
Машинное угадывание (часть II)¹

Д. МИЛЛЕР

Во второй части статьи излагается некоторая форма критического рационализма Поппера. Предполагается, что рациональность науки состоит исключительно в дедуктивных процессах, посредством которых предположения критикуются и улучшаются. Однако крайние формы дедуктивизма отвергаются. Статья завершается решительным отвержением взгляда, согласно которому работы по искусственному интеллекту, включая ДСМ-метод, активно разрабатываемый В.К. Финном, в какой-либо мере опровергает критический рационализм. Машинное обучение проливает мало света как на проблему индукции, так и на роль, которую логика играет в развитии научного знания.

In the second part of the paper a form of Popper's critical rationalism is described. It is supposed that rationality of science consists exclusively in deductive processes by which conjectures are criticized and improved. Nevertheless, extreme forms of rationalism are rejected. The paper ends with decisive rejection of the view that works on artificial intellect, including the DSM method of hypotheses generation actively developed by V.K. Finn, in any degree refute critical rationalism. Machine learning sheds very little light both on the problem of induction and on the role of logic in the development of scientific knowledge.

КЛЮЧЕВЫЕ СЛОВА: критический рационализм, рациональность науки, дедуктивизм, порождение гипотез, искусственный интеллект, ДСМ-метод, машинное обучение, индукция.

KEY WORDS: critical rationalism, rationality of science, deductivism, hypotheses generation, artificial intelligence, JSM-method, machine learning, induction.

§ 4. Критический рационализм

Сейчас подходящий момент высказаться более полно о критическом рационализме и объяснить, почему, несмотря на видную роль, уделяемую им объяснительным гипотезам [Поппер 19576; Дойч 2011], он не признает абдукцию как процесс вывода. Учитывая, как часто критический рационализм излагался самим Поппером, его приверженцами вроде меня и его противниками, изложение его основных идей может показаться здесь излишним. Дело, однако, в том, что критический рационализм снова и снова представляли так, как будто он включает джастификационистские тенденции, ненужные и нежеланные, несмотря на некоторую их поддержку в отдельных высказываниях Поппера – например, в

его рассмотрении правдоподобности (*verisimilitude*) во второй главе [Поппер 1972] (§ 7) и [Поппер 1983] (часть I, § 2), обсуждаемом в моем [Миллер 2006, 126]. В главе 2 моей книги [Миллер 1994] я разобрал по заслугам многие критические замечания, сводящиеся к тому, что критический рационализм (или фальсификационизм) в его применении к эмпирической науке не может не прибегать к индуктивному выводу или к некоторому принципу индукции и потому должен быть объявлен потерпевшим неудачу. С другой стороны, под влиянием Лакатоса некоторые авторы, поощренные отдельными более примирительными формулировками Поппера, стали стряпать недоваренные версии критического рационализма с богатым (но на мой вкус неприемлемым) джастификационистским привкусом. Противники и сторонники в равной мере “по-моему, не желая того свидетельствуют о его [в данном случае Поппера. – Д.М.] оригинальности: ведь ... им так и не удалось понять главного в его философии”, как сам Поппер писал в примечании 2 к главе II [Поппер 1945] о некоторых из тех, кто умалял новизну учения Гераклита о “всеобщей изменчивости”. У критического рационализма есть и важные вторичные моменты, такие как то, что хорошие гипотезы решают подлинные проблемы, но его дух может быть сведен к трем принципам.

(а) Гипотезы могут угадываться свободно (как говорил Хьюэлл). Даже если и существует какая-то процедура для вывода гипотез из фактов, такие выводы не необходимы.

(б) Должно быть предпринято все возможное для опровержения выдвинутого предположения путем дедуктивного вывода из него следствий, способных противоречить фактам (или иным способом выявить неадекватность данного предположения).

(с) Не опровергнутое предположение остается принятым. Никакие дальнейшие действия не нужны (кроме продолжения проверок). Даже если “подтверждение” возможно, оно ничему не служит.

