

ПОСТТЮРИНГОВАЯ МЕТОДОЛОГИЯ: РАЗРУШЕНИЕ СТЕНЫ НА ПУТИ К ОБЩЕМУ ИСКУССТВЕННОМУ ИНТЕЛЛЕКТУ

А. Р. Ефимов

Лаборатория робототехники Сбербанка, Институт философии РАН, Национальный исследовательский технологический университет «МИСиС», Москва, Россия
e-mail: makkawity@gmail.com

Аннотация. В данной статье предпринята попытка комплексной критики теста Тьюринга и выработки критериев качества новых тестов для оценки искусственного интеллекта.

Показано, что предпосылки редуцирования личности и сознания человека к определенным интеллектуальным способностям, использовавшиеся А. Тьюрингом, отражали уровень развития техники того периода. В то же время демонстрируется, что взаимодействие между человеком и машиной можно рассматривать в четырех вариантах: вербальное взаимодействие в физическом мире, невербальное взаимодействие в физическом мире, невербальное взаимодействие в виртуальном мире, вербальное виртуальное взаимодействие. Тест Тьюринга фактически описывает только последний вариант. Чисто вербальное общение между субъектами, подразумевающее мышление только с использованием символьных систем, — это «стена», которая исключает возможность перехода от сложных наблюдаемых явлений к абстрактному образу или понятию. Но в первые десятилетия развития компьютерной техники тест Тьюринга был простым и понятным образцом идеала программирования, соответствовал уровню технических возможностей того времени.

В 2010-х гг. в программировании систем искусственного интеллекта произошли качественные изменения. Во-первых, тест Тьюринга был формально пройден без создания полного искусственного интеллекта. Сегодня машина лучше человека обращается с символьными структурами, однако не обладает возможностью сопоставлять их с реальным миром. Во-вторых, появились новые программные инструменты (нейронные сети глубокого обучения), которые позволяют частично абстрагировать когнитивные процессы (распознавание речи, лица, некоторых образов), но не формулировать новые понятия.

Это дает возможность выдвинуть новые требования к тестовой оценке искусственного интеллекта: он должен поддерживать все формы общения с человеком, абстрагировать образы и конкретизировать понятия, а также участвовать в социальных практиках. Логичным представляется сравнение искусственного интеллекта с возможностями детей разного возраста. При этом антропоморфность робота сама по себе не будет показателем создания искусственного интеллекта.

Ключевые слова: робототехника, искусственный интеллект, Тьюринг, тест Тьюринга, философия искусственного интеллекта, символьные системы, вербальное взаимодействие, нейронные сети.

Для цитирования: Ефимов А. Р. Посттюринговая методология: разрушение стены на пути к общему искусственному интеллекту // Интеллект. Инновации. Инвестиции. — 2020. — № 2. — С. 74–80. DOI: 10.25198/2077-7175-2020-2-74.

POST-TURING METHODOLOGY: DESTRUCTION OF THE WALL ON THE WAY TO GENERAL ARTIFICIAL INTELLIGENCE

A. R. Efimov

Robotics Laboratory of Sberbank, RAS Institute of Philosophy, National Research Technology University «MISIS», Moscow, Russia
e-mail: makkawity@gmail.com

Abstract. This article attempts to comprehensively criticize the «Turing test» and develop quality criteria for new tests for assessing artificial intelligence.

It is shown that the prerequisites for reducing personality and human consciousness, which A. Turing used, reflected the level of development of technology of that period. It is shown that communication between a person and a machine can be considered in four ways: verbal interaction in the physical world, non-verbal interaction in the physical world, non-verbal interaction in the virtual world, and verbal virtual interaction. The Turing test, in fact, describes only the last option. Purely verbal communication between subjects, which involves thinking only using symbolic systems, is a «wall» that excludes the possibility of a transition from complex observable phenom-

ena to an abstract image or concept. But in the first decades of the development of computer technology – the Turing test was a simple, understandable example of the ideal of programming, consistent with the level of technical capabilities.

