса ходьбы от обезьяны к человеку. Современные роботы ходят только благодаря точной математической модели процесса ходьбы, а не в результате самообучения» [11, с. 251]. Таким образом, искусственный интеллект должен быть наделен способностью к выполнению сложных двигательных актов, для чего придется оснащать его системой хранения информации о моторных командах, подчиняющейся принципам ассоциативного, иерархического и инвариантного кодирования этой информации.

### 3.5. Проблема связности

Ввиду того, что чувственное восприятие, понимание речи, выполнение сложных двигательных актов, реализация других функций невозможны без иерархической системы памяти, главной задачей при создании разумной машины является создание подобной системы хранения и извлечения информации. В свою очередь, основной трудностью при создании указанной системы памяти будет решение проблемы связности, подобной связности миллиарда нервных клеток (нейронов) живого мозга. Дело в том, что наша уникальная способность запоминать, хранить и извлекать по принципу ассоциации ту или иную информацию определяется координированной работой десятка миллиардов нервных клеток (нейронов), каждая из которых может быть связана с 5 или 10 тысячами других таких же клеток. Используя интегральные схемы (кремниевые транзисторы), являющиеся элементной базой сегодняшних компьютеров, нельзя даже надеяться достичь такого же масштабного параллельного соединения, характерного для структур нашего мозга. Следовательно, нужно отказаться от традиционных интегральных схем (полупроводниковых микрочипов) и искать новые подходы. Одним из способов решения проблемы связности (проблемы масштабного параллельного соединения) может быть создание «молекулярных» компьютеров. Недавно сотрудники американской корпорации «IBM» обнаружили, что базовым компонентом будущих компьютеров может быть молекула органического соединения нафтаलोцианина или похожего по свойствам соединения. Исследователям удалось показать возможность переключения одиночной молекулы нафталоцианина в положения «включено — выключено» без изменения ее формы. Сделано это бы-

ло с помощью двух атомов водорода из молекулы нафтаलोцианина. В более ранних экспериментах при переключении состояний молекулы изменяли свою форму, что делало их непригодными для построения логических вентилях элементов памяти компьютера.

Таким образом, полноценные процессы самообучения станут доступны компьютеру, когда он приобретет мощную автоассоциативную систему памяти, а для создания такой памяти нужно решить техническую проблему связности. Как только будут найдены средства ее решения, компьютеры реально приблизятся к возможностям человеческого интеллекта.

## **4. Анализ факторов, исключающих возможность алгоритмизации творческой деятельности в сфере науки**

### **4.1. Вероятностная природа индуктивного вывода**

Несмотря на то, что исследователями разработано множество различных стратегий, приемов и правил, позволяющих решать сложнейшие научные задачи, генерировать новые идеи и даже совершать открытия в области науки и техники, в творческой деятельности человека имеются факторы, исключающие возможность строгого алгоритмического описания этой деятельности. Эти факторы не позволяют рассчитывать на то, чтобы однажды был придуман алгоритм (система предписаний), всегда гарантирующий правильное решение.

С первым из таких факторов мы сталкиваемся при анализе логических средств (логических процедур), используемых человеком в научном творчестве. Разумеется, речь идет не о дедукции, ведь дедуктивные рассуждения допускают формализацию, то есть разработку эффективного алгоритма, превращающего определенный тип рассуждений в вычислительный процесс. Если же известен алгоритм (последовательность действий, устойчиво приводящих к решению), то на его основе всегда можно создать компьютерную программу. В данном случае речь идет об индуктивном способе обработки информации. При этом мы вкладываем в понятие индукции не только привычное восхождение от единичных фактов к общим заключениям, но и движение от частного к частному (аналогии). В большинстве случаев ученые используют неполную индукцию, в которой заклю-

чение не следует с необходимостью из посылок и может приводить к ошибкам. Можно указать несколько причин, почему неполная индукция находит более широкое применение в научной деятельности (то есть почему полная индукция часто оказывается недостижимой). Во-первых, элементы (составные части) многих множеств столь многочисленны, что ученому часто не хватает времени жизни на перебор и анализ этих элементов. Если же множество бесконечно (то есть составляющие его элементы неперечислимы), анализ этих элементов, необходимый для реализации полной индукции, становится недостижимой целью. Во-вторых, предметы (объекты) того или иного множества оказываются недоступными, а материальные ресурсы, способные обеспечить их доступность для исследования, ограниченными. Ввиду ограниченности ресурсов мы вновь не можем использовать полную индукцию, поэтому удовлетворяемся неполной. В-третьих, объекты могут быть недоступными для исследования, так как наука еще не достигла требуемого уровня развития, то есть уровня совершенства экспериментального оборудования, с помощью которого можно исследовать эти объекты. Высокая частота использования неполной индукции, которая не охватывает всех элементов множества, о котором делается обобщающий вывод, определяет справедливость высказываний многих авторов (Г. Лейбница, Б. Рассела, Д. Пойа и т. д.) о том, что индукция дает не истину, а ее вероятность. Это же обстоятельство определяет корректность отнесения индукции и аналогии к категории правдоподобных рассуждений.

