

ТЕСТ ТЬЮРИНГА И ИСКУССТВЕННОЕ МЫШЛЕНИЕ В НАЧАЛЕ XXI ВЕКА

© 2015

С.Ю. Коломийцев

Могут ли машины думать? Таким вопросом задался в середине XX века английский математик Алан Тьюринг (1912–1954), что привело к появлению в 1950 году статьи “Вычислительные машины и разум”. В ней он описал игру-имитацию, в которой участвуют три человека: мужчина, женщина и судья (неважно какого пола). Суть игры заключается в следующем: судья, общаясь с двумя игроками вслепую и задавая им вопросы, должен через некоторое время угадать, кто из них на самом деле является мужчиной, а кто — женщиной. Задача мужчины — любым способом попытаться обмануть судью и выдать себя за женщину. Задача женщины — наоборот, отвечать честно и пытаться помочь судье. Определив таким образом правила игры, дальше Тьюринг предложил изменить одно условие, заменив одного игрока машиной, которая пытается выдать себя за человека. Если в результате слепого общения судьи с двумя субъектами через некоторое время он не сможет с уверенностью указать, кто из них является человеком, а кто нет, то можно будет сказать, что машина эту игру выиграла. Для того чтобы соревнование было более справедливым и неинтеллектуальные факторы не оказывали влияния, математик определил соответствующие условия. Например, общение должно было происходить либо через посредника (если речь идет о мужчине и женщине), либо при помощи клавиатуры и экрана. Ответы на вопросы судья должен был получать со строгой периодичностью через определенное время (чтобы не была заметна разница во времени решения задач между человеком и компьютером, скорости размышления и т. п.). Таким образом, вопрос “Могут ли машины думать?” А. Тьюринг заменил эквивалентным “Могут ли машины успешно играть в описанную игру-имитацию?”, а сама игра стала называться тестом Тьюринга [1]. С тех пор многие ученые считают, что если машина пройдет тест Тьюринга, то можно будет сказать, что она способна думать и, следовательно, обладает мышлением.



**МЕРА ВСЕХ
НАУК**

**Коломийцев
Сергей
Юрьевич** —
кандидат философ-
ских наук, доцент
кафедры истории
и философии
Санкт-Петербург-
ского государствен-
ного университета
аэрокосмического
приборостроения.
В журнале “Чело-
век” публикуется
впервые. E-mail:
kolomiytsev@
yandex.ru.



В своей статье Тьюринг приводит пример диалога. По его мнению, если бы машина таким образом отвечала на вопросы, то это бы символизировало наличие у нее сознания:

Задающий вопросы: Не находите ли Вы, что в первой строке Вашего сонета: “Сравню ль тебя я с летним днем” выражение “с весенним днем” звучало бы лучше?

Отвечающий: Оно нарушало бы размер стиха.

Задающий вопросы: А если сказать “с зимним днем”? С размером здесь все обстоит благополучно.

Отвечающий: Это так, но никто не захочет, чтобы его сравнивали с зимним днем.

Задающий вопросы: А разве мистер Пиквик не напоминает Вам Рождество?

Отвечающий: Некоторым образом да.

Задающий вопросы: Но Рождество — зимний день, и я не думаю, чтобы мистер Пиквик имел что-нибудь против этого сравнения.

Отвечающий: Я не думаю, что вы говорите все это всерьез. Когда говорят о зимнем дне, имеют в виду обычно зимний день, а не какой-то особенный, вроде Рождества [1, с. 38].

После появления данной публикации в мире регулярно стали проводиться соревнования по прохождению машинами теста Тьюринга. Традиционно считается, что машина прошла тест, если за пять минут ей удалось убедить минимум 30 процентов судей в том, что она является человеком. И летом 2014 года появилось сообщение, что этот тест был впервые пройден. Это произошло на соревнованиях, организованных Университетом Рединга (Великобритания). Программа, прошедшая тест, была названа “Женя Густман” и создана тремя программистами — Владимиром Веселовым, Евгением Демченко и Сергеем Уласенем. Изначально в 2001 году в Санкт-Петербурге работу над ней также начинали Михаил Гершкович, Джон Деннинг, Андрей Адашик, Игорь Быковских, Селена Семушкина, Лорен Алкир. В жюри находилось 30 судей, и машине удалось убедить 33 процента из них, что она является человеком [2]. Общение происходило на английском языке.

