

Может ли машина мыслить?

Вопрос «может ли машина мыслить?» после перевода его обсуждения с эмоционального уровня, преобладавшего в начале 50-х гг. прошлого века, на рациональный, казалось бы, был решен в пользу признания некорректности его формулировки. Однако успехи в области искусственного интеллекта и других областей знания вновь выдвигают его для обсуждения (о чем, в частности, свидетельствует его включение в программу работы конференции). Это вновь обращает нас к необходимости уточнения философских и методологических оснований его обсуждения.

Обратимся, прежде всего, к истории его возникновения. Появление ЭВМ, способных получать результаты, традиционно ассоциировавшиеся с творческой человеческой деятельностью (сочинение стихов, музыки, доказательство теорем и т.д.), привело ряд исследователей к выводу о том, что достаточно полная модель мыслящего существа должна называться мыслящим существом [5, с. 57]. Иными словами, в основу решения данного вопроса был положен чисто функциональный подход (у А. Тьюринга этот подход получил название «игры в имитацию»). Отсюда был сделан вывод об обладании ЭВМ свойством мышления со всеми вытекающими из этого последствиями.

Постановка вопроса «Может ли машина мыслить?» была в известной степени обусловлена использованием антропоморфной терминологии. Мы часто говорим: «Машина работает», «Дождь идет», «Усталость металла» и т.д. До поры это не вызвало никаких затруднений. Однако использование этой терминологии при обсуждении мировоззренческих проблем оказалось делом не столь безобидным. Человеческие языковые средства, используемые при описании реальности, таковы, что из этого описания можно сделать вывод о существовании целенаправленных стимулов у объектов, хотя в действительности их нет. Дж. Рамсей обратил внимание на то, что при описании, например, биологических систем практически невозможно избежать телеологии, т.е. приписывания им осознанной целенаправленности действия (имея в виду, что животные формулируют для себя цели, к которым стремятся, — описать поведение живой системы таким образом, чтобы такая идея не подразумевалась, трудно). Такие описания получили название «телеографической стенографии». По этому поводу Ст. Бир писал: «Модель языка системы, которым пользуется кибернетика, вносит в описание такие смысловые оттенки, от которых нам хотелось бы избавиться. Но мы не в силах этого сделать и потому всегда должны помнить о возможности неверных толкований при обсуждении нервных систем, мозга и машин. Существуют ли люди, которые готовы отрицать роль этики и эстетики только потому, что создана механическая черепаха? Как это ни абсурдно, действительно существуют. Пусть же тот, кто хоть в какой-то мере понимает истинное значение кибернетики и ее возможности, не следует их примеру» [2, с. 60]. Использование терминов в их точном значении — не пустая формальность. Подведение под определенный термин не соответствующего ему понятия порождает «языковые ловушки», благодаря которым приходят к выводу о машиноподобности психики, мышления.

Решение вопроса «Может ли машина мыслить?» предполагает уточнение понятий «машина» и «мышление». Определений понятия машины много. В данном контексте речь идет об ЭВМ или различных ее «гибридах» типа роботов, т.е. системах, перерабатывающих информацию. У.Р. Эшби предложил исключить из ее

определения «материальность», т.е. необходимость изготовления ее из реальных материалов. С его точки зрения, существенно «только то обстоятельство, является ли поведение системы детерминированным и машиноподобным, а не то, как она построена, будь она хоть из божественной субстанции» [10, с. 64]. Точно также он предлагает исключить из определения машины и ссылку на энергию на том основании, что для ЭВМ имеет значение только регулярность поведения, а не выигрыш или потеря энергии.

Каковы же последствия игнорирования материального и энергетического факторов при определении машины? Всякая реальная система, перерабатывающая информацию, является системой материальной, причем способность перерабатывать информацию во многом зависит от субстрата этой системы. Конечно, энергетические процессы в ЭВМ отступают на второй план по сравнению с информационными процессами, но передача сигнала всегда осуществляется за счет затраты энергии (хотя бы и очень малой). Так что в понимании информационных машин отказ от материального и энергетического параметров в обсуждаемом контексте неправилен — при сравнении человека и машины это важно учитывать.

В плане сравнения человека и машины важен еще один существенный параметр, отличающий информационную систему, моделирующую мышление, от живого организма. Таким параметром является потеря машиной структурности в процессе ее работы. С учетом этого обстоятельства можно сказать, что причина работы машины выступает и как причина ее гибели, что машина в процессе работы увеличивает энтропию за счет потери структурности. Л. Фон Берталанфи подчеркивал отличие живого организма от машины, заключающееся в том, что живой организм развивается в направлении увеличения дифференциации и негомогенности и может корректировать «шум» в более высокой степени, чем это имеет место в коммуникационных каналах неживых систем, — оба эти свойства организма являются результатом того, что организм представляет собой открытую систему [1, с. 67]. Еще один важный параметр, отличающий машину от человека, — это, в конечном счете, производительность первой от второй.