Поскольку индуктивные стратегии мышления гарантируют не истинность, а ее вероятность, предпринимались попытки применить к индукции математическую теорию вероятности (статистики). Однако эти попытки не имели успеха, так как перед началом исследования того или иного множества его структура (количество его типичных элементов) не известна, а в момент завершения исследования (когда структура множества становится изученной) ретроспективная оценка степени правдоподобности индуктивных выводов теряет свою ценность. Касаясь этого вопроса, Б. Рассел в [12] пишет: «Со времени Лапласа делались различные попытки показать, что вероятная истинность индуктивного вывода вытекает из математической теории вероятности. Теперь всеми признается, что все эти попытки были безуспешными. . . » [12, с. 352]. «Во-первых, — аргументирует Б. Рассел, — в математической теории вероятности нет ничего, что оправдывало бы наше понимание как общей, так и частной индукции как

вероятной, как бы при этом ни было велико установленное число благоприятных случаев» [12, с. 361]. Об этом же говорит Д. Пойа в [13]: «Никто еще не предложил ясного и убедительного метода вычисления правдоподобностей в нетривиальных случаях, и если мы ясно себе представим конкретные ситуации, в которых важна правильная оценка правдоподобностей (что мы уже сделали), то мы легко сможем понять, что любое приписывание правдоподобностям определенных числовых значений подвергается большой опасности показаться глупым» [13, с. 368].

Возможность ошибиться (прийти к неправильным заключениям) при использовании индукции и аналогии лишает нас перспективы формализовать эти виды рассуждений так, как это может быть сделано в отношении дедукции. Для индуктивных умозаключений нельзя разработать эффективный алгоритм, некий вычислительный процесс, устойчиво предоставляющий в наше распоряжение достоверные результаты. Соответственно, компьютерная программа, реализующая индукцию, в одних случаях будет выдавать релевантную (соответствующую реальности) информацию, а в других случаях — умозаключения, не отражающие эту реальность.

Конечно, наука располагает средствами, позволяющими устранять наши ошибки. Для того, чтобы убедиться в истинности или ложности того или иного индуктивного заключения, необходимо провести дополнительное исследование (например, непосредственно обратиться к опыту, практике, к сопоставлению полученного заключения с другими, уже доказанными положениями в науке). Если опыт (наблюдение, эксперимент), призванный проверить индуктивную догадку, подтверждает ее, то она сохраняется в арсенале научного знания, а если опровергает, то идея отбрасывается, открывая простор для других идей (гипотез). В ряде случаев ученый, выдвинувший ошибочную гипотезу, впоследствии сам же ставит опыт (эксперимент), доказывающий или разрушающий ее содержание. Или, по крайней мере, заимствует схему постановки эксперимента у своих коллег. Если же условия постановки эксперимента выходят за рамки имеющихся у него материальных ресурсов или уровня развития науки его времени, то возникшую однажды идею проверяют следующие поколения исследователей.

Необходимость в проведении дополнительных исследований (опытов, экспериментов) для выяснения истинности или ложности индуктивных заключений означает, что искусственный интеллект, наде-

ленный индуктивной логикой, одновременно должен обладать способностью к проведению подобных дополнительных исследований. Любой ученый-экспериментатор, отвечая на вопрос о том, как ставятся эксперименты, доказывающие или фальсифицирующие ту или иную концепцию, ответит, что придумать новый эксперимент — это творческая работа. История науки показывает, что условия постановки важных экспериментов обнаруживаются так же, как все новое и значимое: а) методом проб и ошибок (методом последовательного перебора), б) в результате обобщения фактов эффективности определенных экспериментальных схем, найденных другими исследователями, в) благодаря применению аналогии, дополненной приемом синтеза разрозненных технических идей, почерпнутых из разных областей. Изложенное свидетельствует о том, что искусственный интеллект, проверяя достоверность своих индуктивных идей, должен проводить дополнительные эмпирические исследования, предполагающие постановку новых экспериментов, изобретение таких опытов, которые позволят подтвердить или поставить под сомнение те или иные теоретические построения. Поскольку не существует алгоритма изобретения экспериментальных (технических) средств проверки новых идей, вычислительная машина должна создавать эти средства методом проб и ошибок (методом последовательного перебора), индуктивно обобщая факты эффективности уже разработанных экспериментальных схем, осуществляя перенос технических идей из одной области в другую на основе аналогии. Если время жизни вычислительной машины будет ограничено, как в случае человека, задачу постановки новых экспериментов придется решать другим поколениям вычислительных машин, в распоряжении которых окажутся более значительные материальные ресурсы и более высокий уровень научного знания.