Согласно этой позиции наука на самом деле состоит не более чем из позитивных предположений, контролируемых негативными опровержениями. Индукция и абдукция не играют никакой роли в этой лихорадочной диалектике. Значение принципа (а) состоит в том, что в сфере изобретения *индукция и абдукция не нужны*. Значение (б) состоит в том, что в сфере оценки *дедукция нужна*. Значение (с) в том, что, опять-таки в области оценки, *индукция и абдукция тоже не нужны*. С предположением, пережившим суровые испытания, ничего не происходит; оно не доказано, не опровергнуто и не поддержано. Но рассуждение играет большую роль – в отвержении предположений, но не в их построении. Критический рационализм вполне заслуживает обоих своих определений – *критический и рационализм*. Хотя критический рационалист надеется распознать среди имеющихся объяснительных гипотез ту, которая лучше всего объясняет явления в исследуемой области, он не будет выводить ее не демонстративным образом из опытных данных, а попытается вывести ее на сцену, дисквалифицировав ее соперниц. И чем более он отвержен эмпиризму, тем более эти дисквалификации будут поддержаны эмпирическими опровержениями, а не метафизическими возражениями и прочими критическими залпами. (Ничто в этой статье не должно рассматриваться как пропаганда простого (*unaffected*) эмпиризма.) Если опровержения избежит более чем одна гипотеза, критический реалист будет считать свой проект незавершенным. С этой точки зрения прибегать для его завершения к недемонстративному выводу выглядит как немпирическое жульничество. Как полагал уже Рассел, эмпиризму грозит нетерпеливое стремление к окончательности. Но у него не было ясного понимания того, что эмпиризм и рациональность можно спасти, выбросив за борт обоснование. Многие авторы думали, что предоставить формирование предположений необузданному и недисциплинированному воображению значит оставить слишком многое на долю случая. Пирс, например, спрашивал [Пирс 1903, 172–173]: “Как, однако, получается, что на всю эту истину проливает свет процесс, в котором нет ничего обязательного и даже нет никакой тенденции к обязательности? ... Это что – случайность? Подумайте о множестве теорий, которые можно было бы предложить в поисках истины. ... Физик в своей лаборатории натывается на какое-то новое явление. ... Подумайте о том, сколько триллионов триллионов гипотез можно было бы высказать, из которых истинна только одна; и все-таки после двух-трех, самое большее – дюжины догадок физик

почти без промаха попадает на правильную гипотезу. С помощью случая он, вероятно, не сделал бы этого и за все время, прошедшее с тех пор, как Земля затвердела. Вы можете представить то или другое прекрасное психологическое объяснение этого. Но позвольте мне сказать вам, что вся психология на свете оставит логическую проблему там, где она была... Вы можете сказать, что это объясняется эволюцией. Я не сомневаюсь, что это эволюция. Но чтобы объяснить эволюцию случайностью, просто не хватает времени”.

Верно, что догадки не случайны и не находятся целиком на милости случая (как и эволюция). Догадки слепы. Даже тезис, который говорит разве что о том, что из известного не следует ничего о неизвестном, слишком радикален для многих сведущих людей, которым часто не нравится утверждение, будто остроумные идеи, на которые они затратили столько сил, – не более чем слепые (хотя, возможно, и блестящие) догадки. Необъяснимым образом они ценят весомость своей окопавшейся науки больше, чем вспышки своего ничем не связанного воображения. Новые гипотезы вполне могут быть догадками, но они не *просто* догадки; это *информированные*, или *просвещенные*, догадки. Куайн [Куайн 1986, 332] сформулировал это так: “Перескакивать к выводам – наша повседневная практика. Перескакивать к разумным заключениям – образ действий занятого своим делом ученого”. Но информированные догадки тоже могут быть слепыми. Информированными информированные догадки делают более ранние догадки, ничуть не менее слепые, но пережившие критическое исследование, которому они подверглись. Иначе говоря, существуют догадки, информированные тем, что предположительно известно, но слепые к тому, что неизвестно (т.е. слепые к тому, что выходит за пределы прежних догадок). Как мудро сказал Кэмпбелл [Кэмпбелл 1974, 422]: “Выходя за границы уже известного, можно двигаться только вслепую. Если бы можно было двигаться разумно, то это указывало бы на наличие некоей уже накопленной мудрости общего характера”. Некоторые догадки при рассмотрении могут оказаться вытекающими из того, что уже известно, и соответственно не продвигать наше объективное знание. Но и в этом случае перед нами слепой, ведущий слепого [Миллер 20056, 79]. В угадывании нет рационального компонента. Как решительно заявил Пирс во фрагменте, процитированном выше в § 3 первой части этой статьи, “для нее [абдукции] нельзя дать никакого основания (reason) ... да ей и не нужно никаких оснований, поскольку она только высказывает предположения”. Поппер, в то время мало знавший о философии Пирса, высказал примерно то же в § 12 работы “Назад к досократикам” [Поппер 1958]: “Есть лишь один элемент рациональности в наших попытках познать мир: критическое рассмотрение наших теорий. Сами эти теории – догадки”. Этот факт тоже порождает немалое беспокойство. Например, Бернайс [Бернайс 1974], процитировав эти слова Поппера, добавил (§ 14): “начинает казаться, что ... [в этих словах есть] скрытое предположение, что рациональность должна быть *знанием*”, и призвал распространить роль разума за пределы критического процесса. В статье Бернайса я не нашел никаких существенных аргументов против мысли, что угадывание не содержит никакого рассуждения (reasoning), а именно это Поппер, несомненно, имел в виду. Ответ Поппера был уважительным, но он отверг приписываемое ему предположение и признал только, что предположения предшествуют рациональной критике – которая имеет смысл, только если критикуемые высказывания предположительны [Поппер 1974]. То, что предположения и опровержения тесно переплетаются, – я полагаю, никто не оспаривает. Позиция критического рационализма, что чистое угадывание не содержит никаких элементов рассуждения, отличается от обыденной точки зрения, возможно, только в отношении того, на каком уровне сменяющие друг друга предположения анализируются и выделяются. Неосознанные предположения, отвергаемые “сходу” и исключаемые из рассмотрения даже не будучи полностью сформулированными, остаются невидимыми на том уровне, когда карандаш прикасается к бумаге. Об этом ср. замечания Поппера о той легкости, с какой Рассел, судя по всему, составлял свои рукописи [Поппер 1978, 245], его рассказ о методе работы Эйнштейна (там же) и утверждение, что именно “осознанно критическое отношение ... позволяло Эйнштейну быстро отвергать как неадекватные сотни гипотез еще до того, как рассматривать ту или другую из них более внимательно, если она казалась способной выдержать более серьезную критику” [Поппер 1966] (см. § xxi). Пренебрежение к тому,

