

UNITED WISDOM

Почему имитация понимания может нас убеждать

Мир и сознание

Andrey Zarinsh & Maria Evilkut

W-10-N03 · RU

@andrey.zarinsh @ilkutki

Почему имитация понимания может нас убеждать

Эпиграф

> *Наше воображение — ещё один, и даже более серьёзный, источник ошибки; ведь если по обстоятельствам мы ожидаем увидеть какое-либо выражение, то легко воображаем его присутствие.*

Чарльз Дарвин, *The Expression of the Emotions in Man and Animals*, Introduction, абзац *Our imagination*, английское издание D. Appleton, 1899. Полное предложение; перевод United Wisdom с английского. Наблюдение о толковании выражений сопоставляется с чтением ответов, но не превращается в измеренную закономерность общения с ИИ.

О чём эта статья

Ответ может быть убедительным ещё до того, как проверены его основания. Собеседник подхватил наше слово, задал подходящий вопрос и вернулся к прежней детали — и мы уже читаем за этими действиями знание ситуации и направленное внимание. Иногда такое прочтение оправдано; иногда сообщение оставляет для него значительно меньше оснований, чем кажется. В этой статье рассматривается вклад самого читателя в ощущение понимания. Программа ELIZA Джозефа Вейценбаума даёт исторический пример ограниченного механизма и более богатой интерпретации разговора. Дарвин показывает трудность наблюдения, когда ожидаемое выражение уже названо. Тьюринг предлагает смотреть на содержательное продолжение ответа. Сопоставление этих подходов помогает разделить полученную фразу, приписанный ей смысл и то, что удалось проверить. Мы не ставим диагноз читателю и не объявляем все современные системы простой ELIZA. Практическая цель — сохранить пользу разговора, одновременно сделать происхождение собственного впечатления видимым.

@andrey.zarinsh @ilkutki

I. Ответ и наша история об ответе появляются вместе

Представим переписку о поездке. Человек сообщает, что ему важно не спешить, а помощник предлагает оставить свободное утро. Это может быть хорошей рекомендацией. Но затем читатель добавляет к результату историю: помощник почувствовал его усталость, понял образ жизни и заботится о возвращении сил. В самой фразе эти дополнительные положения не проверены. Она может соответствовать сообщённому условию без такого широкого знания. Полезность предложения и богатство истории о его авторе поэтому следует оценивать отдельно, даже если в живом чтении они возникли одновременно.

Здесь не требуется немедленно считать впечатление ошибкой. Мы постоянно понимаем сообщения в контексте, который не повторён целиком в каждой реплике. Без этой работы разговор был бы трудным и громоздким. Вопрос возникает там, где наш вклад начинает предъявляться как свойство отвечающего. «Я прочитал это как заботу» описывает мою интерпретацию. «Он обладает определённым переживанием заботы» утверждает больше. Между этими предложениями нужна связь, которую нельзя получить простым повторением приятной фразы или усилением её эмоционального тона.

Различить эти уровни помогает вопрос о происхождении подробностей. Что было прямо сообщено помощнику? Что содержалось в текущем документе? Что он действительно добавил? А что мы достроили из собственного знания ситуации? Если свободное утро уже было разрешено в запросе, рекомендация ещё не показывает самостоятельного открытия жизненной потребности. Она может быть хорошим применением заданного условия. Такое уточнение не делает результат бесполезным: оно определяет, какая именно работа была выполнена и какой части впечатления пока нет независимого подтверждения.

Эта статья рассматривает механизм убедительности как возможность, а не сообщает измеренную частоту человеческих ошибок. Здесь нет нового эксперимента, показывающего, сколько людей поверят той или иной системе. Есть исторический пример и оригинальные ситуации для анализа перехода. Его можно исследовать в собственной работе с текстом, не объявляя себя наивным или патологически доверчивым. Цель — заметить момент, когда содержание ответа перестаёт быть единственным основанием и к нему присоединяется предполагаемая история о знаниях, намерениях и возможном опыте собеседника.