#### 4.2. Метод проб и ошибок

Очередным фактором, исключающим возможность алгоритмизации творческой деятельности в науке, является метод проб и ошибок (метод последовательного перебора). Метод проб и ошибок или, другими словами, стратегия пошагового исключения играет колоссальную роль в научном исследовании. Когда нет информации, то есть исходных посылок для индуктивного обобщения, а также когда отсутствуют идеи (теоретические конструкции), позволяющие реа-

лизовать аналогию, перенести эти идеи и способы решения из одной области в другую, остается использовать метод последовательного перебора. В зависимости от информационной обеспеченности данный прием исследования можно подразделить на две категории: 1) абсолютно слепой перебор, предполагающий широкомасштабную проверку всех возможных вариантов без опоры на какую-либо предварительную информацию, 2) последовательный перебор, учитывающий прошлые знания, то есть сочетающийся с различными догадками и эвристиками, применяемыми для сокращения перебираемых вариантов (альтернатив). Абсолютно слепой перебор означает отсутствие видимых указателей. В этом случае любое возможное направление поиска выглядит столь же многообещающим или, наоборот, ничего не сулящим, как и все другие. В слепом методе проб и ошибок сплошной просмотр вариантов осуществляется без каких-либо намеков (сведений) о том, окажется ли один из вариантов достойным отбора и какой именно. В противоположность этому перебор, учитывающий прошлые знания, не избавляет от ошибок, но дает возможность корректировать их в свете тех целей, которых нужно достичь. Запоминание неудачных проб предоставляет в наше распоряжение информацию об областях поиска, в которых уже не следует пытаться найти решение, что приводит к экономии материальных ресурсов и времени. Сопоставление информации, полученной на каком-то этапе перебора, с уже имеющимися фактами и идеями, повышает эффективность поиска за счет многочисленных петель обратной связи (обратную связь можно описать схемой: проба → сопоставление с идеями → проба → сопоставление с идеями). В реальном научном исследовании метод последовательного перебора, опирающийся хотя бы на минимум предварительной информации, встречается гораздо чаще, нежели абсолютно слепой перебор. В свое время на это обратил внимание английский философ и методолог науки Карл Поппер, который в одной из статей, содержащихся в [14], пишет: «... Движения ищущего не будут полностью случайными. Тому есть разные причины, как позитивные, так и негативные. Позитивные в основном сводятся к тому, что у ищущего есть проблема, которую он должен решить, а это означает, что у него есть какое-то знание — пусть сколь угодно туманное, — приобретенное ранее тем же по существу методом проб и ошибок; это знание служит ему руководством, что исключает полную случайность» [14, с. 149].

Метод проб и ошибок применяется не только в эмпирической (экспериментальной) области науки, но и на уровне генерации и развития сложных теорий и концепций, представляющих собой обобщение и синтез различных экспериментальных фактов и идей. Говоря словами Дональда Кэмпбелла, глубоко проанализировавшего роль метода проб и ошибок в процессе человеческого познания в очерке «Эволюционная эпистемология» [14], на одном конце шкалы — экспериментатор, использующий эвристику сплошного перебора в рамках возможностей данного лабораторного оборудования, пробуя каждый параметр и перебирающий все мыслимые сочетания (комбинации) без оглядки на теорию. На противоположном конце шкалы мы видим «естественный» отбор научных теорий, которые в режиме проб и ошибок соревнуются друг с другом в адекватности решения различных проблем, то есть в адекватности (соответствии) этих теорий общей совокупности накопленных фактических данных.

Ученые, изучавшие возможности логики (индукции и дедукции) в обобщении исходных посылок, часто подчеркивали то обстоятельство, что для открытия самих посылок не предусмотрено никакого метода (процедуры). Такой точки зрения, в частности, придерживался М. Бунге, который в [15] отмечал, что логические приемы обработки информации оставляют открытым вопрос о том, имеется ли какая-либо отработанная и стандартизированная стратегия поиска фактов, которые в дальнейшем подвергаются обобщению. Как ни странно, такой стратегией является метод проб и ошибок (метод последовательного перебора). Широкомасштабная проверка всех альтернатив, лежащая в основе этого метода, часто оказывается тяжелой, требующей значительных затрат материальных ресурсов и времени, но именно она обеспечивает ученых исходным знанием, необходимым для решения проблем, знанием, запускающим ментальные процессы индукции и аналогии.

История науки изобилует примерами открытий и изобретений, сделанных благодаря стратегии последовательного перебора (скрининга), в котором ошибки совершались не менее часто, чем правильные шаги. К числу открытий, сделанных благодаря методу проб и ошибок, можно отнести открытие закона эллиптического движения планет (Иоганн Кеплер), изобретение нити накаливания для электрической лампы (Томас Эдисон), синтез препарата для лечения сифилиса сальварсана (Пауль Эрлих, Нобелевская премия за 1908 год), промышленный синтез аммиака (Фриц Габер, Нобелевская премия за

1918 год), открытие антибиотика актиномицина (Зельман Ваксман, Нобелевская премия за 1952 год), открытие химического мутагеназа (Иосиф Рапопорт), открытие фактора роста нервов (Рита Леви-Монтальчини, Нобелевская премия за 1986 год), создание баллистических ракет (Сергей Королев), доказательство математической теоремы о четырех красках (Кеннет Аппель и Вольфганг Хакен).