что так и осталось не записанным, – единственное извинение (не говоря о глубоком незнакомстве с тем, как работают математики) для неспособности признать, что в новом математическом доказательстве “сначала возникает интуитивное, основанное на воображении представление о том, как могло бы выглядеть доказательство; затем следует критическая проверка отдельных шагов этого доказательства – критическая ревизия, которая (чаще, чем наоборот) показывает нам, что доказательство неверно” [Поппер 1974] (см. § 28.iv).

§ 5. Чрезмерный дедуктивизм

Критический рационализм иногда кратко называют *дедуктивизмом*, и это, конечно, неплохое название для доктрины, согласно которой все подлинные выводы дедуктивны. Но здесь я избегаю употребления этого термина, поскольку некоторые авторы распространили гегемонию дедуктивной логики еще дальше и стали рассматривать как результаты дедуктивного рассуждения даже догадки, с которых, по мнению критических рационалистов, должен начинаться рост знания. Уоррал [Уоррал 1995, 91] писал: «Факт, что ученые не просто угадывают свои теории; они не делают “смелых” попперовских предположений. Вместо этого они приходят к своим теориям таким способом, который, хотя, конечно, он включает интуицию и творчество, тем не менее можно реконструировать как систематическую и логическую аргументацию, основанную на предшествующих успехах науки и на частях “фонового знания”, взятых в качестве посылок... [Но] даже в отсутствие общего анализа построения теорий нетрудно показать, изучая *детали конкретных исторических эпизодов*, что такой анализ должен существовать».

Захар [Захар 2007] (см. гл. VI) и Масгрейв [Масгрейв 2009, 218–222] также предложили своего рода дедуктивистскую *эвристику*, или логику открытия. Похоже, они согласны, что таким образом открытия могут совершаться только на психологическом уровне, поскольку дедуктивные следствия гипотезы, рассматриваемые объективно, уже присутствуют внутри нее. Дедукция не является творческой (см. об этом “Введение” в первой части статьи). Разоблачая претензии силлогистического рассуждения, Милль справедливо отверг [Милль 2011, 177] попытки “придавать какое-либо серьезное научное значение такой увертке, какую представляет собой различие между тем, что *подразумевается* в посылках, и тем, что в них прямо утверждается” (см. кн. II, гл. iii, § 2). Но мы не должны слишком много уступать сторонникам дедуктивной эвристики. Как минимум необычно, если в процессе дедукции обнаруживается какое-то удивительное заключение. Гораздо более обычно, если сначала возникает догадка, что заключение (или его отрицание) представляет собой теорему, а потом доказывается, что оно действительно теорема. Заключение, которое в типическом случае неожиданно обнаруживается в ходе дедукции или вычисления, – это тождество $0=0$. Для меня действительно остается тайной, в чем заключается смысл тезиса Уоррала. Нет сомнения, что первоначально несформулированные гипотезы могут иногда включаться в дедуктивную сеть, основанную на «предшествующих успехах науки и на частях “фонового знания”». Это не так уж редко встречается в современной математике (большая часть которой может быть включена в аксиоматическую теорию множеств). Но это не может быть нормой, поскольку эта дедуктивная сеть (или одна из ее предшественниц) должна быть одним из тех «“смелых” попперовских предположений», которые избежали внимания Уоррала. Не может быть сомнения и в том, что работа, необходимая для встраивания гипотезы в “структуру уже существующих научных доктрин” [Поппер 1934], представляет собой скорее “пробы, ошибки, экспериментирование, догадки” [Халмош 1985, 321], нежели “систематическую и логическую аргументацию”. Предшествующие доктрины управляют последующими примерно так, как берега реки управляют ее течением. Они не направляют, а перенаправляют их. Я не отрицаю исторической и культурной ценности такого прослеживания связей, но я отвергаю вывод о том, что дедуктивная организация законченного продукта раскрывает логику его открытия. Образовательный культ, известный как *критическое мышление*, или *неформальная логика*, – еще один пример затрачивания чрезмерных усилий на навязывание дедуктивной структуры