Что же с учетом сказанного следует называть машиной в обсуждаемом контексте? Не стремясь дать полное и точное определение машины, обратим внимание на ряд факторов, которые необходимо учитывать при сравнительном анализе человека и машины. В определении машины, способной получать результаты, ассоциирующиеся с творческой человеческой деятельностью, необходимо исходить из того, что она представляет собой систему, искусственно созданную, в конечном счете, человеком, а не возникшую путем естественной эволюции, выполняющую операции в соответствии с целями, поставленными, в конечном счете, человеком, могущую работать без постоянного вмешательства человека, теряющую в процессе своей работы структурность, служащую средством усиления мыслительных функций человека и являющуюся орудием умственного труда человека, а не субъектом психической и социальной деятельности. Тот факт, что машина может изготавливаться не непосредственно человеком, а системой других машин, принципиально положения не меняет, так как и в этом случае создателем машины является человек. Такие машины — орудия умственного труда — могут быть соединены с машинами — орудиями физического труда. Тот факт, что человек в указанном смысле не является машиной, вовсе не означает, что его вообще нельзя представлять в виде машины. Однако при этом важно иметь в виду, какую именно сторону его деятельности мы представляем в виде машины: информационную, энергетическую и т.д.

Некорректности формулировки вопроса «Может ли машина мыслить?» способствовало смешение терминов «моделирование» и «искусственное воспроизведение»: под моделированием стали понимать воспроизведение. Однако, хотя эти понятия тесно связаны, между ними имеется существенное различие. Термин «воспроизводить»

означает производить снова то, что было или есть. Иными словами, воспроизведение предполагает воссоздание какой-либо системы с сохранением ее качественной специфики. При воспроизведении воссоздаются стороны, сохраняющие в совокупности сущность системы, ее природу. Для воспроизведения необходимо тождество не только результатов функционирования воспроизводимой системы, но и процессов, лежащих в основе функционирования, структуры и субстрата. Моделирование же предполагает воссоздание каких-то отдельных сторон моделируемой системы. Н. Винер и А. Розенблют не случайно определили моделирование через *representation* (подобие), а не через *reproduction* (воспроизведение) [12, р. 317]. Поэтому по отношению к живым системам точнее говорить о моделировании как имитации определенных функций. Это относится и к моделированию мышления — оно означает имитацию, а не воспроизведение мыслительной деятельности.

Вместе с тем, было бы неверно абсолютно противопоставлять воспроизведение и имитацию. О воспроизведении можно говорить в широком и узком смысле. В первом случае имеется в виду полное воспроизведение с учетом качеств, свойств и отношений. Во втором случае воспроизводятся какие-то интересующие исследователя стороны объекта — в этом случае предпочтительнее пользоваться термином «имитация» (или «частичное воспроизведение»), поскольку термин просто «воспроизведение» не содержит в себе указания на степень полноты воспроизведения. Если такое указание на степень полноты воспроизведения имеет место, то употребление термина «воспроизведение» становится оправданным. Так, например, в случае частичного воспроизведения объекта на иной субстратной основе используется термин «функциональное воспроизведение».

Чем важна субстратная основа мышления и сознания? Эпиморфоз (процессы, связанные с эволюционными преобразованиями и перестройкой морфологической структуры и психической организации), как наиболее яркое выражение биологического прогресса на его высшем уровне, обусловил становление человека, дал ему необходимую базу для развития нового типа отношений с окружающей средой, обеспечив тем самым выход за рамки действия биологических закономерностей. В этом же русле лежит концепция о существовании на определенных стадиях антропогенеза промежуточной биосоциальной формы отбора [7, с. 81]. Эти соображения отнюдь не умаляют роль труда в становлении человека и не биологизируют процесс антропогенеза. Дело заключается в том (и только в том), что не может быть игнорирован тот факт, что становление человека с его сознанием, определяемое социальным фактором, осуществлялось на специфической биологической основе. Иначе говоря, способ функционирования мозга определяется и его субстратом, и способом организации, вырастая как бы «изнутри» под влиянием социальной среды. В моделирующих мышление устройствах способ этой связи принципиально иной: здесь субстрат не играет важной роли, поскольку ведущая роль отводится способу организации логических сетей, т.е. здесь функция искусственно привязывается человеком к определенной структуре. Таким образом, свойства мышления и сознания — это специфически человеческая способность отражения реальности, для переноса которой на системы, сходные лишь по функционированию, т.е. для отрыва функции от субстрата и произвольного переноса ее на субстрат иной природы, нет оснований. Абсолютизация функционального подхода, отрыв функции от субстрата не продвигают нас к пониманию сущности сознания. Функциональный подход не раскрывает собственно биологических качеств и не снимает вопроса о связи живого с характером его носителя. Поэтому отвлекаться от субстрата можно лишь в определенных пределах в рамках задач, для решения которых отвлечение от субстрата допустимо.