In the 2010s, there have been qualitative changes in programming.

Firstly, the Turing test was formally passed without creating full artificial intelligence. Today, a machine is better than a person in symbolic systems, but does not have the ability to compare them with the real world. Secondly, new software tools (deep learning neural networks) have appeared that allow one to partially abstract processes (recognize speech, faces, some images), but not formulate new concepts.

This allows us to put forward new requirements for the test evaluation of artificial intelligence: it must support all forms of communication with a person, abstract images and specify concepts, and also have the opportunity to participate in social practices. It seems logical to compare artificial intelligence with the capabilities of children of different ages. At the same time, the anthropomorphism of the robot in itself will not be an indicator of the creation of artificial intelligence.

Keywords: robotics, artificial intelligence, Turing, Turing test, philosophy of artificial intelligence, symbolic systems, verbal interaction, neural networks.

Cite as: Efimov, A. R. (2020) [Post-Turing Methodology: Wall Destruction on the Way to General Artificial Intelligence]. *Intellekt. Innovatsii. Investitsii* [Intellect. Innovations. Investments]. Vol. 2, pp. 74–80. DOI: 10.25198/2077-7175-2020-2-74.

Введение. Возведение «стены» Тьюринга

Британский математик Алан Тьюринг своими работами (1937–1952) заложил фундаментальные направления исследований области науки, которую мы сейчас называем «искусственный интеллект». Опираясь на только что возникшую на тот момент теорию вычислимости и информации и на первые инженерные воплощения машин, предназначенных для универсальных вычислений, А. Тьюринг смело подошел к решению сложнейшего вопроса: «Могут ли машины мыслить?». Естественно, он не мог создать модель, которая бы целиком описывала мышление человека и тем более работу головного мозга как основу мышления: тогда для этого категорически не хватало данных нейробиологии. Поэтому он свел человека как субъекта исключительно к сознанию, а сознание представил в виде системы абстрактных символов. Это упрощение стало основой тезиса А. Тьюринга об изоморфности мышления и вычисления: «Если считать результат труда вычислителей (то есть людей, работающих над вычислениями) интеллектуальным, то почему нельзя сделать сходное предположение относительно машин, которые выполняют эти операции быстрее людей» [15].

Продолжая логическую аналогию, А. Тьюринг предположил «бумажную машину», представляющую собой математика, имеющего лишь бумагу, карандаш и ластик. Такая машина, по мнению Тьюринга, могла бы имитировать различные человеческие функции всего лишь за счет выполнения серии вычислений. В работе «Intelligent Machine» 1948 г. А. Тьюринг предложил «сравнение» человека и «бумажной машины», способных осуществлять лишь символическую коммуникацию, так как находятся в разных комнатах и разделены **стеной**.

Именно в этой работе был заложен фундамент того, что через два года в работе 1950 г. «Computing

machinery and intelligence» будет описано как «игра в имитацию» [14]. Дополнительно, в этой работе впервые проводится анализ возможностей и ограничений «воплощенного искусственного интеллекта» имитировать человека. В частности, Тьюринг указывал, что «отсутствие доступа этого создания к простым человеческим радостям, таким как еда, секс, спорт и многие другие, создаст препятствия для имитации интеллекта» [15].

По замыслу А. Тьюринга, исследователям надлежало сосредоточиться на имитации интеллектуальной деятельности человека в пяти областях:

1. Различные игры, такие как шахматы, крестики-нолики, покер, бридж.
2. Изучение языков.
3. Переводы с одного языка на другой.
4. Криптография.
5. Математика.

Из этих пяти областей Тьюринг считал криптографию наиболее полезной в практическом смысле для искусственного интеллекта [15]. Выделение именно этих областей для исследований сказалось на всей последующей истории развития искусственного интеллекта: постановка сравнительно однородных задач, частично решавшихся уже компьютерами 40-х гг., позволяла добиваться все новых результатов простым наращиванием вычислительных мощностей. Сложилась определенная инерция развития, когда громадные усилия стали тратить на решение узкого спектра задач. Но человеческое мышление и социум оперируют куда более широким спектром «головоломок» [2]. В результате уже созданные программные инструменты применяются в разнообразных прикладных областях, которые сейчас не могут быть целиком математизированы. Это создает завышенные ожидания от использования компьютеров или, напротив, порождает скептицизм.

