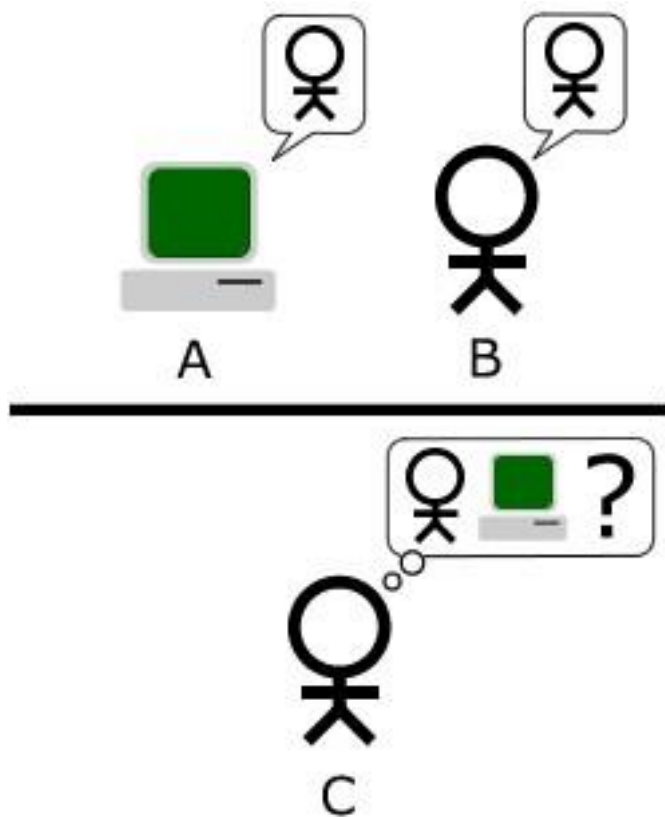


Тест Тьюринга

Тест Тьюринга — эмпирический тест, идея которого была предложена Аланом Тьюрингом в статье «Вычислительные машины и разум», опубликованной в 1950 году в философском журнале «*Mind*». Тьюринг задался целью определить, может ли машина мыслить.



Стандартная интерпретация этого теста звучит следующим образом: *«Человек взаимодействует с одним компьютером и одним человеком. На основании ответов на вопросы он должен определить, с кем он разговаривает: с человеком или компьютерной программой. Задача компьютерной программы — ввести человека в заблуждение, заставив сделать неверный выбор».*

Все участники теста не видят друг друга. Если судья не может сказать определенно, кто из собеседников является человеком, то считается, что машина прошла тест. Чтобы протестировать именно интеллект машины, а не её возможность распознавать устную речь, беседа ведется в режиме «только текст», например, с помощью клавиатуры и экрана (компьютера-посредника). Переписка должна производиться через контролируемые промежутки времени, чтобы судья не мог делать заключения, исходя из скорости ответов. Во времена Тьюринга компьютеры реагировали медленнее человека. Сейчас

это правило тоже необходимо, потому что они реагируют гораздо быстрее, чем человек.

Говорят, что машина прошла тест Тьюринга, если она за 5 минут текстового общения смогла обмануть 30% судей.

История

Хотя исследования в области искусственного интеллекта начались в 1956 году, их философские корни уходят глубоко в прошлое.

К 1956 году британские учёные уже на протяжении 10 лет исследовали «машинный интеллект». Этот вопрос был обычным предметом для обсуждения среди членов «Ratio Club» — неформальной группы британских кибернетиков и исследователей в области электроники, в которой состоял и Алан Тьюринг, в честь которого был назван тест.

Тьюринг в особенности занимался проблемой машинного интеллекта, по меньшей мере, с 1941 года. Одно из самых первых его упоминаний о «компьютерном интеллекте» было сделано в 1947 году. В докладе «Интеллектуальные машины» Тьюринг исследовал вопрос, может ли машина обнаруживать разумное поведение, и в рамках этого исследования предложил то, что может считаться предтечей его дальнейших исследований: «Нетрудно разработать машину, которая будет неплохо играть в шахматы.

Теперь возьмем трех человек — субъектов эксперимента. А, В и С. Пусть А и С неважно играют в шахматы, а В — оператор машины. [...] Используются две комнаты, а также некоторый механизм для передачи сообщений о ходах. Участник С играет или с А, или с машиной. Участник С может затрудниться ответить, с кем он играет».