Применительно к искусственному интеллекту это означает, что для изучения внешнего мира, постижения его законов, изобретения предметов и вещей, не противоречащих этим законам, искусственный интеллект должен использовать метод проб и ошибок. Использовать прием решения проблем, лишенный алгоритмических свойств, не гарантирующий легкость достижения цели. При отсутствии предварительных знаний в новой области (за границами уже известного) вычислительной машине не удастся избежать абсолютно слепого перебора, то есть ширококомасштабной проверки всех возможных вариантов без опоры на какую-либо значимую информацию. Если же последовательный перебор будет сочетаться с различными догадками и эвристиками, применяемыми для сокращения перебираемых вариантов, то и тогда вычислительная машина не сумеет полностью избавиться от ошибок. Запоминание неудачных проб позволит искусственному интеллекту иметь информацию о неперспективных областях поиска, экономить материальные ресурсы. Но в любой своей форме (даже в случае оснащенности эвристиками) метод проб и ошибок является дорогостоящим процессом исследования. Наука, созданная человеком, преодолевает некоторые негативные аспекты метода проб и ошибок за счет параллельности поисков. Параллельность поисков означает, что ежедневно (и даже ежеминутно) над одной и той же проблемой работают сотни и тысячи людей, разделенных друг от друга языками, странами, континентами, исследовательскими лабораториями, разбросанными по всей планете. Значительное количество ученых, решающих одни и те же задачи, часто приводит к одновременным и независимым открытиям (история знает немало примеров таких повторных открытий). Чтобы исключить высокую степень «повторности», организаторы науки проводят многочисленные научные конференции, преследующие цель своевременно информировать представителей разных научных коллективов о полученных результатах, согласовать и скоординировать дальнейшие действия. Излишне говорить, что если бы перед искусственным интеллектом была поставлена задача сравниться с человеком в получении важных

научных и технических результатов, для этого потребовалось бы создать не тысячу вычислительных машин, наделенных индуктивной логикой и способностью извлекать новое знание из экспериментов и наблюдений, а гораздо больше. В противном случае параллельность поисков не привела бы к заметным успехам.

Возвращаясь к вопросу о методе проб и ошибок, укажем, что одним из первых, кто осознал необходимость наделить искусственный интеллект этим методом, весьма далеким от совершенства алгоритмических программ, был уже упоминавшийся нами американский психолог, социолог и философ Дональд Кэмпбелл. В очерке «Эволюционная эпистемология», представленном в [14], Д. Кэмпбелл подчеркивает: «... Компьютер, который генерировал бы свою собственную эвристику, должен был бы делать это путем слепых проб и ошибок при нащупывании эвристических принципов, а отобранные принципы представляли бы накопленное общее знание» [14, с. 113]. Нам остается сожалеть, что Д. Кэмпбелл впервые высказал эту мысль в 1974 году, когда еще не была поставлена 18-я проблема С. Смейла, к решению которой эта мысль имеет прямое отношение.

### 4.3. Фактор случая в научном открытии

Давно установлено, что многие природные процессы нельзя описать без использования вероятностных представлений. Это обстоятельство приводит к тому, что становится невозможным точное предсказание поведения большинства динамических систем. Степень стохастичности исследуемых процессов определяет так называемый горизонт прогноза (пределы предсказуемости) для таких систем. Нечто аналогичное имеет место в творческой деятельности. Например, нельзя с высокой степенью точности предсказать будущее той или иной научной дисциплины, то есть предвосхитить открытия, которые будут сделаны в этой дисциплине, даже если мы обладаем обширной информацией о том, каковы ее успехи в настоящее время. Неожиданные, незапланированные события, которые могут произойти в науке, не позволяют предвидеть все детали развития конкретной отрасли знания. Один из изобретателей лазеров, лауреат Нобелевской премии по физике, Чарльз Таунс в [16] пишет: «Элемент неожиданности — постоянная составная часть технического прогресса, и это как раз то, что невероятно трудно совместить с любым из обычных принципов планирования» [16, с. 160]. «Можно ли, — спрашивает Ч. Та-

унс, — запланировать новую идею и новое, пока еще не известное техническое изобретение? Конечно, нет. Мы не можем доказать, что данное научное направление приведет к новым техническим достижениям, если мы пока не знаем даже сути этих достижений» [16, с. 160]. Об этом же пишет израильский физик, разделяющий с М. Гелл-Манном честь разработки классификации элементарных частиц, Ювал Нееман в [17]. Обсуждая вопрос о невозможности точного предсказания и планирования достижений науки, израильский физик обращается к теме распределения грантов: «Обычно фонд, предоставляющий грант, требует подачи заявки, включающей план предполагаемых исследований и их цели. Очевидно, что открытие, совершаемое благодаря везению, не может быть предсказано. Таким образом, наиболее важные результаты никогда не будут фигурировать в заявках. Следовательно, тот, кто предоставляет гранты, не должен относиться к заявкам слишком серьезно» [17, с. 86–87].