В своей дальнейшей работе А. Тьюринг показал и обосновал важнейший тезис об эквивалентности мышления и вычисления и положил начало тому, что в дальнейшем стало машинным функционализмом, компьютеризмом и тестовым функционализмом [1, 11]. В своей наиболее часто цитируемой работе [13] А. Тьюринг предложил провести «игру в имитацию», которая явилась инженерным решением проблемы ответа на вопрос «Может ли машина мыслить?». Вместо поиска определений того, что есть интеллект машины или интеллект человека, Тьюринг предложил «слепое» сравнение базовой интеллектуальной способности человека, заключающейся в возможности рассуждать и использовать воображение, с действиями компьютера. «Игра в имитацию» – сравнение способности человека и машины вести диалог на естественном языке – стала основой тьюринговой методологии построения искусственного интеллекта, известной как «тест Тьюринга» [7, 8].

В дальнейшем исследователи искусственного интеллекта (ученые и философы) разрабатывали различные мысленные эксперименты, которые могли бы стать основой более совершенной методологии, нежели тест Тьюринга. К сожалению, в погоне за созданием более адекватного теста исследователи упускали из виду даже целостный образ предложенной Тьюрингом методологии («зависимость от органов чувств»). Это упущение предполагается исправить в настоящей работе.

Методология критического анализа «теста Тьюринга»

Дав такое пространное описание, необходимо указать на основные *методологические сложности* в современной оценке «теста Тьюринга»:

1. Он чрезвычайно разрекламирован, поэтому многие исследователи используют популярный и упрощенный пересказ теста: «За пять минут телефонного разговора нужно понять, говорит с вами машина или человек».

2. Любое научное исследование нуждается в простых и понятных проверках, а все инженерные изделия стремятся проверять именно тестами (простыми проверками). Так как «искусственный интеллект» чаще всего представляют именно в виде компьютерных программ, то тест сводится к общению с программой, что сформировало инерцию восприятия «разумных машин».

Преодолевая эти сложности будут на основе применения диалектической методологии. Диалектический принцип соотношения исторического и логического поможет раскрыть, какие модели эффективного анализа сознания могли применять исследователи во времена А. Тьюринга, а какие используются сегодня. Это позволит понять, каким образом критерии интеллекта вырабатывались, исходя из объема доступных в определенный период

данных. Также будут использоваться принципы абстрагирования и конкретизации, позволяющие редуцировать многообразие свойств объекта к набору простых черт модели.

Если вернуться к тьюринговой методологии, то необходимо обратить внимание на три аспекта, имеющих важность для последующих рассуждений.

Во-первых, все пять исследовательских областей, изначально предложенных Тьюрингом, основаны исключительно на символьных коммуникациях. В самом деле, игры, вербальные коммуникации, переводы, решение математических задач и криптография суть лишь обработка символов.

Во-вторых, методология Тьюринга всегда подразумевала *стену*, разделяющую между собой двух ключевых участников-людей (третий участник-компьютер). Эта стена выполняет роль барьера при трансформации информации между компьютером и человеком.

В-третьих, Тьюринг верил, что проблема состоит «главным образом в программировании», и не рассматривал необходимость ускорения быстрого действия цифровых вычислителей – компьютеров – для решения задачи «игры в имитацию». Установленный им предел памяти в 10^{7-9} бит, необходимый для этой игры, был достигнут уже в конце 80-х гг. прошлого века, однако никаких значимых прорывов в это время не произошло. Более того, именно этот период совпал с так называемой «второй зимой» (1987–1993 гг.) искусственного интеллекта.