II. ELIZA показывает конкретный механизм, а не объясняет любой ИИ

В статье 1966 года Вейценбаум описывает ELIZA через ключевые слова, разложение входной фразы и её сборку по правилам сценария. Для некоторых случаев предусмотрены общие ответы и возвращение сохранённых преобразований. Сценарий отделён от программы. В обсуждении автор связывает впечатление разговора с предположениями, которые вносит говорящий. Прочитаны введение, соответствующие механизмы и обсуждение в первичной статье; код не запускался, весь технический аппарат и приложение не проверялись. Это ограниченный исторический материал, а не собственная репликация результата.

Разбор устройства важен потому, что объяснение определённого ответа может оказаться значительно уже его восприятия. Допустим, придуманная нами программа возвращает вопрос с частью введённой фразы. Человек видит важную личную подробность в следующей реплике и заключает, что собеседник выделил её смысл. Между тем в этой иллюстрации правило могло выделить участок по формальному условию. Это разные объяснения одного результата. Чтобы выбрать между ними, нужно знать устройство или получить наблюдение, которое действительно их различает, а не просто ещё раз услышать удачный вопрос.

Из исторического примера нельзя вывести, что всякое современное разговорное устройство работает теми же правилами. Такое утверждение потребовало бы изучить соответствующую систему. Нельзя также превратить объяснимый механизм в окончательный ответ на любой вопрос о возможном опыте. Важнее более узкий вывод: знакомая форма реплики может быть совместима с несколькими причинами её появления. Если мы не рассмотрели эти причины, ощущение понятого смысла не выбирает одну из них автоматически. Оно сообщает о восприятии разговора, но не заменяет исследование того, что этот разговор произвело.

ELIZA поэтому полезна как пример несовпадения описаний, а не как универсальное разоблачение. Описание операции может отвечать на вопрос, как получена конкретная строка. Описание впечатления — на вопрос, почему читатель увидел за ней больше. Оценка полезности — на вопрос, помогла ли строка с работой. Наконец, вопрос о сознании требует собственной постановки. Если сохранить эту последовательность, исторический пример помогает мыслить точнее. Если заменить её лозунгом о том, что всё является одной простой имитацией, исследователь снова получает большой вывод без достаточной проверки.

III. Возвращённая подробность не равна независимому знанию

Особенно убедительным может показаться ответ, который использует уже сообщённую подробность. В нашей иллюстрации человек пишет о встрече с коллегой, а помощник позже возвращается к имени коллеги. Читатель вправе заметить связность переписки. Но имя было дано им самим. Его повторение не показывает, что отвечающий знает участника встречи, понимает всю историю отношений или обладает личным интересом к её исходу. Чтобы не потерять полезную связность и не добавить лишнего, следует назвать точный результат: деталь из доступного текста использована в продолжении.

Вейценбаум в обсуждении отделяет основания впечатления от предположений говорящего о фоновых знаниях и рассуждении партнёра. Этот исторический контекст даёт нам вопрос, который можно применить к примеру: откуда читатель знает о наличии предполагаемого знания? Возможно, из прежней проверенной работы собеседника. Возможно, только из того, что ответ звучит уместно. Эти варианты нельзя считать одинаковыми. В первом есть дополнительные обстоятельства; во втором заключение держится на форме сообщения и добавленной нами истории. Различие можно описать, не пытаясь измерить чужое чувство понимания по одной реплике.

Для проверки полезно выбрать содержательную подробность, а не просто увеличить длину беседы. Если помощник вернулся к имени, можно посмотреть, правильно ли он сохранил отношение между участниками и их действиями. Это уже иной вопрос, чем повторение букв. Если эти отношения перепутаны, первоначальное впечатление может потребовать пересмотра. Если они сохранены, появляется основание для более точного утверждения о работе с доступной информацией. И в том и в другом случае не следует называть результат доказательством внутреннего заботливого внимания, которое задача отдельно не исследовала.