диффузным сегментам разговоров или письменных текстов. Основная задача, предписываемая изучающим “критическое мышление” (я чувствую, что это название должно все время сопровождаться неодобрительными кавычками), – распознавание и исправление того, что считается плохими аргументами, что зачастую состоит в распознании скрытых или опущенных посылок, способных восстановить дедуктивную верность рассуждения. Редко осознается, что подозрительные пассажи, объективно говоря, могут вообще не быть аргументами, даже плохими, несмотря на наличие в них таких маркеров, как *если*, *так* и *потому что*. Поскольку предполагаемые заключения признаются плохо связанными с предполагаемыми посылками, было бы лучше рассматривать эти выводы как простые предположения, возможный привкус аргументации которым придавали бы такие слова, как *если не*, *но* и *несмотря на*, указывающие на наличие оговорок и негативных соображений. С объективной точки зрения “критическое мышление”, следовательно, трижды плохо задумано [Миллер 2005а]. Во-первых, оно неправильно рассматривает предположения как заключения энтимем; во-вторых, оно уводит критическое внимание от того, что говорится в предположении, к кажущимся аргументам, которые якобы должны выдвигаться в его поддержку; и, в-третьих, оно создает впечатление, что если бы эти аргументы были сделаны дедуктивно законными, все было бы хорошо. Короче говоря, “критическое мышление” ошибочно представляет как цель, так и возможности критики. В конце концов, единственная функция аргументации (кроме как, возможно, в математике) – помочь нам выяснить, что верно (или ценно, или мудро, или благоразумно), а что – нет. Несостоятельный аргумент, стремящийся оправдать свое заключение, каким бы провоцирующим он ни был, заслуживает быть проигнорированным, поскольку он ничего не говорит нам о том, что мы хотим знать, – об истинности или ложности заключения. Даже если истинность других посылок не оспаривается, “отсутствующая посылка”, которая должна закрепить верность вывода, ничего нам не дает, поскольку мы можем судить об истинности заключения только в том случае, когда эта посылка воспринимается нами как истинная. Масгрейв, конечно, прав, когда говорит, что “не каждый [несостоятельный] аргумент может рассматриваться как имеющий недостающую посылку” [Миллер 2005а, 206]. Но только если бесплодную задачу критики аргументов сочетать с плодотворной задачей критики их заключений, он будет прав в том, что дедуктивная логика лишится всякой критической функции, если все несостоятельные аргументы рассматривать как энтимемы.

§ 6. Вступают машины

Выше, в § 3, я предложил использовать термин “вывод” применительно только к тем переходам к новым гипотезам, которые каким-то образом связаны правилами, а другие переходы рассматривать как догадки. Я был осторожен и не говорил, что недоказательные (индуктивные или абдуктивные) правила вывода не могут быть сформулированы, поскольку нельзя отрицать, что это возможно. Так же легко сформулировать *контриндуктивные* правила, позволяющие нам из сообщения, что во всех встретившихся случаях за X следовало Z , вывести, что ни в одном случае в будущем за X не будет следовать Z . Действующий человек, конечно, может позволить себе руководствоваться такими индуктивными правилами, и можно спроектировать машину, работающую в соответствии с ними. Нам надо разобраться в том, опровергает ли возможность таких индуктивных машин, как их можно назвать, точку зрения Поппера, что индукции не существует ни логически, ни психологически. Ибо, согласно Гиллису [Гиллис 1996, 53], сообщающему об “успехах машинного обучения”, эту точку зрения “уже нельзя поддерживать в свете таких программ, как ID3 или GOLEM, которые делают индуктивные выводы, основанные на многочисленных наблюдениях, и стали частью научных процедур”. Не отрицая изобретательности, проявленной при разработке таких программ и других им подобных, нам следует, по-моему, проявить осторожность, оценивая методологическое значение таких претензий. Прежде всего, как указывал Тамбуррини в своем детальном исследовании этого вопроса [Тамбуррини 2006, 268]: “Исследования обучающихся систем, проводимые в

рамках искусственного интеллекта, не вынуждают нас отказаться от попперовского радикального скептицизма по отношению к индукции. Правильное понимание результатов как теории обучения, так и машинного обучения не требует никакого обращения к так называемым принципам индукции, которые предположительно дают частичное обоснование гипотезам, эффективно порождаемым вычислительными агентами на основе имеющихся данных”.