Получается, что само по себе наращивание вычислительных мощностей не приводит к способности учитывать подтексты человеческого общения, на необходимость которой указал Тьюринг, отмечая сложность задачи перевода: «...она кажется наиболее зависимой от осуществимости работы органов чувств и движения» [14]. К сожалению, это его замечание было оставлено без внимания многими последующими поколениями исследователей, считавшими достаточным показателем интеллекта лингвистическое поведение и способность играть в различные игры. Но эта методология не могла быть *достаточной* для создания системы, выполняющей все задачи естественного интеллекта.

Методология Тьюринга стала основой огромного семейства самых различных тестов на искусственный интеллект, решающих частные задачи. В данной статье после рассмотрения частных тестов Тьюринга [1] будут выявлены предпосылки для отказа от использования теста Тьюринга в качестве критерия создания полноценного искусственного интеллекта.

Пространство тьюрингоподобных тестов и его ограничения

Более 70 лет прошло с тех пор, как А. Тьюринг высказал свои революционные философские идеи

о возможности создания «мыслящих машин» в основополагающей работе в журнале «Mind» [13]. Несколько поколений математиков, философов, исследователей искусственного интеллекта посвятили множество статей его мысленным экспериментам. В результате создана целая группа тьюрингоподобных тестов, в пространстве которых можно выделить две ортогональных друг другу оси. Назовем эти оси измерениями пространства тьюрингоподобных тестов. Вокруг этих осей можно сгруппировать

все идеи, как высказанные ранее, так и возможные в будущем. Это оси от вербального к невербальному взаимодействию с искусственным интеллектом и от виртуального взаимодействия к физическому.

На рисунке 1 показаны различные тьюрингоподобные тесты, соотнесенные по предложенным выше осям «вербальное-невербальное» и «виртуальное-физическое». Далее предлагается рассмотреть в деталях четыре области данного пространства.