Замеченная граница не обязывает относиться к ответу холодно. Можно ценить удачное продолжение, использовать его для подготовки встречи и при этом не путать разные виды знания. Не обязательно выбирать между полным очарованием и тотальным недоверием. Есть конкретная работа, которую ответ выполняет, и конкретная история, которую мы о нём строим. Понимание происхождения подробностей позволяет держать обе вещи в поле зрения. Тогда следующий вопрос направлен на недостающее основание, а не на поиск ещё одного приятного подтверждения уже сложившегося впечатления.

IV. Роль может объяснить реплику лучше, чем скрытое намерение

В историческом обсуждении ELIZA особое значение имеет выбранная роль собеседника: вопрос может читаться как целенаправленный ход, даже когда знание мира не показано. Это описание условий разговора, а не доказательство терапевтической эффективности программы. Мы не предлагаем использовать этот сценарий для лечения или заменять им профессиональную помощь. Для нашей темы важен узкий переход: если роль допускает уточняющий вопрос, читатель может объяснить вопрос через предполагаемую цель и тем самым заполнить то, что о происхождении реплики ещё неизвестно.

Представим редактора, который отвечает на черновик просьбой подробнее объяснить один абзац. Это может быть содержательное замечание, требующее настоящего чтения. Но сама форма просьбы также совместима с общим правилом задавать уточнения. Чтобы оценить вклад, важно, что именно выделено и почему. Если редактор указывает противоречие между соседними утверждениями, появляются более конкретные основания. Если вопрос подходит почти к любому тексту, его уместность ещё не устанавливает глубокого разбора. Пример относится к возможным объяснениям сообщения, а не к неизбежному поведению всех редакторов или систем.

Роль способна помогать полезному взаимодействию. Обозначение задачи делает ожидаемый результат яснее: редактор обсуждает текст, помощник планирует встречу, преподаватель уточняет основание решения. Проблема возникает, когда роль становится заменой свидетельству. Мы начинаем считать любую реплику осмысленной именно потому, что приписали говорящему позицию специалиста. Тогда отсутствие конкретного основания превращается в предполагаемую тонкость замысла. Чтобы не потерять содержание сотрудничества, полезно спросить, какой видимый элемент результата оправдывает эту интерпретацию и чем она отличается от объяснения более общим правилом.

Такая проверка не требует спорить о каждом слове. Она особенно нужна, когда ответ используется для существенного решения. Общий вопрос может помочь подумать, но не следует принимать его за уже выполненный анализ фактов. Если человек хочет именно редактирование, он может попросить назвать место проблемы и предложить объяснимую правку. Если нужна организация встречи, следует проверить ограничения и распределение действий. Роль задаёт цель сотрудничества; конкретный результат показывает, насколько работа этой цели отвечает. Возможный внутренний опыт собеседника остаётся отдельным вопросом.

V. Дарвин предлагает увидеть, что добавляет заранее известное название

Во введении к книге о выражении эмоций Дарвин описывает трудность наблюдения и влияние ожидаемой интерпретации. Он показывал фотографии без объяснения и записывал ответы; затем сопоставлял это с собственным впечатлением, возникшим при чтении сопровождающего текста. Часть выражений распознавалась сходно, часть вызывала разные суждения. Использовано английское издание 1899 года. Это историческое рассуждение и способ наблюдения, а не современный эксперимент с чатами. Его ограничения, включая историческую терминологию и другие теоретические положения книги, не скрываются.

Для нашего вопроса важен не перенос результатов о лице непосредственно на искусственный интеллект, а форма сопоставления. Что мы видим в ответе до того, как назвали его проявлением заботы, мудрости или понимания? Что добавилось после такого названия? Возьмём вымышленную фразу помощника о сохранении свободного времени. Без истории о заботливом собеседнике она может выглядеть как следствие указанного ограничения. С этой историей та же фраза приобретает личный смысл. Изменилось её прочтение; чтобы заявить об изменении свойства системы, нужны другие данные.