Почему невозможно запланировать то или иное открытие с высокой степенью точности? Другими словами, что является причиной определенной непредсказуемости новых научных достижений? Ответ на эти вопросы подсказывает история науки: многие открытия являются случайными или, по крайней мере, обусловленными определенной долей случайности. Уже обсуждавшийся нами метод проб и ошибок (метод последовательного перебора), предполагающий методический просмотр вариантов с целью решения конкретной задачи, часто дает возможность обнаружить нечто, не входившее в планы и намерения исследователя. Эта ситуация хорошо иллюстрируется судьбой Христофора Колумба, который искал морской путь в Индию, а в действительности открыл Америку. Если называть подобные незапланированные (непреднамеренные) открытия побочными результатами основной линии исследований, то следует отметить, что эти побочные результаты обычно подсказывают решение совсем других задач, не связанных непосредственно с теми проблемами, которые стимулировали начальные поиски. Примечательно, что случайные открытия невозможно исключить никаким количеством информации, которой вы владеете, принимаясь осваивать новую область. Ведь эта информация представляет собой уже оформленное и зафиксированное знание, а творчество предполагает выход за границы известного, за границы, где нет ориентиров и указателей.

Мы уже говорили о том, что метод проб и ошибок является поставщиком исходных посылок (единичных фактов) для логической

операции индукции, посредством которой обрабатываются эти послышки. Поскольку последовательный перебор вариантов, реализуемый с определенной целью, часто порождает на свет побочные находки (случайные открытия), дающие ключ к решению проблем, которые изначально не находились в поле нашего зрения, необходимо подчеркнуть, что фактор случая также является поставщиком (источником) исходных посылок для индукции. Другими словами, можно выделить индукцию, основанную на методе проб и ошибок, и индукцию, базирующуюся на факторе случая. Таким образом, элемент вероятности, свойственный неполной индукции, дополняется элементом вероятности, определяемым случайными находками, которые обеспечивают индукцию (вне зависимости от степени ее полноты) фактами, подлежащими обобщению. Перед нами двойная вероятность индуктивной логики. Однако не следует проникаться чрезмерным пессимизмом, ведь наука успешно развивается, несмотря на эти неалгоритмические аспекты индуктивного мышления. Такие ученые, как Дж. Максвелл, Л. Больцман, А. Пуанкаре, В. Гейзенберг, М. Борн, Э. Лоренц и И. Пригожин, научили нас спокойно воспринимать вероятностный (стохастический) характер закономерностей окружающего мира. Элементы вероятности, присущие методу проб и ошибок и индукции, наделяющие эти творческие стратегии качеством неалгоритмичности, можно воспринимать аналогичным образом.

Среди научных открытий, в возникновении которых фактор случая играл не последнюю роль, мы находим открытие связи между электричеством и магнетизмом (Христиан Эрстед, 1820), изобретение фотографии (Луи Дагер, 1835), изобретение телефона (Александр Белл, 1875), открытие электромагнитных волн (Генрих Герц, 1886), обнаружение рентгеновских лучей (Вильгельм Рентген, 1895) открытие радиоактивности (Анри Беккерель, 1896) и т. д. Примером случайного открытия, сделанного совсем недавно и удостоенного Нобелевской премии, является открытие полимеров, способных проводить электрический ток подобно металлам. Существенную роль в этом открытии, совершенном японским химиком Хидеки Широкавой (Нобелевская премия по химии за 2000 год), сыграла ошибка, допущенная его аспирантом при постановке эксперимента. В. А. Марихин в [18] пишет: «А в 1971 году профессор Токийского технологического института Хидеки Широкава дал своему аспиранту задание синтезировать полимер ацетилена. Впервые полиацетилен был получен еще в 1955 году в виде темного порошка, не обладающего никакими осо-

бо выдающимися свойствами. Однако аспирант по ошибке добавил в реакционную смесь в 1000 раз больше катализатора, чем требовалось по методике (наверное, перепутал граммы с миллиграммами), в результате чего вместо темного порошка получил роскошную пленку с металлическим блеском. Едва взглянув на эту пленку, Широкава подумал, что она может послужить основой для создания полимеров, обладающих свойствами металлических проводников» [18]. Об этом же сообщает М. Рыбалкина в [19]: «Как это часто бывает в истории науки, открытию помогла случайность. Студент Широкавы как-то по ошибке добавил слишком много катализатора, в результате чего бесцветный пластик вдруг стал отражать свет подобно серебру, и это навело на мысль, что он перестал быть изолятором. Дальнейшие исследования привели к открытию полимера с проводимостью, в десятки миллионов раз превосходящей обычный пластик. Это открывает путь к новой электронике XXI века, основанной на органических материалах» [19, с. 187]. Разумеется, с точки зрения методологии научного исследования, подобные ошибки, приводящие к открытиям, нельзя предвидеть и заранее планировать.

Решая проблему С. Смейла о том, каковы пределы естественного и искусственного интеллектов, мы вынуждены постоянно сравнивать когнитивную деятельность человека и вычислительных машин. К каким результатам можно прийти, проводя такое сравнение, если нам известно, что фактор случая играет существенную роль в научном (творческом) поиске? Предлагаемый нами ответ может показаться парадоксальным, но он логически вытекает из реального положения вещей: искусственный интеллект должен уметь делать случайные открытия. Ему необходимо приобрести способность получать незапланированные (непреднамеренные) научные результаты, являющиеся побочным продуктом тех или иных исследований. Решая определенную задачу и сталкиваясь с ситуацией, когда случайно обнаруживаются факты (сведения), содержащие решение совсем другой задачи, он должен радикально менять направление поисков, отказываться от прежних гипотез, если они не согласуются с тем, что удалось узнать благодаря случайному, никем не ожидаемому открытию. Фактор случая, присутствующий в научном поиске и нередко определяющий его успешность (эффективность) — кульминация неалгоритмической (невычислимой) природы творческого процесса.