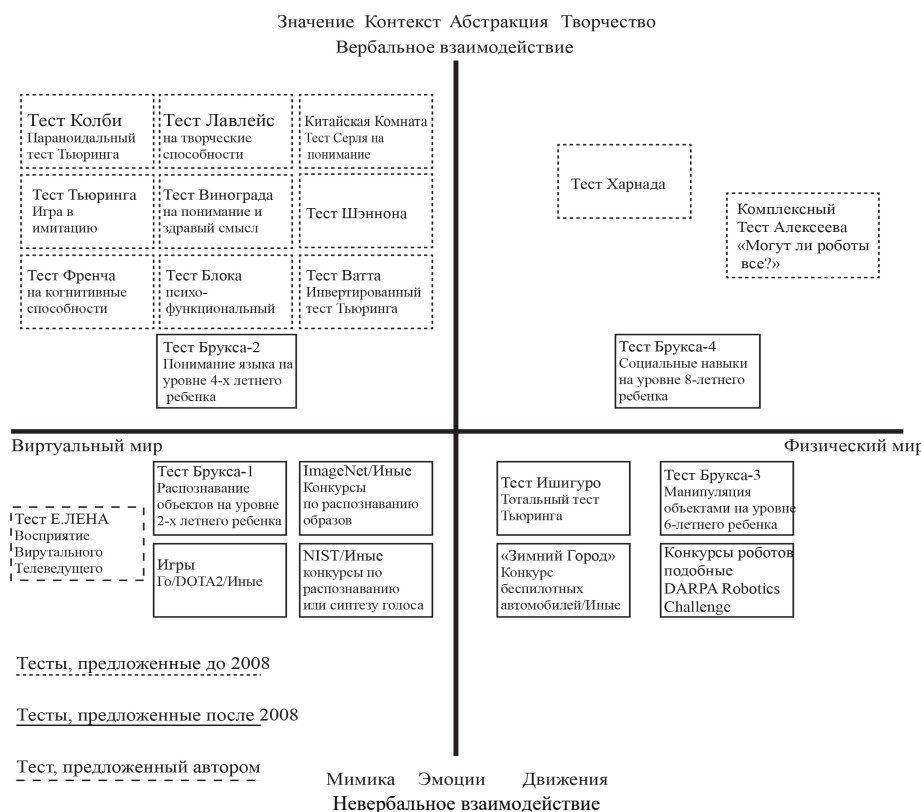


Рисунок 1. Пространство тьюрингоподобных тестов

Вербальное взаимодействие в виртуальном мире. В силу исторических причин большинство тестов (мысленных экспериментов), разработанных до 2008 г., относится именно к этой области. В самом деле, классический тест Тьюринга, тест леди Лавлейс на творческие способности, параноидальный тест Колби, социальный тест Шэннона, тест Серля, психофункциональный тест Блока ориентированы на разного рода проверку вербальных способностей при общении человека и искусственного интеллекта. При этом средой, с которой взаимодействует человек, является виртуальный мир (экран, клавиатура, мышка). Даже если компьютер оценивает предвзятость собеседника и его социальные мотивы, все равно важнейшим основанием ответов выступают каузальные связи в уже созданных системах понятий [4].

Вербальное взаимодействие в физическом мире. Эта область не была популярна среди исследовате-

лей, т.к. с самого начала Тьюринг отверг ее. Лишь С. Харнад [5] и А. Алексеев [1] предложили проведение комплексных тестов, демонстрирующих вербальное взаимодействие человека и искусственного интеллекта в физическом мире.

Невербальное взаимодействие в виртуальном мире. Эта область пространства тьюрингоподобных тестов долгое время оставалась малозамеченной, хотя на ее важность для искусственного интеллекта впервые обратил внимание именно А. Тьюринг. Ярким примером тестов на невербальное взаимодействие в виртуальном мире являются тесты на распознавание образов и на распознавание и синтез человеческой речи. Искусственный интеллект в данном случае никак не изменяет физический мир, никакого смыслового, вербального взаимодействия при этом не происходит даже в случае распознавания речи, поскольку машина лишь верно определяет сами слова, но не их смысл [16].

Невербальное взаимодействие в физическом мире. Эта область является самой сложной, т.к. более всего зависит от уровня развития технологий робототехники и искусственного интеллекта. Если виртуальный мир обладает стандартными характеристиками внешней среды, то реальность неисчерпаема, в ней высока роль случайности, а абстрагирование максимально затруднено. С самого начала эта область игнорировалась исследователями, начиная с А. Тьюринга, хотя ее значимость в коммуникации людей подчеркивается всеми исследователями коммуникаций. Х. Исигуро [6] предлагает проверять уровень технологической зрелости робототехники и искусственного интеллекта за счет сравнения робота-андроида и человека при простом общении: робот говорит исключительно предзаписанные фразы человека, но сам обладает максимальным внешним сходством с человеком.

Такие тесты, как тест Р. Брукса или тест Э. ЛЕНА, направлены на выявление интеллектуальных способностей на основе невербальных взаимодействий. Брукс предложил компетентностный подход, предполагающий, что роботы должны научиться распознавать любые объекты окружающего мира и фразы естественного языка, а также обладать моторикой и социальными навыками хотя бы на уровне ребенка. Тест Э. ЛЕНА основан на разработанной автором, руководителем творческого коллектива, в Лаборатории робототехники Сбербанка в 2019 г. специализированной платформе, позволяющей преобразовать текст в видеоизображение теледиктора, обладающего внешностью и мимикой реальной актрисы.

Рассматривая многообразие тестов, можно, во-первых, заметить, что большинство из них подражает работе лишь в одной области, которая, по мнению исследователей, в наибольшей степени отвечает задаче создания искусственного интеллекта. Во-вторых, стена Тьюринга, отделяющая субъекта теста (человека-судью) от объекта тестирования (компьютера, робота) только продолжала укрепляться.

Эмпирическое выявление неадекватности теста Тьюринга

За последние десять лет произошло два важных события, настолько потрясших стену Тьюринга, что она дала сильную трещину и готова упасть.