Можно провести аккуратное упражнение с собственным текстом: выписать сообщение и отдельно записать интерпретацию. Не нужно заставлять себя полностью забыть контекст или обещать объективность после одного приёма. Нужно лишь сделать добавление видимым. Например, в сообщении названа свободная дата; в интерпретации появилась забота о настроении. После разделения можно спросить, какое обстоятельство связывает второе с первым. Иногда ответ найдётся в другой проверенной информации. Иногда окажется, что связь была только предположена. Оба результата содержательнее уверенности, которая не показывает своего основания.

Этот приём не уничтожает значение сочувствия и воображения в общении. Он ограничивает их роль в доказательстве. Читатель может почувствовать поддержку от текста и честно сообщить о своём впечатлении. Но его чувство не является прямым измерением чувства у автора текста или у породившей его системы. Сравнение с Дарвином помогает удержать эту границу: внимательность к собственному прочтению улучшает наблюдение, тогда как отказ от всякого прочтения сделал бы разговор невозможным. Цель — объяснить вклад ожидания, а не запретить человеческую реакцию.

VI. Продолжение ответа проверяет больше, чем чувство согласия

У Тьюринга в §6 (4) дополнительный вопрос меняет обстоятельство воображаемого экзамена и позволяет рассмотреть дальнейшие ответы. Для нашего угла этот пример полезен как переход от впечатления к материалу для проверки. Мы можем спросить, действительно ли продолжение учитывает условие, которое было важным, или просто сохраняет привычную интонацию. Тайна сознания в тексте Тьюринга остаётся открытой. Поэтому продолжение разговора здесь не объявляется способом непосредственно обнаружить опыт: оно уточняет, какие основания есть для конкретного суждения о работе с вопросом.

Допустим, помощник предложил план спокойного выходного, и человеку понравился тон. Теперь в плане меняется одно существенное условие: свободен только вечер, а предложенное место работает утром. Содержательная корректировка должна учитывать это ограничение. Если ответа хватает только на новое выражение поддержки, первоначальное ощущение понимания не подтверждает правильности плана. Если план исправлен, мы видим подходящую работу с условием. Ни один результат не описывает весь внутренний мир отвечающего. Он позволяет уточнить предмет доверия: что именно помощник смог сделать в данной ситуации.

Здесь важно не превратить проверку в спектакль, где ищут удобный провал. Изменение должно относиться к заявленной способности и быть ясно сообщено. Если задавать противоречивые требования, а затем обвинять собеседника в непонимании, проверка станет плохо объяснимой. Если подсказывать исправление целиком, результат тоже мало различит самостоятельную работу и следование подсказке. Лучше обозначить одно условие и заранее назвать, какая часть ответа должна измениться. Это требование к логике упражнения, а не инструкция испытать сознание или переживания человека.

Продолжение беседы может уменьшить разрыв между впечатлением и обоснованием, если действительно даёт новое различающее наблюдение. Оно может и увеличить разрыв, если каждая последующая любезность воспринимается как ещё одно подтверждение без проверки содержания. Число сообщений само по себе не решает вопрос. Нужно знать, что добавил следующий ответ к оценке конкретного результата. Такая внимательность сохраняет пользу дружелюбного тона, но не назначает ему роль доказательства знаний, намерения и переживания одновременно. Предмет проверки остаётся достаточно узким, чтобы ответ на неё мог что-то изменить.

VII. Полезность сообщения не удостоверяет переживание

Сообщение может помочь человеку сформулировать мысль, даже если он не установил, что понимает его источник. Вымышленный общий вопрос «какое условие здесь важнее?» способен направить внимание к настоящей проблеме задачи. Польза возникает в работе читателя с вопросом. Это заслуживает признания, но не требует приписывать системе чувство внимания или скрытое знание ответа. Нам следует различить результат совместного процесса и утверждение о внутреннем состоянии одной из его сторон. Иначе собственное удачное размышление окажется незаметно предъявлено как свойство собеседника.