Таким образом, использование терминов в их исконном значении — не пустая формальность. Термины «мышление», «сознание» возникли для обозначения

свойств человека как системы, прошедшей длительное биологическое и социальное развитие, т.е. развитие на определенном биологическом субстрате со свойственной ему организацией и в определенной социальной среде. Переносить свойства «мышления» и «сознания» на системы иной природы, состоящие из другого субстрата, имеющего иную организацию, методологически некорректно. Перефразируя слова Г. Цопфа, можно сказать, что, зачисляя человека в разряд машин на основании лишь внешнего подобия функционирования, мы «пытаемся затолкать слона в чайник вместо того, чтобы поискать более пригодное для слона вместилище» [9, с. 76]. Сравнительный анализ работы мозга и ЭВМ свидетельствует в пользу этого вывода. Для обозначения того, что «делает» ЭВМ, введен и получил права «гражданства» термин «искусственный интеллект». Сам вопрос «может ли машина мыслить?» должен быть переведен в методологически более корректную форму: каковы возможности искусственного интеллекта (каковы возможности моделирования мышления). Проблема заключается, таким образом, не в том, «может ли машина мыслить?», а в том, как соотносятся между собой естественный и искусственный интеллект, есть ли предел моделированию мышления.

Сейчас становится очевидным, что априорное установление каких бы то ни было пределов для моделирования мышления — дело рискованное. Такие ограничения выдвигались не раз, и столько же раз они снимались с развитием устройств, моделирующих мышление. Так, на первых этапах моделирования считалось невозможным создать устройство для решения задач, не имеющих жесткого алгоритма. После создания таких устройств было сформулировано другое ограничение: нельзя создать устройства, изменяющего свое поведение в зависимости от ситуации, т.е. способного к обучению и самообучению. Когда и такие устройства были созданы, ограничение приняло другую форму: моделирующее мышление устройство не может расширить число заложенных в программу человеком исходных положений. Но и это ограничение было снято с появлением вероятностных моделирующих устройств. Наиболее решительные возражения против информационного подхода к психике высказываются некоторыми психологами, ставящими под сомнение роль моделирования в объяснении психических явлений. Однако данное возражение справедливо лишь по отношению к возможности воспроизведения психических функций в ЭВМ, а не моделирования (имитации) их. Противники возможности моделирования творческих мыслительных функций выдвинули ряд аргументов в пользу обоснования своей позиции.

Аргумент первый: существование алгоритмически неразрешимых задач делает невозможным моделирование творческих мыслительных процессов. К таким задачам относят: распознавание выводимости (А. Черч), установление тождества теории групп (П.С. Новиков), распознавание эквивалентности слов в любом исчислении (А.А. Марков и Э. Пост). Однако данный аргумент уязвим. Человек способен у алгоритмически неразрешимых задач находить разрешимые частные случаи. Любой мыслительный процесс можно моделировать, если он доступен описанию.

Аргумент второй: для моделирования процесса решения задачи его необходимо формализовать, а поскольку полная формализация невозможна, то решение не всех задач поддается моделированию. Действительно, теорема К. Геделя показывает, что на основе формального исчисления не может быть изложено даже учение о целых числах. Если имеется достаточно мощная непротиворечивая формальная система, то в ней при помощи математических средств, выражающих эту систему, можно сформулировать такие утверждения, которые в ее рамках нельзя ни доказать, ни опровергнуть, т.е. в рамках данной системы эти утверждения считаются неразрешимыми. На этом основании был сделан вывод о том, что не все математические операции могут быть моделированы и что теоремы неполноты и неразрешимости вообще выдвигают