И далее он говорит, что “интеллектуальное поведение, демонстрируемое обучающимися системами, вполне можно объяснить в терминах процессов проб и исправления ошибок”. Поэтому, если и используется что-то, что можно назвать индукцией, она не регулятивна в том смысле, который был рассмотрен во “Введении” в первой части статьи. Кроме того, мы должны отметить неизбежную черту всех подходов, вводящих правила порождения гипотез или извлечения следствий: гипотеза, что это правило подходит для ситуации, в которой оно принимается, не входит в область действия этого правила и потому, если только она не порождается слепым угадыванием, должна порождаться или выводиться по какому-то другому правилу. Я отмечал в другом месте, что принятие некоторого правила принятия решений не избавляет принимаемые решения от неопределенности, внутренней присущей тому затруднению, которое делает решение необходимым. *Исход* принятого решения, конечно, остается неопределенным, но я не это имею в виду. Ибо не менее неопределенным остается целесообразность подчинения этому правилу. В байесовской философии, например, считается почти аксиомой, что хорошее принятие решений – это принятие решений согласно Бернулли с его правилом максимизации ожидаемой полезности. Однако правило Бернулли не является неоспоримым “правилом рациональности”, особенно если речь идет о субъективных вероятностях, и решение руководствоваться им может быть гарантировано только полностью предположительным суждением, что руководствоваться этим правилом целесообразно [Миллер 1998]. По ходу дела задействуется немалая доля проб и ошибок, но результаты, получаемые ID3, задача которой – “индуцировать правила классификации” [Гиллис 1996, 33], и GOLEM’ом, созданным для индуцирования обобщений [Там же, 41–44], не выводятся дедуктивно из входных данных. Эти программы (как и другие) действительно делают индуктивные выводы. В каждую из этих программ встроено правило вывода, и сообщается, что выводы, производимые по этому правилу, – хорошие выводы в некотором согласованном смысле. Это не значит, что все эти выводы оказываются истинными или точными, но (как я полагаю) только то, что они по большей части оказывались успешными. Пока что в этом нет никакой проблемы. Но если есть гипотеза, что само это правило – хорошее правило, которое будет продолжать, в будущем и в других областях, производит хорошие выводы, то она не была получена с использованием этого правила. Аргумент Юма не так-то легко обойти. А если такой гипотезы нет, то успехи, записанные на счет этого правила, остаются не более чем интересным историческим курьезом. Возможность индуктивных машин, успешных при подходящих условиях, была объяснена самим Поппером в конце § v [Поппер 1957a] и более подробно в части II [Поппер 1983]. Попперовская машина порождает поток данных (последовательность шаров различного материала и цвета) и использует то, что он называет *простым индуктивным правилом* (иногда его называют *прямым правилом*), для вывода из наблюдаемых частот различных событий вероятностей их появления и универсальных обобщений, если есть какие-либо достоверные обобщения. Например: «Таким образом можно обнаружить, что вероятность появления вслед за парой “стальной-медный” события “медный”, или пары “стальной-медный”, или любой тройки, кроме тройки “стальной-стальной-медный”, равна нулю... Машина открыла “закон”, согласно которому события “стальной” имеют тенденцию происходить ... как последовательности точно трех таких событий». Поппер замечает, что в конструкции машины заложено предположение, что мир, подлежащий исследованию, таков, что “к нему может быть успешно применено простое индуктивное правило”. Как мы знаем из игры “*Красные и синие*” (описанной Поппером там же, часть II, § 8), это условие нетривиально. Поппер заключает: “Я никогда не говорил, что ... мы не можем успешно использовать простое индуктивное правило, если это позволяют объективные условия. Но я утверждаю, что с помощью индукции мы не мо-

жем обнаружить, позволяют ли объективные условия использовать простое индуктивное правило ... создатели машины ... должны решить, что составляет ее мир, что будет для нее индивидуальными событиями, *что составляет повторение*".

Можно, конечно, построить машину с задачей узнавать что-то о мире безо всякого представления о том, при каких условиях она должна работать. Невозможно оценить эффективность такой машины, если только она не окажется эффективной при любых условиях. Известно, что индуктивные машины не таковы. Поэтому, чтобы претензии на создание успешной индуктивной машины стоило рассматривать, существенно, чтобы проектировщик такой машины оговорил сферу ее применения. Мысль Поппера состоит в том, что индуктивная машина сама не может помочь нам обнаружить условия, при которых она будет работать (хотя она может помочь нам обнаружить условия, при которых она выйдет из строя). За каждой успешной индуктивной машиной или правилом индуктивного вывода стоит вольная догадка. Чем точнее мы описываем необходимые и достаточные условия ее успешного действия, тем ближе мы подходим к дедуктивному объяснению того, что делает машина. Тамбуррини в § 4 [Тамбуррини 2006] поставил по существу тот же диагноз программе ID3: если "*исходные предпосылки ... или предубеждения, воплощенные в самой ID3 (которые определяют как язык выражения понятий, так и построение деревьев решений) ... можно успешно сформулировать в декларативной форме, то алгоритм обучения понятиям, такой как ID3, можно переписать как доказательство теорем*". На это Гиллис [Гиллис 2009] дал два ответа. Первый ответ, выражающий неспецифицированные сомнения в том, что ID3 на практике можно так переписать, можно проигнорировать. Второй ответ утверждает, что ID3, "переописанная как доказательство теорем, была бы гораздо сложнее, чем исходная ID3, представленная как система индуктивного обучения. Зачем нам вводить все эти ненужные усложнения, которые никогда не будут приняты на практике? [Если мы] допускаем введение индуктивных правил вывода, мы получаем простые индуктивные компьютерные системы, практически успешные... [Если мы] допускаем только дедуктивные правила ... мы вынуждены пытаться преобразовать эти системы в эквивалентные доказательства теорем... Это трудная, вероятно безнадежная задача, увеличивающая сложность безо всякой практической выгоды".

Такой ответ можно было бы ожидать от инженера, заботящегося только об эффективности машины, или от бухгалтера, занятого только затратами и выгодами. Саймон тоже кажется гораздо меньше заинтересованным эпистемологическим статусом, нежели сравнительной эффективностью различных методов обработки данных [Саймон 1973]. Более философским ответом было бы спросить, почему ID3 (или любая другая индуктивная машина) успешна в тех случаях, в которых она успешна? Есть ли в мире нечто объективное, что объясняет успех машины, или же ее успех есть артефакт устройства машины и способа ее применения? Невозможно отрицать, что мир явлений частично регулярен, а частично нет, и что одна из основных задач науки – распознавать регулярности, лежащие под поверхностью. Если исследователи искусственного интеллекта (ИИ) обнаружили еще более глубокие и более общие регулярности, объясняющие успехи, приписываемые обучающимся машинам, им надо было бы дать нам об этом знать.