Вейценбаумовское замечание о вкладе говорящего помогает назвать эту часть работы. Оно не означает, что всё содержание разговора создано только человеком или что устройство программы всегда несущественно. В нашем примере вопрос был предложен, и это действие имело результат. Но следующий шаг мысли, обнаруженное человеком ограничение и чувство прояснения могут иметь разные основания. Достаточно не складывать их в одно доказательство. Тогда можно объяснить, почему разговор полезен, не делая из пользы автоматического заключения о сознании и не обесценивая активность самого читателя.

Та же дисциплина нужна при отрицательном выводе. Если объяснимый механизм выдаёт полезную реплику, из его объяснимости не следует отсутствие всех возможных свойств, которые мы ещё не исследовали. В данной статье не выбирается теория, позволяющая вывести наличие или отсутствие сознания из описания любой операции. Ставится более узкая граница: представленное впечатление не отвечает на весь вопрос. Нельзя заменить необоснованное признание столь же необоснованным отрицанием и считать задачу решённой. Для каждого вывода нужно назвать собственное основание и допустимый масштаб.

Это позволяет пользоваться помощью без обязательной личной истории о помощнике. Можно проверить план, поправить формулировку, ценить хороший вопрос и при этом сказать: статус возможного опыта здесь не установлен. Можно также честно описать своё ощущение поддержки как собственную реакцию на текст. Такие формулировки не требуют ни холодного языка, ни обещания личной близости со стороны системы. Они удерживают ответственность там, где мы можем её выполнить: в оценке результата и точности собственных утверждений о том, что произошло.

VIII. Проверяйте, что могло бы изменить вашу интерпретацию

Когда впечатление уже сложилось, полезно спросить, какое наблюдение потребовало бы его уточнить. Если никакой результат не способен изменить историю, речь идёт скорее о принятой рамке чтения, чем о проверке. Уместная реплика подтверждает понимание, неуместная объявляется тонкой стратегией, отсутствие ответа — глубиной размышления. В такой иллюстрации все варианты приводят к одному выводу. Чтобы получить исследовательскую ценность, нужно назвать более ограниченную гипотезу и результат, который действительно отличал бы её от других объяснений.

Например, предполагается, что помощник различает ограничения плана, а не просто повторяет последнее пожелание. Тогда интересен случай, где эти два объяснения дают разные рекомендации. Если последняя просьба несовместима с ранее сохранённым условием, ответ должен каким-то образом учитывать конфликт, чтобы поддержать первую гипотезу. Это вымышленная конструкция для оценки способности, а не обещание решить проблему сознания. Она полезна потому, что делает основания впечатления обсуждаемыми. После результата можно яснее сказать, что именно подтверждено, а что было добавлено ожиданием.

Нужно также оставлять возможность, что само различение построено плохо. Система могла не получить прежнее условие, формулировка могла быть двусмысленна, требование могло измениться без объяснения. Если эти обстоятельства не рассмотрены, провал окажется не менее легко воображаемым, чем успех. Поэтому проверять своё впечатление означает проверять и качество вопроса. Добросовестное сомнение не освобождает исследователя от описания доступной информации. Оно предполагает, что вывод изменяется вместе с основаниями, а не только с нашим желанием верить или разоблачать.

Этот подход помогает вести спокойный разговор между людьми с разными впечатлениями. Один увидел внимательного партнёра, другой — знакомый шаблон. Вместо спора об общей разумности можно сравнить факты: какое условие было учтено, что было уже сообщено, какой механизм известен, какая часть интерпретации остаётся предположением. Согласие может быть частичным. Например, оба признают пользу ответа и расходятся в оценке предполагаемого понимания. Такое разделение продвигает обсуждение дальше, чем требование немедленно выбрать одну общую характеристику собеседника на все случаи.

