

УДК 16

## МЫШЛЕНИЕ МАШИН ПО АЛАНУ ТЬЮРИНГУ

Хорин П. А., Нестеров А. Ю.

Самарский национальный исследовательский университет  
имени академика С. П. Королёва, г. Самара

Для рассмотрения тезиса о возможности мышления машин необходимо определить смысл терминов «мышление» и «машина».

Разработка проблемы искусственного интеллекта (ИИ) связана с результатами исследования естественного интеллекта (ЕИ). Что касается вопроса о сознании – он эквивалентен вопросу об отношении сознания к главному мозгу.

Сознание обладает качеством субъективной реальности (СР). Качество СР является источником основной проблемы связи сознания и головного мозга. [1]

Явление СР связано с мозговыми процессами, как информация – с носителем. Связь явления СР со своим нейродинамическим кодом – функционал. Основная задача заключается в расшифровке нейродинамического кода, т.е. в преобразовании чуждого кода в естественный код. Когда чуждый код становится естественным – это означает развитие системы.

Рассмотрим тезис об изофункционализме систем, развитый Аланом Тьюрингом (один и тот же набор функций может быть воспроизведен системами, различными по своим субстратным, физическим свойствам). Для обоснования тезиса о возможности обретения искусственным интеллектом качества СР важное теоретическое значение имеет принцип инвариантности (ПИ) информации по отношению к физическим свойствам ее носителя. Из него следует возможность возникновения различных вариантов кодовой самоорганизации. ПИ означает, что одна и та же информация может иметь носители с разными физическими свойствами. Теоретически были возможны иные варианты, но в доступном нам мире существует лишь тот вариант самоорганизующейся системы, наделенной СР, который был изобретен, создан биологической эволюцией. Он является наиболее ёмким, быстрым и наиболее экономичным в энергетическом отношении (код ДНК, код в нервной системе). Этот тезис так или иначе показывает пути сближения ИИ с ЕИ.

Под «машиной» мы понимаем только цифровые вычислительные машины. Предполагается, что эти машины могут выполнять любую операцию, которую мог бы выполнить человек-вычислитель, у которого есть некий набор инструкций – «книга правил». Можно считать, что цифровая вычислительная машина состоит из трех частей: запоминающее, исполнительное и контролирующее устройство.

Впоследствии такая абстрактная машина получила название «Машина Тьюринга» (МТ). Она была предложена Аланом Тьюрингом для формализации понятия алгоритма.

Для работы машины необходимо заполнить таблицу команд, т.е. запрограммировать машину. К сожалению, МТ – это машина с дискретными состояниями. На основе этого тезиса можно расширить основной вопрос следующим образом: существует ли машина с дискретными состояниями, которая могла бы хорошо играть в «игру в имитацию»? [2]

Хорошей игрой в имитацию будем называть случай, когда человек больше чем в 70% случаев ответил, что машина – это человек, хотя на самом деле им не является.

Возможными путями решения поставленной задачи могут являться самообучаемые машины, нейронные сети, перцептроны и другие попытки реализации ИИ. Но означает ли хорошая игра в имитацию акт мышления? Ведь продублировать

одного из носителей естественного интеллекта не значит изучить основополагающие принципы интеллекта, а, следовательно, и не значит понять процесс мышления.

Рассмотренная МТ прекрасно решает алгоритмические задачи, однако исчерпывается ли мышление алгоритмическими задачами? И можно ли назвать машину такого типа ИИ-ом?

Для разрешения возникших вопросов нам понадобится разобраться с определением ИИ, а точнее, развести понятия слабого и сильного ИИ.

ИИ в слабом смысле – это инструмент, который помогает изучить сознание. Сильный ИИ – это само сознание (которое обладает когнитивными состояниями).

МТ можно рассматривать в рамках ИИ в слабом смысле. Анализ результатов, полученных в результате игры в имитацию, может помочь в улучшении модели сознания человека. Изменение тренировочной выборки в случае обучаемых машин даст рано или поздно так называемую «хорошую игру в имитацию». Но вышеперечисленные факторы не гарантируют наличия «понимания» у ИИ такого типа.

Хорошим аргументом может послужить мысленный эксперимент «Китайская комната», сформулированный Джоном Сёрлом [3].

Человек за пределами комнаты, который знает китайские иероглифы, передаёт в комнату иероглифы с вопросами, а на выходе ожидает получить осознанный ответ. В такой ситуации человек в комнате может выступать в роли компьютера. Ни такой человек, ни компьютер на его месте не сможет научиться и создать новые правила, если не будет знать значение символов. Он не будет понимать ни вопроса, ни ответа, который сам составил.

Что касается человека вне комнаты, то он в свою очередь может быть уверен, что в комнате находится человек, который знает и понимает иероглифы. В результате будет пройден тест Тьюринга, но не будет совершен акт мышления.

Данная ситуация хорошо показывает, что «машинное мышление» - это метафора, внятно определённая как задача имитации.

Современные искусственные информационные системы далеки от самоорганизации такого типа, т.к. на входе и на выходе у них стоит человек, задающий программу и использующий результат их деятельности.

Значительным вкладом Алана Тьюринга является формализация теории алгоритмов. Именно она пока является основополагающим фактором в программировании любого рода кибернетических систем.

Единственный способ развития искусственного мышления – это исследования в области ЕИ и попыток применения механизмов ЕИ к ИИ. Другой вариант разрешения данной проблемы возможен в случае, если существует другая форма СР или если человек сможет изобрести некую самоорганизующуюся систему с качественно отличающейся от известной нам моделью СР.

#### Библиографический список

1. Дубровский, Д.И. Сознание, мозг, искусственный интеллект [Текст]: сб. статей / Д.И. Дубровский. – Москва, 2007.
2. Turing, A. Computing machinery and intelligence [Текст] // Mind: – Oxford: Oxford University Press, 1950. – No. 59. – P. 433 - 460.
3. Searle, John. R. (1980) Minds, brains, and programs [Текст] // Behavioral and Brain Sciences 3 (3): P. 417-457.