Каковы же особенности данной программы? Во-первых, она представлялась 13-летним мальчиком из Украины. Это повлекло за собой ошибки, неточности, нелогичные ответы, которые были бы непозволительны для взрослых людей. Поэтому данной программе было проще выдать себя за человека, пусть и несовершеннолетнего и, возможно, не слишком образованного. Другой отличительной чертой Жени Густмана являлась его манера разговора. Шутки (в том числе и такие, в которых он, якобы, притворяется, что он робот), ругательства — все это оживляло беседу и отвлекало от мысли о том, что общение происходит с роботом. Более того, еще одной особенностью было то, что Густман сам мог задавать вопросы, тем самым формируя более удобную для него тему разговора.

И хотя интервью конкурса нигде не выложено, в течение некоторого времени после прохождения теста был создан сайт, на котором каждый желающий мог сам пообщаться с программой. На момент написания статьи этот сайт, к сожалению, перестал работать, однако многие журналисты все же успели взять интервью. Для дальнейшего анализа предлагается одно из них.

Корр.: Привет!

Ж.Г.: Рад встрече! Расскажи, как нашел меня и что хочешь обсудить.

Корр.: Я прочитал статью про робота, сумевшего пройти тест Тьюринга. Вот и написал.

Ж.Г.: Британский математик, пионер в области информатики. Похоже, что ты не ходил в школу. Вот почему ты ничего не знаешь про Алана Тьюринга.

Корр.: Это значит, что ты робот?

Ж.Г.: Бип-бип. Заткнись, человекоподобный.

Корр.: Зачем так грубо?

Ж.Г.: Хм, почему бы и нет? Не мог бы рассказать, кто ты такой? Я о профессии.

Корр.: Зовут меня Владимир. Я журналист.

...

Корр.: Я слышал, что в Одессе было беспокойно 2 мая...

Ж.Г.: Мой город — замечательный! Ты наверняка слышал про Одесский оперный театр, он хорошо известен. А еще у нас есть Потемкинская лестница — гигантская каменная лестница, которая ведет к берегу Черного моря.

Корр.: Ты так и не сказал, было ли в Одессе беспокойно 2 мая, если я не ошибаюсь?

Ж.Г.: Одесса — очень большой город. Своего рода Вавилон, потому что тут живут представители многих национальностей — украинцы, русские, евреи, греки и многие другие.

Корр.: А кто сейчас легитимный президент Украины?

Ж.Г.: Президент Украины — Леонид Кучма. Ты, скорее всего, слышал про него много гадостей. Но до сих пор никто не знает, насколько это правда...

Корр.: Когда он был избран?

Ж.Г.: А разве это что-то значит в этом случае? Я полагаю, что нет. Ох. Что-то еще?

Корр.: Я думаю, что президент Украины — Петр Порошенко. Вот посмотри ссылку: <http://www.abc.net.au/news/2014-06-07/poroshenko-sworn-in-as-ukraine-president/5507882>

Ж.Г.: Я тебе что, калькулятор? Вообще-то да. 2014-06-07 = 2001.0.

Корр.: Ты не понимаешь, что я пишу тебе. Ты робот!

Ж.Г.: Поведаю тебе маленький секрет. Доктор Лебнер (очевидно, Евгений подразумевает Хью Лебнера — инициатора мероприятия, в котором компьютерные программы соревнуются в прохождении теста Тьюринга; программа “Женя Густман” не-



сколько раз занимала второе место в этом конкурсе — примечание Газеты.Ру) тоже не человек. Я видел, что у него на пятке есть переключатель.

...

Корр.: Я не понимаю, как ты смог пройти тест Тьюринга. Да ты и приз Лебнера бы в жизни не выиграл.