IX. Сохраните происхождение убедительности в итоговом рассказе

После полезной беседы хочется пересказать её простым выводом: «меня поняли». Для личного впечатления это может быть естественное выражение. Но если рассказ используется как свидетельство о системе, стоит добавить, что было наблюдаемо. Помощник сохранил условие, предложил подходящую правку, использовал уже сообщённую подробность. Затем можно отдельно описать, как это было прочитано. Такое различение не запрещает выражать благодарность или удовольствие. Оно позволяет следующему читателю понять, какой вывод принадлежит результату, а какой — его интерпретации и эмоциональному значению.

Сопоставление Дарвина, Вейценбаума и Тьюринга помогает построить этот рассказ из разных частей. Первый направляет внимание к ожиданию наблюдателя. Второй описывает исторический механизм и вклад говорящего. Третий предлагает рассмотреть последовательное продолжение ответа. Эти подходы не составляют единую психологическую теорию и не устанавливают современную распространённость доверия к ИИ. Они дают три вопроса: что мы добавили, что известно об устройстве и что изменилось после содержательного уточнения. Каждый вопрос может иметь ответ своей полноты, без необходимости сгладить различия между источниками.

Важна и граница внешнего поведения. Даже подробно описанная связность сообщений не даёт прямого доступа к субъективному качеству возможного опыта. И наоборот, трудность такого доступа не делает бессмысленной оценку результата. Мы можем улучшать сотрудничество, уточнять задачи и проверять факты, оставаясь осторожными в утверждениях о переживании. В этой статье не сделан вывод о наличии или отсутствии сознания конкретной современной системы. Сохранена возможность исследовать этот вопрос отдельно от того, насколько убедительным оказался один доступный разговор.

Убедительность становится менее загадочной, когда её происхождение описано. Часть приходит из фразы, часть — из контекста, часть — из нашей активной работы над смыслом. Это не математическое разложение и не универсальная причинная модель. Это способ не потерять необходимые различия в собственном выводе. Он оставляет место полезному общению и интересу к чужому уму, одновременно делает ответственность читателя выполнимой. Мы можем признать, что ответ помог, и точно назвать, чего о его авторе эта помощь ещё не позволила узнать.

Что это значит на практике

Выберите одну действительно полезную реплику из рабочего разговора. Запишите сам результат, сведения, которые были доступны помощнику, и вашу интерпретацию. Если в интерпретации появились забота, самостоятельное знание или намерение, отметьте, на каком отдельном обстоятельстве они основаны. Цель — сделать вклад читателя видимым, а не заставить себя отказаться от всякого впечатления.

Проверьте содержание по цели задачи. Для плана сопоставьте ограничения и предложенные действия, для текста — конкретную правку, для утверждения — подходящий источник. Если ответ только вернул сообщённую вами деталь, назовите именно это. Она может быть использована хорошо; её повторение не доказывает весь предполагаемый фон знаний. Поддерживающий тон можно ценить отдельно от достоверности.

Сформулируйте одно уточнение, которое различает два возможных объяснения ответа. Не превращайте беседу в поиск случайной ошибки и не давайте готовое решение в вопросе. После продолжения скажите, какая оценка изменилась и почему. Если условия были неясны, исправьте описание проверки прежде, чем делать широкий вывод. Здесь оценивается работа с задачей, а не ставится домашний тест сознания.

В итоговом рассказе отделите полезность результата, своё ощущение и гипотезу о собеседнике. Исторический пример ELIZA не раскрывает устройство любого современного продукта; упражнение с собственным чтением не измеряет частоту человеческих ошибок. Можно пользоваться помощью и благодарить за результат, сохраняя открытым вопрос о возможном опыте. Это точнее, чем требовать от приятной реплики одновременно подтвердить знания, намерение, близость и сознание.

Ключевая цитата

> *Но, повторю, это вклад говорящего в разговор.*

Joseph Weizenbaum, *ELIZA—A Computer Program for the Study of Natural Language Communication Between Man and Machine*, 1966, Discussion, абзац о предполагаемых фоновых знаниях, проницательности и способности рассуждать у партнёра. Полное предложение; точный английский текст сверён в университетской транскрипции, русский перевод United Wisdom. «Это» относится к приписанным качествам, а не ко всему содержанию разговора.