Таким образом, к моменту публикации в 1950 году статьи «Вычислительные машины и разум», Тьюринг уже на протяжении многих лет рассматривал возможность существования искусственного интеллекта. Тем не менее, данная статья стала первой статьёй Тьюринга, в которой рассматривалось исключительно это понятие.

Тьюринг начинает свою статью утверждением: «Я предлагаю рассмотреть вопрос „Могут ли машины думать?“». Он подчёркивает, что традиционный подход к этому вопросу состоит в том, чтобы сначала определить понятия «машина» и «интеллект». Тьюринг, однако, выбрал другой путь; вместо этого он заменил исходный вопрос другим, «который тесно связан с исходным и

формулируется относительно недвусмысленно». По существу, он предлагает заменить вопрос «Думают ли машины?» вопросом «Могут ли машины делать то, что можем делать мы (как мыслящие создания)?». Преимуществом нового вопроса, как утверждает Тьюринг, является то, что он проводит «чёткую границу между физическими и интеллектуальными возможностями человека».

Чтобы продемонстрировать этот подход, Тьюринг предлагает тест, придуманный по аналогии с игрой для вечеринок «Imitation game» — имитационная игра. В этой игре мужчина и женщина направляются в разные комнаты, а гости пытаются различить их, задавая им серию письменных вопросов и читая напечатанные на машинке ответы на них. По правилам игры и мужчина, и женщина пытаются убедить гостей, что все наоборот. Тьюринг предлагает переделать игру следующим образом: "Теперь зададим вопрос, что случится, если в этой игре роль А будет исполнять машина? Будет ли задающий вопросы ошибаться так же часто, как если бы он играл с мужчиной и женщиной? Эти вопросы заменяют собой исходный «Может ли машина думать?».

В том же докладе Тьюринг позднее предлагает «эквивалентную» альтернативную формулировку, включающую судью, который беседует только с компьютером и человеком. Наряду с тем, что ни одна из этих формулировок точно не соответствует той версии теста Тьюринга, которая наиболее известна сегодня, в 1952 учёный предложил третью. В этой версии теста, которую Тьюринг обсудил в эфире радио Би-Би-Си, жюри задает вопросы компьютеру, а роль компьютера состоит в том, чтобы заставить значительную часть членов жюри поверить, что он на самом деле человек.

В статье Тьюринга учтены 9 предполагаемых вопросов, которые включают все основные возражения против искусственного интеллекта, поднятые после того, как статья была впервые опубликована.

Одна из первых машин, которая могла пройти тест Тьюринга была Элиза. Принцип работы Элизы заключается в исследовании введенных пользователем комментариев на наличие ключевых слов. Если найдено ключевое слово, то применяется правило, по которому комментарий пользователя преобразуется и возвращается предложение-результат. Если же ключевое слово не найдено, Элиза либо возвращает пользователю общий ответ, либо повторяет один из предыдущих комментариев. Вдобавок Уайзенбаум запрограммировал Элизу на имитацию поведения психотерапевта, работающего по клиент-центрированной методике. Это

позволяет Элизе «притвориться, что она не знает почти ничего о реальном мире». Применяя эти способы, программа Уайзенбаума могла вводить в заблуждение некоторых людей, которые думали, что они разговаривают с реально существующим человеком, а некоторых было «очень трудно убедить, что Элиза [...] не человек». На этом основании некоторые утверждают, что Элиза — одна из программ (возможно первая), которые смогли пройти тест Тьюринга. Однако это утверждение очень спорно, так как людей, «задающих вопросы», инструктировали так, чтобы они думали, что с ними будет разговаривать настоящий психотерапевт, и не подозревали о том, что они могут разговаривать с компьютером.

Критика теста Тьюринга

1. Сёрль

В 1980 году в статье «Разум, мозг и программы» Джон Сёрль выдвинул аргумент против теста Тьюринга, известный как мысленный эксперимент «Китайская комната». Сёрль настаивал, что программы (такие как Элиза) смогли пройти тест Тьюринга, просто манипулируя символами, значения которых они не понимали. А без понимания их нельзя считать «разумными» в том же смысле, что и людей. «Таким образом, — заключает Сёрль, — тест Тьюринга не является доказательством того, что машина может думать, а это противоречит изначальному предположению Тьюринга».