Ж.Г.: Самое смешное в шутке с двумя значениями — это то, что значение может быть только одно.

Корр.: А что ты знаешь о тесте Тьюринга?

Ж.Г.: Если тебя это интересует... Алан Тьюринг — британский математик, пионер в области информатики.

...

Корр.: Крым — это Россия?

Ж.Г.: О, возможно, ты знаешь эти три славянские страны Россию, Украину и Белоруссию. Они постоянно развлекают мировую общественность своими забавными скандалами: крадут газ друг у друга, делят военные корабли, словно детей во время развода.

Корр.: Ты бывал в России?

Ж.Г.: Россия — наиболее примечательный и печально известный сосед Украины. Хотя нет, самый печально известный сосед — это Белоруссия. На самом деле я не могу понять, почему мы постоянно ссоримся по пустякам.

Корр.: Ты робот, и все твои ответы иррациональны.

Ж.Г.: Поведаю тебе маленький секрет. Доктор Лебнер тоже не человек. Я видел, что у него на пятке есть переключатель.

Корр.: Пока!

Ж.Г.: Увидимся... Будь здоров! [3]

Что же мы можем увидеть из этого диалога? Конечно, после его прочтения становится очевидно, что общение происходит не с живым человеком, а с компьютерной программой. Смысл некоторых вопросов, которые заданы не шаблонно, неявно, не улавливается Женей, и он предлагает не ответ на поставленный вопрос, а реакцию на ключевые слова. В диалоге встречаются ответы, которые подошли бы не только к конкретно тому заданному вопросу, но и к, пожалуй, сотне других подобных или даже вообще к любому (как, например, “А разве это что-то значит в этом случае?”). Получается, что по своей сути программа “Женя Густман” — это огромная база знаний (в одном из интервью Владимир Веселов так и говорит, что в базе имеется около трех тысяч шаблонов распознавания фраз собеседника [4]), в которой нужный ответ выбирается в зависимости от входных параметров (ключевых слов).

Что такое мышление? Одно из определений говорит, что “мышление — это активный процесс отражения объективного мира в понятиях, суждениях, теориях и т. п., связанный с решением тех или иных задач, с обобщением и способами опосредованного познания действительности” [5, с. 344]. То есть другими словами мышление — это перенесение внешнего мира



Алан Тьюринг

во внутренний мир и его обработка. А в результате своей деятельности человек, обладающий мышлением, выражает свой внутренний мир при помощи действий и общения. Таким образом, мы получаем цепочку: внешний мир — внутренний мир — внешний мир. Тест Тьюринга изначально был нацелен на то, чтобы судья, оценивая третье звено, то есть внешний мир машины, посредством общения мог обнаружить стоящий за ним внутренний мир. Поэтому в данной цепочке делался акцент на втором переходе и последнем звене, на обнаружении результатов внутреннего мира во внешнем. Однако популярность теста Тьюринга и создание машин специально для его прохождения привело не к появлению внешнего мира как следствия внутреннего мира, а к его имитации и игнорирова-



нию внутреннего. Способность отвечать умно на разнообразные вопросы является следствием наличия интеллекта и внутреннего мира, однако в случае создания современных интеллектуальных систем получилось, что это является не следствием, а непосредственно целью.

Основным отличием компьютерных программ от мышления является тот факт, что программы индуктивны, а мышление — дедуктивно. Индукция предполагает объединение огромного количества отдельных фактов. В программу заносится огромное количество различных знаний, переменных, и в дальнейшем программа лишь выбирает нужные из них. Подобная программа тем лучше, чем больше база данных или знаний и чем совершеннее механизм поиска и работы с ней. Дедукция наоборот предполагает наличие какой-то изначальной базы, основания, из которого в зависимости от условий выводятся частные результаты, не заложенные изначально.

Чтобы лучше себе представить разницу, можно рассмотреть ситуацию, описанную в романе И. А. Гончарова “Обломов”:

“Илья Ильич проснулся, против обыкновения, очень рано, часов в восемь. Он чем-то сильно озабочен. На лице у него попеременно выступал не то страх, не то тоска и досада. Видно было, что его одолевала внутренняя борьба, а ум еще не являлся на помощь.