§ 7. ДСМ-метод

ДСМ-рассуждения – метод автоматического перехода от более или менее неорганизованных данных к соответствующей каузальной объяснительной теории. В моих кратких замечаниях я буду опираться на прекрасное краткое изложение Бёрча [Бёрч 2000], описывающего ДСМ-метод как метод *слияния*, или синтеза, *данных*, и в какой-то мере на более техническое изложение Финна [Финн 2011] и его же с соавторами [Аншаков, Финн и Виноградов 2005]. Подробного изложения не требуется для простых философских замечаний, которые я хочу сделать.

Использование ДСМ-метода начинается с плохо определенных и неполных данных о хорошо определенной области объектов с некоторым числом (монадических) свойств.

В типическом случае в базе данных (БД) есть информация, что некоторые из исследуемых объектов имеют, а другие не имеют свойства P . Для других объектов информации нет, а еще для других объектов есть противоречивая информация: в БД говорится, что эти объекты обладают как P , так и $\neg P$. Как я понимаю, первая задача состоит в том, чтобы расширить и скорректировать исходную информацию, экстраполируя любые выполняющиеся ограниченные обобщения и используя аналогии. Далее в БД вносится порядок с помощью того, что выглядит как метод проб и ошибок, постепенно заполняя пробелы и преодолевая избыточности. На некотором шаге процедура останавливается и ничего нового не возникает. Этот момент отмечает начало того, что Финн называет абдуктивной стадией автоматического исследования. План его состоит в том, чтобы найти для любого интересующего нас P то, что можно предположительно считать причиной P , используя метод, близкий к миллевскому методу согласия: рассматриваются все возможные каузальные гипотезы, и та из них, которая лучше всего подходит к определению *максимальной причины*, принимается как объяснение P .

В Заключении своей работы Бёрч [Бёрч 2000] говорит о ДСМ-методе, что он “может давать полезные результаты на БД, содержащих только небольшое количество информации”. По словам Финна [Финн 2002, 407 сл.], этот метод “способен решать проблемы определенного типа, требующие использования автоматизированного рассуждения ... [и] не могут быть решены людьми за реальное время ... [Это] синтез когнитивных процедур – индукции, аналогии и абдукции ... в смысле Ч.С. Пирса ... [который был] испытан экспериментально в различных предметных областях (фармакология, биохимия, техническая диагностика, социология)”. Насколько я понимаю, некоторые из этих успешных экспериментов использовали ДСМ-метод для установления каузальных связей в областях, независимо от этого хорошо понимаемых. Но мы уже должны были научиться, основываясь на замечании Поппера в [Поппер 1957а], см. § 1, не слишком обольщаться успехами такого рода. Мне хотелось бы больше узнать об областях, в которых этот метод не мог установить никаких каузальных связей, а также о том, на основе каких независимых критериев решался вопрос о том, считать ли такой результат успехом или неудачей. Особенно мне хотелось бы знать о результатах ДСМ-метода в малоперспективных случаях, таких как астрологические данные, собранные Гокленом и другими, в которых, как кажется, проявляются зависимости между датой рождения выдающихся личностей и областями, в которых они достигли известности [Гоклен 1988; Эйзенк и Найас 1982]. Но прежде всего мне хотелось бы знать об условиях, при которых ДСМ-метод претендует на эффективность. Финн [Финн 2002, 413] перечисляет ряд синтаксических “условий применимости ДСМ-метода АПГ [автоматического порождения гипотез]”. Но, как было объяснено в предыдущем параграфе, если мы хотим эмпирически оценить достоинства ДСМ-метода, нам требуется еще кое-что. Финн, как Саймон [Саймон 1973], Михальский [Михальский 1983, 87 сл.] (цит. в [Гамбуррини 2006, 268]) и многие другие, по-видимому, полагают, что философская проблема индукции нуждается в косметическом ремонте: “Проблему индукции в отношении универсальных теорий ... следует заменить проблемой адекватности формализованных эвристик в ИС [интеллектуальных системах] соответствующим классам проблем” [Финн 2002, 408]. Если эту задачу замены следует рассматривать как чисто инженерную, то мы можем признать, что было сделано несколько шагов к ее решению. Но если проблема “адекватности формализованных эвристик в ИС” есть философская проблема о том, как мы можем получать знание о мире, то я не вижу, чтобы был достигнут какой-либо прогресс. Не понимаю я и того, почему проблему Юма, которая не решается заменой ее на другую проблему, нужно заменить; т.е. почему ей надо позволить исчезнуть из рассмотрения нерешенной, как это рекомендовали даже философы (что мы видели в § 1). Можно, конечно, жить и без ее решения, т.е. без теории рациональности. Но я не знаю, что, по мнению тех, кто отказался от надежды на непротиворечивую теорию рациональности, мы должны понимать под такими логическими выражениями, как “правдоподобные рассуждения” [Финн 2011] (см. Прил. 4). Портреты Милля, Пирса и Поппера, героев ДСМ-кампании, украшают обложку книги Финна [Финн 2011]. Я попробовал показать, в каком смысле сторонники ДСМ-метода не использовали в полной мере идеи