#### 4.4. Теорема Геделя о неполноте

В уже упоминавшейся лекции «Математические проблемы следующего столетия» [1] Стивен Смейл формулирует проблему определения границ человеческого и искусственного интеллектов, ссылаясь на монографию Р. Пенроуза «Новый ум короля» [20]. С. Смейл обращает, прежде всего, внимание на анализ теоремы Геделя о неполноте, проведенный Р. Пенроузом. Р. Пенроуз напоминает о том, что теорема Геделя о неполноте дала отрицательный ответ на вопрос Д. Гильберта, возможно ли доказать непротиворечивость аксиом арифметики, исходя из самих аксиом арифметики. К. Гедель (1931) показал, что любая формальная арифметическая система аксиом и правил вывода должна включать в себя утверждения, которые не являются ни доказуемыми, ни опровержимыми в рамках данной системы. Изначально геделевский результат имел отношение к арифметике, но впоследствии он был обобщен на все формальные системы, не обязательно арифметические. Заслуга такого обобщения принадлежит А. Тарскому и его школе, о чем пишет Л. Д. Беклемишев в [21]. Следует отметить, что и сам К. Гедель догадывался о том, что его теорема о неполноте имеет отношение не только к арифметике. А. М. Хазен в [22] приводит абстрактную формулировку теоремы о неполноте, автором которой является сам К. Гедель: «Полное эпистемологическое описание языка  $A$  нельзя осуществить на том же языке  $A$ , ибо в языке  $A$  нельзя определить понятие истинности предложений языка  $A$ » [22, с. 281].

Рассматривая теорему Геделя о неполноте в ее обобщенной формулировке, Р. Пенроуз усматривает в ней свидетельство того, что не существует механической процедуры (совокупности правил) для решения всех задач, независимо от степени их сложности. Учитывая тот факт, что математики открывают теоремы и ищут их доказательства, не обладая каким-либо универсальным алгоритмом, в самом себе содержащим критерии истинности, Р. Пенроуз приходит к выводу, что в работе нашего сознания присутствует существенно неалгоритмическая составляющая. Следовательно, человеческий интеллект не может надлежащим образом моделироваться (описываться) строгими алгоритмами, в которых нет места неопределенности и вероятности. С точки зрения Р. Пенроуза, теорема Геделя о неполноте является ограничением не только для человеческого интеллекта, но и для тех вычислительных машин, которые мы надеемся наделить однажды признаками «разумности».

В каком смысле следует понимать это ограничение? Чтобы ответить на этот вопрос, необходимо обратить внимание на то, что теорема Геделя имеет отношение к замкнутым формальным системам (к закрытым алгоритмам). Поскольку основной задачей науки является исследование окружающего мира путем проведения разнообразных опытов и экспериментов, легко понять, что замкнутой формальной системой окажется любая система, не заимствующая необходимую информацию из опыта и эксперимента. По сути дела, К. Гедель, изложив свою теорему о неполноте, давшую отрицательное решение проблемы Д. Гильберта, привлек наше внимание к тому, что любые алгоритмы (алгоритмические системы), направленные на исследование внешнего мира, не могут быть эффективными, если они являются закрытыми.

О том, что ограничения, вытекающие из результата Геделя, распространяются лишь на замкнутые (закрытые) формальные системы, пишут многие авторы. А. С. Потапов в книге [23] указывает: «... Многие аргументы перестают работать применительно к открытым безостановочным алгоритмическим системам (или воплощенному интеллекту). В частности, для таких систем некорректным оказывается доказательство неразрешимости проблемы останова (и множество следствий из него), а также неприменимыми оказываются теоремы Геделя, накладывающие ограничения на возможности формальных систем. Обоснования недостижимости функции понимания, отчасти справедливые для замкнутых систем, оказываются некорректными по отношению к системам, осуществляющим формирование понятий в результате взаимодействия с реальным миром» [23, с. 702]. Об этом же сообщает А. Н. Кочергин в статье [24]: «Существует обоснованное мнение о том, что ограничения, вытекающие из теоремы Геделя, относятся лишь к машинам Тьюринга, не получающим из внешней среды никакой информации. При условии, если машина получает информацию из внешней среды, являющейся бесконечной информационной системой, она оказывается способной решать неконструктивные проблемы, относительно которых можно было бы доказать их алгоритмическую неразрешимость» [24, с. 215].