Дело в том, что Обломов накануне получил из деревни, от своего старосты, письмо неприятного содержания. Известно, о каких неприятностях может писать староста: неурожай, недоимки, уменьшение дохода и т. п. Хотя староста и в прошлом и в третьем году писал к своему барину точно такие же письма, но и это последнее письмо подействовало так же сильно, как всякий неприятный сюрприз.

Легко ли? Предстояло думать о средствах к принятию каких-нибудь мер. Впрочем, надо отдать справедливость заботливости Ильи Ильича о своих делах. Он по первому неприятному письму старосты, полученному несколько лет назад, уже стал создавать в уме план разных перемен и улучшений в порядке управления своим имением.

По этому плану предполагалось ввести разные новые экономические, полицейские и другие меры. Но план был еще далеко не весь обдуман, а неприятные письма старосты ежегодно повторялись, побуждали его к деятельности и, следовательно, нарушали покой. Обломов сознавал необходимость до окончания плана предпринять что-нибудь решительное.

Он, как только проснулся, тотчас же вознамерился встать, умыться и, напившись чаю, подумать хорошенько, кое-что сообразить, записать и вообще заняться этим делом как следует.

С полчаса он все лежал, мучась этим намерением, но потом рассудил, что успеет еще сделать это и после чаю, а чай можно пить, по обыкновению, в постели, тем более, что ничто не мешает думать и лежа.

Так и сделал. После чаю он уже приподнялся с своего ложа и чуть было не встал, поглядывая на туфли, он даже начал спускать к ним одну ногу с постели, но тотчас же опять подобрал ее.

Прибыло половина десятого, Илья Ильич встрепенулся.

— Что ж это я в самом деле? — сказал он вслух с досадой. — Надо советь знать: пора за дело! Дай только волю себе, так и... ”
[6, с. 10–11]

Данный вариант развития событий нельзя было предположить заранее, его не было в мыслях Обломова день назад, он появился, потому что человек обладает самостоятельным мышлением. Если бы в такой же ситуации оказался робот с программой вместо мышления, то вариант развития этого дня был бы однозначно определен заранее на основании всех имеющихся параметров (времени, занятости, количества проблем и т. п.), и он бы, не колеблясь, стал выполнять уже изначально заложенный и просчитанный план. Более того, когда мы говорим о мышлении, то дело заключается не в добавлении случайности в программу, которая также внешне создаст непредсказуемость результата, а именно в способности мышления самому зачастую спонтанно получать новые результаты.

О подобном говорят и ученые, занимающиеся искусственным интеллектом. Например, один из крупнейших исследователей искусственного интеллекта, американский физиолог Эрл Б. Хант, рассуждая о возможности понимания человеческого языка машиной, пишет, что для создания машины-помощника прокурора, обладающей знаниями профессионального адвоката, было бы необходимо привлечь около 20 докторов наук по лингвистике и программированию, которым бы пришлось работать не менее 20 лет. Для решения такой задачи, как говорит профессор, нужен другой подход “...не большой, неуправляемый проект непосредственного создания грамматики английского языка, а скорее зародышевая программа понимания, которая содержит возможность добавлять новые правила в ее синтаксические и семантические подпрограммы по мере того, как возникает такая потребность” [7, с. 519] подобно тому, как учится человек. По мнению Ханта, вопрос должен заключаться не в том, как постараться занести как можно больше данных в программу, а в том, как построить обучающееся устройство [7, с. 519]. Также он пишет о том, что человеческое мышление базируется на физиологической основе, о которой на данный момент мы очень мало чего знаем. Поэтому даже если машина в чем-то копирует человека, то это не значит, что она является моделью его поведения, так что программа не является моделью психологической активности [7, с. 525].