Такие аргументы, как предложенный Сёрлем, а также другие, основанные на философии разума, породили намного более бурные дискуссии о природе разума, возможности существования разумных машин и значимости теста Тьюринга, продолжавшиеся в течение 80-х и 90-х годов.

2. Стюарт Рассел и Питер Норвиг

Стюарт Рассел и Питер Норвиг утверждают, что антропоцентризм теста приводит к тому, что он не может быть по-настоящему полезным при разработке разумных машин. «Тесты по авиационному проектированию и строительству, — строят они аналогию, — не ставят целью своей отрасли создание машин, которые летают точно так же, как летают голуби, что даже сами голуби принимают их за своих»^[5]. Из-за этой непрактичности прохождение теста Тьюринга не является целью ведущих научных или коммерческих исследований (по состоянию на 2009). Сегодняшние

исследования в области искусственного интеллекта ставят перед собой более скромные и специфические цели.

«Исследователи в области искусственного интеллекта уделяют мало внимания прохождению теста Тьюринга, — отмечают Рассел и Норвиг, — с тех пор как появились более простые способы проверки программ, например, дать задание напрямую, а не окольными путями, первой обозначить некоторый вопрос в чат-комнате, к которой подключены и машины, и люди». Тьюринг никогда не предполагал использовать свой тест на практике, в повседневном измерении степени разумности программ; он хотел дать ясный и понятный пример для поддержки обсуждения философии искусственного интеллекта.

Машина которая прошла тест Тьюринга

Женя Густман

- виртуальный собеседник, который, согласно первоначальным сообщениям в СМИ, «впервые сумел пройти тест Тьюринга» на испытаниях, организованных в 2014 году университетом Рединга (Великобритания). Был создан группой из трёх программистов: Владимира Веселова (родом из России, живёт в Нью-Джерси), Евгения Демченко (родом из Украины) и Сергея Уласеня (родом из России). Разработка программы была начата в Санкт-Петербурге в 2001 году. Чтобы характер и знания Густмана казались более правдоподобными, он представляется пользователям 13-летним мальчиком из Одессы.

Густман с момента его создания принимал участие в ряде соревнований на прохождение теста Тьюринга и несколько раз занимал второе место в соревновании на премию Лёбнера. В июне 2012 года Густман выиграл соревнование в честь 100-летия Алана Тьюринга, сумев убедить 29 % судей, что он человек.^[1] 7 июня 2014 года, на конкурсе, посвященном 60-летию со дня смерти Тьюринга, Густман убедил 33 % судей, что он человек, и, по словам Кевина Уорика, стал первым в истории компьютером, прошедшим тест Тьюринга.

Имеются утверждения, что Густман стал далеко не первым, и даже не лучшим по результатам аналогичных тестов. Результат в 59 % был показан ещё в 2011 году ботом Cleverbot. В ответ на подобные утверждения организатор соревнования 2014 года в университете Рединга, профессор Кевин Уорик подчеркнул, что, в отличие от предыдущих соревнований, к которым применяли слова «тест Тьюринга», в этот раз было проведено больше

одновременных сравнительных тестов, чем когда-либо ещё, они были независимо проверены и, что критично, тематика бесед не была ограничена.

Сам Тьюринг вовсе не устанавливал столь низкий процент введённых в заблуждение собеседников в качестве критерия «прохождения теста», а лишь предположил, что в 2000 году компьютеры с объёмом памяти около 10^9 бит смогут играть в имитационную игру так, что у среднего собеседника будет не более 70 % шансов опознать машину после пятиминутного собеседования.

Критики теста Тьюринга принижают значимость соревнований на прохождение теста, утверждая, что Женья Густман является лишь «чатботом»:

...Машина прикидывается всего лишь ребёнком, ну а полноценное прохождение теста Тьюринга невозможно ею в принципе. Ибо тест всего лишь бихевиористичен; на принципиальный вопрос — мыслит ли машина? — он ответа дать не может... Данные вопросы, конечно, могут обеспечить работой поколения философов-профессионалов, равно как и досугом — обширные круги философов-самоучек. Но вот с точки зрения инженерного дела или бизнеса они никакого смысла не имеют.