Летом 2014 г. Королевское научное общество Великобритании провело тест Тьюринга в точности так, как было предписано в его работе. Тест проводился среди диалоговых интерфейсных программ, иногда называемых чат-ботами. Победителем оказался чат-бот Eugene Gustman, имитировавший личность тринадцатилетнего мальчика из Одессы. Программа сумела обмануть более 30% судей.

Проведенный тест Тьюринга вызвал лавину критики в адрес организаторов этого теста, так как тест показал не то, что искусственный интеллект можно считать созданным, а всего лишь доказал, что мы с легкостью можем обмануть других. Философ А. Сломэн справедливо указал на нерелевантность теста Тьюринга как бихевиористического подхода для оценки степени интеллектуальности какой-либо системы [12]. Формальное прохождение теста Тьюринга не говорит о том, что программа обладает полноценной субъектностью.

Второе событие, повлиявшее на актуальность Тьюринговой методологии: нейронные сети стали основной научной и технологической парадигмой всей области исследований искусственного интеллекта. Предполагается, что наращивание вычислительных мощностей и объемов данных, доступных искусственным нейронным сетям, приведет к прорыву в области искусственного интеллекта. Однако прохождение теста Тьюринга такой интеллектуальной системой демонстрирует лишь способность такой системы быстро находить релевантную реплику, основываясь на обучающей выборке.

Прохождение теста Тьюринга в 2014 г. означает, что стена Тьюринга не является гарантией «слепого» тестирования. Компьютер может имитировать рассуждения человека, не имея вообще никакого понятия о рассуждении, не умея переходить от конкретики к абстракции, идеализировать качества предметов. В такой ситуации стена Тьюринга, отделяющая объект тестирования от субъекта, теряет смысл.

Предлагаемые принципы посттьюринговой методологии исследований искусственного интеллекта

Представляется совершенно логичным становление новой методологии оценки достижений в искусственном интеллекте, учитывающей как опыт последних семидесяти лет, так и новые технологические возможности.

Во-первых, это отказ от антропоморфизма в представлении о разумном компьютере. Стена, воздвигнутая Тьюрингом и отделяющая испытателя и испытуемого (человека или робота), по сути стимулирует человека оценивать искусственный интеллект в сравнении с собой, порождая избыточный технологический антропоморфизм. Создание искусственного интеллекта, рассуждающего подобно человеку, общающегося подобно человеку, возможно, является не самым эффективным способом ответить на вопрос А. Тьюринга о мышлении машин [16].

Во-вторых, можно говорить о разнообразии форм и методов познания, доступных компьютерам. Абстрагирование и конкретизация в самом широком спектре должны использоваться искусст-

венным интеллектом. Тут идеалом становится самостоятельная формулировка новых понятий и моделирование собственной картины мира [10].

В-третьих, необходимо в компьютерных системах реализовать разнообразие всех форм общения, доступных человеку. Общение в символьных структурах компьютерами во многом освоено, но моторика роботов остается несовершенной. «Виртуальное-невербальное», «физическое-невербальное», «физическое-вербальное» взаимодействие пока затруднены. Вероятно, идеал, к которому должны стремиться машины, – это эмоционально окрашенное общение с помощью «пяти чувств».

В-четвертых, искусственный интеллект должен участвовать в социальных практиках человека как младший, но наделенный субъектностью партнер. Компьютерная система с искусственным интеллектом в таких отношениях может выступать в качестве своеобразного помощника, выполняющего вспомогательные функции в человеческой деятель-

ности, основываясь на «считывании» потребностей человека.

Заключение

«Тест Тьюринга» фактически утратил релевантность: компьютерные программы, не являющиеся полноценными искусственными интеллектами, могут его проходить в символьных системах коммуникаций. При этом абстрагирование освоено программами лишь в минимальных формах, что ограничивает их познавательные возможности.