Если существуют способы снятия ограничений, налагаемых теоремой Геделя о неполноте на человеческое мышление и деятельность вычислительных машин, то почему же результат Геделя все-таки является одним из принципов, позволяющих дать правильный ответ на вопрос американского математика С. Смейла о границах (пределах)

человеческого и искусственного интеллектов? Потому что теорема Геделя о неполноте относится к числу факторов, исключаящих возможность строгой алгоритмизации творческой деятельности в сфере науки (и не только в этой сфере). Если бы не существовало указанной теоремы, отражающей ключевые особенности человеческого познания, то можно было бы постигать законы природы без обращения к опыту и эксперименту, то есть изучать внешний мир на основе универсального алгоритма, в самом себе содержащего гарантии достоверности получаемых результатов. Между тем такой универсальный метод никому не известен, поэтому ученым приходится изучать окружающий мир методом проб и ошибок (методом последовательного перебора), часто наталкиваясь на непреднамеренные и незапланированные открытия, индуктивно обобщая единичные факты до гипотез и идей, которые постоянно проверяются наблюдением и экспериментом. Необходимость подобных проверок, как мы уже сказали, связана с вероятностной природой индуктивного вывода, то есть с тем, что в индуктивном умозаключении часто не охватывается вся совокупность элементов, относящихся к тому множеству, относительно которого мы делаем обобщающий вывод.

Искусственному интеллекту, который будет работать не на основе жесткой компьютерной программы, а благодаря процессам самообучения, потребуется овладеть теми же процедурами получения нового знания: широкомасштабным перебором вариантов, индуктивной стратегией обработки информации, искусством находить одну вещь, когда ищешь другую (методом случайных открытий). Без процессов самообучения, непрерывного накопления опыта искусственному интеллекту не удастся преодолеть ограничений, накладываемых теоремой Геделя о неполноте, запрещающей эффективное функционирование закрытых алгоритмических (не обучающихся) систем. Образно выражаясь, сама теорема Геделя о неполноте, точнее, ограничения, связанные с ней, заставляют нас пойти по пути создания вычислительных машин, способных к обучению.

В современной информатике выделяют два способа обучения: с учителем и без него. Конечно, вычислительная машина должна освоить способ обучения без учителя, так как этот способ, являющийся наиболее сложным, обладает и наиболее широкими возможностями. Ведь именно этот тип обучения позволяет получить что-то действительно новое. По свидетельству А. С. Потапова, «по сути, значительную часть научных исследований можно трактовать как обучение

без учителя (или, в крайнем случае, как обучение с подкреплением). Откуда возьмется информированный учитель, когда решения не известны никому из людей и лишь природе можно задать вопрос о правильности решения, и только тогда, когда есть возможность выйти за рамки чистых наблюдений и поставить правильный эксперимент?» [23, с. 224].

## 5. Четыре аспекта методологии научного творчества, проливающие свет на 18-ю проблему С. Смейла

Итак, что ожидает искусственный интеллект, когда он приблизится к совершенству человеческого мозга по таким параметрам, как сенсорное восприятие, память, владение языком, выполнение сложных двигательных актов? Нетрудно догадаться, что он столкнется с теми пределами, которые характерны для человеческого познания. Как мы уже убедились, этими пределами являются четыре фактора: теорема Геделя о неполноте, вероятностная природа индуктивного вывода, метод проб и ошибок и фактор случая в научном открытии. Указанные факторы, исключающие возможность алгоритмизации творческой деятельности в сфере науки, дают решение 18-й проблемы американского математика С. Смейла, сформулированной им в уже упоминавшейся лекции «Математические проблемы следующего столетия» [1].

Теорема Геделя о неполноте свидетельствует о том, что алгоритмы (процедуры, методы, приемы) исследования реальности не могут содержать в самих себе критериев истинности. Эти критерии истинности содержатся в практике (наблюдении, эксперименте). Результат Геделя запрещает эффективное функционирование закрытых (замкнутых) алгоритмических стратегий. Именно поэтому Гедель обнаружил, что поставленная Д. Гильбертом проблема обоснования непротиворечивости арифметики средствами самой арифметики неразрешима.

Вероятностная природа индуктивного вывода означает, что логическая операция индукция в одних случаях дает истину, а в других способна приводить к ошибкам. В свое время на этот аспект индукции обратил внимание Г. Лейбниц, который отказался от индукции и стал искать универсальный метод открытия («универсальную харак-

теристику»), способный гарантировать правильный результат. Но такого метода (правила применения всех правил) не существует, о чем догадывался еще И. Кант, скептически отозвавшийся в адрес мечты Г. Лейбница об «универсальной характеристике». Индукция уступает дедуктивному мышлению в строгости, надежности и неоспоримости, но, тем не менее, обеспечивает нас новыми знаниями. Следует также учитывать, что благодаря индуктивному обобщению опыта творческой деятельности открыты десятки сотен различных эвристик, расширяющих наши возможности в решении научных и технических задач.

Метод проб и ошибок (метод последовательного перебора), как мы уже отмечали, играет колоссальную роль в научном исследовании. Когда нет информации, то есть исходных посылок для индуктивного обобщения, а также когда отсутствуют идеи (теоретические конструкции), позволяющие реализовать аналогию, остается использовать метод последовательного перебора. Метод проб и ошибок применяется не только в эмпирической (экспериментальной) области науки, но и на уровне генерации и развития сложных теорий и концепций, представляющих собой обобщение и синтез различных экспериментальных фактов и идей.