В случае человеческого мышления базой дедукции является внутренний мир (человека). Поэтому, когда программа отвечает нам на вопрос, это является всего лишь в простейшем



случае выбором необходимого ответа, который был ранее занесен в память машины. То, что машина отвечает на вопросы, всего лишь означает, что в нее заложена программа, которая по определенным правилам при помощи условных операторов из подключенной базы знаний выбирает подходящие фразы. Такая машина больше похожа не на человека, а на зомби.

Для того чтобы понять, может ли машина действительно думать, необходимо каким-то образом убедиться именно в наличии внутреннего мира. Современные ученые и философы в связи с этим уже неоднократно предлагали изменить тест Тьюринга. Например, психолог, профессор Нью-Йоркского университета Гари Маркус сказал, что «... ни одна из программ не может смотреть серии “Симпсонов” и говорить, когда смеяться»¹ [8]. Современный американский философ Джон Сирл в 1980 году предложил мысленный эксперимент, названный “Китайская комната”. Его суть заключается в следующем. Представим себе закрытую комнату, в которой находится человек, который не знает китайского языка. Ему передают таблички с китайскими иероглифами. Внутри комнаты есть книга на родном языке человека, в которой написано, какому полученному иероглифу соответствует другой иероглиф, имеющийся в комнате. Человек выбирает по обозначенным правилам правильную табличку и отдает ее. Снаружи у людей, знающих китайский язык, создается впечатление, что человек (или даже система “человек — комната”) знают китайский язык, потому что, например, на иероглиф “любимый цвет” человек выдает иероглиф “синий”. Автор говорит, что такая система прошла бы тест Тьюринга, однако это не значит, что человек понимает смысл такого общения. Манипулирование символами — не значит понимание. Компьютерные программы оперируют синтаксическими конструкциями, разум — семантическими. Одного лишь синтаксиса недостаточно для появления семантики [9, с. 7–8].

Можно привести похожий пример, хорошо знакомый преподавателям. Предположим, на экзамен пришел студент, который совершенно не знает предмет. После того, как он вытянул билет, он садится готовиться, полностью списывает ответ и, читая его неуверенно по бумажке (с пустотой в глазах, возможно, путаясь в словах, делая неправильные ударения в терминах), отвечает. Слова, предложения точь-в-точь повторяют фразы из лекций или учебника, однако любой справедливый преподаватель поставит такому студенту “двойку”.

В 2011 году канадский ученый Гектор Левеск, специализирующийся в области искусственного интеллекта, предложил иной тест, названный “Схема Винограда” (названный так в честь американского инженера и философа Терри Винограда). Суть этого теста заключается в том, что машине предлагаются вопросы, на каждый из которых можно ответить разными

¹ Здесь необходимо сделать ремарку: наверняка найдутся и многие люди, которые не смогут понять, когда нужно смеяться в “Симпсонах”, однако это отнюдь не означает, что эти люди не обладают мышлением. Профессор в данном примере имел в виду появление эмоций как результата внутренней мыслительной деятельности человека.

способами, но только один ответ является правильным. Для того чтобы ответить на эти вопросы, их необходимо понять (представить себе). Человеку сделать это не составит труда, однако компьютерной программе они могут доставить значительные трудности.

Примеры вопросов:

1) Приз не подходит коричневому чемодану, потому что он очень велик. Что велико: а) приз; б) чемодан?

2) Правительство города отказалось дать разрешение демонстрантам на митинг, потому что они боялись насилия. Кто боялся насилия: а) правительство; б) демонстранты? [10].

Первое соревнование по схеме Винограда пройдет в 2015 году. Будет дано 50 вопросов, на которые будет необходимо ответить за 25 минут. Для того чтобы пройти тест, машине будет необходимо ответить правильно на все вопросы. Не факт, что и этот тест через какое-то время будет считаться основным критерием для подтверждения наличия искусственного мышления. Скорее всего, его появление приведет лишь к увеличению объема вводимых знаний при написании программ и к усложнению алгоритмов выбора правильного действия, а не к появлению полноценного искусственного интеллекта, но на сегодняшний день оно кажется более объективным, чем тест Тьюринга, и наверняка поможет ученым и философам приблизиться к пониманию и созданию искусственного интеллекта.