Переход к посттьюринговому подходу в оценке искусственного интеллекта позволит нам сосредоточиться на создании систем, обладающих возможностями демонстрировать различные навыки в четырех основных областях: формирование облика системы под трудовые операции, полноценная формулировка новых понятий и их использование, общение с человеком на основе всех пяти чувств и, наконец, социальная субъектность [3].

Литература

1. Алексеев А. Ю. Комплексный тест Тьюринга: философско-методологические и социо-культурные аспекты. – М.: ИнтелЛ, 2013. – 304 с.
2. Кузнецов О. П. Ограниченная рациональность и принятие решений // Искусственный интеллект и принятие решений. – 2019. – № 1. – С. 3–15.
3. Adams Sam S., Banavar G., Campbell M. I-athlon: Toward a Multidimensional Turing Test // AI Magazine. – 2016. – Vol. 37. – pp. 78–84.
4. Baumeister D., Neugebauer D., Rothe Jörg Schadrack H. Verification in incomplete argumentation frameworks // Artificial Intelligence. – 2018. – Vol. 264. – pp. 1–26.
5. Harnad S. Minds, Machines and Turing: The Indistinguishability of Indistinguishables // Journal of Logic, Language and Information. – 2000. – Vol. 9. – pp. 425–445.
6. Ishiguro H. Android Science. Toward a New Cross-Interdisciplinary Framework. // Robotics Research. – 2007. – Vol. 28. – pp. 118–127.
7. Jarrold W., Yeh P.Z. The Social-Emotional Turing Challenge // AI Magazine. – 2016. – Vol. 37 (1). – pp. 31–38.
8. Kitano H. Artificial Intelligence to Win the Nobel Prize and Beyond: Creating the Engine for Scientific Discovery // AI Magazine. – 2016. – Vol. 37 (1) – Pp. 39–49.
9. Müller V. C., Bostrom N. Future progress in artificial intelligence: A Survey of Expert Opinion // Müller V.C. (ed.) Fundamental Issues of Artificial Intelligence. – Berlin: Springer (Synthese Library), 2016. – Pp. 555–572.
10. Pease A., Lawrence J., Budzynska K., Corneli J., Reed C. Lakatos-style collaborative mathematics through dialectical, structured and abstract argumentation // Artificial Intelligence. – 2017. – Vol. 246. – Pp. 181–219.
11. Putnam H. Minds and machines // Hook S (Ed.) Dimensions of mind. – London: Collier-Macmillan, 1960. – Pp. 138–164.
12. Sloman A. Judging Chatbots at Turing Test 2014 [Электронный ресурс]. – Режим доступа: <https://www.cs.bham.ac.uk/research/projects/cogaff/misc/turing-test-2014.html> (дата обращения: 10.02.2020).
13. Turing A. Computing machinery and intelligence // Mind. – 1950. – Vol. 59. – Pp. 433–460.
14. Turing A. Digital Computers Applied to Games // Bowden B.V. (ed.) Faster Than Thought (A Symposium on Digital Computing Machines). – London: Sir Isaac Pitman & Sons Ltd., 1953. – Pp. 286–310.
15. Turing A.M. «Intelligent Machinery». National Physical Laboratory Report (1948) // Meltzer B., Michie D. (eds) Machine Intelligence 5. – Edinburgh: Edinburgh University Press, 1969. – Pp. 3–23.
16. Zitnick C. Lawrence [et al.] Measuring Machine Intelligence Through Visual Question Answering // AI Magazine. – 2016. – Vol. 3. – Pp. 63–72.

References

1. Alekseev, A. Yu. (2013) *Kompleksnyy test T'yuringa: filosofsko-metodologicheskiei socio-kul'turnye aspekty* [Comprehensive Turing test: philosophical, methodological and socio-cultural aspects]. Moscow: Intell, 304 p.