Наконец, фактор случая в научном открытии — еще одна неустрашимая особенность человеческого познания. Случайные или «полуслучайные» открытия невозможно исключить никаким количеством информации, которой вы владеете, принимаясь осваивать новую область. Ведь эта информация представляет собой уже оформленное и зафиксированное знание, а творчество предполагает выход за границы известного, за те границы, где нет ориентиров и указателей. Фактор случая (элемент непредсказуемости отдельных аспектов научного поиска), наряду с методом проб и ошибок, является поставщиком исходных посылок для индукции.

Непредсказуемость научных открытий, обусловленная фактором случая, имеет свои эстетические аспекты. В мире, в котором все события фатально предопределены (предначертаны) и заранее известно, что произойдет через десяток или сотню лет, невозможно творчество и рождение нового. Напомним, что негативные последствия зависимости научного исследования от элемента случая преодолеваются фактором времени и параллельностью поисков, о которых мы уже говорили. Чем больше времени затрачивается на решение той или иной проблемы, тем выше вероятность получения результата.

## 6. Заключение

Удивляет тот факт, что С. Смейл, будучи математиком и формулируя в своем списке проблем сугубо математические проблемы, которые до сих пор остаются нерешенными, поставил вопрос об установлении границ (пределов) естественного и искусственного интеллектов, который имеет гораздо большее отношение к информатике, нежели к математической дисциплине. Конечно, теорема Геделя о неполноте, являющаяся одним из факторов, исключающих алгоритмический характер творческой деятельности, безусловно, относится к сфере математики, но одна эта теорема не дает решения 18-й проблемы С. Смейла. Существуют три других аналогичных фактора (перечисленные выше), которые наделяют творческую деятельность признаками невычислимости и неформализуемости. Выяснение роли этих факторов в научном творчестве требует использования достижений таких сфер знания, как когнитивная психология, нейрофизиология, логика, история научных открытий, теория искусственного интеллекта. Впрочем, не стоит удивляться тому, что проблема, поставленная в одной области знания, часто решается методами, заимствованными из совсем других областей. Например, Г. Перельман (2002) нашел доказательство гипотезы Пуанкаре, решив тем самым 2-ю проблему С. Смейла (чисто топологическую проблему), используя метод потоков Риччи–Гамильтона, а этот метод относится не к топологии, а к области математической физики.

## Список литературы

- [1] Смейл С. Математические проблемы следующего столетия // Современные проблемы хаоса и нелинейности. — Ижевск: Институт компьютерных исследований, 2002.
- [2] Ясницкий Л. Н. Введение в искусственный интеллект. — М.: Академия, 2008.
- [3] Холланд Д. Генетические алгоритмы // В мире науки. — 1992, № 9–10.
- [4] Рассел С., Норвиг П. Искусственный интеллект: современный подход. — М.: Вильямс, 2006.
- [5] Хокинс Дж., Блейкли С. Об интеллекте. — М.: Вильямс, 2007.
- [6] Ушакова Т. Н. Психолингвистика. — М.: ПЕР СЭ, 2006.

- [7] Люгер Дж. Искусственный интеллект. — М.: Вильямс, 2003.
- [8] Дмитриев А. Хаос, фракталы и информация // Наука и жизнь. — 2001, № 5.
- [9] Солсо Р. Когнитивная психология. — СПб.: Питер, 2002.
- [10] Гейм А. Случайные блуждания: непредсказуемый путь к графену // Успехи физических наук. — 2011. Т. 181, № 12.
- [11] Жданов А. А. Автономный искусственный интеллект. — М.: Бинном, 2008.
- [12] Рассел Б. Человеческое познание: его сфера и границы. — М.: Республика, 2000.
- [13] Пойа Д. Математика и правдоподобные рассуждения. — М.: Наука, 1975.
- [14] Эволюционная эпистемология и логика социальных наук. — М.: Едиториал УРСС, 2000.
- [15] Бунге М. Интуиция и наука. — М.: Прогресс, 1967.
- [16] Таунс Ч. Квантовая электроника и технический прогресс // Успехи физических наук. — 1969. Т. 98, вып. 1.
- [17] Нееман Ю. Счастливый случай, наука и общество. Эволюционный подход // Путь. — 1993. № 4.
- [18] Марихин В. А. Синтетические металлы // Химия и жизнь. — 2000. № 6.
- [19] Рыбалкина М. Нанотехнологии для всех. — М.: Nanotechnology News Network, 2005.
- [20] Пенроуз Р. Новый ум короля. — М.: Едиториал УРСС, 2003.
- [21] Беклемишев Л. Д. Теоремы Геделя о неполноте и границы их применимости // Успехи математических наук. — 2010. Т. 65, вып. 5 (395).
- [22] Хазен А. М. О возможном и невозможном в науке. — М.: Наука, 1988.
- [23] Потапов А. С. Искусственный интеллект и универсальное мышление. — СПб.: Политехника, 2012.
- [24] Кочергин А. Н. Искусственный интеллект, психика и творчество // Искусственный интеллект: междисциплинарный подход / ред. — Д. И. Дубровский и В. А. Лекторский. — М.: ИИнтелл, 2006.