Пирса и Поппера. Что же касается Милля, нижеследующий отрывок из гл. II его трактата “О свободе” [Милль 1859] предлагает более мудрую философию познания, чем вся “Система логики” 1843 г.: “Для нас не существует никакого другого ручательства в истинности какого бы то ни было мнения, кроме того, что каждому человеку представляется полная свобода доказывать его ошибочность, а между тем ошибочность его не доказана. Если вызов на критику не принят, или если принят, но критика оказалась бессильной, то это еще нисколько не значит, что мы обладаем истиной, – мы можем быть еще очень далеко от истины, но, по крайней мере, мы сделали все для ее достижения, что только могло быть сделано при настоящем состоянии человеческого понимания, – мы по крайней мере не пренебрегли ничем, что могло раскрыть нам истину, и если поле для критики остается открытым, то мы можем надеяться, что ошибки, какие есть в нашем мнении, будут раскрыты для нас, как только ум человеческий сделается способен к их раскрытию, а покамест имеем основание думать, что настолько приблизились к истине, насколько это возможно для нас в данную минуту. Вот только до какой степени человек достигает знания истины, и вот единственный путь, которым он может достигать этого знания”².

Пер. с англ. Д.Г. Лахути

ЛИТЕРАТУРА

Аншаков, Финн, Виноградов 2005 – *Anshakov O.M., Finn V.K., Vinogradov D.W.* Logical Means For Plausible Reasoning of JSM-Type; см.: <http://www.ifispan.waw.pl/studialogica/AnshakovFinnVinogradov.pdf/>.

Бернайс 1974 – *Bernays P.I.* Concerning Rationality // [Шилпп 1974]. Рус. пер.: [Эволюционная эпистемология 2000, 154–162].

Бёрч 2000 – *Burch R.W.* Semeiotic Data Fusion // Proceedings of the Third International Conference on Data Fusion. P., 2000. Vol. II. См. также: <http://www.student.nada.kth.se/~tessy/RobertBurch.pdf/>. Рус. пер.: [Финн 2011].

Гиллис 1996 – *Gillies D.A.* Artificial Intelligence and Scientific Method. Oxford, 1996.

Гиллис 2009 – *Gillies D.A.* Problem-Solving and the Problem of Induction // Rethinking Popper / Ed. by Z. Parusnikova, R.S. Cohen. Dordrecht; L., 2009.

Гоклен 1988 – *Gauquelin M.* Written in the Stars. Wellingborough, 1988.

Дойч 2011 – *Deutsch D.* The Beginning of Infinity. Explanations that Transform the World. L., 2011.

Захар 2007 – *Zahar E.G.* Why Science Needs Metaphysics. Chicago (IL), 2007.

Койт 1998 – *Popper K.R.* Logik der Forschung / Hrsg. H. Keuth. Berlin, 1998.

Куайн 1986 – *Quine W.V.* The Sensory Support of Science // *Discursos*. Granada, 1986. Перепеч. в: Quine W.V. Confessions of a Confirmed Extensionalist / Ed. by D. Føllesdal, D.B. Quine. Cambridge (MA), 2008.

Кэмпбелл 1974 – *Campbell D.T.* Evolutionary Epistemology // [Schilpp 1974]. Рус. пер.: [Эволюционная эпистемология 2000, 92–146].

Масгрейв 2009 – *Musgrave A.E.* Popper and Hypothetico-Deductivism // Handbook of the History of Logic / Ed. by S. Hartmann. Vol. 10: Inductive Logic. Amsterdam, 2009. См.: http://stephanhartmann.org/HNL10_Musgrave.pdf/.

Миллер 1994 – *Miller D.W.* Critical Rationalism: A Restatement and Defence. Chicago; La Salle (IL), 1994.

Миллер 1998 – *Miller D.W.* On Methodological Proposals // [Койт 1998].

Миллер 2005а – *Miller D.W.* Do We Reason When We Think We Reason, or Do We Think? // Learning for Democracy. 2005. Vol. 1. N 3. См.: <http://go.warwick.ac.uk/dwmiller/lfd/>.pdf.

Миллер 2005б – *Miller D.W.* A Reply to Trudy Govier and Philip Adey // [Миллер 2005а].

Миллер 2006 – *Miller D.W.* Out of Error: Further Essays on Critical Rationalism. Aldershot; Burlington (VT), 2006.

Милль 1843 – *Mill J.S.* A System of Logic, Ratiocinative and Inductive. L.; N. Y.; Bombay, 1843.

Милль 1859 – *Mill J.S.* On Liberty. L., 1859. Цитируется по рус. пер.: Милль Дж.С. О свободе / Пер. с англ. А.Н. Неведомского // Антология мировой либеральной мысли (I половины XX века). М., 2000.