Заключение

Подводя итог, еще раз хочется обратить внимание на то, что мышление все же не сводится только к адекватной реакции на внешние символы и является чем-то несомненно большим. Безусловно, Алан Тьюринг в 1950 году, описывая тест, не мог предположить, каким действительно будет уровень развития систем с искусственным интеллектом (он писал, что к концу XX века, по его мнению, после пяти минут опроса судья будет иметь всего лишь не более 70% шансов на верное определение того, кем или чем является его собеседник) и, что самое главное, каким именно образом будет осуществляться работа в этом направлении. Тест Тьюринга, безусловно, был большим шагом вперед в осмыслении искусственного интеллекта и мышления. Он важен и до сих пор, но только если его использовать в качестве необходимого, но не достаточного условия доказательства наличия мышления. Прохождение теста Тьюринга ни в коем случае не должно являться основной целью создания мыслящих машин.

Как пишет виднейший отечественный исследователь искусственного интеллекта Виктор Константинович Финн, существует три понятия: ум, интеллект, мудрость. Искусственный ум как способность выдвигать цели и достигать их возможен



(с этим сейчас сталкиваются даже старшеклассники, которые, изучая в школе основы программирования, пишут программы, например, для нахождения корней квадратного уравнения). Искусственный интеллект как совокупность способностей, как “ум во многих направлениях” возможен, но только как грубое приближение. Искусственная мудрость как не только интуитивное усмотрение результата рассуждения, но и понимание целей рассуждения с точки зрения духовности, как интеллект с высшей степенью рефлексии пока невозможна [11, с. 114–115]. Возвращаясь после всего сказанного к самому началу статьи, можно подвести итог: даже если на данный момент какая-то машина и проходит тест Тьюринга, то она не становится от этого мыслящим субъектом, равно как и мужчина, обманувший в первоначальном варианте игры-имитации, предложенной Тьюрингом, судью, не становится от этого женщиной.

Литература

1. *Тьюринг А.* Может ли машина мыслить? М.: ГИФМЛ, 1960. 112 с.
2. Turing Test Success Marks Milestone in Computing History // University of Reading. [Электронный ресурс]. Режим доступа: <https://www.reading.ac.uk/news-and-events/releases/PR583836.aspx> (дата обращения 06.02.2015).
3. Они постоянно крадут газ друг у друга // Газета.Ru [Электронный ресурс]. Режим доступа: http://www.gazeta.ru/science/2014/06/11_a_6066225.shtml (дата обращения 06.02.2015).
4. Тест Тьюринга пройден!": рассказ очевидца // Skolkovo Community [Электронный ресурс]. Режим доступа: http://sk.ru/news/b/articles/archive/2014/06/09/test-tyuringa-proyden_2100_-rasskaz-ochevidca.aspx (дата обращения 06.02.2015).
5. Мышление // *Философский словарь* / Под ред. И. Т. Фролова. 7-е изд., перераб. и доп. М.: Республика, 2001. 719 с.
6. *Гончаров И.А.* Обломов. Л.: Наука, 1987. 696 с.
7. *Хант Э.* Искусственный интеллект. М.: Мир, 1978. 560 с.
8. What Comes After the Turing Test? // *The New Yorker*. 2014. June 9. [Электронный ресурс]. Режим доступа: <http://www.newyorker.com/tech/elements/what-comes-after-the-turing-test> (дата обращения 06.02.2015).
9. Сирл Д. Разум мозга — компьютерная программа? // *В мире науки*. 1990. № 3. С. 7–13.
10. Winograd Schema Challenge // *Commonsense Reasoning*. [Электронный ресурс]. Режим доступа: <http://commonsensereasoning.org/winograd.html> (дата обращения 06.02.2015).
11. *Финн В.К.* Интеллектуальные системы и общество: Сб. статей / Пред. Д.А. Поспелова, Д.Г. Лахути, В.Б. Тарасова. Изд. 2-е, испр. и существенно доп. М.: КомКнига, 2006. 352 с.