2. Kuznetsov, O. P. (2019) [Limited rationality and decision making]. *Iskusstvennyj intellekt i prinyatie reshenij* [Artificial intelligence and decision making]. Vol. 1, pp. 3–15. (In Russ.).
3. Adams Sam S., Banavar G., Campbell M. (2016) I-athlon: Toward a Multidimensional Turing Test. *AI Magazine*. Vol. 37., pp. 78–84. (In Eng.).
4. Baumeister D., Neugebauer D., Rothe Jörg Schadrack H. (2018) Verification in incomplete argumentation frameworks. *Artificial Intelligence*. Vol. 264, pp. 1–26. (In Eng.).
5. Harnad, S. (2000) Minds, Machines and Turing: The Indistinguishability of Indistinguishables. *Journal of Logic, Language and Information*. Vol. 9, pp. 425–445. (In Eng.).
6. Ishiguro, H. (2007) Android Science. Toward a New Cross-Interdisciplinary Framework. *Robotics Research*. Vol. 28, pp. 118–127. (In Eng.).
7. Jarrold, W., Yeh, P. Z. (2016) The Social-Emotional Turing Challenge. *AI Magazine*. Vol. 37 (1), pp. 31–38. (In Eng.).
8. Kitano, H. (2016) Artificial Intelligence to Win the Nobel Prize and Beyond: Creating the Engine for Scientific Discovery. *AI Magazine*. Vol. 37 (1), pp. 39–49. (In Eng.).
9. Müller, V. C., Bostrom, N. (2016) Future progress in artificial intelligence: A Survey of Expert Opinion. In: Müller V. C. (ed.) *Fundamental Issues of Artificial Intelligence*. Berlin: Springer (Synthese Library), pp. 555–572 (In Eng.).
10. Pease, A., Lawrence, J., Budzynska, K., Corneli, J., Reed, C. (2017) Lakatos-style collaborative mathematics through dialectical, structured and abstract argumentation. *Artificial Intelligence*. Vol. 246, pp. 181–219. (In Eng.).
11. Putnam, H. (1960) Minds and machines. In: Hook S. (Ed.) *Dimensions of mind*. London: Collier-Macmillan, pp. 138–164. (In Eng.).
12. Sloman, A. Judging Chatbots at Turing Test 2014. Available at: <https://www.cs.bham.ac.uk/research/projects/cogaff/misc/turing-test-2014>. (accessed 10.02.2020) (In Eng.).
13. Turing, A. (1950) Computing machinery and intelligence. *Mind*. Vol. 59, pp. 433–460. (In Eng.).
14. Turing, A. (1953) Digital Computers Applied to Games. In: Bowden B. V. (ed.) *Faster Than Thought (A Symposium on Digital Computing Machines)*. London: Sir Isaac Pitman & Sons Ltd., pp. 286–310. (In Eng.).
15. Turing, A. M. (1948) “Intelligent Machinery”. National Physical Laboratory Report. In: Meltzer B., Michie D. (eds) *Machine Intelligence 5*. Edinburgh: Edinburgh University Press, 1969, pp. 3–23.
16. Zitnick, C. (2016) Lawrence [et al.] Measuring Machine Intelligence Through Visual Question Answering. *AI Magazine*. Vol. 3, pp. 63–72.

Информация об авторе:

Альберт Рувимович Ефимов, руководитель Лаборатории робототехники Сбербанка, соискатель Института философии РАН, старший преподаватель Национального исследовательского технологического университета «МИСиС», Москва, Россия

SPIN-код: 2880-6351, **Author ID:** 7638791, **ORCID ID:** 0000-0001-6857-8659
e-mail: makkawity@gmail.com

Статья поступила в редакцию: 11.02.2020; принята в печать: 28.02.2020.

Автор прочитал и одобрил окончательный вариант рукописи.

Information about author:

Albert Ruvimovich Efimov, Head of Sberbank Robotics Laboratory, PhD candidate, Institute of Philosophy RAS, senior lecturer NRTU «MISiS», Moscow, Russia

SPIN-код: 2880-6351, **Author ID:** 7638791, **ORCID ID:** 0000-0001-6857-8659
e-mail: makkawity@gmail.com

The paper was submitted: 11.02.2020.

Accepted for publication: 28.02.2020.

The author has read and approved the final manuscript.