преграду исчерпывающему познанию систем, описываемых аксиоматическим формальным языком. Однако ограничения, вытекающие из теоремы Геделя, относятся лишь к машинам Тьюринга, не получающим из внешней среды информацию. Как было показано В.М. Глушковым, если машина получает информацию из внешней среды, она «оказывается способной решать неконструктивные проблемы, относительно которых можно было доказать их алгоритмическую неразрешимость» [3, с. 21]. Неразрешимость, следовательно, относится к абстрактному мышлению, а не к процессу познания в целом. Поскольку из теоремы Геделя вытекает не только неполнота той или иной системы, но также указание на то, что не охватывается данной системой, то в определенных рамках можно формализовать процесс перехода от одной системы к другой, более мощной. Формальная система, получив «толчок» для своего существования, приобретает известную самостоятельность. Но поскольку характер такой системы чисто формальный, то эта система не позволяет определить, в каком направлении она должна быть расширена, если какая-либо задача, содержащаяся в ней, неразрешима. Для этого необходимо учитывать ее содержательное отношение с другой, более мощной системой. Если процесс творчества рассматривать как такое «расширение», то следует признать, что он больше относится к психологии, нежели логике — направление этому «расширению» дает, в конечном счете, человек.

Аргумент третий: творческие функции мышления не могут быть описаны с помощью математических средств. Поскольку устройства, моделирующие мышление, имеют дело с формально-логическими отношениями, а для обширных областей человеческого мышления характерны содержательные связи, то эти области не могут быть познаны с помощью математических методов. Однако дело не в том, что нельзя в принципе математически описать некоторые мыслительные операции, а в том, что это нельзя сделать с помощью существующих в настоящее время математических средств. Как известно, существующий математический аппарат возник преимущественно для нужд физики и инженерной техники, где переменность и число переменных незначительны. Речь, стало быть, должна идти о «выращивании» новых глав математики применительно к новым объектам, о создании исчислений, способных описывать мыслительные процессы.

Аргумент четвертый: моделирование мыслительных процессов возможно лишь в рамках дедукции, т.е. выведения следствий из какой-то суммы знаний, а творческий процесс не ограничивается дедукцией. Однако уже существующие диагностические системы с успехом моделируют процесс остановки диагноза, предполагающий индуктивные выводы, использование аналогии и т.д.

Аргумент пятый: можно моделировать лишь решение сформулированной задачи, но не постановку цели и интерпретацию достигнутого результата, которые относятся к числу творческих операций. Однако этот аргумент внутренне противоречив, поскольку и решение задачи содержит элементы творчества, к числу которых относятся выработка критериев поиска и критериев отбора, что успешно моделируется.

Рассмотренные аргументы не могут служить доказательством принципиальной невозможности моделирования сложных форм творческой деятельности. По нашему мнению, вопрос о моделировании творческой деятельности целесообразно перевести в другую плоскость — сосредоточить внимание на поисках конкретных путей моделирования тех или иных видов творческой деятельности. Уже и сейчас ясно, что ЭВМ может решать задачи, которые человек без ее помощи решить не в состоянии [12]. О возможностях будущих поколений ЭВМ и роботов мы можем только догадываться. Но важно быть готовыми к любому повороту событий. Наиболее перспективным направлением обещает быть создание систем интегрального интеллекта и различного рода роботов (особенно наноботов). Человек должен держать под контролем меру их «самостоятельности», ибо в реальности законы робототехники А. Азимова не работают.

Литература:

1. *Берталанфи Л. фон.* Общая теория систем — критический обзор // Исследования по общей теории систем. М., 1969.
2. *Бир Ст.* Кибернетика и управление производством. М., 1965.
3. *Глушков В.М.* Мышление и кибернетика. М., 1966.
4. *Вулридж Д.* Механизмы мозга. М., 1965.
5. *Колмогоров А.Н.* Жизнь и мышление с точки зрения кибернетики // *Опарин А.И.* Жизнь, ее соотношение с другими формами движения материи. М., 1962.
6. *Кочергин А.Н.* Искусственный интеллект и мышление // *Философия искусственного интеллекта.* М., 2005.
7. *Семенов Ю.И.* Как возникло человечество. М., 1966.
8. *Тринчер К.С.* Существование и эволюция живых систем и второй закон термодинамики // *Вопросы философии.* 1962. № 6.
9. *Цопф Г.* Отношение и контекст // *Принципы самоорганизации.* М., 1966.
10. *Эшби У.Р.* Конструкция мозга. М., 1962.
11. *Appel K., Haken W.* The solution of the colour map problem // *Sci. Amer.* 1977. Vol. CXXXVII, № 10.
12. *Rosenblueth A., Wiener N.* The Role of Models in Science // *Philosophy of Science.* 1945. Vol. 12, № 4.