Милль 2011 – *Милль Д.С.* Система логики силлогистической и индуктивной. М., 2011. (Рус. пер. [Милль 1843]).

Михальский 1983 – *Michalski R.S.* A Theory and Methodology of Inductive Learning // Machine Learning, An Artificial Intelligence Approach / Ed. by R.S. Michalski, J. Carbonell, T.M. Mitchell. Berlin, 1983.

Пирс 1903 – *Peirce C.S.* Lectures on Pragmatism. Цит. по изд.: *Peirce C.S.* Collected Papers / Ed. by C. Hartshorne, P. Weiss. Vol. 5. Cambridge (MA), 1934.

Поппер 1934 – *Popper K.R.* Logik der Forschung. Vienna, 1934. Расширенное англ. изд.: The Logic of Scientific Discovery. L., 1959. Рус. пер.: *Поппер К.Р.* Логика научного исследования: Пер. с англ. / Под общ. ред. В.Н. Садовского. М., 2005.

Поппер 1945 – *Popper K.R.* The Open Society and Its Enemies. L., 1945. Рус. пер.: *Поппер К.* Открытое общество и его враги: В 2 т. М., 1992.

Поппер 1957a – *Popper K.R.* Philosophy of Science: A Personal Report // British Philosophy in the Mid-Century / Ed. by C.A. Mace. L., 1957. Перепеч.: [Поппер 1963].

Поппер 1957b – *Popper K.R.* The Aim of Science // Ratio. 1957. Vol. 1. N 1. Перепеч.: [Поппер 1972] (гл. 5); [Поппер 1983] (§ 15).

Поппер 1958 – *Popper K.R.* Back to Presocratics // Proceedings of the Aristotelian Society. N. S. Vol. LIX (1958–1959). L., 1959.

Поппер 1963 – *Popper K.R.* Conjectures and Refutations. L., 1963. Рус. пер.: *Поппер К.* Предположения и опровержения: Рост научного знания. М., 2004.

Поппер 1966 – *Popper K.R.* Of Clouds and Clocks: An Approach to the Problem of Rationality and the Freedom of Man. St Louis (MO), 1966. Перепеч. [Поппер 1972] (гл. 6).

Поппер 1972 – *Popper K.R.* Objective Knowledge. An Evolutionary Approach. Oxford, 1972. Рус. пер.: *Поппер К.* Объективное знание: Эволюционный подход / Пер. с англ. Д.Г. Лахути; Отв. ред. В.Н. Садовский. М., 2002.

Поппер 1974 – *Popper K.R.* Replies to my critics // [Шилпп 1974].

Поппер 1978 – *Popper K.R.* Natural Selection and the Emergence of Mind // Dialectica. 1978. Vol. 32. N 3–4. Ссылки даются на сокращенное изд.: Popper Selections / Ed. by D.W. Miller. Princeton (NJ), 1985.

Поппер 1983 – *Popper K.R.* Realism and the Aim of Science / Ed. W.W. Bartley. L., 1983.

Поппер 1998 – *Popper K.R.* The World of Parmenides. Essays on the Presocratic Enlightenment / Ed. by A.F. Petersen. L.; N. Y., 1998.

Саймон 1973 – *Simon H.A.* Does Scientific Discovery Have a Logic? // Philosophy of Science. 1973. Vol. 40. N 4. Перепеч. в [Койт 1998] (гл. 11).

Тамбуррини 2006 – *Tamburrini G.* Artificial Intelligence and Popper's Solution to the Problem of Induction // Karl Popper: A Centenary Assessment / Ed. by I.C. Jarvie, K.M. Milford, D.W. Miller. Vol. II: Metaphysics and Epistemology. Aldershot; Burlington (VT), 2006.

Уоррал 1995 – *Worrall J.* 'Revolution in Permanence': Popper on Theory-Change in Science // Karl Popper: Philosophy and Problems / Ed. by A. O'Hear. Cambridge [u.a.], 1995.

Финн 2002 – *Finn V.K.* The Use of Induction in Plausible Reasoning in Intelligent Systems // K. Popper 2002 Centenary congress. Vienna 3 July – 7 July 2002. Abstracts. Vienna, 2002. Ссылки даются на перепечатку в [Финн 2011].

Финн 2011 – *Финн В.К.* Искусственный интеллект. Методология, применения, философия. М., 2011.

Халмош 1985 – *Halmos P.R.* I Want to be a Mathematician. An Automathography. N. Y., 1985.

Шилпп 1974 – The Philosophy of Karl Popper / Ed. by P.A. Schilpp. La Salle (IL), 1974.

Эволюционная эпистемология 2000 – Эволюционная эпистемология и логика социальных наук. Карл Поппер и его критики. М., 2000.

Эйзенк и Найас 1982 – *Eysenck H.J., Nias D.K.B.* Astrology: Science or Superstition?. L., 1982.

Примечания

¹ См. первую часть статьи: Вопросы философии. 2012. № 7.

² Цитируемый русский перевод [Милль 1859] изменен только в одном месте: “Если вызов на критику не принят, или если принят” вместо “Если вызов на критику не принят, или если принять...”. – *Прим. пер.*