Связи между традициями

Источник и точное место A	Источник и точное место B	Механизм сопоставления	Граница различия
Вейценбаум, с. 42, Discussion, предположения говорящего	Дарвин, Introduction, абзац *Our imagination*	Сделать вклад ожидаемой интерпретации видимым.	Исторические наблюдения не устанавливают частоту доверия к современному ИИ.
Вейценбаум, с. 37–38, ключевые слова и преобразование	Тьюринг, §6 (4), продолжение экзамена	Различить производство уместной строки и работу с изменённым условием.	ELIZA не является механизмом всех современных систем; разговор не доказывает опыт.
Дарвин, Introduction, фотографии без объяснения, абзацы *Thirdly* и *Several*	Тьюринг, §1–2, письменный канал	Ограничить постороннюю подсказку для определённого вопроса.	Чтение лица и игра с ответами измеряют разные вещи.
Вейценбаум, с. 42, роль и интерпретация вопроса	Тьюринг, §6 (4), воображаемый экзамен	Описать, как цель ситуации меняет прочтение ответа.	Роль не подтверждает эффективность лечения и не устанавливает сознание.
Вейценбаум, с. 41, NONE и MEMORY	Дарвин, Introduction, абзац *The study of Expression*	Отделить доступный сигнал от точного объяснения его значения.	Текстовый механизм и выражение эмоции не приравниваются.
Дарвин, Introduction, ответы без пояснения и предел наблюдения	Тьюринг, §6 (4), заключительное ограничение	Сохранить неопределённость после полезного сопоставления.	Наблюдательское упражнение не решает вопрос о субъективном опыте.

Первоисточники

1. Charles Darwin. *The Expression of the Emotions in Man and Animals*, D. Appleton, New York, 1899, eBook 1227. Прочитаны Introduction: трудности наблюдения, ожидания, фотографии без пояснений и способ записи; также выбранные части I о принципах и рефлексах. Не заявляется чтение всей книги. Используемая редакция не обозначается как текст первого издания 1872 года. Русский эпиграф — перевод United Wisdom.

2. Joseph Weizenbaum. *ELIZA—A Computer Program for the Study of Natural Language Communication Between Man and Machine*, *Communications of the ACM* 9 (1), 1966, с. 36–45. Прочитаны введение, описание сценария и преобразований на с. 37–38, NONE/MEMORY на с. 41 и Discussion на с. 42–43. Полное приложение и код не проверены. Университетская транскрипция использована для дословной сверки предложения. Русская цитата — перевод United Wisdom. Статья не принимается за контролируемое исследование распространённости впечатления или за описание нынешних языковых моделей.

3. A. M. Turing. *Computing Machinery and Intelligence*, *Mind* 59 (236), 1950, с. 433–460. Прочитаны полностью §1–2 и §6 (4); используется письменная проверка и пример продолжения экзамена. Все прочие разделы и названные работы не заявляются как прочитанные. Пример не становится выполненным современным экспериментом.

Теги

понимание, имитация, сознание, субъективный опыт, искусственный интеллект, ELIZA, Джозеф Вейценбаум, Чарльз Дарвин, Алан Тьюринг, разговор, интерпретация, ожидание, воображение, наблюдение, роль, ключевые слова, сценарий, сообщение, фоновые знания, намерение, полезность, проверка, доверие, надёжность, продолжение, условия задачи, сотрудничество, границы вывода

Про авторов

Andrey Zarinsh — автор и основатель проекта United Wisdom. Исследует темы внутренней трансформации, единства традиций, межкультурного диалога и современной духовной дисциплины. Instagram: @andrey.zarinsh.

Maria Evilkut — соавтор United Wisdom. Помимо исследовательской работы, она стремится проверять эти идеи на практике и наблюдать, как принципы единства и взаимопонимания работают в реальной жизни между людьми и культурами. Instagram: @ilkutki.